APPLICATIONS OF NONPARAMETRIC STATISTICS TO

MULTICOMPONENT SOLIDS MIXING

by

JUI-RZE TOO

B. S. in Chemical Engineering, National Taiwan University, 1972

M. S. in Chemical Engineering, Kansas State University, 1978

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1979

Approved by:

R. M. Rubison

Major Professor

TABLE OF CONTENTS

I. INTRODUCTION

Solids mixing or blending is an operation by which two or more particulate solid materials are scattered randomly in a mixer by the random movement of the particles. Solids mixing tends to eliminate existing inhomogeneities, or to reduce gradients. Although it is one of the oldest industrial operations, it is still one of the most widely employed. For example, it is essential in plastic processing, ore smelting, pharmaceutical preparation, fertilizer production, food manufacture, and catalytic synthesis of chemicals. Generally, solids mixing operations are often multicomponent in nature in that each of such operations involves blending of more than two ingredients. Thus, the study of solids mixing for multicomponent mixtures is of practical importance.

Statistical analysis has been a major tool in solids mixing investigations because of the stochastic nature of mixing processes. The statistical properties of a multicomponent heterogeneous solids mixture have been of intense interest to researchers in the field of solids mixing, yet a systematic approach to this problem is still lacking.

The theory of nonparametric methods is essentially concerned with the development of statistical inference procedures without the explicit assumptions regarding the functional form of the probability distribution of the sample observations. Since the distributions of the components during mixing are usually unknown, nonparametric statistical methods should provide a class of appropriate and effective techniques for the analysis of mixing systems. The applications of certain nonparametric tests for solids mixing for binary mixtures has been previously demonstrated by Lai, Wang and Fan [1].

The object of this study is to demonstrate the applicability of non-parametric statistics to the analysis of mixing processes of multicomponent mixtures and the characterization of such mixtures. As specific examples, the mixing processes carried out in a drum mixer and several mixtures generated by the processes are considered.

II. NONPARAMETRIC STATISTICS

In most statistical problems, a class of distributions or states of nature assumed as possible models is defined by a probability density function of given form, which depends on a finite number of real parameters. In other words, if the basic distribution is known, one may be able to derive optimal tests of hypotheses and confidence intervals based on the distribution. In many case an experimenter does not know the form of the basic distribution and is in need of statistical techniques which are applicable regardless of the form of the density. These techniques are called nonparametric or distribution-free methods.

The term "distribution free" refers to the fact that no assumptions are made about the underlying distribution except that the distribution function being sampled is absolutely continuous or purely discrete. The term "nonparametric" refers to the fact that there are no paramenters involved in the traditional sense. The restriction to absolutely continuous distribution function is a simplifying assumption that allows us to use the fact that ties occur with probability 0. They apply to very wide families of distributions rather than only to families specified by a particular functional form.

In nonparametric statistics, the measurement scale need not be numerical. Usually measurements can be classified as one of four levels depending on the precision represented by the measurement procedure. They are

(1). Nominal scale: In nominal (scale) measurements no physical meaning is attached to the values of the numbers. We simply assign numerical names to the types of outcomes, however the principle of order in real number system is not relevant. Of the four measurement scales, nominal is the least precise.

(2). Ordinal or ranking scale: When measurements are made on an ordinal scale, the elements can be arranged in a meaningful order, which corresponds to their relative positons or sizes. In a taste test for five different brands of beer, the tasters may rank beers as 1, 2, 3, 4, and 5 according to their preferences. Note that the rank does not indicate how much better one beer is preferred.

(3). Interval scale: When the elements can be ordered and the arithmetic difference between the elements is meaningful, the data are measured on an interval scale. Thus, we can say not only that one element is larger than or smaller than another, but also by how much. This scale of measurement is much more informative than either of the scales above, since the fact that the distance between elements can be determined implies that there is a fixed unit of measurement and a zero point, the latter being arbitrary. Thus, interval scale data are quantitative in the sense, that the numbers have a true meaning.

(4). Ratio scale: For ratio scale measurements we have not only the order property, a unit of measurement and a meaningful arithmetic difference between elements, but also a fixed origin or zero point as opposed to an arbitrary origin. The term "ratio scale" is used because the ratio of two measurements on this highest scale is meaningful.

The validity of the nonparametric statistical inference does not rest on a specific probability model in the population. Nonparametric procedures exist for data from all four scales of measurements. Such procedures are very useful in many different areas of application. Nonparametric methods can be applied to test a variety of hypotheses. According to types of inferences, the major nonparametric statistical tests are summarized below:

1. Goodness-of-Fit Tests

   Chi-square test, Kolmogorov-Smirnov test

2. Tests of Location or Central Tendency

   Sign test, Wilcoxon signed rank test, Mann-Whitney-Wilcoxon test, Normal scores Van der Waerden test, Kruskal-Wallis test, Friedman test.

3. Tests of Scale or Dispersion

   Siegel-Tukey test, Klotz test, Ansari-Bradley test, Mood test.

4. General Distribution Tests

   Equality of k proportions test, Chi-square test, Kolmogorov-Smirnov test.

5. Association Analysis

   Spearman test, Kendall Tau test, Chi-square test.

6. Tests for Randomness or Trend

   Ordinary runs test, Runs up and down test.

Table 1 gives some available nonparametric tests depending on types of samples obtained and types of measurements involved. For further details, we may be referred to some nonparametric statistics texts such as Gibbons [2,3], Conover [4], Lehmann [5], Hájek [6] or Hollander and Wolfe [7]. Note that we may be referred to Puri and Sen [8] for nonparametric multivariate methods.

Besides the advantage of robustness against distributional assumptions, nonparametric statistical methods often involve less computational work, and therefore, are easier and quicker to apply than other statistical methods. Another advantage of nonparametric statistical techniques is that much of their theory may be developed rigorously using elementary combinatorial mathematics.

THIS BOOK CONTAINS NUMEROUS PAGES WITH DIAGRAMS THAT ARE CROOKED COMPARED TO THE REST OF THE INFORMATION ON THE PAGE.

THIS IS AS RECEIVED FROM CUSTOMER.

Table 1. A brief chart for testing hypotheses and the appropriate nonparametric test

		Type of measurement involved		
Type of sample obtained	Hypothesis test involving...	Nominal (observations may be separated according to categories)	Ordinal* (observations may be arranged from smallest to largest)	Interval** (the numerical value of the observation has meaning)
One random sample $X_1, X_2, \ldots, X_n$	Means (medians)	Binomial test	Quantile test	Wilcoxon test
	(Conf. int. for means)	Conf. int. for p	Conf. int. for x_p	Conf. int. for mean
	Nonrandomness	Wald-Wolfowitz test	Cox and Stuart test Spearman's rho	
	Goodness-of-fit	Chi-square test	Kolmogorov test Cramer-von Mises test Lilliefors test	
	Cond. band for F(x)		Cond. band for F(x)	
Paired observations, or two matched samples $(X_1, Y_1), \ldots, (X_n, Y_n)$	Means (medians)	McNemar test	Sign test	Wilcoxon test Randomization test
	(Conf. int. for diff between means)	Conf. int. for p	Conf. int. for x_p	Conf. int. for diff.
	Independence	Chi-square test Fisher's test	Sign test Bell-Doksum test Olmstead-Tukey test	

Table 1---continued

Multivariate observations, or the randomized complete block design	Means	Cochran test	Friedman test Bell-Doksum test	
Two random samples, $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_m$. (see also tests for several random samples)	Means (medians) (Conf. int. for diff. between means) Variances Identical distributions	Chi-square test	Mann-Whitney test Tukey's quick test Mann-Whitney Conf. int. Tukey's conf. int. Siegel-Tukey test Smirnov test Cramer-von Mises test Wald-Wolfwitz test	Randomization test
Several random samples	Means (medians) Identical distributions	Chi-quare test	Median test Kruskal-Wallis test Bell-Doksum test Slippage test Birnbaum-Hall test Smirnov test	
Other types	Means (medians)	Many-way contingency table	Median test extended Durbin test for BIBD	

*The methods listed under Nominal may also be used here.

**The methods listed under Nominal and Ordinal may also be used here.

III. THEORY

Consider a mixture which has (m+1) components. For the trivial case (m=1), the mixture is called binary. In this paper, we are particularly interested in the nontrivial case where m>1. Let X_{ij} be a random variable denoting the weight fraction of the i-th component in the j-th spot sample ($i = 1, 2, \ldots, m+1$; $j = 1, 2, \ldots, n$). Since

$$\sum_{i=1}^{m+1} X_{ij} = 1,$$

only m of (m+1) weight fractions need to be determined. Thus,

$$\underline{X}_j = [X_{1j}\ X_{2j}\ \cdot\ \cdot\ \cdot\ X_{mj}]', \qquad j = 1, 2, \ldots, n \tag{1}$$

will denote an arbitrary selection of m weight fractions for a given sample. Several nonparametric statistical methods, which can be applied in analyzing a variety of sampling results of multicomponent solids mixing, are presented in this section.

1. One-Sample Location Problem

In many mixing problems, the true component proportions in a mixture are known. The problem of interest then becomes a test of the sampling procedure. If sampling is random throughout the mixture, the sample mean vector should be representative of the population and the sample mean vector should not be significantly different from the specified component proportions. Multivariate rank tests for the one-sample location problem [8] are thus appropriate for a test of the sampling procedure.

Suppose that n spot samples are taken from a mixture. Let $\underline{X}_j$ ($j = 1, 2, \ldots, n$) be a random sample (vector-valued) with a continuous cumulative distribution $F(\underline{x})$, $\underline{x} \varepsilon R^m$, where R^m is the set of all m-tuples $\underline{x} = [x_1\ x_2 \ldots x_m]'$. $F(\underline{x})$ may be written as

$$F(\underline{x}) = F(\underline{x}, \underline{\mu}) \tag{2}$$

where $\underline{\mu} = [\mu_1\ \mu_2\ \ldots\ \mu_m]'$ is a location (vector) parameter. The random (m×n) matrix takes the form

$$[\underline{X}_1\ \ \underline{X}_2\ \ \cdots\ \ \underline{X}_n]$$

$$= \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m1} & X_{m2} & \cdots & X_{mn} \end{bmatrix} \tag{3}$$

We now derive a test of the null hypothesis

$$H_0:\ \underline{\mu} = \underline{\mu}_0 \tag{4}$$

against the alternative hypothesis

$$H_1:\ \underline{\mu} \neq \underline{\mu}_0$$

where $\underline{\mu}_0 = [\mu_{10}\ \mu_{20}\ \cdots\ \mu_{m0}]'$ is a specified vector. Let

$$Y_{ij} = X_{ij} - \mu_{i0} \tag{5}$$

denote the adjusted sample values, then the random sample matrix becomes

$$\underline{Z}_n = \begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1n} \\ Y_{21} & Y_{22} & \cdots & Y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{m1} & Y_{m2} & \cdots & Y_{mn} \end{bmatrix} \tag{6}$$

Ranking the n elements in each row of $\underline{Z}_n$ in increasing order of their absolute value, we obtain an (m×n) rank matrix

$$\underline{R}_n = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1n} \\ R_{21} & R_{22} & \cdots & R_{2n} \\ \vdots & \vdots & & \vdots \\ R_{m1} & R_{m2} & \cdots & R_{mn} \end{bmatrix} \tag{7}$$

where R_{ij} is the rank of $|Y_{ij}|$ $(j = 1, 2, \ldots, n)$ among the set $\{|Y_{i1}|, |Y_{i2}|, \ldots, |Y_{in}|\}$. Since the populations are assumed continuous, the probability of a tie is zero. Then for each i $(i = 1, 2, \ldots, m)$, we replace the ranks $1, 2, \ldots, n$ in the i-th row of $\underline{R}_n$ by a set of general scores denoted by

$$\{E_j^{(i)}, \quad j = 1, 2, \ldots, n\}$$

Hence, we obtain an $(m \times n)$ matrix of general scores $\underline{E}_n$ corresponding to $\underline{R}_n$:

$$\underline{E}_n = \begin{bmatrix} E_{R_{11}}^{(1)} & E_{R_{12}}^{(1)} & \cdots & E_{R_{1n}}^{(1)} \\ E_{R_{21}}^{(2)} & E_{R_{22}}^{(2)} & \cdots & E_{R_{2n}}^{(2)} \\ \vdots & \vdots & & \vdots \\ E_{R_{m1}}^{(m)} & E_{R_{m2}}^{(m)} & \cdots & E_{R_{mn}}^{(m)} \end{bmatrix} \tag{8}$$

We refer the reader to Chapter 4 of Puri and Sen [8] for a detailed discussion of the regularity conditions on a score generating function which determines the constants $E_j^{(i)}$.

We now consider a univariate rank order ststistic for each coordinate (component) of the form

$$T^{(i)} = \sum_{j=1}^{n} E_{R_{ij}}^{(i)} C_{ij} \tag{9}$$

where $C_{ij} = 1$ or -1 according as $Y_{ij} > 0$ or $Y_{ij} < 0$, respectively. Therefore, $T^{(i)}$ is the difference of the sum of the scores $E_j^{(i)}$ for which $Y_{ij} > 0$ and the sum of those for which $Y_{ij} < 0$.

Let

$$\underline{T} = [T^{(1)} \; T^{(2)} \; . \; . \; . \; T^{(m)}] \tag{10}$$

Under the null hypothesis, eqn. (4), the mean and dispersion matrix of $\underline{T}$ are

$$E[\underline{T}] = \underline{0} \tag{11}$$

and

$$E[\underline{T}'\underline{T}] = n\underline{V} = n[v_{i\ell}]_{m\times m} \tag{12}$$

where

$$v_{i\ell} = \frac{1}{n} \sum_{j=1}^{n} E_{R_{ij}}^{(i)} \, E_{R_{\ell j}}^{(\ell)} \, C_{ij} \, C_{\ell j}, \tag{13}$$

$$i, \ell = 1, 2, \ldots, m$$

Note that

$$v_{ii} = \frac{1}{n} \sum_{j=1}^{n} [E_{R_{ij}}^{(i)}]^2 , \qquad i = 1, 2, \ldots, m \tag{14}$$

The test statistic, S, formed by

$$S = \frac{1}{n} [\underline{T} \, \underline{V}^{-1} \, T'] \tag{15}$$

is asymptotically distributed as a chi-square random variable with m degrees of freedom for large samples. If $\underline{T}$ is stochastically different from $\underline{0}$, S will be large which will lead us to reject the null hypothesis. The appropriate P-value [see Appendix A] is the probability that a chi-square variable is greater than or equal to the observed value of S, that is, a right tail probability.

Two special cases are considered in this paper for the one-sample location problem. First, setting

$$E_j^{(i)} = 1, \qquad i = 1, 2, \ldots, m; \qquad j = 1, 2, \ldots, n \tag{16}$$

the test statistic, S, is the multivariate sign test. Second, if

$$E_j^{(i)} = \frac{j}{n+1}, \qquad i = 1, 2, \ldots, m; \qquad 1 \leq j \leq n \tag{17}$$

then S reduces to the multivariate generalization of the one sample Wilcoxon signed rank statistic.

2. Test of Homogeneity of Dispersion Matrices

The degree of mixedness is used to judge the difference between various experimental situations (treatments). For example, we may be interested in comparing different types of mixers, mixing speeds or mixing times. The covariance (dispersion) matrix characterizes the degree of dispersion of each component proportion. Therefore in testing for the homogeneity (or equality) of several dispersion matrices, we may be able to judge whether their degrees of mixedness are significantly different. In other words, we may determine if the variation in composition among spot samples of each treatment is identical. A multivariate nonparametric test of the equality of dispersion matrices discussed in [8] is used to assess treatment effects.

Suppose that we wish to compare the effects among t treatments. All t treatments are assumed to be mutually independent. The number of spot samples for the treatment k is denoted by n_k. Let $X_{ij}^{(k)}$ be a random variable representing the weight fraction of the i-th component in the j-th sample for the treatment k. Also let

$$\underline{X}_j^{(k)} = [X_{1j}^{(k)} \; X_{2j}^{(k)} \; \ldots \; X_{mj}^{(k)}]', \qquad j = 1, 2, \ldots, n_k \tag{18}$$

be n_k independent and identically distributed (vector-valued) random variables having a m-variate absolutely continuous cumulative distribution function (c.d.f.) $F^{(k)}(\underline{x})$ for k = 1, 2, ..., t.

Assuming the identity of locations, we test the hypothesis of the equality of dispersion matrices, i.e.

$$H_0: \underline{\Sigma}^{(1)} = \underline{\Sigma}^{(2)} = \ldots = \underline{\Sigma}^{(t)} \tag{20}$$

against the alternative hypothesis

$$H_1: \underline{\Sigma}^{(1)}, \underline{\Sigma}^{(2)}, \ldots, \underline{\Sigma}^{(t)} \text{ are not all identical.}$$

Let

$$N = \sum_{k=1}^{t} n_k$$

denote the total number of observations

$$\underline{X}_j^{(k)}, \qquad j = 1, 2, \ldots, n_k; \qquad k = 1, 2, \ldots, t$$

and define

$$\underline{Z}_N = \begin{bmatrix} X_{11}^{(1)} & X_{12}^{(1)} & \ldots & X_{1n_1}^{(1)} & X_{11}^{(2)} & X_{12}^{(2)} & \ldots & X_{1n_2}^{(2)} & \ldots & X_{11}^{(t)} & X_{12}^{(t)} & \ldots & X_{1n_t}^{(t)} \\ X_{21}^{(1)} & X_{22}^{(1)} & \ldots & X_{2n_1}^{(1)} & X_{21}^{(2)} & X_{22}^{(2)} & \ldots & X_{2n_2}^{(2)} & \ldots & X_{21}^{(t)} & X_{22}^{(t)} & \ldots & X_{2n_t}^{(t)} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ X_{m1}^{(1)} & X_{m2}^{(1)} & \cdot & X_{mn_1}^{(1)} & X_{m1}^{(2)} & X_{m2}^{(2)} & \ldots & X_{mn_2}^{(2)} & \ldots & X_{m1}^{(t)} & X_{m2}^{(t)} & \ldots & X_{mn_t}^{(t)} \end{bmatrix} \tag{21}$$

a random matrix of dimension $(m \times N)$. Ranking the N elements in each row of $\underline{Z}_N$ in increasing order of magnitude, we obtain an $(m \times N)$ rank matrix

$$\underline{R}_N = \begin{bmatrix} R_{11}^{(1)} & R_{12}^{(1)} & \ldots & R_{1n_1}^{(1)} & R_{11}^{(2)} & R_{12}^{(2)} & \ldots & R_{1n_2}^{(2)} & \ldots & R_{11}^{(t)} & R_{12}^{(t)} & \ldots & R_{1n_t}^{(t)} \\ R_{21}^{(1)} & R_{22}^{(1)} & \ldots & R_{2n_1}^{(1)} & R_{21}^{(2)} & R_{22}^{(2)} & \ldots & R_{2n_2}^{(2)} & \ldots & R_{21}^{(t)} & R_{22}^{(t)} & \ldots & R_{2n_t}^{(t)} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ R_{m1}^{(1)} & R_{m2}^{(1)} & \ldots & R_{mn_1}^{(1)} & R_{m1}^{(2)} & R_{m2}^{(2)} & \ldots & R_{mn_2}^{(2)} & \ldots & R_{m1}^{(t)} & R_{m2}^{(t)} & \ldots & R_{mn_t}^{(t)} \end{bmatrix} \tag{22}$$

where the possibility of ties may be ignored in probability by virtue of the continuity of the c.d.f.'s. For each i (i = 1, 2, ..., m), we replace the ranks in the i-th row of $\underline{R}_N$ by a set of general scores

$$E_j^{(i)} = \sqrt{12}\left(\frac{j}{N+1} - \frac{1}{2}\right), \qquad j = 1, 2, \ldots, N \tag{23}$$

and obtain the corresponding score matrix

(see the following page) (24)

Now, let us define

$$U_{i\ell}^{(k)} = \frac{1}{n_k - 1}\left\{\sum_{j=1}^{n_k} E_{R_{ij}^{(k)}}^{(i)} E_{R_{\ell j}^{(k)}}^{(\ell)} - \frac{1}{n_k}\left[\sum_{j=1}^{n_k} E_{R_{ij}^{(k)}}^{(i)}\right]\left[\sum_{j=1}^{n_k} E_{R_{\ell j}^{(k)}}^{(\ell)}\right]\right\} \tag{25}$$

$$i \leq \ell = 1, 2, \ldots, m; \qquad k = 1, 2, \ldots, t$$

and

$$U_{i\ell}^{*} = \frac{1}{N-1}\left\{\sum_{k=1}^{t}\sum_{j=1}^{n_k} E_{R_{ij}^{(k)}}^{(i)} E_{R_{\ell j}^{(k)}}^{(\ell)} - N\,\bar{E}^{(i)}\bar{E}^{(\ell)}\right\} \tag{26}$$

where

$$\bar{E}^{(i)} = \frac{1}{N}\sum_{k=1}^{t}\sum_{j=1}^{n_k} E_{R_{ij}^{(k)}}^{(i)} \tag{27}$$

Furthermore, let

$$v_{i\ell,i'\ell'}(\underline{R}_N) = \frac{1}{N}\sum_{k=1}^{t}\sum_{j=1}^{n_k} E_{R_{ij}^{(k)}}^{(i)} E_{R_{\ell j}^{(k)}}^{(\ell)} E_{R_{i'j}^{(k)}}^{(i')} E_{R_{\ell' j}^{(k)}}^{(\ell')}$$

$$- U_{i\ell}^{*} U_{i'\ell'}^{*}, \qquad i, i', \ell, \ell' = 1, 2, \ldots, m \tag{28}$$

Setting

$$r = \frac{1}{2}(i-1)(2m-i) + \ell \qquad \text{for } i \leq \ell = 1, 2, \ldots, m$$

we rewrite

$$\{U_{i\ell}^{(k)}, \quad i \leq \ell = 1, 2, \ldots, m\}$$

$$\underline{E}_N = \begin{bmatrix} E^{(1)}_{R^{(1)}_{11}} & E^{(1)}_{R^{(1)}_{12}} & \cdots & E^{(1)}_{R^{(1)}_{1n_1}} & E^{(1)}_{R^{(2)}_{11}} & E^{(1)}_{R^{(2)}_{12}} & \cdots & E^{(1)}_{R^{(2)}_{1n_2}} & \cdots & E^{(1)}_{R^{(t)}_{11}} & E^{(1)}_{R^{(t)}_{12}} & \cdots & E^{(1)}_{R^{(t)}_{1n_t}} \\ E^{(2)}_{R^{(1)}_{21}} & E^{(2)}_{R^{(1)}_{22}} & \cdots & E^{(2)}_{R^{(1)}_{2n_1}} & E^{(2)}_{R^{(2)}_{21}} & E^{(2)}_{R^{(2)}_{22}} & \cdots & E^{(2)}_{R^{(t)}_{2n_2}} & \cdots & E^{(2)}_{R^{(t)}_{21}} & E^{(2)}_{R^{(t)}_{22}} & \cdots & E^{(2)}_{R^{(t)}_{2n_2}} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ E^{(m)}_{R^{(1)}_{m1}} & E^{(m)}_{R^{(1)}_{m2}} & \cdots & E^{(m)}_{R^{(1)}_{m2}} & E^{(m)}_{R^{(2)}_{m1}} & E^{(m)}_{R^{(2)}_{m2}} & \cdots & E^{(m)}_{R^{(2)}_{mn_2}} & \cdots & E^{(m)}_{R^{(t)}_{m1}} & E^{(m)}_{R^{(t)}_{m2}} & \cdots & E^{(m)}_{R^{(t)}_{mn_t}} \end{bmatrix}$$

$$= \sqrt{12}\left(\frac{1}{N+1}\,\underline{R}_N - \frac{1}{2}\,\underline{J}_N\right) \tag{24}$$

where

$$\underline{J}_N = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix}_{m \times N}$$

as

$$\underline{U}^{(k)} = \{U_r^{(k)}, \qquad r = 1, 2, \ldots, \frac{1}{2}m(m+1)\} \tag{29}$$

and

$$\{U_{i\ell}^*, \qquad i \leq \ell = 1, 2, \ldots, m\}$$

as

$$\underline{U}^* = \{U_r^*, \qquad r = 1, 2, \ldots, \frac{1}{2}m(m+1)\} \tag{30}$$

and

$$\underline{V}_N(\underline{R}_N) = [v_{rs}\ (\underline{R}_N)],\ r,s = 1, 2, \ldots, \frac{1}{2}m(m+1) \tag{31}$$

Thus, the test statistic can be expressed as

$$L = \sum_{k=1}^{t} n_k\ [\underline{U}^{(k)} - \underline{U}^*]\ \underline{V}_N^{-1}\ (\underline{R}_N)\ [\underline{U}^{(k)} - \underline{U}^*]' \tag{32}$$

Under the null hypothesis, eqn. (20), the test statistic L (for large samples) is asymptotically distributed as a chi-square random variable with ν degrees of freedom, where

$$\nu = \frac{1}{2}\ m(m+1)(t-1) \tag{33}$$

The P-value for this test is a right tail probability from a chi-square distribution table with the appropriate degrees of freedom.

3. Distribution-free Tests of Fit

Besides testing hypotheses concerning parameters of location and dispersion we are often interested in the validation of a specified distribution.

The goodness-of-fit problem in this multivariate setting may be described as follows:

Let $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_n$ be independent multivariate random variables with the distribution function $F(\underline{x})$. We wish to test the hypothesis

$$H_0:\ F(\underline{x}) = F_0(\underline{x}) \tag{34}$$

against the alternative

$$H_1: \quad F(\underline{x}) \neq F_0(\underline{x})$$

where $F_0(\underline{x})$ is some particular distribution function (either continuous or discrete). We can distinguish two special cases for tests of fit:

(i) Simple null hypotheses

Under a simple null hypothesis, the distribution of the random variable is completely specified by $F_0(\underline{x})$.

(ii) Composite null hypotheses

Under a composite null hypothesis, the distribution of the random variable is not completely determined by $F_0(\underline{x})$. If a composite null hypothesis depends upon unknown parameters, their maximum likelihood estimators [9] are usually used to derive the appropriate test.

When samples are obtained in a multicomponent solids mixing problem, the data can be expressed as

component / sample	1	2	. . .	m	m+1
1	f_{11}	f_{12}	. . .	f_{1m}	$f_{1(m+1)}$
2	f_{21}	f_{22}	. . .	f_{2m}	$f_{2(m+1)}$
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
n	f_{n1}	f_{n2}	. . .	f_{nm}	$f_{n(m+1)}$

where f_{ij} ($i = 1, 2, \ldots, (m+1)$; $j = 1, 2, \ldots, n$) denotes the number of particles of the i-th component in the j-th spot sample. If the samples are taken from a specified distribution, the expected number of particles of the i-th component in the j-th sample will be known and denoted by e_{ij}.

Furthermore, under the null hypothesis, eqn. (34), there should be close agreement between these corresponding frequencies. The deviations $(f_{ij} - e_{ij})$ measure lack of agreement. We eliminate the signs by squaring each difference, and reduce that value to original units by dividing by the respective e_{ij}. Thus

$$(f_{ij} - e_{ij})^2/e_{ij}$$

measures lack of agreement for the i-th component in the j-th sample. An overall measure of the lack of agreement is the sum of these individual measures. Thus, the test statistic Q is defined as

$$Q = \sum_{i=1}^{m+1} \sum_{j=1}^{n} \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \tag{35}$$

A small value of Q supports the null hypothesis H_0, whereas a large value reflects a general incompatibility between the frequencies observed and those expected under H_0.

The asymptotic distribution (large n) of Q is independent of the underlying distribution. For a simple null hypothesis, Q is asymptotically distributed as a chi-square random variable with mn degrees of freedom under H_0.

As mentioned earlier, it is sometimes necessary to estimate some parameter values before the test can be performed. Once the parameters are estimated and subsequently used to estimate e_{ij}, Q is calculated according to eqn. (35) as before. For a composite null hypothesis, again the distribution of Q is approximately chi-square but with n(m-w) degrees of freedom, where w denotes the number of independent unspecified parameters. Reduction of the number of degrees of freedom shifts the boundary of the critical region so that Q has to be smaller for acceptance at a given level.

4. Binomial Test

A common problem in solids mixing involves the blending of an active ingredient with several diluents. The homogeneity of this active ingredient in the entire mixture is of primary importance. Given a prescribed quality standard we are interested in testing the hypothesis that the proportion of mixture which meets the quality standard exceeds a fixed level.

In general, the hypothesis may take one of the following forms for some specified value of θ_0 $(0 < \theta_0 < 1)$

(1) One-sided alternatives

(a) H_{01} : $\theta \leq \theta_0$ versus H_+: $\theta > \theta_0$ (36)

(b) H_{02} : $\theta \geq \theta_0$ versus H_-: $\theta < \theta_0$ (37)

(2) Two-sided alternative

H_{03} : $\theta = \theta_0$ versus H_a: $\theta \neq \theta_0$ (38)

We first consider a test of the hypothesis (1a). Suppose that n spot samples are drawn from the mixture with each spot sample being classified as satisfactory or unsatisfactory. Denote the numbers of satisfactory and unsatisfactory samples by S_+ and S_-, respectively. The hypothesis H_{01} is rejected at the α level whenever

$$S_+ > C_{\alpha 1}$$

where the critical value $C_{\alpha 1}$ is determined such that

$$Pr_{\theta_0}[\ S_+ > C_{\alpha 1}] = \alpha$$

Note that the distribution of S_+ , when $\theta = \theta_0$, is binomial with parameters n and θ_0 , hence

$$E_{\theta_0}[\ S_+] = n\theta_0$$

Similarly, we reject the hypothesis H_{02} at the α level whenever

$$S_- > C_{\alpha 2}$$

where

$$Pr_{\theta_0}[\ S_- > C_{\alpha 2}] = \alpha$$

Here the distribution of S_- , when $\theta = \theta_0$, is binomial with parameters n and $(1-\theta_0)$, hence

$$E_{\theta_0}[\ S_-\ ] = n\ (\ 1 - \theta_0)$$

Since

$$S_+ + S_- = n$$

the rejection region of an α level tests of the hypothesis H_{03} is determined by

$$S_+ < C_{\alpha 1} \quad \text{or} \quad S_+ > C_{\alpha 2}$$

where

$$Pr_{\theta_0}[\ S_+ < C_{\alpha 1}] + Pr_{\theta_0}[\ S_+ > C_{\alpha 2}] = \alpha$$

An equal tails test selects critical values $C_{\alpha 1}$ and $C_{\alpha 2}$ such that

$$Pr_{\theta_0}[\ S_+ < C_{\alpha 1}]$$

and

$$Pr_{\theta_0}[\ S_+ > C_{\alpha 2}]$$

are approximately equal.

For large samples, we define the standardized variables (with a continuity correction of 0.5) to be

$$z_+ = \frac{S_+ - n\theta_0 - 0.5}{n\theta_0\ (1 - \theta_0)} \tag{39}$$

and

$$z_- = \frac{S_- - n(1-\theta_0) - 0.5}{n\theta_0\ (1-\theta_0)} \tag{40}$$

The P-value associated with the tests of the above three hypotheses are obtained from the standard normal table [2] as:

Hypothesis	P-value
H_{01} vs H_{+}	$Pr[\ Z > z_{+}]$
H_{02} vs H_{-}	$Pr[\ Z > z_{-}]$
H_{03} vs H_{a}	$2\ (\max(\ Pr[\ Z > z_{+}]\ ,\ Pr[\ Z > z_{-}]))$

where Z has a standard normal distribution.

IV. EXPERIMENT

The experimental apparatus, materials and procedures employed are described in this section. To minimize experimental and computational effort, ternary particles systems were chosen to demonstrate the analysis of multicomponent solids mixing by nonparametric statistical methods.

1. Apparatus and Materials

The apparatus used in this experiment was a cylindrical plexiglass mixer of the following dimensions: internal length 38.1 cm (15 in), diameter 14.0 cm (5.5 in) and end flanges diameter 25.4 cm (10 in). The tube was split axially so that the upper portion could be removed for loading and sampling. The end flanges were accurately made to insure that during mixing the axis of rotation coincided with the geometric axis of the mixer. The plexiglass cylinder was set horizontally on a jar mill whose rotational speed was accurately maintained at a speed between 10 and 50 r.p.m.. Particles used in this experiment were Lucite spheres with an average diameter of 0.16 cm (small), 0.32 cm (medium) and 0.48 cm (large) with an average density of 1.156 g/cm^3.

2. Procedure

Prior to mixing, two thin semi-circular partitions were placed between the two ends of mixer normal to the mixer axis dividing it into three equal compartments. One hundred and seventy grams each of 3 types of particles were loaded in each compartment, respectively. Approximately 30% of the overall volume of the mixer was occupied by the particles. The bed was then leveled, the partitions were removed, the cover was put in place, and the mixer was rotated. Two types of systems were created:

(1) Heterogeneous

Three types of particles of different sizes (small, medium and large) were employed in this system.

(2) Homogeneous

Three types of (large) particles, which had identical properties except color, were used in this system.

Table 2 summarizes the experimental conditions of each run. After a predetermined mixing time, twelve spot samples were randomly drawn from the mixture for each experimental run and the weight fractions of three types of particles in the sample were recorded. For the homogeneous system, the number of particles of each type in the sample was also counted.

Table 2. Summary of particle system and experimental conditions

(1) Heterogeneous system (small, medium and large particles)		
experimental run	rotating speed (r.p.m.)	mixing time (min.)
1	30	2
2	30	10
3	30	30
4	20	10
5	45	10
(2) Homogeneous system (large particles)		
experimental run	rotating speed (r.p.m.)	mixing time (min.)
6	30	30
7	45	30
8	45	60
9	45	150

V. RESULTS AND DISCUSSION

The theories and procedures presented in Section III. are employed to analyze the experimental data obtained. Implications of various tests are explained.

1. Test of sampling techniques

In a multicomponent solids mixing problem, the sample mean vector should not deviate greatly from its known population mean vector. On the other hand, the mean vector by itself should not be used as a measure of the degree of mixedness, since, if the batch is properly sampled, the only variation between sample mean vectors should be the sampling variation, regardless of how well mixed the batch is. If the mean vector $\underline{\mu}$ differs significantly from the population mean vector $\underline{\mu}_0$, the sampling may have been biased due to location or method [10]. If so, this bias should be eliminated before further sampling.

To accomplish this, we have to test the hypothesis that the mean vector is specified, e.g., to test the null hypothesis

$$H_0: \ \underline{\mu} = \mu_0 = [\tfrac{1}{3} \quad \tfrac{1}{3}]' \tag{41}$$

against the alternative hypothesis

$$H_1: \underline{\mu} \neq [\tfrac{1}{3} \ \tfrac{1}{3}]'$$

An example of this calculation is shown below for the first experimental run. The experimental data for the 3 particle sizes (small, medium and large) in 12 random spot samples are tabulated in Table 3. By selecting the small and medium sized particles, we express the sample data matrix, $\underline{x}$, as

Table 3 Experimental data expressed in weight fraction for experimental runs 1 through 3.

spot sample	experimental run 1 (2 min, 30 r.p.m.)			experimental run 2 (10 min, 30 r.p.m.)			experimantal run 3 (30 min, 30 r.p.m.)		
	small	medium	large	small	medium	large	small	medium	large
1	.127	.797	.076	.012	.259	.729	.035	.426	.539
2	.968	.032	.000	.435	.565	.000	1.000	.000	.000
3	.004	.370	.626	.115	.606	.279	.029	.493	.477
4	.000	.410	.590	.017	.771	.212	.000	.049	.951
5	.022	.695	.283	.271	.611	.118	.109	.482	.409
6	.992	.008	.000	.146	.836	.017	.085	.507	.408
7	.869	.123	.008	.089	.250	.661	.008	.200	.792
8	.241	.752	.007	.929	.071	.000	.064	.463	.473
9	.000	.000	1.000	.272	.517	.211	.064	.583	.353
10	.018	.658	.324	.600	.225	.175	.136	.529	.336
11	.987	.013	.000	.067	.202	.731	.052	.504	.444
12	.969	.031	.000	.431	.419	.150	.050	.474	.476

$$\underline{x} = \begin{bmatrix} 0.127 & 0.797 \\ 0.968 & 0.032 \\ 0.004 & 0.370 \\ 0.000 & 0.410 \\ 0.022 & 0.695 \\ 0.992 & 0.008 \\ 0.869 & 0.123 \\ 0.241 & 0.752 \\ 0.000 & 0.000 \\ 0.018 & 0.658 \\ 0.987 & 0.013 \\ 0.969 & 0.031 \end{bmatrix}'$$

which, after adjustment for $\underline{\mu}_0$, yields

$$\underline{y} = \underline{x} - \underline{\mu}_0 = \begin{bmatrix} -0.206 & 0.464 \\ 0.635 & -0.301 \\ -0.329 & 0.037 \\ -0.333 & 0.077 \\ -0.311 & 0.362 \\ 0.659 & -0.325 \\ 0.536 & -0.210 \\ -0.092 & 0.419 \\ -0.333 & -0.333 \\ -0.315 & 0.325 \\ 0.654 & -0.320 \\ 0.636 & -0.302 \end{bmatrix}'$$

Since ties occur in the application of rank tests, we use a midrank procedure that assigns the simple average of the ranks which would have been assigned to the observations if they were not tied. Thus, ranking the elements of each row of $\underline{y}$ in increasing order of their absolute values, we obtain

$$\underline{R}_n = \begin{bmatrix} 2 & 9 & 5 & 6.5 & 3 & 12 & 8 & 1 & 6.5 & 4 & 11 & 10 \\ 12 & 4 & 1 & 2 & 10 & 7.5 & 3 & 11 & 9 & 7.5 & 6 & 5 \end{bmatrix}$$

Two multivariate rank tests were developed to test the hypothesis of a prescribed mean vector:

(1) A multivariate sign test

The score matrix takes the form

$$\underline{E}_n = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

According to eqn. (9), we have

$$T^{(1)} = -2, \quad T^{(2)} = 0$$

and,

$$\underline{T} = [\ -2 \quad 0\ ]$$

Thus, from eqns. (13) and (14), we have

$$\underline{V} = \begin{bmatrix} 1 & -\frac{10}{12} \\ -\frac{10}{12} & 1 \end{bmatrix}$$

and

$$\underline{V}^{-1} = \begin{bmatrix} \frac{36}{11} & \frac{30}{11} \\ \frac{30}{11} & \frac{36}{11} \end{bmatrix}$$

Therefore, the test statistic is calculated as

$$S = \frac{1}{n}\ [\underline{T}\ \underline{V}^{-1}\ \underline{T}']$$

$$= \frac{1}{12}\ [\ -2 \quad 0\ ] \begin{bmatrix} \frac{36}{11} & \frac{30}{11} \\ \frac{30}{11} & \frac{36}{11} \end{bmatrix} \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

$$= 1.0909$$

Since S is asymptotically distributed as a chi-square random variable with 2 degrees of freedom, we can calculate the P-value as

$$P = 0.5796$$

Such a large P-value supports H_0; hence, the sampling technique is judged to be representative of the mixture.

(2) A multivariate generalization of the Wilcoxon signed-rank test

In this case, the score matrix

$$\underline{E}_n = \frac{1}{13}\, \underline{R}_n$$

By eqn. (9),

$$T^{(1)} = 1.692, \quad T^{(2)} = 0.692$$

and

$$\underline{T} = [\, 1.692 \quad 0.692 \,]$$

According to eqns. (13) and (14), we have

$$\underline{V} = \begin{bmatrix} 0.320 & -0.158 \\ -0.158 & 0.320 \end{bmatrix}$$

Thus, the test statistic, eqn. (15), is

$$S = \frac{1}{n} \, [\, \underline{T}\, \underline{V}^{-1}\, \underline{T}' \,]$$

$$= 1.5476$$

The associated P-value is

$$P = 0.4613$$

Therefore, use of the multivariate Wilcoxon signed-rank procedure leads to the same conclusion as the multivariate sign test and the sampling technique is judged to be representative of the mixture.

Table 4 lists the P-values for all pairs of particles considered and the two multivariate rank tests. Calculations have been carried out for experimental runs 2 through 5. The results from these runs are also shown in Table 4.

2. Test of Treatment Effects

The data from the first three experimental runs are used to illustrate a test of significance of treatment effects. This is accomplished by testing the homogeneity of their covariance matrices. Thus, we test the hypothesis

Table 4. Results of testing the sampling technique

	Experimental run	Calaulation based on the pair of (small, medium)	(medium, large)	(small, large)	Inference: H_0, eqn. (41), is rejected
multivariate sign test	1	0.5796	0.2138	0.1054	no
	2	0.1653	0.4758	0.0765	no
	3	0.1054	0.1054	0.2231	no
	4	0.2231	0.1353	0.1353	no
	5	0.1653	0.7788	0.1653	no
multivariate Wilcoxon signed-rank test	1	0.4613	0.3285	0.2712	no
	2	0.3293	0.3196	0.2366	no
	3	0.3392	0.3443	0.7649	no
	4	0.3013	0.2808	0.3488	no
	5	0.5732	0.6985	0.8583	no

$$H_0 : \underline{\Sigma}^{(1)} = \underline{\Sigma}^{(2)} = \underline{\Sigma}^{(3)} \tag{42}$$

against the alternative

$$H_1 : \underline{\Sigma}^{(1)}, \underline{\Sigma}^{(2)} \text{ and } \underline{\Sigma}^{(3)} \text{ are not identical}$$

In other words, H_0 hypothesizes no significant difference in dispersion for mixing times of 2, 10 and 30 minutes.

The experimental data are listed in Table 3. Using the small and medium sized particles for illustration, we define

$$\underline{x}_N = \begin{bmatrix} \underline{x}^{(1)} & \underline{x}^{(2)} & \underline{x}^{(3)} \end{bmatrix}_{2x36}$$

where

$$\underline{x}^{(1)} = \begin{bmatrix} 0.127 & 0.797 \\ 0.968 & 0.032 \\ 0.004 & 0.370 \\ 0.000 & 0.410 \\ 0.022 & 0.695 \\ 0.992 & 0.008 \\ 0.869 & 0.123 \\ 0.241 & 0.752 \\ 0.000 & 0.000 \\ 0.018 & 0.658 \\ 0.987 & 0.013 \\ 0.969 & 0.031 \end{bmatrix}', \quad \underline{x}^{(2)} = \begin{bmatrix} 0.012 & 0.259 \\ 0.435 & 0.565 \\ 0.115 & 0.606 \\ 0.017 & 0.771 \\ 0.271 & 0.611 \\ 0.146 & 0.836 \\ 0.089 & 0.250 \\ 0.929 & 0.071 \\ 0.272 & 0.517 \\ 0.600 & 0.225 \\ 0.067 & 0.202 \\ 0.431 & 0.419 \end{bmatrix}'$$

and

$$\underline{x}^{(3)} = \left[\begin{matrix} 0.035 & 0.426 \\ 1.000 & 0.000 \\ 0.029 & 0.493 \\ 0.000 & 0.049 \\ 0.109 & 0.482 \\ 0.085 & 0.507 \\ 0.008 & 0.200 \\ 0.064 & 0.463 \\ 0.064 & 0.583 \end{matrix}\right.$$

$$\begin{bmatrix} 0.136 & 0.529 \\ 0.052 & 0.504 \\ 0.050 & 0.474 \end{bmatrix}$$

Ranking the 36 elements of each row of the matrix $\underline{x}_N$ in increasing order of magnitude, we obtain the rank matrix

$$\underline{R}_N = \left[\underline{R}^{(1)} \ \underline{R}^{(2)} \ \ \underline{R}^{(3)} \right]_{2x36}$$

where

$$\underline{R}^{(1)} = \begin{bmatrix} 16 & 31 & 3 & 1.5 & 7 & 36 & 29 & 21 & 1.5 & 6 & 35 & 32 \\ 35 & 7 & 16 & 20 & 32 & 2 & 9 & 33 & 1 & 31 & 3 & 6 \end{bmatrix}$$

$$\underline{R}^{(2)} = \begin{bmatrix} 4 & 26 & 15 & 5 & 22 & 17 & 12 & 30 & 23 & 27 & 10.5 & 25 \\ 13 & 28 & 29 & 34 & 30 & 36 & 12 & 8 & 27 & 11 & 10 & 22 \end{bmatrix}$$

and

$$\underline{R}^{(3)} = \begin{bmatrix} 33 & 13 & 20 & 14 & 18 & 10.5 & 8 & 19 & 9 & 34 & 28 & 24 \\ 5 & 23 & 17 & 26 & 21 & 18 & 19 & 25 & 24 & 4 & 15 & 14 \end{bmatrix}$$

According to eqn. (24), the general score matrix is of the form

$$\underline{E}_N = \sqrt{12}\left(\frac{1}{N+1}\underline{R}_N - \frac{1}{2}\underline{J}_N\right)$$

where

$$\underline{J}_N = \begin{bmatrix} 1 & 1 & \ldots\ldots & 1 \\ 1 & 1 & \ldots\ldots & 1 \end{bmatrix}_{2x36}$$

Thus, in our example,

$$\underline{E}_N = \sqrt{12}\left(\frac{1}{37}\underline{R}_N - \frac{1}{2}\underline{J}_N\right)$$

Using eqns. (25) through (31), (see Appedix B for a list of computer program)

$$\underline{U}^{(1)} = [\,1.7099 \quad -0.7724 \quad 1.5668\,]$$

$$\underline{U}^{(2)} = [\,0.6743 \quad -0.0612 \quad 0.9201\,]$$

$$\underline{U}^{(3)} = [\,0.7034 \quad -0.4645 \quad 0.4565\,]$$

$$\underline{U}^* = [\ 0.9727 \quad -0.4126 \quad 0.9730\ ]$$

$$\underline{V}_N(\underline{R}_N) = \begin{bmatrix} 0.6606 & -0.2705 & -0.3515 \\ -0.2705 & 0.9632 & -0.2442 \\ 0.1870 & -0.2442 & 0.6623 \end{bmatrix}$$

and

$$\underline{V}_N^{-1}(\underline{R}_N) = \begin{bmatrix} 1.7878 & 0.4113 & -0.3515 \\ 0.4113 & 1.2402 & 0.3411 \\ -0.3515 & 0.3411 & 1.7348 \end{bmatrix}$$

According to eqn. (32), the test statistic is

$$L = \sum_{k=1}^{3} n_k \ [\ \underline{U}^{(k)} - \underline{U}^*\ ]\ \underline{V}_N^{-1}(\underline{R}_N)\ [\ \underline{U}^{(k)} - \underline{U}^*\ ]$$

$$= 21.6334$$

Since L is asymptotically distributed as a chi-square random variable with six degrees of freedom, the associated P-value is

$$P \doteq 0.001$$

Therefore, the null hypothesis H_0, eqn. (43), is rejected, and we conclude that there exists a significant difference in dispersion between the mixing times of 2, 10 and 30 minutes. The P-value based on small and large sized particles and that based on medium and large sized particles are 0.001 and 0.006 respectively.

In the second experiment, we test the effect of rotating speeds (20, 30 and 45 r.p.m.) on dispersion for a fixed mixing time of 10 minutes (Experimental runs 2, 4, and 5). The P-values corresponding to test statistics for pairs (small ,medium), (small, large) and (medium, large) are 0.720, 0.360 and 0.753 respectively. Because, the P-values are quite large, we conclude that the difference in dispersion among rotating speeds of 20, 30 and 45 r.p.m. is not significant.

3. Test of the Completely Mixed State

In solids mixing, the completely random (or mixed) state is characterized by the property that the probability of selecting a particle of a given component from anywhere in the mixture is identical. When the population proportions are known for the components of a mixture, the chi-square goodness-of-fit test can be used to test the hypothesis that the mixture is in the completely mixed state. In other words, we test

$$H_0 : \text{the mixture is in the completely mixed state} \tag{44}$$

against the alternative

$$H_1 : H_0 \text{ is not true}$$

The data generated in run 6 of the experiments is shown below.

spot sample	color distribution* Green	Red	White	total number of particles	expected number of particles for each color
1	47	34	22	103	34.33
2	39	28	31	98	32.6
3	26	16	42	84	28
4	24	10	50	84	28
5	38	42	17	97	32.33
6	33	39	35	107	35.67
7	35	47	9	91	30.33
8	31	26	54	111	37
9	32	39	31	102	34
10	36	20	27	83	27.67
11	33	76	11	120	40
12	40	43	37	120	40

* The ratio is 1 : 1 : 1 for categories Green : Red : White, respectively.

From eqn. (35), the test statistic is computed as

$$Q = \sum_{i=1}^{3} \sum_{j=1}^{12} \frac{f_{ij} - e_{ij})^2}{e_{ij}}$$

$$= \frac{(47 - 34.33)^2}{34.33} + \frac{(34 - 34.33)^2}{34.33} + \frac{(22 - 34.33)^2}{34.33} + \ldots + \frac{(40 - 40)^2}{40}$$

$$\frac{(43 - 40)^2}{40} + \frac{(37 - 40)^2}{40}$$

$$= 162.68$$

Under the null hypothesis, eqn. (44), Q is asymptotically distributed as the chi-square random variable with 24 degrees of freedom. Since

$$P << 0.001,$$

the null hypothesis is rejected (at the usual levels); we conclude that mixture has not reached the completely mixed state. The following table summarizes the results of tests of completely mixed state for experimental runs 6 through 9:

experimental run	mixing time (min)	rotating speed (r.p.m.)	expected distribution	Test statistic Q	associated probability P	inference about H_0, eqn. (44)
6	30	30	$\frac{1}{3} : \frac{1}{3} : \frac{1}{3}$	162.28	$<<0.001$	rejected
7	30	45	$\frac{1}{3} : \frac{1}{3} : \frac{1}{3}$	56.25	<0.001	rejected
8	60	45	$\frac{1}{3} : \frac{1}{3} : \frac{1}{3}$	43.43	0.009	rejected
9	150	45	$\frac{1}{3} : \frac{1}{3} : \frac{1}{3}$	13.20	0.963	accepted

4. Test of a Quality Standard

In quality control involving multicomponent solids mixing, we may be concerned with the fraction of a population which meets a quality criterion. The binomial test can be used to solve this multicomponent solids mixing problem. Again let x_{ij} denote the weight fraction of the i-th component in the j-th sample and μ_{i0} the population weight fraction of i-th component. We may set the criterion as

$$\sum_{i=1}^{m+1} \lambda_i \, (x_{ij} - \mu_{i0})^2 \leq \varepsilon \tag{45}$$

where

λ_i = arbitrary positive constant which might reflect the relative importance of the i-th component being mixed.

and

ε = pre-selected positive real number.

We say that a spot sample is satisfactory, if it satisfies this criterion; otherwise, it is unsatisfactory.

Assume that three components are equally important. Hence, let

$$\lambda_1 = \lambda_2 = \lambda_3 = 1$$

We wish to test that the satisfactory proportion of a mixture has reached 95% for a pre-selected value of 0.015 for ε. In the following examples, we test the null hypothesis

$$H_0 : \theta \geq .95 \tag{46}$$

against the alternative hypothesis

$$H_- : \theta < .95$$

The calculation for run 7 of the experiments is shown in Table 5. The results are

$$S_+ = 7, \quad S_- = 5, \quad n = 12$$

The guide indicates that the appropriate P-value is left-tail probability for $S_+ = 7$ with a parameter of .95, which from a binomial table is

$$P = .0002$$

Since this P-value is so small, we conclude that the data reject H_0 in favor

Table 5. The calculation for experimental run 7 for testing the quality standard as defined in eqn. (45)

spot sample	x_{ij}	x_{2j}	x_{3j}	$\sum_{i=1}^{3}(x_j - \mu_{i0})^{2**}$	is it satisfactory?
1	0.3333	0.2564	0.4103	0.0018	yes
2	0.3537	0.4146	0.2317	0.0174	no
3	0.2941	0.3765	0.3294	0.0034	yes
4	0.3196	0.2474	0.4330	0.0175	no
5	0.3113	0.4811	0.2075	0.0382	no
6	0.2857	0.3286	0.3587	0.0050	yes
7	0.2785	0.2405	0.4810	0.0334	no
8	0.3678	0.3908	0.2414	0.0129	yes
9	0.3229	0.2917	0.3854	0.0046	yes
10	0.3182	0.3636	0.3182	0.0014	yes
11	0.3367	0.5000	0.1633	0.0567	no
12	0.3636	0.2987	0.3377	0.0021	yes

** $\mu_{10} = \mu_{20} = \mu_{30} = 0.3333$

of H_-. Hence, we conclude that mixing is not adequate. The results of runs 6 through 9 are summarized in the following table.

experimental run	S_+	S_-	n	P	Inference concerning H_0, eqn. (46).
6	5	7	12	0.0000	rejected
7	7	5	12	0.0002	rejected
8	10	2	12	0.1184	accepted
9	12	0	12	1.0000	accepted

Note that this test may also be used in the analysis of a continuous mixing process.

VI. CONCLUSIONS

Statistical analysis is recognized as a major tool in solids mixing investigations. Traditionally, results of sampling have been analyzed using normal theory statistical techniques [10, 12].

This study proposes the applicability of several nonparametric statistical techniques to problems in multicomponent solids mixing. The most important feature of a nonparametric procedure is its lack of dependence on a particular distribution type, e.g., normal. Since the distributions of components during mixing are usually unknown, nonparametric procedures comprise a substantial collection of alternatives to the classical parametric procedures.

Recently, the extension of nonparametric techniques from the univariate to the multivariate case has been pursued in [8]. The present study demonstrate the applicability of multivariate tests of location and dispersion to test the hypotheses concerning a sampling technique and the significance of treatment effects in multicomponent solids mixing problems.

The proposed nonparametric procedures were tested with actual homogeneous and heterogeneous ternary mixtures generated by a drum mixer. In spite of the small number ($n = 12$) of the sample obtained, the results tend to support the practical significance of nonparametric statistics in the evaluation of mixing systems.

Besides the robustness of the nonparametric methods against the assumption of a specified distributional form, it is important to note their simplicity in application. An effort will be made in the future to study the performance of the proposed nonparametric methods for larger sample sizes.

NOTATIONS

C_α	critical value at the significance level α
C_{ij}	sign indicator of Y_{ij}
$E(X)$	expected value of random variable X
$\underline{E}_N$	score matrix as defined in eqn. (24)
$\underline{E}_n$	score matrix as defined in eqn. (8)
$\bar{E}^{(i)}$	mean score of the i-th variate
e_{ij}	expected number of particles of the i-th component in the j-th sample
$F(x)$	cumulative distribution function (c.d.f.)of X
f_{ij}	number of particles of the i-th component in the j-th sample
H_0	null hypothesis
H_a	two-sided alternative
H_+	one-sided alternative with positive direction
H_-	one-sided alternative with negative direction
L	test statistic for testing homogeneity of dispersion matrices
m	number of variates
N	total number of spot samples for t treatments
n	number of spot samples
n_k	number of samples of treatment k
P	associated probability
Q	test statistic for goodness of fit test
$\underline{R}_N$	(m x N) rank matrix as defined in eqn. (22)
$\underline{R}_n$	(m x n) rank matrix as defined in eqn. (7)
R_{ij}	rank of Y_{ij} among (Y_{i1} , ..., Y_{in})
R^m	set of all order m-tuples (x_1 x_2 ... x_m)
$R_{ij}^{(k)}$	rank of $X_{ij}^{(k)}$ among ($X_{i1}^{(k)}$, ..., $X_{in_t}^{(k)}$) for k = 1, 2, ..., t
S_-	number of unsatifactory samples

S_-	number of unsatisfactory samples
S	test statistic for testing equality of mean vectors
$T^{(i)}$	univariate rank order statistic as defined in eqn. (9)
$\underline{T}$	row vector as defined in eqn. (10)
t	number of treatments
$\underline{U}_{\ell}^{(k)}$	row vector as defined in eqn. (29)
$\underline{U}_{i\ell}^{*}$	row vector as defined in eqn. (30)
$U_{i\ell}^{(k)}$	as defined in eqn. (25)
$U_{i\ell}^{*}$	as defined in eqn. (26)
$v_{i\ell}$	as defined in eqn. (13)
$\underline{V}$	as defined in eqn. (12)
$\underline{V}_N$	as defined in eqn. (31)
w	number of unspecified parameters estimated from data
$\underline{X}_j$	row vector = $[X_{1j}\ X_{2j}\ .\ .\ .\ X_{mj}]$, $j = 1, 2, ..., n$
$\underline{X}_j^{(k)}$	row vector = $[X_{1j}^{(k)}\ X_{2j}^{(k)}\ .\ .\ .\ X_{mj}^{(k)}]$, $j = 1, 2, ..., n$
$X_j^{(k)}$	row vector = $[X_{1j}^{(k)}\ X_{2j}^{(k)}\ .\ .\ .\ X_{mj}^{(k)}]$, $j = 1, 2, ..., n$
$X_{ij}^{(k)}$	a random variable representing the weight fraction of the i-th component in the j-th sample for the k-th treatment
$\underline{x}_j$	realization of $\underline{X}_j$
Y_{ij}	$= X_{ij} - \mu_{i0}$
$\underline{Z}_n$	(m x n) random matrix
$\underline{Z}_N$	(m x N) pooled random matrix
z_+	standardized variable as defined in eqn. (39)
z_-	standardized variable as defined in eqn. (40)
θ	a parameter representing the probability of satisfaction
$\underline{\mu}$	location vector parameter

μ_{i0} weight proportion of the i-th component in the population

ν degrees of freedom as defined in eqn. (33)

χ^2 chi-square distribution

$\underline{\Sigma}^{(k)}$ dispersion matrix of treatment k

REFERENCES

1. F. S. Lai, R. H. Wang and L. T. Fan, An application of nonparametric statistics to the sampling in solids mixing, Powder Technol., 10 (1974) 13.

2. J. D. Gibbons, Nonparametric Methods for Quantitative Analysis, Holt, Rinehart and Winston, New York, 1976.

3. J. D. Gibbons, Nonparametric Statistical Inference, McGraw-Hill, New York, 1971.

4. W. J. Conover, Practical Nonparametric Statistics, Wiley, New York, 1971.

5. E. L. Lehmann, Nonparametrics: Statistical Methods Based on Ranks, McGraw-Hill, New York, 1975.

6. J. Hájek, Nonparametric Statistics, Holden-Day, San Francisco, 1969.

7. M. Hollander and D. A. Wolfe, Nonparametric Statistical Methods, Wiley, New York, 1973.

8. M. L. Puri and P. K. Sen, Nonparametric Methods in Multivariate Analysis, Wiley, New York, 1971.

9. A. M. Mood, F. A. Graybill, and D. C. Boes, Introduction to the Theory of Statistics, McGraw Hill, New York, 3rd edn., 1970.

10. S. S. Weidenbaum, in T. B. Drew and J. W. Goopee, Jr. (eds.), Advances in Chemical Engineering, Vol. II, Academic Press, New York, 1958.

11. P.M.C. Lacey, The mixing of solid particles, Trans. Inst. Chem. Eng., 21 (1943) 53.

12. R. H. Wang, L. T. Fan and J. R. Too, Multivariate statistical analysis of solids mixing, Powder Technol., 21 (1978) 171.

APPENDIX A. On the Use of P-values in Hypothesis Testing

The traditional method of testing a hypothesis is the determination of a rejection region and a corresponding rejection rule for which the probability of making a Type I error does not exceed some preselected value called the level of the test. In many cases, the choice of the level of the test is arbitrary and in some testing situations the chosen level may not even be attainable. These problems are circumvented by the reporting of P-values.

The P-value is defined as the probability under the null hypothesis of a sample outcome equal to or more extreme than that observed. The reporting of P-values clearly contains more information than merely reporting the decision made on a hypothesis at a possibly arbitrary level.

The use of P-value is clear for those tests in which the outcomes can be ordered according to how extreme they are relative to the expected outcome under the null hypothesis. In those unambiguous cases the P-value is the probability associated with a corresponding right or left tail probability. In the more complex situations where "more extreme" is an ambiguous relation, conventions must be defined for the reporting of P-values.

APPENDIX B List of a Computer Program for Calculating the Test Statistic L

A computer program for calcuting the test statistic L in eqn. (32) is developed and listed in the following pages. The input data are the elements of the rank matrix, eqn. (22). The symbols used in this program are listed as below.

M : m; number of variates

NT : T; number of treatments

N(K) : n_k; number of spot samples for treatment k.

E(K,I,J) : $E^{(i)}_{R_{ij}}(k)$; general score of the i-th component in the j-th spot sample for treatment k.

NTT : N; total number of observations.

S(K,I,J) : $U^{(k)}_{ij}$ as defined in eqn. (25).

SS(I,J) : $U^{*}_{.}$ as defined in eqn. (26).

G(I) = $\bar{\bar{E}}^{(i)}$ as defined in eqn. (27).

V(I,J,II,JJ) : $V_{ij,i'j'}$ as defined in eqn. (28)

L : L; test statistic as defined in eqn. (32)

Note that two subroutines are used. Subroutine MINV is used to obtain the inverse matrix of a non-singular matrix. Subroutine GMPRD is used to get the product of two general matrices. The result of calculation of the test statistic for comparison among the first three runs is also shown in page 56.

```
C      THIS PACKAGE IS USED TO CALCULATE THE TEST STATISTIC L, EQN. (32)
C      M = NUMBER OF VARIATES
C      NT = NUMBER OF TREATMENTS
C      N(K) = NUMBER OF SPOT SAMPLES FOR TREATMENT K
C      NTT = TOTAL NUMBER OF OBSERVATIONS
C      E(K,I,J) = GENERAL SCORE OF THE I-TH COMPONENT  IN THE J-TH SPOT
C                 SAMPLE FOR TREATMENT K
       DIMENSION E(3,2,12), S(3,2,2),N(3),D(3,2),G(2),V(2,2,2,2),
     *        IS(3),W(3,3),T(3,3),SS(2,2),A(9),L1(3),L2(3),B(3),R(3)
       M=2
       NT=3
       READ (5,1) (N(I),I=1,NT)
     1 FORMAT (3I10)
       READ (5,2) (((E(K,I,J),J=1,12  ),I=1,M),K=1,NT)
     2 FORMAT (12F6.0)
       DO 90 K=1,3
       WRITE (6,12) K
    12 FORMAT (/,'  THE INPUT DATA,  MATRIX R(',I1,')')
       WRITE (6,11) ((E(K,I,J),J=1,12),I=1,M)
    11 FORMAT (12F7.1)
    90 CONTINUE
       NTT=0
       DO 5  I=1,NT
     5 NTT=NTT+N(I)
       DO 20 K=1,NT
       DO 20  I=1,M
       DO 20 J=1,12
       E(K,I,J)=(E(K,I,J)/(NTT+1)-.5)*(12**.5)
    20 CONTINUE
       DO 10 K=1,NT
       DO 10  I=1,M
       D(K,I)=0.
       DO 15 J=1,12
    15 D(K,I)=D(K,I)+E(K,I,J)
    10 CONTINUE
       DO 25 K=1,NT
       DO 25  I=1,M
```

```
      DO 25 J=1,M
      S(K,I,J)=0.
      DO 30 JI=1,12
   30 S(K,I,J)=E(K,I,JI)*E(K,J,JI)+S(K,I,J)
      S(K,I,J)=S(K,I,J)-D(K,I)*D(K,J)/N(K)
      S(K,I,J)=S(K,I,J)/(N(K)-1)
   25 CONTINUE
      DO 26 I=1,M
      G(I)=0.
      DO 27 K=1,NT
      DO 27 JI=1,12
      G(I)=G(I)+E(K,I,JI)
   27 CONTINUE
      G(I)=G(I)/NTT
   26 CONTINUE
      DO 40 I=1,M
      DO 40 J=1,M
      SS(I,J)=0.
      DO 35 K=1,NT
      DO 35 JI=1,12
      SS(I,J)=SS(I,J)+E(K,I,JI)*E(K,J,JI)
   35 CONTINUE
      SS(I,J)=SS(I,J)-NTT*G(I)*G(J)
      SS(I,J)=SS(I,J)/(NTT-1)
   40 CONTINUE
      DO 45 I=1,M
      DO 45 J=1,M
      DO 45 II=1,M
      DO 45 JJ=1,M
      V(I,J,II,JJ)=0.
      DO 50 K=1,NT
      DO 50 JI=1,12
      V(I,J,II,JJ)=V(I,J,II,JJ)+E(K,I,JI)*E(K,J,JI)*E(K,II,JI)*E(K,JJ,JI
     *   )
   5C CONTINUE
      V(I,J,II,JJ)=V(I,J,II,JJ)/NTT-SS(I,J)*SS(II,JJ)
   45 CONTINUE
```

```
      DO 55 K=1,NT
      T(K,1)=S(K,1,1)
      T(K,2)=S(K,1,2)
      T(K,3)=S(K,2,2)
   55 CONTINUE
      TS(1)=SS(1,1)
      TS(2)=SS(1,2)
      TS(3)=SS(2,2)
      W(1,1)=V(1,1,1,1)
      W(1,2)=V(1,1,1,2)
      W(1,3)=V(1,1,2,2)
      W(2,1)=V(1,2,1,1)
      W(2,2)=V(1,2,1,2)
      W(2,3)=V(1,2,2,2)
      W(3,1)=V(2,2,1,1)
      W(3,2)=V(2,2,1,2)
      W(3,3)=V(2,2,2,2)
      DO 80 K=1,NT
      WRITE (6,6) K
    6 FORMAT (/,'  VECTOR U(',I1,')')
      WRITE (6,3) (T(K,I),I=1,3)
    3 FORMAT (   3F15.6)
   80 CONTINUE
      WRITE (6,7)
    7 FORMAT (/,'  VECTOR U(*)')
      WRITE (6,3) (TS(I),I=1,3)
      WRITE (6,8)
    8 FORMAT (/,'  MATRIX V(R), EQN. (31).',/)
      WRITE (6,3) ((W(I,J),I=1,3),J=1,3)
      NI=3
      A(1)=W(1,1)
      A(2)=W(2,1)
      A(3)=W(3,1)
      A(4)=W(1,2)
      A(5)=W(2,2)
      A(6)=W(3,2)
      A(7)=W(1,3)
```

```
      A(8)=W(2,3)
      A(9)=W(3,3)
      CALL MINV (A,NI,DI,L1,L2)
      WRITE (6,9)
    9 FORMAT (/,'  THE INVERSE MATRIX OF V(R)',/)
      WRITE (6,3) (A(I),I=1,9)
      STAT=0.
      DO 65 K=1,NT
      DO 70 I=1,3
      B(I)=T(K,I)-TS(I)
   70 CONTINUE
      CALL GMPRD (B,A,R,1,3,3)
      CALL GMPRD (R,B,ST,1,3,1)
      STAT=STAT+ST*N(K)
   65 CONTINUE
      WRITE (6,4)  STAT
    4 FORMAT (//,'  THE TEST STATISTIC L =', F10.6)
      STOP
      END
C                                                                       MINV  10
C     ..................................................................MINV  20
C                                                                       MINV  30
C        SUBROUTINE MINV                                                MINV  40
C                                                                       MINV  50
C        PURPOSE                                                        MINV  60
C           INVERT A MATRIX                                             MINV  70
C                                                                       MINV  80
C        USAGE                                                          MINV  90
C           CALL MINV(A,N,D,L,M)                                        MINV 100
C                                                                       MINV 110
C        DESCRIPTION OF PARAMETERS                                      MINV 120
C           A - INPUT MATRIX, DESTROYED IN COMPUTATION AND REPLACED BY  MINV 130
C               RESULTANT INVERSE.                                      MINV 140
C           N - ORDER OF MATRIX A                                       MINV 150
C           D - RESULTANT DETERMINANT                                   MINV 160
C           L - WORK VECTOR OF LENGTH N                                 MINV 170
C           M - WORK VECTOR OF LENGTH N                                 MINV 180
```

```
C                                                                       MINV 190
C        REMARKS                                                        MINV 200
C           MATRIX A MUST BE A GENERAL MATRIX                           MINV 210
C                                                                       MINV 220
C        SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED                  MINV 230
C           NONE                                                        MINV 240
C                                                                       MINV 250
C        METHOD                                                         MINV 260
C           THE STANDARD GAUSS-JORDAN METHOD IS USED. THE DETERMINANT   MINV 270
C           IS ALSO CALCULATED. A DETERMINANT OF ZERO INDICATES THAT    MINV 280
C           THE MATRIX IS SINGULAR.                                     MINV 290
C                                                                       MINV 300
C     ..................................................................MINV 310
C                                                                       MINV 320

      SUBROUTINE MINV(A,N,D,L,M)                                        MINV 330
      DIMENSION A(1),L(1),M(1)                                          MINV 340
C                                                                       MINV 350
C        ...............................................................MINV 360
C                                                                       MINV 370
C        IF A DOUBLE PRECISION VERSION OF THIS ROUTINE IS DESIRED, THE  MINV 380
C        C IN COLUMN 1 SHOULD BE REMOVED FROM THE DOUBLE PRECISION      MINV 390
C        STATEMENT WHICH FOLLOWS.                                       MINV 400
C                                                                       MINV 410
C     DOUBLE PRECISION A,D,BIGA,HOLD                                    MINV 420
C                                                                       MINV 430
C        THE C MUST ALSO BE REMOVED FROM DOUBLE PRECISION STATEMENTS    MINV 440
C        APPEARING IN OTHER ROUTINES USED IN CONJUNCTION WITH THIS      MINV 450
C        ROUTINE.                                                       MINV 460
C                                                                       MINV 470
C        THE DOUBLE PRECISION VERSION OF THIS SUBROUTINE MUST ALSO      MINV 480
C        CONTAIN DOUBLE PRECISION FORTRAN FUNCTIONS.  ABS IN STATEMENT  MINV 490
C        10 MUST BE CHANGED TO DABS.                                    MINV 500
C                                                                       MINV 510
C        ...............................................................MINV 520
C                                                                       MINV 530
C        SEARCH FOR LARGEST ELEMENT                                     MINV 540
```

```
C                                                      MINV 550
      D=1.0                                            MINV 560
      NK=-N                                            MINV 570
      DO 80 K=1,N                                      MINV 580
      NK=NK+N                                          MINV 590
      L(K)=K                                           MINV 600
      M(K)=K                                           MINV 610
      KK=NK+K                                          MINV 620
      BIGA=A(KK)                                       MINV 630
      DO 20 J=K,N                                      MINV 640
      IZ=N*(J-1)                                       MINV 650
      DO 20 I=K,N                                      MINV 660
      IJ=IZ+I                                          MINV 670
   10 IF( ABS(BIGA)- ABS(A(IJ))) 15,20,20              MINV 680
   15 BIGA=A(IJ)                                       MINV 690
      L(K)=I                                           MINV 700
      M(K)=J                                           MINV 710
   20 CONTINUE                                         MINV 720
C                                                      MINV 730
C        INTERCHANGE ROWS                              MINV 740
C                                                      MINV 750
      J=L(K)                                           MINV 760
      IF(J-K) 35,35,25                                 MINV 770
   25 KI=K-N                                           MINV 780
      DO 30 I=1,N                                      MINV 790
      KI=KI+N                                          MINV 800
      HOLD=-A(KI)                                      MINV 810
      JI=KI-K+J                                        MINV 820
      A(KI)=A(JI)                                      MINV 830
   30 A(JI) =HOLD                                      MINV 840
C                                                      MINV 850
C        INTERCHANGE COLUMNS                           MINV 860
C                                                      MINV 870
   35 I=M(K)                                           MINV 880
      IF(I-K) 45,45,38                                 MINV 890
   38 JP=N*(I-1)                                       MINV 900
      DO 40 J=1,N                                      MINV 910
      JK=NK+J                                          MINV 920
```

```
      JI=JP+J
      HOLD=-A(JK)
      A(JK)=A(JI)
   40 A(JI) =HOLD
C
C        DIVIDE COLUMN BY MINUS PIVOT (VALUE OF PIVOT ELEMENT IS
C        CONTAINED IN BIGA)
C
   45 IF(BIGA) 48,46,48
   46 D=0.0
      RETURN
   48 DO 55 I=1,N
      IF(I-K) 50,55,50
   50 IK=NK+I
      A(IK)=A(IK)/(-BIGA)
   55 CONTINUE
C
C        REDUCE MATRIX
C
      DO 65 I=1,N
      IK=NK+I
      HOLD=A(IK)
      IJ=I-N
      DO 65 J=1,N
      IJ=IJ+N
      IF(I-K) 60,65,60
   60 IF(J-K) 62,65,62
   62 KJ=IJ-I+K
      A(IJ)=HOLD*A(KJ)+A(IJ)
   65 CONTINUE
C
C        DIVIDE ROW BY PIVOT
C
      KJ=K-N
      DO 75 J=1,N
      KJ=KJ+N
      IF(J-K) 70,75,70
   70 A(KJ)=A(KJ)/BIGA
```

```
   75 CONTINUE                                                          MINV1310
C                                                                       MINV1320
C        PRODUCT OF PIVOTS                                              MINV1330
C                                                                       MINV1340
      D=D*BIGA                                                          MINV1350
C                                                                       MINV1360
C        REPLACE PIVOT BY RECIPRUCAL                                    MINV1370
C                                                                       MINV1380
      A(KK)=1.0/BIGA                                                    MINV1390
   80 CONTINUE                                                          MINV1400
C                                                                       MINV1410
C        FINAL ROW AND COLUMN INTERCHANGE                               MINV1420
C                                                                       MINV1430
      K=N                                                               MINV1440
  100 K=(K-1)                                                           MINV1450
      IF(K) 150,150,105                                                 MINV1460
  105 I=L(K)                                                            MINV1470
      IF(I-K) 120,120,108                                               MINV1480
  108 JQ=N*(K-1)                                                        MINV1490
      JR=N*(I-1)                                                        MINV1500
      DO 110 J=1,N                                                      MINV1510
      JK=JQ+J                                                           MINV1520
      HOLD=A(JK)                                                        MINV1530
      JI=JR+J                                                           MINV1540
      A(JK)=-A(JI)                                                      MINV1550
  110 A(JI) =HOLD                                                       MINV1560
  120 J=M(K)                                                            MINV1570
      IF(J-K) 100,100,125                                               MINV1580
  125 KI=K-N                                                            MINV1590
      DO 130 I=1,N                                                      MINV1600
      KI=KI+N                                                           MINV1610
      HOLD=A(KI)                                                        MINV1620
      JI=KI-K+J                                                         MINV1630
      A(KI)=-A(JI)                                                      MINV1640
  130 A(JI) =HOLD                                                       MINV1650
      GC TO 100                                                         MINV1660
  150 RETURN                                                            MINV1670
      END                                                               MINV1680
```

```
C                                                                       GMPR  10
C     ..................................................................GMPR  20
C                                                                       GMPR  30
C        SUBROUTINE GMPRD                                               GMPR  40
C                                                                       GMPR  50
C        PURPOSE                                                        GMPR  60
C           MULTIPLY TWO GENERAL MATRICES TO FORM A RESULTANT GENERAL   GMPR  70
C           MATRIX                                                      GMPR  80
C                                                                       GMPR  90
C        USAGE                                                          GMPR 100
C           CALL GMPRD(A,B,R,N,M,L)                                     GMPR 110
C                                                                       GMPR 120
C        DESCRIPTION OF PARAMETERS                                      GMPR 130
C           A - NAME OF FIRST INPUT MATRIX                              GMPR 140
C           B - NAME OF SECOND INPUT MATRIX                             GMPR 150
C           R - NAME OF OUTPUT MATRIX                                   GMPR 160
C           N - NUMBER OF ROWS IN A                                     GMPR 170
C           M - NUMBER OF COLUMNS IN A AND ROWS IN B                    GMPR 180
C           L - NUMBER OF COLUMNS IN B                                  GMPR 190
C                                                                       GMPR 200
C        REMARKS                                                        GMPR 210
C           ALL MATRICES MUST BE STORED AS GENERAL MATRICES             GMPR 220
C           MATRIX R CANNOT BE IN THE SAME LOCATION AS MATRIX A         GMPR 230
C           MATRIX R CANNOT BE IN THE SAME LOCATION AS MATRIX B         GMPR 240
C           NUMBER OF COLUMNS OF MATRIX A MUST BE EQUAL TO NUMBER OF ROWGMPR 250
C           OF MATRIX B                                                 GMPR 260
C                                                                       GMPR 270
C        SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED                  GMPR 280
C           NONE                                                        GMPR 290
C                                                                       GMPR 300
C        METHOD                                                         GMPR 310
C           THE M BY L MATRIX B IS PREMULTIPLIED BY THE N BY M MATRIX A GMPR 320
C           AND THE RESULT IS STORED IN THE N BY L MATRIX R.            GMPR 330
C                                                                       GMPR 340
C     ..................................................................GMPR 350
C                                                                       GMPR 360
```

```
      SUBROUTINE GMPRD(A,B,R,N,M,L)
      DIMENSION A(1),B(1),R(1)
C
      IR=0
      IK=-M
      DO 10 K=1,L
      IK=IK+M
      DO 10 J=1,N
      IR=IR+1
      JI=J-N
      IB=IK
      R(IR)=0
      DO 10 I=1,M
      JI=JI+N
      IB=IB+1
   10 R(IR)=R(IR)+A(JI)*B(IB)
      RETURN
      END
```

```
THE INPUT DATA,  MATRIX R(1)
 16.0   31.0    3.0    1.5    7.0   36.0   29.0   21.0    1.5    6.0   35.0   32.0
 35.0    7.0   16.0   20.0   32.0    2.0    9.0   33.0    1.0   31.0    3.0    6.0

THE INPUT DATA,  MATRIX R(2)
  4.0   26.0   15.0    5.0   22.0   17.0   12.0   30.0   23.0   27.0   10.5   25.0
 13.0   28.0   29.0   34.0   30.0   36.0   12.0    8.0   27.0   11.0   10.0   22.0

THE INPUT DATA,  MATRIX R(3)
 33.0   13.0   20.0   14.0   16.0   10.5    8.0   19.0    9.0   34.0   28.0   24.0
  5.0   23.0   17.0   26.0   21.0   18.0   19.0   25.0   24.0    4.0   15.0   14.0

VECTOR U(1)
      1.709870      -0.772361       1.566832

VECTOR U(2)
      0.674330      -0.061226       0.920113

VECTOR U(3)
      0.703415      -0.464538       0.456536

VECTOR U(*)
      0.972715      -0.412605       0.972966

MATRIX V(R), EQN. (31).

      0.660559      -0.270491       0.187034
     -0.270491       0.963211      -0.244185
      0.187034      -0.244185       0.662332

THE INVERSE MATRIX OF V(R)

      1.781610       0.411252      -0.351543
      0.411252       1.240149       0.341079
     -0.351543       0.341079       1.734835

THE TEST STATISTIC L = 21.633430
```

ACKNOWLEGMENTS

The author wishes to express his sincere gratitude to Dr. R. M. Rubison for his excellent guidance and creative advice in directing this work. To Dr. L. T. Fan goes a special debt of gratitude for his constant encouragement and financial support during the course of this study. Appreciation is also extended to Dr. R. F. Nassar and Dr. F. S. Lai for serving on the author's graduate comittee and for reviewing this report. Special gratitude is extended to his wife, Bernice M. Y. Wu, and his parents for their steady encouragement.

APPLICATIONS OF NONPARAMETRIC STATISTICS TO

MULTICOMPONENT SOLIDS MIXING

by

JUI-RZE TOO

B. S. in Chemical Engineering, National Taiwan University, 1972

M. S. in Chemical Engineering, Kansas State University, 1978

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1979

ABSTRACT

This study demonstrates the applicablity of nonparametric procedures to the analysis of mixing processes. In particular, multivariate nonparametric methods are used to evaluate the properties of a multicomponent solids mixture. Specific problems considered are:

(1) test of sampling techniques,

(2) a test of treatment effects,

(3) a test of the completely mixed state, and

(4) a test of a quality standard.

The usefulness of the proposed nonparametric techniques is amply demonstrated with both homogeneous and heterogeneous mixtures generated by a drum mixer. The techniques presented in this paper are also applicable to any other mixers.