DISCRIMINATION OF SEDIMENTARY ENVIRONMENTS BASED ON PARTICLE SIZE STATISTICS

by

DAH CHENG WU

B. S., National Taiwan University, 1962
M. S., University of Oklahoma, 1966
Ph. D., University of Oklahoma, 1969

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

1973

Approved by

Major Professor

TABLE OF CONTENTS

PAGE

LIST OF TABLES

LIST OF ILLUSTRATIONS

1. INTRODUCTION

1.1 Purpose and Scope of Investigation

Particle size distribution is a very significant textural property of clastic sediments because it may shed light on the genesis of sediments.

Though it is widely known that the particle size distribution of a substance obtained from mechanical disintegration follows either the lognormal law or the Rosin's law, sedimentary processes, transporting modes, and multiple sources cause the particle size of sediments to deviate from any known probability distribution.

This report essentially summarizes several published articles dealing with the use of particle size statistics for discrimination among sedimentary environments. Theoretical consideration of particle size distributions is followed by the discussion of problems concerning their application to sediments. A comparison of the graphic and moment methods for the estimation of the population moments of the particle size distribution is also presented. The description of particle size statistics in discriminating sedimentary environments is emphasized. Finally, the multivariate method of linear discriminant analysis is introduced and used to differentiate sedimentary environments for a set of real observed data. A discussion on tests of hypotheses of means and covariance matrices between two multivariate normal populations is also included. Discriminatory power and probability of

misclassification are briefly described as well.

1.2 Definition of Individual Particle and Working Units

The American Society for Testing Materials (ASTM), defines "individual particle" as " a minute unit of matter whose size and shape depends on the forces of cohesion". It is normally only a single crystal or a particle of regular shape with a specific gravity approximating that of a single crystal. "Working units" and "ultimate particles" are terms for the actual discrete units which may be individual particles or aggregates; they do not change their state of dispersion with time nor by the addition of peptizing agents throughout the determination procedure.

1.3 Principles of Particle Size Classification

The principles for the classification of particle sizes and the corresponding methods (Herdan, 1953, p. 35) are:

Principles	Methods
Geometrical Similarity	Sieving, Microscopic Examination
Similarity in Hydrodynamic Behavior	Sedimentation, Elutriation
Similarity in Optical Density	Light Extinction
Similarity in Surface Properties of Particle Aggregates	Permeability, Adsorption Methods

According to the principle of geometrical similarity, particles whose two dimensional images or whose three dimensional forms have

the same nominal diameter or nominal sectional diameter, respectively, are said to be of one size. Nominal diameter is the diameter of a sphere of the same volume as the particle. Nominal sectional diameter is the diameter of a circle of the same area as the projected image of the particle on its largest face. According to the principle of similarity in hydrodynamical behavior, particle size is calculated from Stokes' law; and particles with the same settling velocity in a fluid medium are considered to be of one size. These two principles and the corresponding methods of measurement are used by geologists to determine the particle size of sediments and sedimentary rocks.

2. PARTICLE SIZE DISTRIBUTIONS

2.1 Normal Distribution for Describing Particle Size

The probability law most popular in nature is the Gaussian (normal) law. One might be inclined to think that nature, though obeying very nicely the Gaussian law in the region of the very large objects and Maxwell's law in that of the small, has not learned its lesson too well in the intermediate region. Particle sizes of materials rarely show a normal distribution. As stated by Herdan (1953), normal distributions of particle size are found chiefly among the particulate substances produced by chemical processes such as condensation and precipitation.

The normal distribution of particle size is given by

$$f(x) = (\sigma\sqrt{2\pi})^{-1} \exp\left(-\frac{(x-M)^2}{2\sigma^2}\right)$$

where f(x) is the probability density, x the diameter of a particle, M the true arithmetic mean of particle diameter, and σ the true standard deviation of particle diameter.

2.2 Lognormal Distribution of Particle Size

If the logarithm of x is normally distributed, the distribution of x is said to be lognormal. The lognormal distribution is a more realistic representation of particle size of a substance than the normal distribution because large variations in particle size of substance are not uncommon. The logarithmic transformation

of observations of particle diameter would increase the accuracy and scope of estimating results.

According to the discussion of Kottler (1950a), the lognormal law can be derived from Galton's (1879) law of growth, $dx/dt = kx$, where dx/dt is the rate of growth of crystal size in diameter, t is time and k is the velocity constant. Because the scale of time has an arbitrary zero, we may assign $t = -\infty$ to the initial stage of crystallization and $t = +\infty$ to the end of the crystallization process. Integration of the above differential equation yields

$t = 1/k \log x - 1/k\ C$, where C is the integration

constant. If we assume t is a unit normal variable, i.e. $\log x \sim N(C,k^2)$, the probability density of the lognormal distribution of x is

$$f(x) = (x\sqrt{2\pi}\ k)^{-1} \exp\ (-1/2(\log x - C)^2/k^2) \qquad x > 0$$

If the law of growth is replaced by the law of decay, $dx/dt = -kx$, the lognormal distribution of size of particles by breakage from large fragments of substance can be derived in the same way. The law of decay can be deduced by assuming the relative rate of change dx/x is independent of x, if the time interval dt is kept constant. In other words, one assumes that the relative rate of decay in breakage is independent of the absolute quantity of material present.

Another advantage of the lognormal distribution over the normal distribution in describing particle size distribution is as follows:

It is often possible to measure size in terms of either diameter or volume of small objects of the same shape. If the distribution of the diameter is normal, then the distribution of volume will necessarily be asymmetrical. The normal law cannot be true in both cases. However, if the logarithms of the diameter are distributed normally with standard deviation, λ, the logarithms of the volume will be distributed normally with standard deviation 3λ.

Pulverized quartz, granite, calcite, limestone, soda, ash, sodium bicarbonate, alumina, and clay have size distributions which can be fitted satisfactorily by the lognormal law (Herdan, 1953).

2.3 Krumbein's Phi (Ø) Normal Distribution

The disparity between the extreme sizes is enormous in sediments; hence, particle sizes of sediments have to be measured on some type of geometric or logarithmic scale. Udden (1898) introduced a geometric scale based on a center of 1 mm and multiplier or divisor of 2. The names proposed by Udden to describe the size grades were modified by Wentworth (1922) to form the commonly used Udden-Wentworth grade scale.

In constructing histograms, we usually use equal width intervals in defining the classes. The wide range of size among sediments makes it generally impossible to use equal intervals without introducing an indefinitely large number of very small classes. As pointed out by Krumbein (1936), in most mechanical size analyses the data are assembled in classes according to the Udden-Wentworth grade scale. Thus, the class 1/2 - 1/4 mm is followed

on the right by the class 1/4 - 1/8 mm, an interval half its size. If the grade scale is plotted on ordinary arithmetic graph paper, the grades rapidly decrease in width as the scale is followed to the right due to the convention of plotting the coarser grades at the left. Similarly, a histogram or frequency curve plotted on such a scale will be quite asymmetrical, with much of the material assembled at one end or the other. However, when the logarithms of the diameters are used, the intervals become equal and the resulting frequency curve increases much in symmetry.

If logarithms to either base 10 or e are used, the class limits are non-integral. Thus, Krumbein (1936) suggested that logs to the base 2 be used. Then $\log_2 2 = +1$, $\log_2 1 = 0$, $\log_2 1/2 = -1$ etc. As a result of this procedure, negative values apply to diameters smaller than 1 mm. There probably are more fine-grained sediments (sands, silts, clays) than there are gravels in nature so that most analyses will lie in the range below 1 mm. This suggests that negative logs be used to avoid negative numbers in this important range and also to convert the grade scale to one which increases to the right as most ordinary scales do. The symbol ϕ (phi) is called the Wentworth Exponent by Krumbein (1936) and defined to be $-\log_2 \xi$, where ξ is the numerical value of the grain diameter in mm. Accordingly, if $-\log_2 X$ is normally distributed, the random variable X is said to be phi normally distributed. As a matter of fact, phi normal distribution is equivalent to the lognormal distribution because

$$\log_e X = \log_2 X / \log_2 e = -1/\log_2 e(-\log_2 X) = K(-\log_2 X);$$

where K is a constant.

2.4 Rosin and Rammler Distribution of Particle Size

A distribution function was proposed and applied to broken coal by Rosin and Rammler (1933). Rosin's distribution was later proposed to be applicable not only to coal but to cement, gypsum, magnesite, clay, dye-stuff, quartz, flint, glass, and ores. The mathematical expression of the law is as follows:

$$R = 100 \exp(-bx^n) \qquad 0 < x < \infty \ldots\ldots\ldots \qquad (1)$$

where R is the cumulative percentage of particle sizes larger than x (mean projected diameter), and b and n are distribution constants. Bennett (1936) defined $b = (1/\bar{x})^n$ to make the exponent a dimensionless number, where $\bar{x}$ is the absolute size constant which is the size for which the residue is $100 \exp - \left(\frac{\bar{x}}{\bar{x}}\right) = 36.78$. Thus, Rosin's equation becomes

$$R = 100 \exp -(x/\bar{x})^n \ldots\ldots\ldots\ldots\ldots\ldots\ldots \qquad (2)$$

Based upon the assumption that rock material is isotropic and that the probability of fracture of any grain in any particular place or direction is equal to the probability of fracture elsewhere, the Rosin's equation was derived theoretically by Bennett (1936). He also provided abundant experimental evidence to verify the validity of Rosin's law as applied to run-of-mine coal and the products of

subsequent washing and handling, etc. He concluded that Rosin's law safely can be applied to sizes from one five-hundredth of an inch up to three inches.

Dividing both sides of equation (2) by 100 and inverting, one obtains

$$100/R = \exp\ (x/\bar{x})^n \qquad (3)$$

If the logarithm of both sides of equation (3) be taken twice, one obtains

$$\text{loglog}\ 100/R = n(\log x - \log \bar{x}) \qquad (4)$$

$\log \bar{x}$ and n are constant for any particular distribution, so equation (4) may be written

$$\text{loglog}\ 100/R = C + n \log x \qquad (5)$$

where C is a constant.

It is possible to construct graph paper graduated logarithmically in accordance with the linear form of Rosin's equation (5) but calibrated with values of x (grain size) on the abscissa and values of R on the ordinate. Then data obtained from a population of particles whose sizes are distributed according to Rosin's law will lie on a straight line when plotted on the special graph paper. This graph paper was used by Kittleman (1964) to demonstrate hypothetical, artificial, and natural size frequency distributions. It forms the basis of a statistical test for goodness of fit to Rosin's distribution.

2.5 Discussion of Particle Size Distribution of Sediments

The sample statistics: mean, standard deviation, skewness and kurtosis are extensively used by geologists to describe and compare particle size distributions. As pointed out by Krumbein (1938), symmetrical phi size frequency curves are not uncommon among certain types of sediments, but curves having values of kurtosis close to 3.0 are comparatively rare among samples of sediments.

Folk and Ward (1957) stated "Non-normal kurtosis and skewness values are held to be the identifying characteristics of bimodal sediments even where such modes are not evident in the frequency curves." They are saying that non-normal skewness and kurtosis values are probably indications of two populations of grain sizes each of which is lognormally distributed.

After having examined some 260 grain size analyses from the Mississippi Delta, Spencer (1963) found that all sediments in the Mississippi Delta may be adequately described by two lognormal grain size populations either mixed together in different proportions or kept separate. The two populations are sand and clay. Thus, Spencer (1963), hypothesized that all clastic sediments are essentially mixtures of three or less fundamental populations of lognormal grain size distributions. These three populations are "gravel", "sand", and "clay". Accordingly, deviation from lognormality (phi normality) may be well-explained as a result of mixing two or more size populations, according to those authors.

Roger and Schubert (1963), on the other hand, proposed that populations of different sized material have different types of size

distribution because they are produced by different abrasional and fracturing mechanisms. References to experimental evidences and natural sediments were cited by them to support the following statements:

> "Pebble- or gravel-sized material formed by single-stage crushing without extensive transport should theoretically obey Rosin's size distribution. Pebble- and sand-sized material that have undergone extensive repeated abrasion should follow a log-normal size distribution. The fine-grained silt and clay that are produced by chipping from large grains are presumed to have a normal size distribution."

Middleton (1962) comments on size distribution and moment measures of geologic samples as follows:

> "In attempting to define a natural probability distribution, one may proceed in one of two ways: inductively or deductively. The inductive approach is to measure either the entire distribution or a sufficiently large part of it, and then to fit a curve either by method of least squares or some other method. The deductive way, is to deduce the distribution law by making certain initial assumptions about the mechanism which controls the natural processes."

As discussed in the proceeding chapter, Kottler (1950a and 1950b) deduced the lognormal distribution of particle sizes in photographic emulsions and Bennett (1936) deduced Rosin's distribution of broken coal.

A hypothesis based on deduction from principles should not be readily discarded simply because of a few exceptions. However, the hypothesis of a lognormal distribution of particle size of sediments based purely on convenience and general experience may be rejected if departure of observations from theory can be demonstrated.

With regard to the method of moments, Middleton (1962)

pointed out two drawbacks. First, it is possible to obtain different probability functions which have the same moments (even if one considers more than the first four moments). Second, for many distributions (e.g. highly skewed distributions), the sample moments are relatively inefficient estimators of the population moments.

In general, spot samples, such as used in Friedman's (1962) study, are collected for the size analysis of sediments. However, a spot sample is far from a random sample. Thus, no valid inference can be drawn from the size distribution of the spot sample about the size distribution of the whole population.

3. STATISTICAL ANALYSIS OF PARTICLE SIZE OF SEDIMENTS

3.1 Computation of Sample Moments of Size Analysis Data

In mathematical statistics, we may describe the properties of a set of data under study in terms of population moments. The kth moment of a distribution is defined as $n_k = \int_{-\infty}^{\infty} x^k f(x)\, dx$,

where f(x) is the density function of a continuous variate.

In practice, the density function f(x) is usually unknown. Thus, it is customary to take a sample of size N and classify the data by dividing sample space of x into equal class intervals in a manner similar to that used in constructing a histogram. The sample moment measures are then obtained by summing the individual moments for each class interval. A measure obtained in this manner is referred to as the sample moment and is defined as

$$n_k = 1/N \sum_{i=1}^{h} [(x_i)^k f_i]$$

where x_i is the value of the midpoint of the ith class interval, f_i is the class frequency, and h is the number of classes. If both N and h are very large, the sample moments approach the value of the corresponding theoretical moments. Sample moments about the mean are defined as

$$m_k = 1/N \sum_{i=1}^{h} [(x_i - M)^k f_i]$$

where $M = n_1$ is the arithmetic mean.

3.2 Distribution of a Percentile

Consider a sample of size N drawn from a one-dimensional continuous distribution with the cumulative distribution function $F(x)$ and the frequency function $f(x)$. Let λ_p denote the percentile of order p of the distribution; that is, $F(\lambda_p) = p$, where $0 < p < 1$. We also assume that in some neighborhood of $x = \lambda_p$, the frequency function is continuous and has a continuous derivative $f'(x)$.

Let $\phi = \phi_p$, a random variable, denote the corresponding sample percentile and $u = [Np]$ represents the largest integer smaller than Np if Np is not an integer. We define the probability that ϕ_p will lie in the infinitesmal region between x and x+dx to be $g(x)dx$, which is identical with the probability that, among the N sample values, $u = [Np]$ are $< x$ and $N-u-1$ are $> x+dx$, while the remaining value falls between x and x+dx. This is the sample distribution of a percentile which Cramér (1946, p. 368) expresses as follows:

$$g(x)dx = \binom{N}{u}(N-u)(F(x))^{u}(1-F(x))^{N-u-1}f(x)dx,$$

where $g(x)$ is the density of the random variable of order statistic $\zeta_{u+1}^{(N)}$. The distribution of $\phi = \phi_p$, the sample percentile, is asymptotically normally distributed with λ_p and variance $(p(1-p)/N(f(\lambda_p))^2)$ (see Cramér, 1946, p. 368-369). For a phi normal distribution with the parameters M_ϕ and σ_ϕ, the median is M_ϕ and we have $f(M_\phi) = 1/(2\pi\sigma_\phi^2)^{1/2}$. Thus, the median $\phi_{1/2}(\phi_{50})$ of a sample of N from the

distribution is asymptotically normal $(M_\phi, \sigma_\phi^2 \pi/2N)$.

3.3 Approximate Graphic Analogues to Moment Measures Based on Percentile

Statistical parameters such as standard deviation, skewness etc. obtained from moments may be the most relevant numerical descriptions of size frequency distributions because they are based on standard statistical relations and take into consideration the entire sediment distribution. The dependence of moment measures on the entire distribution is a limitation to their practical application to particle size analysis of sediments since mechanical analyses of sediments frequently result in open-ended curves, and do not give the coarse and fine limits of the distribution. In fact, the very fine-grained fractions of some sediments are technically unmeasurable by conventional sedimentation methods. Another disadvantage of the moment measures is the complex and time-consuming procedure required to compute them. However, this is now a less serious drawback considering the availability of high-speed computers.

Graphic moment measures were first introduced by Trask (1932). He defined the sorting coefficient and the skewness coefficient as $S_o = (Q_3/Q_1)^{1/2}$, $(Q_3 > Q_1)$ and $S_k = Q_1 Q_3/Md^2$, respectively, where Q_1 and Q_3 are the diameters in millimeters corresponding to the 25th and 75th percentiles respectively of a cumulative percent curve and Md is the median diameter. The Trask's quartile's measures and their equivalent in phi notation are easy to compute and hence have received rather wide usage. However, quartile measures are limited in value

because they are based on the central 50 percent of the sediment distribution and have no particular significance in the geometry of a normal curve.

The 16th and 84th percentiles of a normal cumulative frequency curve of particle size represent diameters one standard deviation either side of the mean and the 2 1/2th and 97 1/2th percentiles represent diameters two standard deviations either side of the mean. As many sediment distributions approximate a phi normal distribution or one of the family of curves derived from it, the use of percentiles based upon standard deviations seems more meaningful for the description of size frequency distribution than those based on quartiles used by Trask (1932).

As discussed in the last section, the sample percentile is asymptotically normally distributed with mean λ_p and variance $p(1-p)/N(f(\lambda_p))^2$. Errors of sampling are minimum at the median (50 percentile) and increase symmetrically away from the medium toward the higher and lower percentile for this normal distribution because $f(\lambda_p)$ is maximum at $\lambda_p = \lambda_{1/2}$. This theoretical result was confirmed by plotting the range in phi units for various percentiles from the cumulative curves of repeated analyses of each of several sediments (Figure 1).

Inspection of Figure 1 indicates that percentiles one standard deviation either side of the median (ϕ_{16} and ϕ_{84}) can be determined with almost the same accuracy as the median, and that there is appreciably greater inaccuracy in percentile measurements two standard deviations either side of the median ($\phi_{2\ 1/2}$ and $\phi_{97\ 1/2}$). However, the errors

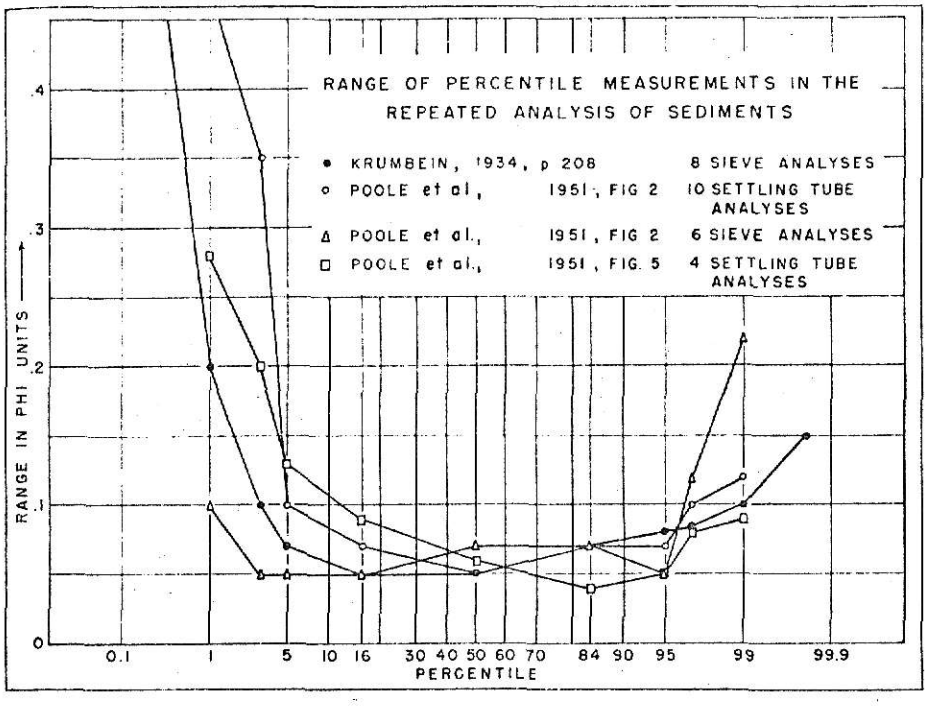

Figure 1. Range in Phi Units for Various Percentiles from the Cumulative Curves of Repeated Analyses of Each of Several Samples (after Inman, 1952).

in measuring the 5th and 95th percentiles are appreciably less than those for the 2 1/2th and 97 1/2th percentiles. In general, it would be extremely difficult to obtain the 97 1/2 percent diameter in analysis of fine-grained sediments. For these two reasons, the 5th and 95th percentile diameter, obtained from cumulative frequency curve is used by Inman (1952) as a working approximation to two standard deviations. The considerations discussed above led him to select five parameters to describe the frequency distribution of sediments (Table 1).

Because sample moments of distributions are more variable than theoretical measures, Inman (1952) decided to compare the approximate graphic measures with theoretical moment measures of known distributions. Inman's graphic parameters (M_ϕ, σ_ϕ, α_ϕ, $\alpha_{2\phi}$, and β_ϕ) of the Pearson Type I and Type II, Gram-Charlier Series and chi square curves were computed and compared with the corresponding theoretical moment measures. The reasons for selecting these four distributions are related to their frequent application to sediments. For each distribution, Inman (1952) varied the parameter of skewness and noted the effect on the phi deviation (σ_ϕ). He observed in every example that the ratio of the phi deviation to the standard deviation decreases as the absolute value of skewness increases. In other words, the degree of correlation of the graphic parameters to the corresponding moment measures decreases as the distribution becomes more skew. However, Inman (1952) said "For a fairly wide range of distributions, the first three moment measures can be ascertained from the graphic parameters with about the same degree of accuracy as is obtained by

Table 1. Descriptive Measure of Sediment-Size Distribution According to Several Authors (after Blatt et al., 1972).

Measure	Trask*	Inman	Folk and Ward
Median	$Md = P^{**}_{50}$	$Md_\phi = \phi^{***}_{50}$	$Md_\phi = \phi_{50}$
Mean	$M = \frac{P_{25} + P_{75}}{2}$	$M_\phi = \frac{\phi_{16} + \phi_{84}}{2}$	$M_z = \frac{\phi_{16} + \phi_{50} + \phi_{84}}{3}$
Dispersion (sorting)	$So = \frac{P_{75}}{P_{25}}$	$\sigma_\phi = \frac{\phi_{84} - \phi_{16}}{2}$	$\sigma_I = \frac{\phi_{84} - \phi_{16}}{4} + \frac{\phi_{95} - \phi_5}{6.6}$
Skewness	$Sk = \frac{P_{25}P_{75}}{Md^2}$	$\sigma_\phi = \frac{M_\phi - Md_\phi}{\sigma_\phi}$ $\sigma_{2\phi} = \frac{1/2(\phi_5 + \phi_{95}) - Md}{\sigma_\phi}$	$Sk_I = \frac{\phi_{16} + \phi_{84} - 2\phi_{50}}{2(\phi_{84} - \phi_{16})}$ $+ \frac{\phi_5 + \phi_{95} - 2\phi_{50}}{2(\phi_{95} - \phi_5)}$
Kurtosis	$K = \frac{P_{75} - P_{25}}{2(P_{90} - P_{10})}$	$\beta_\phi = \frac{1/2(\phi_{95} - \phi_5) - \sigma_\phi}{\sigma_\phi}$	$K_G = \frac{\phi_{95} - \phi_5}{2.44(\phi_{75} - \phi_{25})}$

*The formula for kurtosis was proposed by Krumbein and Pettijohn. Many workers have used the square root of Sk, rather than Sk itself, as a measure of skewness.

**P indicates a percentile measure, measured in millimeters.

***ϕ indicates a ϕ percentile.

computing rough moment measures".

Because sample percentiles are asymptotically normally distributed, a similar relation exists for a graphic measure which is a linear combination of percentiles. By assuming that the grain size is lognormally distributed, McCammon (1962) calculated the statistical efficiency of the graphic measures of mean and standard deviation. The statistical efficiency of a graphic measure is the ratio of the variance of the distribution of the corresponding efficient estimate (minimum variance unbiased estimate) and the variance of the limiting distribution of the graphic measure. The latter is discussed by Mosteller (1946, p. 377-408) in detail. The efficiencies of graphic mean and sorting proposed by Inman (1952), Folk and Ward (1957) and others are compared with those proposed by McCammon (1962) in Table 2.

Table 2, Efficiencies of Graphic Measures

Graphic Measures by McCammon	Efficiency	Graphic Measures by others	Efficiency
MEAN SIZE			
$(\phi_{20} + \phi_{50} + \phi_{80})/3$	88	$(\phi_{16} + \phi_{84})/2$	74
$(\phi_{10} + \phi_{30} + \phi_{50} + \phi_{70} + \phi_{90})/5$	93	$(\phi_{25} + \phi_{75})/2$	81
$(\phi_5 + \phi_{15} + \phi_{25} + \phi_{35} + \phi_{45} + \phi_{55} + \phi_{65} + \phi_{75} + \phi_{85} + \phi_{95})/10$	97	$(\phi_{16} + \phi_{50} + \phi_{84})/3$	88
SORTING			
$(\phi_{85} + \phi_{95} - \phi_5 - \phi_{15})/5.4$	79	$(\phi_{75} - \phi_{25})/1.35$	37
$(\phi_{70} + \phi_{80} + \phi_{90} + \phi_{97} - \phi_3 - \phi_{10} - \phi_{20} - \phi_{30})/9.1$	87	$(\phi_{84} - \phi_{16})/2$	54
		$(\phi_{95} - \phi_5)/3.3$	64
		$(\phi_{84} - \phi_{16})/4 + (\phi_{95} - \phi_5)/6.6$	79

Despite the higher efficiencies, McCammon's (1962) scheme has not been widely used by geologists. Some of the reasons could be as follows:

(1) Comparatively more percentiles are involved in the computation of graphic measures.

(2) His method for computing efficiency is valid only for phi normal distributions.

(3) The schemes of Inman (1952) and Folk and Ward (1957) have already been widely accepted and adopted in the geologic literature

Folk and Ward (1957) modify Inman's ϕ deviation measure by including in their parameters the 5th and 95th percentiles. Their measure is defined by the equation (see Table 1)

$$\sigma_I = (\phi_{84} - \phi_{16})/4 - (\phi_{95} - \phi_5)/6.6$$

Friedman (1962) did empirical correlation studies between three graphic sorting measures (Trask's, Inman's, and Folk and Ward's) and the standard deviation computed by the sample moment method. Correlation diagrams show that both the Trask coefficient and the Inman sorting measure approximate the standard deviation and that a high correlation exists between the standard deviation and sorting measure of Folk and Ward (1957).

4. PARTICLE SIZE STATISTICS AND SEDIMENTARY ENVIRONMENTS

4.1 Application of Particle Size Statistics to Differentiate Sedimentary Environments

A distinction between beach, dune, and river sands was shown numerically by Friedman (1961) by computing the moments. In terms of phi units, the skewness of beach sand is generally negative and those of dune and river sands are generally positive. A plot of mean grain size against skewness results in an almost complete separation of the fields representing dune sands and beach sands. The mineralogical composition of the sands does not affect the sign of skewness. In addition, dune sands tend to have smaller standard deviations than river sands. A plot of skewness and standard deviation also results in a good separation of the fields representing beach sands and river sands (Figure 2).

The upper size range of grains carried in suspension or by saltation during the transportation of river sands is governed by the competency of the transporting medium. But fine particles in transport are not affected by such a limitation. Accordingly, the lack of a "tail" at the coarse-grained end of the frequency distribution curve results in positive skewness of river sands. In beach sands, the fine-grained particles of sands are winnowed away by wave actions and coarse-grained particles are left behind. Thus, the distribution of beach sand appears to have a "chopped off" tail at the fine-grained end; that is, it resembles a truncated normal curve, with negative skewness. In the formation of dunes, the wind leaves

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.

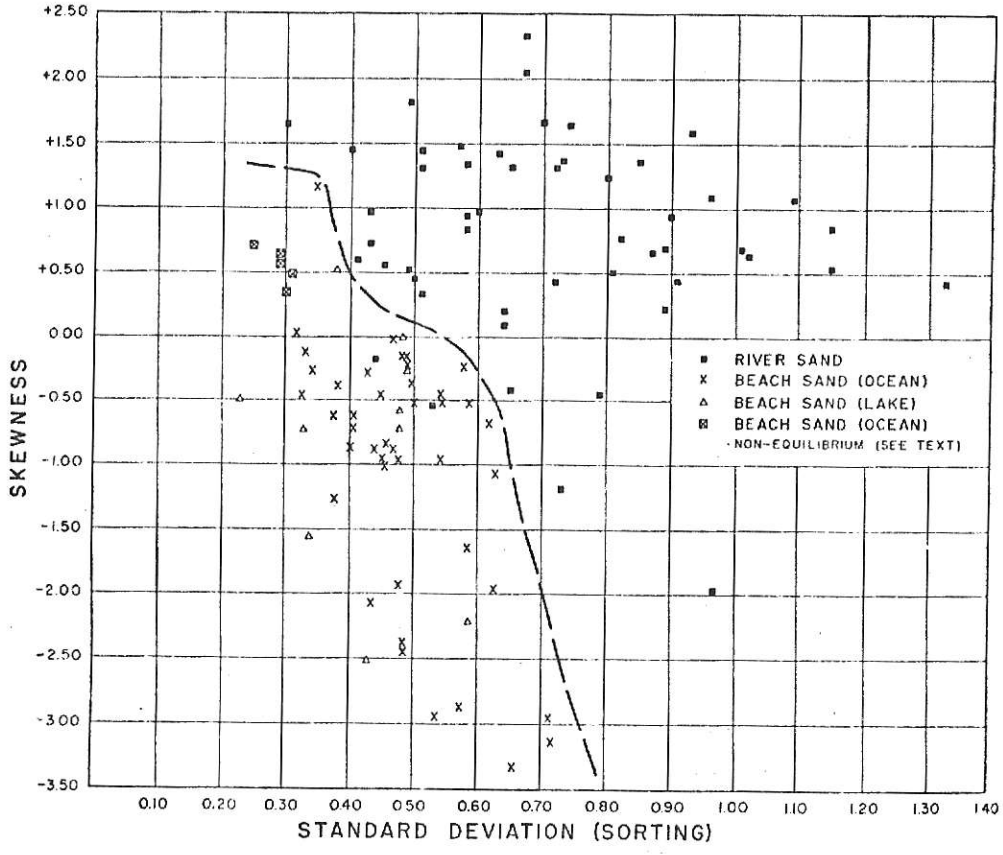

Figure 2. Plot of Skewness and Standard Deviation, Using Phi (Ø) Scale, for Beach and River Sands (after Friedman, 1961).

coarser sand grains behind because its strength is generally too low to move them. The truncation of the coarse tail results in a positive skewness of distribution of the dune sands. One would expect the dune sands and beach sands to be better sorted than river sands because of winnowing effect of wind and wave.

Moiola and Weiser (1968) analyzed textural parameters (mean diameter, standard deviation, skewness, kurtosis) calculated from a total 120 samples of modern beach, coast dune, inland dune, and river sands. These parameters were calculated by using the graphic method following formulae of Folk and Ward (1957) (Table 1). Plots of combinations of textural parameters were presented to demonstrate how effective they are in differentiating between samples from two different environments. The conclusions given by them are:

(1) The combination of mean diameter vs. skewness is most effective in differentiating between beach and inland dune sands and coastal dune sands.

(2) The combination of mean diameter vs. standard deviation is most effective in differentiating between beach and river sands, and between river and coastal dune sands.

(3) Skewness vs. kurtosis is never the most effective combination of parameters in discriminating environments.

Visher (1969) wrote on grain size distributions and depositional processes. The cumulative frequency percentage expressed as probability is plotted against the phi grain size of sands. The resulting plot usually exhibits two or three straight line segments (Figure 3). Visher (1969) stated that such segments represent

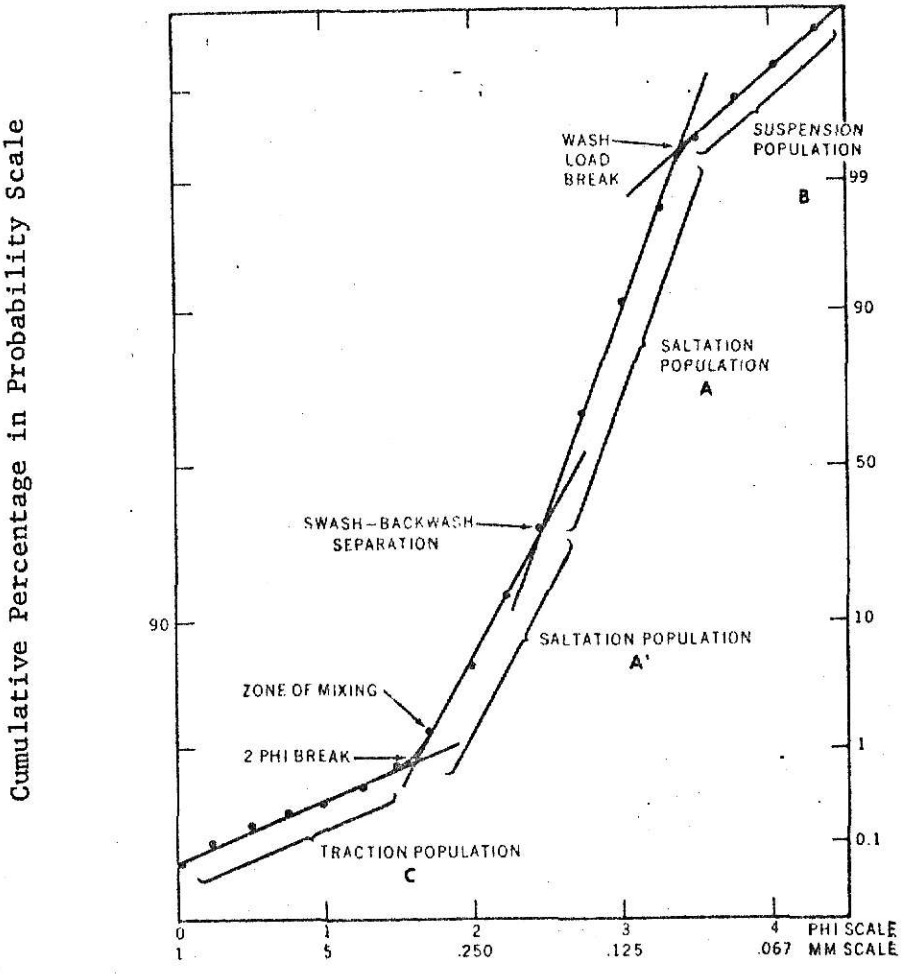

Figure 3. Example Showing Four Truncated Lognormal Populations in a Grain Size Distribution from Lower Swash Zone (after Visher, 1969).

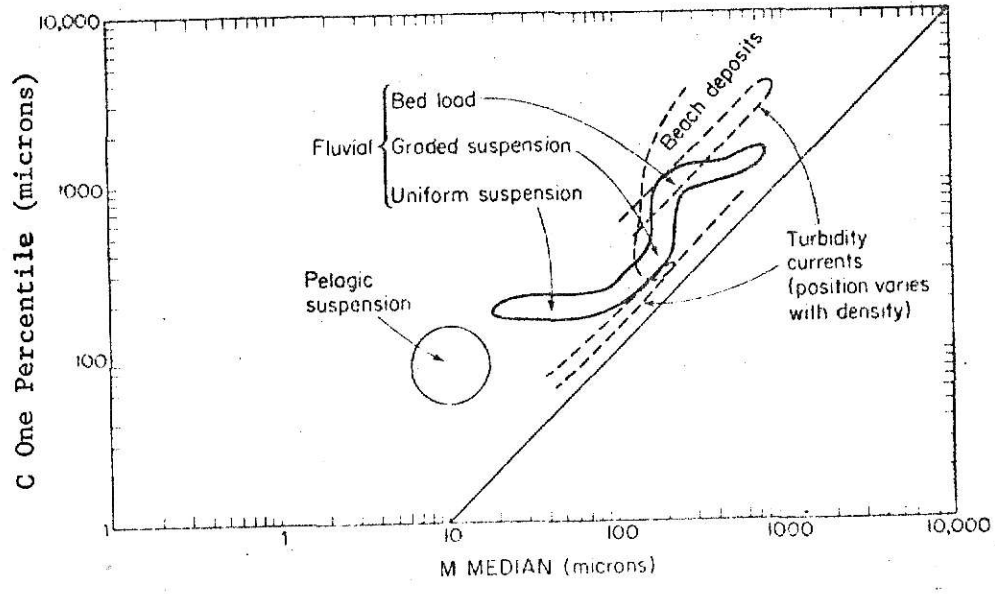

Figure 4. C-M Pattern of Sediments Deposited in

separate lognormal populations. Each population is truncated and jointed with the next population to form a single distribution. This means that a grain size distribution is composed of several lognormal populations each with a different mean and standard deviation. He also shows that this interpretation can be well substantiated by fluid mechanics concepts and observations. Transportation processes such as surface creep, saltation, and suspension may be reflected in a separate lognormal subpopulation within a single grain size distribution (Figure 3). Therefore, the sorting, size range, degree of mixing, and the points of truncation of these subpopulations may provide information on provenance, currents, waves, and rates of deposition. Using 1500 samples, these parameters were found by Visher (1969) to vary in a predictable and systematic manner and to have significance in terms of transport and deposition. He also claims that the ancient sands and sandstones can be compared with the modern sands of known environments of deposition (beach, river, channel, dune etc.) on the basis of similarity in curve shapes. The consistency of curve shapes from sample to sample produced by similar processes and that between ancient and modern analogues enable us to obtain an independent determination of the processes of formation of ancient sands.

It is known that the loads of fine- and coarse-grained sediments are largely independent of each other in such rivers as the Mississippi and Enoree. Thus, Passega (1957, 1964) stated that parameters such as the coefficients of sorting and skewness which measure attributes of the total sediment (the combination of coarse and fine

fractions), being apt to be a function of independent variables, cannot be chosen to express the character of the deposition. Therefore, he proposed two parameters C and M. The parameter C is defined as the diameter in microns corresponding to one percent of cumulative size frequency curve and is used as an approximation of the maximum grain size (a measure of "competency" of transporting medium). The parameter M is the median in microns or the 50 percentile and represents the average coarseness of the sediment. The C-M pattern (Figure 4) is very useful in determining the depositional environments of sediments.

By assuming the amount of sediment in each size class is a unique attribute of a particular sediment sample; that is, that there are as many variables as class intervals, Klovan (1966) applied factor analysis on the size analysis data of sixty-nine recent sediment samples collected from Barataria Bay, Louisiana. Three factors, which were interpreted as wind-wave energy, current energy, and gravitational energy, adequately described the grain size variations in these samples.

This technique offers several advantages over the method of plotting statistical (textural) parameters to determine sedimentary environments, namely,

(1) It makes use of the entire spectrum of grain size distribution.

(2) It is more objective.

(3) A priori knowledge of the environmental and geographic locations of the sediment samples is not required for classifying them into environmentally distinct facies.

4.2 Applicability of Grain Size Statistics in Determining Environments of Ancient Sediments

The main purpose of the study of modern sediments is to use information from known environments in classifying environments of ancient sediments. But the original size distribution of consolidated ancient sedimentary rocks can hardly be obtained, especially in the clay and silt size range, because of the problems of disaggregation, flocculation, dispersion, crushing, recrystallization, cleavage, etc. For example, disaggregation of lithified sands tends to break soft grains, producing spurious contributions to "fines". One advantage of the Passega's (1957, 1964) scheme, as discussed in the last section, is that it makes use mainly of the coarse and medium sizes, both of which may be fairly accurately estimated in lithified sandstones. As most ancient sandstones are more or less lithified, Passega's scheme may be better applied in the study of sedimentary environments of ancient sandstones.

Nearly one-thousand size distribution curves of ancient sediment samples have been studied by Visher (1969). He found that the major difference observed between ancient and modern grain size distribution is the amount of particles less than 44 microns in size. Some reasons suggested for this are:

(1) Diagenetic addition of clays

(2) Post-depositional mixing

(3) Sediment settling downward through the pores

(4) Possible transportation by interstitial fluids

By ignoring fines, Visher (1969) is still able to match ancient sediments with modern analogues deposited by similar processes on the basis of comparisons of their general grain-size distribution curves.

5. LINEAR DISCRIMINANT FUNCTION AND ITS APPLICATION TO DIFFERENTIATING SEDIMENTARY ENVIRONMENTS

5.1 Introduction

A linear discriminant function was introduced by Fisher (1936) as one solution for the problem of classifying an observation into one of several populations. In the case of two populations, an experimeter has an observation that he knows is from either population 1 or population 2, and measures p characteristics of the single sample. On the basis of these p measures, he decides from which population it probably was drawn. The discriminant function which serves for this purpose is obtained from two sets of data from known populations.

The discriminant function has been used fairly often by geologists. Emery and Griffiths (1954), and Griffiths (1957), used the method to distinguish oil-bearing and ore-bearing sediments from barren ones. McIntyre (1961) discriminated among three sedimentary environments by using heavy minerals. Porter et al. (1963) discriminated between marine and fresh-water argillaceous sediments by use of trace elements. Mellon (1964) distinguished between calcite and silica cement in sandstone by the use of six measured properties of the rock. Tien (1968) applied the discriminant function to differentiate among Pleistocene deposits of Kansas based on the relative abundance of three kinds of clay minerals. Sahu (1964) showed that sand from several environments can be distinguished on the basis of the four sample moments of particle size distribution by

discriminant analysis.

5.2 The Technique of Discrimination

Let $\underline{x}_{11}, \underline{x}_{12}, \ldots\ldots \underline{x}_{1N_1}$ be p-dimensional observation vectors from population π_1 and $\underline{x}_{21}, \underline{x}_{22}, \ldots\ldots \underline{x}_{2N_2}$ be p-dimensional observation vectors from population π_2, and let $\underline{x}$ be an observation vector of unknown classification but from either π_1 or π_2. Assume that the two populations are multivariate normally distributed with known means $\underline{\mu}_1$ and $\underline{\mu}_2$ and a common covariance matrix $\underline{\Sigma}$. As discussed in Anderson (1958, p. 133-134), the density of population i is

$$P_i(\underline{x}) = \frac{1}{(2\pi)^{p/2}|\underline{\Sigma}|^{1/2}} \exp[-1/2(\underline{x}-\underline{\mu}_i)'\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu}_i)] \ldots\ldots\ldots (1)$$

The ratio of densities for π_1 and π_2 is

$$\frac{P_1(\underline{x})}{P_2(\underline{x})} = \frac{\exp[-1/2(\underline{x}-\underline{\mu}_1)'\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu}_1)]}{\exp[-1/2(\underline{x}-\underline{\mu}_2)'\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu}_2)]}$$

$$= \exp\{-1/2[(\underline{x}-\underline{\mu}_1)'\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu}_1) - (\underline{x}-\underline{\mu}_2)'\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu}_2)]\} \ldots\ldots (2)$$

The region of classification into π_1 is the set of $\underline{x}$'s for which (2) is larger than, or equal to, a suitably chosen K according to a lemma by Neyman and Pearson (1933). Because the logarithmic function is monotonic increasing, the inequality can be written in terms of the logarithm of (2) as

$$-1/2[(\underline{x}-\underline{\mu}_1)'\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu}_1) - (\underline{x}-\underline{\mu}_2)'\underline{\Sigma}^{-1}(\underline{x}-\underline{\mu}_2)] \geq \log K \ldots\ldots\ldots (3)$$

The left side of equation (3) can be expanded as

$$-1/2[\underline{x}'\Sigma^{-1}\underline{x} - \underline{x}'\Sigma^{-1}\underline{\mu}_1 - \underline{\mu}_1'\Sigma^{-1}\underline{x} + \underline{\mu}_1'\Sigma^{-1}\underline{\mu}_1 - \underline{x}'\Sigma^{-1}\underline{x} + \underline{x}'\Sigma^{-1}\underline{\mu}_2 + \underline{\mu}_2'\Sigma^{-1}\underline{x} - \underline{\mu}_2'\Sigma^{-1}\underline{\mu}_2]$$

By rearranging the terms, we obtain

$$\underline{x}'\Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2) - 1/2(\underline{\mu}_1 + \underline{\mu}_2)'\Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2)$$

The first term is the discriminant function which is a linear function of the components of the observation vector. The decision rule is to classify x into population π_1 if

$$\underline{x}'\Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2) - 1/2(\underline{\mu}_1 + \underline{\mu}_2)'\Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2) \geq \log K \quad \ldots\ldots(4)$$

Otherwise, classify it into population π_2.

Assume that a priori probabilities h and 1-h are known such that a unit randomly chosen from the pooled population has a probability h of belonging to population π_1 and a probability of 1-h of belonging to population π_2 and $\ell(i;j)$ is the loss from classifying an observation from the jth population as one from the ith. Then, K is given by $K = (1-h)\ell(1;2)/h\ell(2;1)$. Thus, the classification rule states that the vector x should be assigned to population π_1 if

$$\underline{x}'\Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2) - 1/2(\underline{\mu}_1 + \underline{\mu}_2)'\Sigma^{-1}(\underline{\mu}_1 - \underline{\mu}_2) \geq \log \frac{(1-h)\ell(1;2)}{h\ell(2;1)} \quad \ldots\ldots(5)$$

In the particular case, if $h = 1/2, \ell(1;2) = \ell(2;1)$

$$\log[(1-h)\ell(1;2)/h\ell(2;1)] = \log 1 = 0$$

The equation (5) becomes

$$\underline{x}'\underline{\Sigma}^{-1}(\underline{\mu}_1 - \underline{\mu}_2) - 1/2(\underline{\mu}_1 + \underline{\mu}_2)'\underline{\Sigma}^{-1}(\underline{\mu}_1 - \underline{\mu}_2) \geq 0$$

for classifying $\underline{x}$ into π_1, otherwise, into π_2.

If $\bar{\underline{x}}_1$ and $\bar{\underline{x}}_2$ are the sample mean vectors for the two groups and $\underline{S}$ is the pooled estimate of $\underline{\Sigma}$, we wish to determine the linear combination of the responses $\underline{a}'\underline{x}$ which produces the largest critical t^2-ratio.

$$t^2(\underline{a}) = \frac{[\underline{a}'(\bar{\underline{x}}_1 - \bar{\underline{x}}_2)]^2 N_1 N_2/(N_1 + N_2)}{\underline{a}'\underline{S}\underline{a}}$$

where

$$\bar{\underline{x}}_1 = \frac{1}{N_1}\sum_{i=1}^{N_1} \underline{x}_{1i}, \quad \bar{\underline{x}}_2 = \frac{1}{N_2}\sum_{i=1}^{N_2} \underline{x}_{2i}$$

$$\underline{S} = \frac{\sum_{i=1}^{N_1} (\underline{x}_{1i} - \bar{\underline{x}}_1)(\underline{x}_{1i} - \bar{\underline{x}}_1)' + \sum_{i=1}^{N_2} (\underline{x}_{2i} - \bar{\underline{x}}_2)(\underline{x}_{2i} - \bar{\underline{x}}_2)'}{N_1 + N_2 - 2}$$

As pointed out by Morrison (1967, p. 130), the coefficient vector $\underline{a}$ is given by the homogeneous system of equations

$$\left[\frac{N_1 N_2}{N_1 + N_2}(\bar{\underline{x}}_1 - \bar{\underline{x}}_2)(\bar{\underline{x}}_1 - \bar{\underline{x}}_2)' - \lambda \underline{S}\right] \underline{a} = \underline{0},$$

where λ can be shown to be the maximum $t^2(\underline{a})$ or merely the sample Hotelling's T^2. Only one solution exists because the rank of the coefficient matrix is (p-1). The solution is $\underline{a} = \underline{S}^{-1}(\bar{\underline{x}}_1 - \bar{\underline{x}}_2)$ which maximizes $t^2(\underline{a})$ and thus $y = \underline{x}'\underline{S}^{-1}(\bar{\underline{x}}_1 - \bar{\underline{x}}_2) = c_1x_1 + c_2x_2 + \ldots c_px_p$ is the linear discriminant function.

The decision rule is as follows:

If an observation vector $\underline{x}$ is such that

$$\underline{x}'\underline{S}^{-1}(\bar{\underline{x}}_1 - \bar{\underline{x}}_2) - 1/2(\bar{\underline{x}}_1 + \bar{\underline{x}}_2)'\underline{S}^{-1}(\bar{\underline{x}}_1 - \bar{\underline{x}}_2) \geq \log \frac{(1-h)\ell(1;2)}{h\ell(2;1)}$$

classify it into π_1; otherwise into π_2.

5.3 Test of Equality of Covariance Matrices

Since the equality of covariance matrices among populations is presumed for the aforementioned discriminant function, it may be necessary to test this assumption prior to the discriminant analysis. Test procedures are discussed in Morrison (1967, p. 152-153).

Let $\underline{S}_1$, $\underline{S}_2$ be sample covariance matrices of $MVN_p(\mu\ \Sigma_i)$ based upon n_i degrees of freedom where $i = 1,2$.

To test H_o: $\underline{\Sigma}_1 = \underline{\Sigma}_2 = \Sigma$

versus H_1: $\underline{\Sigma}_1 \neq \underline{\Sigma}_2$

Let $\bar{\underline{S}} = \dfrac{n_1\underline{S}_1 + n_2\underline{S}_2}{n_1 + n_2}$ be the pooled estimate of $\underline{\Sigma}$ when H_o is true.

The test statistic $u = MC^{-1}$ is approximately distributed as a chi-square variate with $1/2p(p+1)$ degrees of freedom as n_i tends to infinity, where

$$M = (n_1 + n_2)\ \log|\bar{\underline{S}}| - (n_1\ \log|\underline{S}_1| + n_2\ \log|\underline{S}_2|)$$

$$C^{-1} = 1 - \frac{(2p^2 + 3p - 1)}{6(p+1)}\left(\frac{1}{n_1} + \frac{1}{n_2} - \frac{1}{(n_1 + n_2)}\right)$$

Reject H_o if the test statistic

$$u > \chi^2_{[\alpha;1/2p(p+1)]}$$

5.4 Test of Equality of Means

It may be desirable to test the hypothesis that two vector means between two populations are equal. If they are in fact equal or nearly equal, then a discriminant function based on these attributes probably will not be successful.

Let $\underline{x}_{11}, \underline{x}_{12}, \ldots\ldots \underline{x}_{1N_1}$ be a random sample of size N_1 from $MVN_p(\underline{\mu}_1, \underline{\Sigma})$, and $\underline{x}_{21}, \underline{x}_{22} \ldots\ldots \underline{x}_{2N_2}$ be a random sample of size N_2 from $MVN_p(\underline{\mu}_2, \underline{\Sigma})$.

Let's define the notations as follows:

$$\underline{X}_1 = \begin{bmatrix} \underline{x}'_{11} \\ \underline{x}'_{12} \\ \vdots \\ \underline{x}'_{1N_1} \end{bmatrix}_{N_1 \times p} \qquad \underline{X}_2 = \begin{bmatrix} \underline{x}'_{21} \\ \underline{x}'_{22} \\ \vdots \\ \underline{x}'_{2N_2} \end{bmatrix}_{N_2 \times p}$$

$$\hat{\underline{\mu}}_1 = \bar{\underline{x}}_1 = \begin{bmatrix} \bar{x}_{11} \\ \bar{x}_{12} \\ \vdots \\ \bar{x}_{1p} \end{bmatrix}_{p \times 1} \qquad \hat{\underline{\mu}}_2 = \bar{\underline{x}}_2 = \begin{bmatrix} \bar{x}_{21} \\ \bar{x}_{22} \\ \vdots \\ \bar{x}_{2p} \end{bmatrix}_{p \times 1}$$

Obtain the pooled estimate $\underline{S}$ of $\underline{\Sigma}$ as follows:

$$\underline{A}_1 = \underline{X}'_1 (\underline{I}_{N_1} - \frac{1}{N_1} \underline{J}_{N_1}) \underline{X}_1$$

$$\underline{A}_2 = \underline{X}'_2 (\underline{I}_{N_2} - \frac{1}{N_2} \underline{J}_{N_2}) \underline{X}_2$$

where $\underline{J}_{N_1}$ and $\underline{J}_{N_2}$ are $N_1 \times N_1$ and $N_2 \times N_2$ matrices with elements of unity respectively, $\underline{I}_{N_1}$ and $\underline{I}_{N_2}$ are identity matrices of order N_1 and N_2.

$$\underline{S} = \frac{\underline{A}_1 + \underline{A}_2}{N_1 + N_2 - 2}$$

Since $\bar{\underline{x}}_1 \sim MVN_p(\underline{\mu}_1, \frac{\underline{\Sigma}}{N_1})$ $\quad$ $\bar{\underline{x}}_2 \sim MVN_p(\underline{\mu}_2, \frac{\underline{\Sigma}}{N_2})$, $(\bar{\underline{x}}_1 - \bar{\underline{x}}_2) \sim MVN_p$

$$(\underline{\mu}_1 - \underline{\mu}_2, \frac{\underline{\Sigma}}{N_1} + \frac{\underline{\Sigma}}{N_2}) = MVN_p(\underline{\mu}_1 - \underline{\mu}_2, \frac{N_1 + N_2}{N_1 N_2} \underline{\Sigma})$$

To test $H_o(\underline{a}) : \underline{a}'(\underline{\mu}_1 - \underline{\mu}_2) = 0$

$$H_1(\underline{a}) : \underline{a}'(\underline{\mu}_1 - \underline{\mu}_2) \neq 0$$

They are equivalent to test $H_o : \underline{\mu}_1 = \underline{\mu}_2$ vs. $H_1 : \underline{\mu}_1 \neq \underline{\mu}_2$

because the original multivariate hypothesis is true if and only if $H_o(\underline{a})$ holds for all nonnull $\underline{a}$.

The test statistic is

$$t(\underline{a}) = \frac{\underline{a}'(\bar{\underline{x}}_1 - \bar{\underline{x}}_2)\sqrt{\frac{N_1 N_2}{N_1 + N_2}}}{\sqrt{\underline{a}'\underline{S}\,\underline{a}}}$$

The acceptance region is of course

$$t^2(\underline{a}) \leq t^2_{\beta/2,(N_1 + N_2 - 2)}$$

The multivariate acceptance region is the intersection of all the univariate acceptance regions or

$$\bigcap_{\underline{a}} [t^2(\underline{a}) \leq t^2_{\beta/2,(N_1 + N_2 - 2)}].$$

The above intersection is equivalent to

$$\underset{\underline{a}}{\text{Max}}\ t^2(\underline{a}) \leq t^2_{\beta/2,(N_1 + N_2 - 2)}$$

Rao's (1965, p. 48) theorem states as follows:

> "Let $\underline{A}$ be a positive definite $p \times p$ matrix and $\underline{u} \in E_p$, that is $\underline{u}$ is an p vector. Then Sup $(\underline{u}'\underline{x})^2/\underline{x}'\underline{A}\underline{x} = \underline{u}'\underline{A}^{-1}\underline{u}$"

$$t^2(\underline{a}) = \frac{N_1 N_2}{N_1 + N_2} \frac{[(\bar{\underline{x}}_1 - \bar{\underline{x}}_2)'\underline{a}]^2}{\underline{a}'\underline{S}\underline{a}}$$

where $\underline{\mu} = (\bar{\underline{x}}_1 - \bar{\underline{x}}_2)$

$\underline{x} = \underline{a}$

$\underline{A} = \underline{S}$

Thus,

$$\text{Sup}\ t^2(\underline{a}) = \text{Max}\ t^2(\underline{a}) = \frac{N_1 N_2}{N_1 + N_2} (\bar{\underline{x}}_1 - \bar{\underline{x}}_2)'\underline{S}^{-1}(\bar{\underline{x}}_1 - \bar{\underline{x}}_2) = T^2$$

where T^2 is called the two-sample Hotelling's T^2 statistic.

When the null hypothesis is true, the quantity $\frac{(N_1 + N_2 - p - 1)}{(N_1 + N_2 - 2)p}\ T^2$

has F distribution with degrees of freedom p and $N_1 + N_2 - p - 1$

(Morrison, 1967, p. 126).

Reject the $H_o: \underline{\mu}_1 = \underline{\mu}_2$, if and only if

$$T^2 \geq T^2_{\alpha;p,N_1 + N_2 - p - 1} = \frac{(N_1 + N_2 - 2)p}{(N_1 + N_2 - p - 1)} F_{\alpha;p,N_1 + N_2 - p - 1}$$

The ellipsoidal confidence region of $100(1 - \alpha)\%$ confidence is specified by those vector $\underline{\int}$ satisfying the inequality

$$(\bar{\underline{x}}_1 - \bar{\underline{x}}_2 - \underline{\delta})' \underline{S}^{-1} (\bar{\underline{x}}_1 - \bar{\underline{x}}_2 - \underline{\delta}) \leq \frac{N_1 + N_2}{N_1 N_2} T_\alpha^2; p, N_1 + N_2 - p - 1$$

where

$$\underline{\delta} = \underline{\mu}_1 - \underline{\mu}_2$$

$100(1 - \alpha)\%$ simultaneous confidence intervals about all linear combination of $\underline{\delta}$ say $\underline{a}'\underline{\delta}$ for any choice of $\underline{a}$ are $\underline{a}'(\bar{\underline{x}}_1 - \bar{\underline{x}}_2) - \sqrt{\underline{a}'\underline{S}\underline{a}(\frac{1}{N_1} + \frac{1}{N_2})}$

$$T_\alpha; p, N_1 + N_2 - p - 1 \leq \underline{a}'\underline{\delta} \leq \underline{a}'(\bar{\underline{x}}_1 - \bar{\underline{x}}_2) + \sqrt{\underline{a}'\underline{S}\underline{a}(\frac{1}{N_1} + \frac{1}{N_2})}\ T_\alpha; p, N_1 + N_2 - p - 1$$

Simultaneous confidence intervals may show us which of p variables have contributed to the significant T^2, provided the null hypothesis is rejected. Thus, simultaneous confidence intervals may provide information for the purpose of deleting those insignificant variables employed in a discriminant function if it is necessary. But they also are quite conservative because they maintain the α for all decisions.

5.5 Test for Discriminatory Power

After a discrimination procedure has been established, it is of considerable interest to determine whether the discriminator is really useful. As discussed in Press (1972, p. 381-382), confusion matrices, which were defined by Massey (1965) as a table of correct and incorrect classifications, are used for this purpose. The confusion matrix has entries N_{ij} which are the numbers of observations known to belong to population π_i but were classified into population π_j.

$$C = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$$

where $N_{11} + N_{12} = N_1$

$N_{21} + N_{22} = N_2$

Let $N = N_{11} + N_{22}$ be the number of correct classification and $\bar{N} = N_{12} + N_{21}$ be the number of misclassifications. Classifying observations by chance, we expect to have $\frac{N_1 + N_2}{2}$ correctly classified and $\frac{N_1 + N_2}{2}$ misclassified. Compute

$$\chi^2 = \frac{(N - \frac{N_1 + N_2}{2})^2 + (\bar{N} - \frac{N_1 + N_2}{2})^2}{\frac{N_1 + N_2}{2}} = \frac{(N_1 + N_2 - 2N)^2}{N_1 + N_2}$$

Test H_o:Hits took place at random

H_1:The discrimination procedure did better than just chance

Reject H_o iff $\chi^2 > \chi_{\alpha,1}$

5.6 Probability of Misclassification

Under the assumption of multivariate normality, the probability of misclassifying $\underline{x}$ when $\underline{\mu}_1$, $\underline{\mu}_2$ and $\underline{\Sigma}$ are known and used to calculate the discriminant function, is $\Phi(-D/2)$ where Φ is the distribution function of a normal random variable with zero mean and unit variance and

$$D^2 = (\underline{\mu}_1 - \underline{\mu}_2)' \underline{\Sigma}^{-1} (\underline{\mu}_1 - \underline{\mu}_2)$$

which is known as the Mahalanobis distance between two populations.

For large samples, D^2 can be estimated by $(\bar{x}_1 - \bar{x}_2)' S^{-1} (\bar{x}_1 - \bar{x}_2)$

By assuming equal loss and h = 1/2, the estimated probabilities of misclassification are as follows: The probability of misclassifying an observation from population π_2 into population π_1 is $P(z \leq - D/2)$ and the probability of misclassifying an observation from π_1 into π_2 is $P(z \geq D/2)$.

5.7 Differentiation of Sedimentary Environments by Linear Discriminant Analysis of Particle Size

Moiola and Weiser (1969) indicated that the application of the discriminant function is a reliable and effective way to discriminate among size distributions of sands collected from several different environments. The basic goal of this technique is to use two sets of data whose origins are known to seek that linear combination of p measured characteristics that provides criteria to correctly predict the origin of any future observation. Moiola and Weiser (1969) claim that application of discriminant analysis to whole Ø grain size data is even more effective in differentiating between modern beach, coast dune, inland dune, and river sands than textural parameters calculated from quarter Ø data. This technique is also effective in determining depositional environments of ancient sandstone bodies.

Fourteen sand samples, half of which are river sands and half are dune sands, collected at Hunters Island and vicinity, Riley County, Kansas, were used to run the discriminant analysis for demonstration purposes. The results of sieve analyses performed by graduate students, Department of Geology, Kansas State University are shown in Table 3.

Table 3. Mechanical Size Analysis of Data on River Sand and Dune Sand Collected at Hunters Island and Vicinity (in weight percent)

Sample No. \ Size Classes		< -1 ø	-1-0 ø	0-1 ø	1-2 ø	2-3 ø	3-4 ø	> 4 ø
R- 1	River Sand	7.03	13.32	42.13	30.31	5.47	1.24	0.5
R- 3		25.55	22.73	27.59	19.36	3.44	0.92	0.41
R- 4		21.26	23.69	27.45	23.78	2.47	1.1	0.25
R- 5		6.13	15.79	39.37	33.24	2.35	1.97	1.15
R- 6		15.12	17.7	31.15	25.63	3.35	3.11	3.94
R- 7		10.06	19.05	36.25	24.36	2.38	2.94	4.96
R- 8		10.91	17.57	37.91	30.05	2.07	1.14	0.35
D- 2	Dune Sand	0	0	0.5	42.99	51.05	4.6	0.86
D- 4		0	0.04	1.18	32.06	33.97	21.71	11.04
D- 5		0	0	3.13	54.66	29.04	10.14	3.03
D- 6		0	0.09	0.35	2.47	24.17	52.86	20.06
D- 8		0	0	1.72	48.19	29.01	12.89	8.19
D- 9		0	0	0.08	8.27	52.95	30.3	8.4
D-13		0	0.02	0.46	22.12	56.1	17.32	3.98

The original quarter ϕ data are regrouped into whole ϕ data in this Table. The exact sample localities are in Husain (1964).

We assume:

(1) The weight percentage in the seven ϕ size classes has a multivariate normal distribution with different mean vectors and a common covariance matrix between river sand and dune sand.

(2) All sample observations are independent.

(3) A randomly chosen unit from the pooled population has a probability of 1/2 of belonging to river sand or dune sand.

(4) The loss functions, when we classify an object to be river or dune sand, are equal.

The discriminant function calculated by using the APL computing procedure on the data in Table 3 is

$$y = 15.43x_{(<-1)} + 28.23x_{(-1-0)} + 39.29x_{(0-1)} - 4.28x_{(1-2)}$$

$$+ 0.12x_{(2-3)} - 7.11x_{(3-4)} + 10.02x_{(>4)}$$

If $y = \underline{x}'\underline{S}^{-1}(\bar{\underline{x}}_r - \bar{\underline{x}}_d) \geq 1/2(\bar{\underline{x}}_r + \bar{\underline{x}}_d)'\underline{S}^{-1}(\bar{\underline{x}}_r - \bar{\underline{x}}_d) = 914.27$,

we classify it as river sand, otherwise dune sand, where $\underline{x}$ is a vector of observation, $\bar{\underline{x}}_r$ and $\bar{\underline{x}}_d$ are the sample mean vector of river sand and dune sand respectively, and $\underline{S}$ is the pooled sample covariance matrix. The result shows that all the fourteen sand samples of known origin can be correctly classified into the right group by using the discriminant function. By applying Mahalanobis' D^2 method, the estimated probability of misclassifying a river sand or a dune sand into the wrong population is $P(z > 23.12) = P(z < -23.12) \cong 0$ where z is unit normal variate.

Advantages of this method over the statistical parameters method in differentiating sedimentary environments are as follows:

(1) This technique is much more effective because it seeks to accomplish for p measured variables what is accomplished in two variables by plotting a scatter diagram and drawing lines on the diagram that best separate the different groups.

(2) The discriminant function based upon whole Ø data is more reliable and effective in differentiating sedimentary environments than textural parameters computed from quarter Ø (Moiola and Weiser, 1969). Thus, it would save time in size analysis to use this method.

(3) As pointed out by Visher (1969), the main difference observed between modern and ancient grain size distributions is the amount of fine of less than 44 microns in size. By ignoring the fine fractions, the discriminant function probably can effectively differentiate the sedimentary environments as well. Whereas, we may not be able to calculate the required moment measures (statistical parameters) if the fine fractions are ignored. Hence, a discriminant function is an effective technique for ancient environment discrimination.

6. CONCLUSIONS

Lognormal and Rosin's distributions were found to be applicable to distributions of particle size in naturally occurring aggregates. Sedimentary materials that have not been transported theoretically follow Rosin's size distribution law. Transported sediments are generally believed to have an approximate lognormal distribution, even though very few samples show a very close approximation to the ideal distribution. Many authors interpret all observed size distributions in terms of mixtures of lognormal distributions. A fully satisfactory mathematical model for the origin of the lognormal distribution in sediments has not yet been proposed.

Sample moments and their graphic equivalents were widely used to describe and compare observed size distributions of sediments. Plotting a scatter diagram of two of the four sample moments and drawing lines on the diagram that best separate the different environment groups of sands effectively separate modern sand groups of distinct origin as claimed by many authors. The discriminant function, which is based upon p observed variables, is much more powerful and effective than the two dimensional scatter diagram in discriminating sedimentary environments by particle size. Another advantage of discriminant analysis of particle size is its effectiveness in differentiating ancient sediments of different origin.

With the availability of high speed modern computers, the discriminant analysis is going to replace the scatter diagram plotting in differentiating sedimentary environments in the future. As pointed

out by Blatt, Middleton and Murray (1972), the size characteristic of sands on the continental shelf and in offshore bar environments has not been carefully studied yet. More research on size characteristic of modern offshore sands and ancient analogues is expected and discriminant function should play an important role in solving the classification problem.

REFERENCES

Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis, John Wiley & Sons Inc., New York.

Bennett, J. G. (1936). "Broken coal", Journal of Institute of Fuel, 10, 22-29.

Blatt, H., Middleton, G. and Murray,R. (1972). Origin of Sedimentary Rocks, Prentice-Hall Inc., New Jersey.

Cramér, H. (1946). Mathematical Methods of Statistics, Princeton University Press, Princeton.

Davis, J. C. (1972). Statistics and Data Analysis in Geology, John Wiley & Sons Inc., New York.

Emery, J. R., and Griffiths, J. C. (1954). "Differentiation of oil-bearing from barren sediments by quantitative petrographic analysis, Producers Monthly, 19:2, 33-37.

Folk, R. L. and Ward, W. C. (1957). "Brazos River Bar: A study in the significance of grain size parameters", Journal of Sedimentary Petrology, 27, 3-27.

Friedman, G. M. (1961). "Distinction between dune, beach, and river sands from their textural characteristics", Journal of Sedimentary Petrology, 31, 514-529.

_____. (1962). "On sorting coefficients and lognormality of the grain size distribution of sandstones", Journal of Geology, 70, 737-753.

Gaddum, J. H. (1945). "Lognormal distributions", Nature, 156, 463-466.

Galton, F. (1879). "The geometric mean in vital and social statistics", Proceedings of the Royal Society of London, 29, 365-367.

Griffiths, J. C. (1957). "Petrographical investigation of the Salt Wash sediments", U. S. Atomic Energy Comm., RME-3151.

Herdan, G. (1953). Small Particle Statistics, Butterworth's, London.

Husain, A. (1964). Differentiation Between River and Dune Sand, Hunters Island and Vicinity, Riley County, Kansas, Unpublished M. S. thesis, Kansas State University.

Inman, D. L. (1952). "Measures for describing the size distribution of sediments", Journal of Sedimentary Petrology, 22, 125-145.

Kittleman, L. R. (1964). "Application of Rosin's distribution to size frequency analysis of clastic rocks", Journal of Sedimentary Petrology, 34, 483-502.

Kloven, J. E. (1966). "The use of factor analysis in determining depositional environment from grain size distributions", Journal of Sedimentary Petrology, 36, 115-125.

Kottler, F. (1950a). "The distribution of particle sizes, Part I", Journal of the Franklin Institute, 250:4, 339-356.

_____. (1950b). "The distribution of particle sizes, Part II", Journal of the Franklin Institute, 250:5, 419-441.

Krumbein, W. C. (1936). "Application of logarithmic moments to size frequency distribution of sediments," Journal of Sedimentary Petrology, 6, 35-47.

_____. (1938). "Size frequency distribution of sediments and the normal phi curve", Journal of Sedimentary Petrology, 8,84-90.

Krumbein, W. C. and Graybill, F. A. (1965). An Introduction to Statistical Models in Geology, McGraw-Hill Book Co., New York.

Massey, W. F. (1965). "On methods: discriminant analysis of audience characteristics", Journal of Advertising Research, 5, 39-48.

McCammon, R. B. (1962). "Efficiencies of percentile measures for describing the mean size and sorting of sedimentary particles", Journal of Geology, 70, 453-465.

McIntyre, D. D. (1961). "A comparison of three association environments, glacial till, fluvioglacial delta and beach sand in terms of their quartz, garnet, and hornblende grains", Pennsylvania State University, Mineral Industries Experimental Station.

Mellon, G. B. (1964). "Discriminatory analysis of calcite- and silicate-cemented phases of the Mountain Park Sandstone", Journal of Geology, 72, 786-809.

Middleton, G. M. (1962). "On sorting coefficients and the lognormality of the grain size distribution of sandstone: a discussion", Journal of Geology, 70, 754-756.

Moiola, R. J. and Weiser, D. (1968). "Textural paramenters: an evaluation", Journal of Sedimentary Petrology, 38, 45-53.

_____. (1969). "Environmental analysis of ancient sandstone bodies by discriminant analysis", Bulletin of the American Association of Petroleum Geologists, 53, 733 (Abstract).

Morrison, D. F. (1967). Multivariate Statistical Methods, McGraw-Hill Book Co., New York.

Mosteller, F. (1946). "On some useful inefficient statistics", Annals of Mathematical Statistics, 17, 377-408.

Neyman, J. and Pearson, E. S. (1933). "On the problem of the most efficient tests of statistical hypothesis", Philos. Trans. Royal Society of London, Series A, 231, 289-337.

Passega, R. (1957). "Texture as characteristic of clastic deposition", Bulletin of the American Association of Petroleum Geologists, 41:9, 1952-1984.

_____. (1964). "Grain size representation by CM patterns as a geological tool", Journal of Sedimentary Petrology, 34, 830-847.

Porter, P. E., et al. (1963). "Trace elements in marine and fresh water argillaceous sediments", Geochim. Cosmochim. Acta, 27, 669-694.

Press, S. J. (1972). Applied Multivariate Analysis, Holt, Rinehart and Winston Inc., New York.

Rao, S. C. (1965). Linear Statistical Inference and its Applications, John Wiley & Sons, Inc., New York.

Roger, J. J. W. and Schubert, C. (1963). "Size distribution of sedimentary rocks", Science, 141, 801-802.

Rosin, Paul and Rammler, E. (1933). "Laws governing the fineness of powdered coal", Journal of Institute of Fuel, 7, 29-36.

Roy, S. N. (1957). Some Aspects of Multivariate Analysis, John Wiley & Sons, Inc., New York.

Sahu, B. K. (1964). "Depositional mechanisms from the size analysis of clastic sediments", Journal of Sedimentary Petrology, 34, 73-83.

Spencer, D. W. (1963). "The interpretation of grain size distribution curves of clastic sediments", Journal of Sedimentary Petrology, 33, 180-190.

Tien, P. L. (1968). "Differentiation of Pleistocene deposits in Northeastern Kansas by clay minerals", Journal of Clays and Clay Minerals, 16, 99-107.

Trask, P. D. (1932). Origin and Environment of Source Sediments of Petroleum, Gulf Pub. Co., Houston.

Udden, J. A. (1898). "Mechanical composition of wind deposits", Augustana Library Publications, No. 1.

Visher, G. S. (1969). "Grain size distributions and depositional processes", Journal of Sedimentary Petrology, 39, 1074-1106.

Wentworth, C. K. (1922). "A scale of grade and class terms for clastic sediments", Journal of Geology, 30, 377-392.

ACKNOWLEDGEMENTS

The author is indebted to Dr. Holly C. Fryer who directed the report and gave invaluable suggestions and criticisms.

The writer also wishes to express his sincere appreciation to Dr. Ray A. Waller and Dr. Page C. Twiss for reading and correcting the manuscript and offering suggestions for its improvement.

DISCRIMINATION OF SEDIMENTARY ENVIRONMENTS BASED ON PARTICLE SIZE STATISTICS

by

DAH CHENG WU

B. S., National Taiwan University, 1962
M. S., University of Oklahoma, 1966
Ph. D., University of Oklahoma, 1969

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

1973

Probability distributions of particle size and their application to differentiate sedimentary environments of sediments are examined. A lognormal distribution of particle size can be deduced theoretically from the law of growth. Rosin's distribution function of particle size has been derived by assuming that the sediment is isotropic and that the probability of fracture of any grain in any particular place or direction is equal to the probability of fracture elsewhere. Particle size of sediments is generally believed to have a lognormal distribution to the first approximation, even though few samples show a very close approximation to the ideal distribution. Some authors think that populations of different sized sediment have different types of size distribution because the populations are produced by different abrasional and fracturing mechanisms. Other authors claim that all clastic sediments are essentially mixtures of several fundamental populations of lognormal grain size distribution due to different transporting processes for different sized sediments, or to multiple sources. Another problem is that it is not possible to collect a random sample from a sedimentary population (e.g. a beach). Therefore, no valid inference can be drawn from the size distribution of nonrandom samples about the size distribution of the population as a whole.

Sample moments and their graphic equivalents calculated from size analysis data are widely accepted by geologists for describing and comparing observed particle size distributions of sediments. If we assume that the particle size distribution of sediments is lognormal,

efficiency of different graphic moment measures can be computed. However, sample moments cannot provide a complete summary of the data contained in the distribution. It is possible to have distributions that have exactly the same four moment measures but differ significantly in some respect.

Plotting observed values of two sample moments in a two dimensional scatter diagram and drawing lines that best separate the sand groups of different origin have been used by many authors to discriminate among modern sand environments. The application of discriminant functions is much more effective for differentiating among size distributions of samples of sand collected from several different environments because discriminant functions are based on p observed variables instead of two. Another advantage of discriminant analysis of particle size is its applicability to effective discrimination among different environments of ancient sediments.

Size analysis data of dune sands and river sands collected at Hunters Island, Riley County, Kansas, were used to illustrate the discriminant analysis. All samples of known origin can be correctly classified into the right group by the computed discriminant function.