NOV 9 1964

83214

MASTER

ON THE PROPER USE OF TRANSFORMATIONS OF LOG NORMAL FUNCTIONS
IN SMALL-PARTICLE STATISTICS

A. Goldman, H. D. Lewis and R. H. Moore

ABSTRACT

A standard notation is suggested for use in all fields concerned with small particle data analysis. The representation of particle size data by various forms of the log normal function is discussed and related to methods of "particle size" measurement. Transformations among parameters defining the functions are discussed, and the general functional relationships are developed and listed in tabular form. The use of the proper parameters for data comparisons or for distribution definition is emphasized.

This paper was submitted for publication in the open literature at least 6 months prior to the issuance date of this Microcard. Since the U.S.A.E.C. has no evidence that it has been published, the paper is being distributed in Microcard form as a preprint.

LEGAL NOTICE

This report was prepared as an account of Government sponsored work. Neither the United States, nor the Commission, nor any person acting on behalf of the Commission:

A. Makes any warranty or representation, expressed or implied, with respect to the accuracy, completeness, or usefulness of the information contained in this report, or that the use of any information, apparatus, method, or process disclosed in this report may not infringe privately owned rights; or

B. Assumes any liabilities with respect to the use of, or for damages resulting from the use of any information, apparatus, method, or process disclosed in this report.

As used in the above, "person acting on behalf of the Commission" includes any employee or contractor of the Commission, or employee of such contractor, to the extent that such employee or contractor of the Commission, or employee of such contractor prepares, disseminates, or provides access to, any information pursuant to his employment or contract with the Commission, or his employment with such contractor.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

ON THE PROPER USE OF TRANSFORMATIONS OF LOG NORMAL
FUNCTIONS IN SMALL-PARTICLE STATISTICS*

A. Goldman, H. D. Lewis and R. H. Moore

Los Alamos Scientific Laboratory, University of California
Los Alamos, New Mexico

Introduction

The log normal function has been advantageously utilized by many experimenters to represent either the frequency or the functional form of particle size data. The hypothesis of log normality has been of widespread use as a statistical model representing the frequency distribution of a dimension defining the "size" of a particulate material.^{1,2,3} Unfortunately, the nomenclature and notation appearing in many of the articles in the literature obviously could lead to considerable confusion and to complete misuse of experimental results. For example, the term "size-frequency by weight" used by Hatch (see page 30 Ref. 3) and others does not literally refer to the frequency distribution of particle weights. The notations " $\ln M_g$," " $\log \sigma_g$," " $\ln^2 \sigma_g$," " $(\log s_g)^2$ " that appear in Herdan,⁵ Irani and Callis,⁶ and elsewhere^{7,8,9} could be mystifying to those acquainted with the popularly used convention of using Greek letters to denote distribution or population parameters and Roman letters or modified Greek letters such as $\hat{\mu}$, $\tilde{\mu}$, etc. to denote estimates of

*Work performed under the auspices of the United States Atomic Energy Commission.

population parameters. The above notation may also lead to futile attempts on the part of many investigators to give practical meaning to the anti-logs of "parameters" seemingly indicated by such notation. For example, M_g and σ_g are difficult to interpret as being parameters of a "distribution."

It is the intent of this expository article to suggest a standard notation in the development of important log normal transformations in particle size data analysis, and to suggest the use of these transformations in comparison of data obtained through different sizing techniques. It is hoped that this presentation will unify, strengthen, and clarify the fine work of prior authors in this field. Proof of the important transformations will be given in the appendices to augment the text.

No assumptions will be made about the shape of the particles; however, it may be convenient for the reader to assume spherically shaped particles in consideration of the shape factor, C_k .

Notation and Definition

The following notations and definitions will be used throughout the text, and are suggested as a standard for use in all branches of small particle technology.

d: A measurable dimension characteristic of all particles in the population of particles.

k: An integer = 0, 1, 2, 3,..... determined by the particular geometric property of the particle under consideration.

Shape factor C_k : A coefficient constant that relates the k^{th} power

of d to a geometric property of the particle (e.g. for $k = 2$ and spherical particles, d = sphere diameter, $C_2 = \pi$, and $C_2 d^2$ is the surface area of the particle).

"Size": An arbitrary term related to some measured or deduced characteristic property such as a linear dimension, d ; an area, a ; a volume v ; or a weight, w .

$(d_i^*, d_{i+1}^*), (a_i^*, a_{i+1}^*), (v_i^*, v_{i+1}^*), (w_i^*, w_{i+1}^*)$: The range of "sizes" in the i^{th} "size" interval; $i = 1, 2, 3, \dots, I$; where I = total number of size intervals.

d_i, a_i, v_i, w_i : The midpoint values in the i^{th} "size" interval e.g. d_i denotes the midpoint of the i^{th} linear dimension interval, and v_i the midpoint of the i^{th} volume interval.

a_i', v_i', w_i' : The total area, volume, or weight, respectively, of the particles in the i^{th} "size" interval.

n_i : The frequency of occurrence of "sizes" in the i^{th} "size" interval.

$N = \sum_{i=1}^I n_i$: Total number of particles in I "size" intervals.

$A = \sum_{i=1}^I a_i' = \sum_{i=1}^I n_i C_2 d_i^2$: Total surface area of particles in the I "size" intervals.

$V = \sum_{i=1}^I v_i' = \sum_{i=1}^I n_i C_3 d_i^3$: Total volume of particles in the I "size" intervals.

$W = \sum_1^I w_i' = \sum_1^I \rho C_3 n_i d_i^3$: Total weight of particles in the I "size" intervals. (ρ = particle density; weight/unit volume, and $W = \rho V$).

$\frac{a_i'}{A}$: The relative surface area or surface area fraction of particles in the i^{th} "size" interval.

$\frac{v_i'}{V}$: The relative volume or volume fraction of particles in the i^{th} "size" interval.

$\frac{w_i'}{W}$: Relative weight or weight fraction of particles in the i^{th} "size" interval.

$m_k' = \frac{\sum_1^I n_i d_i^k}{N}$: The k^{th} moment about the origin of the frequency histogram of d .

$m_1' = \bar{d} = \frac{\sum_1^I n_i d_i}{N}$: The first moment about the origin of the frequency histogram of d , the arithmetic mean of d .

$\bar{d}_k = \frac{\sum_1^I n_i d_i^{(k+1)}}{\sum_1^I n_i d_i^k}$: An average dimension of sample particles based on a k^{th} moment relationship. (When $k = 0$, $\bar{d}_0 = \bar{d} = \frac{\sum_1^I n_i d_i}{N}$ and $\bar{d}$ is the arithmetic mean which is identically equal to m_1' .)

$m_k = \frac{\sum_1^I n_i (d_i - \bar{d})^k}{N}$: The k^{th} moment about the mean of the frequency histogram of d .

$$m_2 = s^2 = \frac{\sum_1^I n_i (d_i - \bar{d})^2}{N} : \text{The second moment about the mean of the frequency histogram of } d, \text{ the variance of the histogram of } d \text{ or the sample variance of } d.$$

$$s_{d_k}^2 = \frac{\sum_1^I n_i d_i^k (d_i - \bar{d}_k)^2}{\sum_1^I n_i d_i^k} : \text{The sample variance of } d \text{ based on a } k^{\text{th}} \text{ moment relationship.}$$

$$\overline{C_k d^k} = \frac{\sum_1^I n_i C_{ki} d_i^k}{N} : \text{The general expression for an average characteristic property of the particles in a sample. The basic definition of } \bar{d}, \bar{a}, \bar{v}, \text{ or } \bar{w}.$$

$$\bar{a} = \frac{\sum_1^I n_i a_i}{N} = \frac{\sum_1^I n_i C_{2i} d_i^2}{N} : \text{Average individual surface area of particles in the sample.}$$

$$\bar{v} = \frac{\sum_1^I n_i v_i}{N} = \frac{\sum_1^I n_i C_{3i} d_i^3}{N} : \text{Average individual volume of particles in the sample.}$$

$$\bar{w} = \frac{\sum_1^I n_i \rho C_{3i} d_i^3}{N} = \rho \bar{v} : \text{Average individual weight of particles in the sample.}$$

$$s^2(C_k d^k) = \frac{\sum_1^I n_i (C_{ki} d_i^k - \overline{C_k d^k})^2}{N} : \text{The basic definition of the variance of a characteristic property of particles in a sample, i.e., the basic definition of } s_a^2, s_v^2 \text{ and } s_w^2.$$

$$s_a^2 = \frac{\sum_1^I n_1 (a_1 - \bar{a})^2}{N}: \text{ Sample variance of particle surface areas.}$$

$$s_v^2 = \frac{\sum_1^I n_1 (v_1 - \bar{v})^2}{N}: \text{ Sample variance of particle volumes.}$$

$$s_w^2 = \frac{\sum_1^I n_1 (\rho C_3 a_1^3 - \bar{w})^2}{N} = \rho^2 s_v^2: \text{ Sample variance of particle weights.}$$

$$p(h) = \frac{1}{\sqrt{2\pi} \sigma_h} e^{-\frac{(h - \mu_h)^2}{2\sigma_h^2}}: \text{ The definition of the probability density of a normally distributed population of the random variable } h, \text{ having mean } \mu_h, \text{ and variance } \sigma_h^2.$$

$$P[h_0 < h < h_1] = \frac{1}{\sqrt{2\pi} \sigma_h} \int_{h_0}^{h_1} e^{-\frac{(h - \mu_h)^2}{2\sigma_h^2}} dh: \text{ The probability that } h \text{ is found in the interval } (h_0, h_1).$$

$x = \ln d$: The definition of x , $\ln d$ is the natural logarithm of d .

μ_x : The mean value of a normal distribution of the population of x .

σ_x^2 : The variance of a normal distribution of the population of x .

$z_k = \ln C_k d^k$: The definition of z_k .

μ_{z_k} : The mean value of a normal distribution of the population of z_k .

$\sigma_{z_k}^2$: The variance of a normal distribution of the population of z_k .

$$f_k(x) = \frac{e^{kx} e^{-\frac{(x - \mu_{xk})^2}{2\sigma_{xk}^2}}}{\sqrt{2\pi} \sigma_{xk}}: \text{The definition of the } k^{\text{th}} \text{ moment relationship.}$$

μ_{xk} : The first moment about the origin of $f_k(x)$. This parameter is analogous to μ_x but is not a "mean" in the strict technical sense.

σ_{xk}^2 : The second moment about μ_{xk} . This parameter happens to be identically equal to σ_x^2 ; however, it is not a variance in the strict technical sense.

Weight Fraction-"Size" Relationships, and "Size" Distributions

Consider the schematic histograms in Figures 1, 2, and 3. These represent typical data obtained by experimental techniques. The histogram of Figure 1 represents the relative frequencies of occurrence of particles in the dimension "size" intervals. These data are typical of that obtained by microscopic particle count.¹⁰ The "histogram" of Figure 2 represents the weight fraction of particles occurring in the dimension "size" intervals as might be obtained by screen analysis,¹¹ or by use of the Micromerograph.¹² Figure 2 is not really a histogram by the true definition but it has properties analogous to the relative frequency histogram. The histogram of Figure 3 represents the relative frequency of occurrence of particles in the volume "size" intervals as might be obtained using the Coulter Counter.¹³ All three of these Figures could, hypothetically, represent data obtained from one given sample. The calculation of the statistics of the sample, e.g. the estimates of population parameters, would therefore differ for each method of data representation. For example, the statistics calculated for the representation of Figure 1 could be the arithmetic mean,

$$\bar{d}_o = \frac{\sum_{i=1}^I n_i d_i}{N}, \text{ and variance } s_{d_o}^2 = \frac{\sum_{i=1}^I n_i (d_i - \bar{d}_o)^2}{N}.$$

The denominator of the expression for $s_{d_0}^2$ could be written $N-1$ instead of N . Then, $s_{d_0}^2$ is an unbiased estimator of the variance of the distribution of d . However, the sample size in small-particle statistics is such that, in general, $N \gg 10^3$; therefore, the above definition of $s_{d_0}^2$ can be considered proper.[†] A mean for the sample data of Figure 2 could be computed by

$$\bar{d}_3 = \frac{\sum_1^I n_{11} d_1^4}{\sum_1^I n_{11} d_1^3} = \frac{\sum_1^I w_i' d_1}{W}, \text{ which is the mean of } d \text{ based on a third moment, or weight fraction relationship.}$$

$$\text{A variance could be computed as } s_{d_3}^2 = \frac{\sum_1^I n_{11} \rho C_3 d_1^3 (d_1 - \bar{d}_3)^2}{\sum_1^I n_{11} \rho C_3 d_1^3} = \frac{\sum_1^I w_i' (d_1 - \bar{d}_3)^2}{W}.$$

Similarly, the mean, $\bar{v}$, and variance s_v^2 could be computed for the histogram, Figure 3.

Many other definitions of mean "size" could be constructed; for instance, by using moment relationships to define $\bar{d}_k$ or by computing $\bar{a}$. Each of these statistics would be a representation of some property of the sample. This indicates the complexity of "mean value" considerations justifiably discussed in the literature, (see general references 4,5,6,7,8,9), and illustrates the need for caution in using mean values and variances, especially in data comparison.

[†] For discussion of unbiased estimators of variance see: P. G. Hoel, "Introduction to Mathematical Statistics," p. 52, and sections 10.2 and 10.2.1, pp. 196-200, New York, Wiley & Sons; 1954. A. Hald, "Statistical Theory with Engineering Applications," section 4.4, pp. 72-76, New York, Wiley and Sons, 1952.

In general, histograms of particle "size" data are not symmetrical about a mean particle "size." Therefore, the computed sample means and variances may not be sufficient to describe the sample uniquely without consideration of statistics such as skewness, or other higher moments. There exist no known simple, convenient methods of relating the various values of $\bar{d}_k$ and $s_{d_k}^2$ as for example relating $\bar{d}_3$ to $\bar{v}$ or relating $s_{d_3}^2$ to s_v^2 . Consequently, it is very difficult to compare data for the same sample obtained by various measurement methods.

It has been shown in the literature that the log normal function can be used in many cases^{4,14} to represent particle size data of the type shown in Figures 1, 2, and 3.

Certain relationships can be written if it can be established that data such as that schematically shown in Figures 1, 2, and 3 can be respectively represented by the log normal functions¹⁵ depicted by Figures 4, 5, and 6. These relationships are:

$$\frac{n_i}{N} \approx \int_{x_i^*}^{x_{i+1}^*} p(x) dx; \text{ i.e. } \frac{n_i}{N} \text{ can be approximated by the integration of } p(x) \text{ over the } i^{\text{th}} \text{ interval.} \quad (1)$$

$$\frac{w_i'}{W} \approx \int_{x_i^*}^{x_{i+1}^*} f_3(x) dx; \text{ i.e. } \frac{w_i'}{W} \text{ can be approximated by the integration of } f_3(x) \text{ over the } i^{\text{th}} \text{ interval.} \quad (2)$$

$$\frac{n_i}{N} \approx \int_{z_i^*}^{z_{i+1}^*} p(z_3) dz; \text{ i.e. } \frac{n_i}{N} \text{ can be approximated by the integration of } p(z_3) \text{ over the } i^{\text{th}} \text{ interval.} \quad (3)$$

The meaning of the above approximations can be illustrated by comparing the

shaded areas in Figures 1 and 4. Smith and Jordan¹⁷ have attempted to clarify some erroneous statements found in the literature concerning the use of $p(x)$ to approximate $\frac{n_1}{N}$. The functional representations of (1), (2), and (3) above, are assumed to represent an entire, infinite population of all the particles represented by the sample. In order to estimate these functional relationships, computed statistics are used instead of the true but unknown population parameters. Therefore the notation for the sample statistics, i.e. $\hat{\sigma}_x^2$, $\hat{\mu}_x$ etc. should be used to denote estimates of population parameters.

Several important facts should be noted before the representations (1), (2), and (3) are fully defined. First, these relationships are not equalities. It must be emphasized that only the relative frequency relationships of Figures 1 and 3 can be represented by probability densities, $p(x)$ or $p(z)$. The weight fraction (or area fraction, or volume fraction) relationship is represented by a form of the log normal function which is denoted $f(x)$.

The second, and perhaps most important, fact to note is that the parameters which uniquely define each function illustrated by Figures 4, 5, and 6 are themselves related by known linear functions when the relationships (1), (2), and (3) describe the same sample. This is to say that no matter how the data are obtained or represented, the parameters defining $p(x)$ can be obtained from parameters defining $f(x)$ or $p(z)$. These parameters defining the $p(x)$ representation are those which should be obtained and used in all cases of data comparisons. The methods of determining the relations among the defining parameters of log normal functions,

and the proper use of their interrelations have not been rigorously treated in the particle size literature. The third fact worthy of emphasis is that the log normal functions which can define $f(x)$ are moment functions of $p(x)$ and can be used to compute estimates of any desired geometric property such as specific surface, surface area, weight, or volume of particulate material. An additional fact which is emphasized is that if a decision is made to use a log normal function to represent the sample of particulate material, the properties of the continuous function should be utilized in all subsequent statistical computations, especially in data comparison. Some of these properties will be described in the following section.

Log normal Functions, Log normal Distributions and Parametric Transformations

The log normal function representing Figure 4 is written as follows:

$$\frac{n_i}{N} \approx \int_{x_i^*}^{x_{i+1}^*} p(x) dx = \frac{1}{\sqrt{2\pi} \sigma_x} \int_{x_i^*}^{x_{i+1}^*} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} dx \quad (4)$$

where: $p(x)$ is the probability density of the population of $x = \ln d$, with mean, μ_x , and variance σ_x^2 .

The function representing the weight fraction vs $\ln d$ relationship illustrated in Figure 5 is written:

$$\begin{aligned}
\frac{w_i}{W} &\approx \int_{x_i^*}^{x_{i+1}^*} f_3(x) dx = \frac{1}{\sqrt{2\pi} \sigma_{x_3}} \int_{x_i^*}^{x_{i+1}^*} e^{-\frac{(x-\mu_{x_3})^2}{2\sigma_{x_3}^2}} dx \\
&= \frac{1}{\sqrt{2\pi} \sigma_x} \int_{x_i^*}^{x_{i+1}^*} e^{-\frac{[x-(\mu_x + 3\sigma_x^2)]^2}{2\sigma_x^2}} dx \quad (5)
\end{aligned}$$

where $f_3(x)$ is a log normal function with parameters, μ_{x_3} and $\sigma_{x_3}^2$. The parameter $\sigma_{x_3}^2$ is identically equal to σ_x^2 .

The function representing the distribution of volumes illustrated by Figure 6 is written:

$$\begin{aligned}
\frac{n_i}{N} &\approx \int_{x_i^*}^{x_{i+1}^*} p(z_3) dz = \frac{1}{\sqrt{2\pi} \sigma_{z_3}} \int_{x_i^*}^{x_{i+1}^*} e^{-\frac{(z-\mu_{z_3})^2}{2\sigma_{z_3}^2}} dz, \\
&= \frac{1}{\sqrt{2\pi} 3\sigma_x} \int_{x_i^*}^{x_{i+1}^*} e^{-\frac{[x-(\ln C_3 + 3\mu_x)]^2}{2(9\sigma_x^2)}} dx \quad (6)
\end{aligned}$$

where $p(z_3)$ is the probability density of the population of $z_3 = \ln v$, with mean μ_{z_3} and variance $\sigma_{z_3}^2$.

Functions of the general form of (5) are moment relationships, and can be written:

$$\begin{aligned}
\frac{n_1 d_1^k}{\sum_1 n_1 d_1^k} &\approx \int_{x_1^*}^{x_{1+1}^*} f_k(x) dx = \frac{1}{\sqrt{2\pi} \sigma_{xk}} \int_{x_1^*}^{x_{1+1}^*} e^{-\frac{(x-\mu_{xk})^2}{2\sigma_{xk}^2}} dx \\
&= \frac{1}{\sqrt{2\pi} \sigma_x} \int_{x_1^*}^{x_{1+1}^*} e^{-\frac{[x-(\mu_x + k\sigma_x^2)]^2}{2\sigma_x^2}} dx
\end{aligned} \tag{7}$$

where $f_k(x)$ is a log normal function with parameters μ_{xk} and σ_{xk}^2 .

Functions of the general type of (6) are written:

$$\begin{aligned}
\frac{n_1}{N} &\approx \int_{x_1^*}^{x_{1+1}^*} p(z_k) dz_k = \frac{1}{\sqrt{2\pi} \sigma_{z_k}} \int_{x_1^*}^{x_{1+1}^*} e^{-\frac{(z_k - \mu_{z_k})^2}{2\sigma_{z_k}^2}} dz_k \\
&= \frac{1}{\sqrt{2\pi} k\sigma_x} \int_{x_1^*}^{x_{1+1}^*} e^{-\frac{[x-(\ln C_k + k\mu_x)]^2}{2k^2\sigma_x^2}} dx
\end{aligned} \tag{8}$$

where $p(z_k)$ is the probability density of $z_k = \ln C_k d^k$, having mean μ_{z_k} and variance $\sigma_{z_k}^2$.

Note that in (8) z_k is actually the natural logarithm of a geometric relationship such that for $k = 2$, and $C_2 =$ area shape factor, $p(z_2)$ is the probability density of logarithms of particle areas. Also note that for $k = 1$, $C_1 = 1$, then $z_1 = x = \ln d$.

The means and variances of the above log normal functions can be related in the following manner. [It should be recalled that $f(x)$, $p(x)$, and $p(z)$ are here considered to represent various characteristics of the same particle population.] Consider transforming Figure 4 to Figure 6,

i.e. relating μ_{z_k} to μ_x , and $\sigma_{z_k}^2$ to σ_x^2 ; then a linear transformation is used. It can be readily shown that any linear function of a normally distributed variable is also normal. The proof of this is given in Appendix 1. If x is distributed with mean μ_x and variance σ_x^2 , then

$$z_k = \ln C_k + kx \quad (9)$$

is also distributed normally with mean and variance

$$\begin{aligned} \mu_{z_k} &= \ln C_k + k\mu_x \\ \sigma_{z_k}^2 &= k^2 \sigma_x^2 \end{aligned} \quad (10)$$

This transformation is of the type discussed by Kapteyn.¹⁵

Consider transforming $p(x)$ representing Figure 4 to $f(x)$ representing Figure 5, i.e. relating μ_x to μ_{xk} , and σ_x^2 to σ_{xk}^2 . The function, $f(x)$, represented in Figure 5 can be shown to be normal, with first moment

$$\mu_{xk} = \mu_x + k\sigma_x^2 \quad (11)$$

and second moment about μ_{xk}

$$\sigma_{xk}^2 = \sigma_x^2 \quad (12)$$

This proof is given in Appendix 2, by use of a moment transformation.

An analogous theorem has been developed by Atchison and Brown.¹⁸

The relationships among the parameters discussed previously can be illustrated by four typical cumulative percent vs "size" plots schematically shown in Figures 7 and 8. These are so-called log-probability

plots discussed by various authors.[†]

The functions represented in Figure 7 are $P(x) = \int_{-\infty}^x p(x)dx$ and $F_3(x) = \int_{-\infty}^x f_3(x)dx$. These functions respectively describe the cumulative frequency and cumulative weight fraction of particles of diameter less than d . In this plot, the variable $x = \ln d$, plotted as abscissa, is common to both relationships. Note that $\mu_{x3} = \mu_x + 3\sigma_x^2$ and $\sigma_{x3}^2 \equiv \sigma_x^2$. It is important to note that $F_3(x)$ is not a probability function. It is a weight fraction- $\ln$ diameter relationship which is functionally related to the probability function $P(x)$.

The functions represented in Figure 8 are

$$P(z_3) = \int_{-\infty}^{z_3} P(z_3)dz_3 \text{ and } P(x) = \int_{-\infty}^x P(x)dx.$$

These are both cumulative frequency distributions or probability functions of the different, but linearly related variables $x = \ln d$, and

$$z_3 = \ln(C_3 d^3) = \ln C_3 + 3x. \text{ Note that } \mu_{z_3} = \ln C_3 + 3\mu_x \text{ and } \sigma_{z_3}^2 = 9\sigma_x^2.$$

The population mean (μ_x, μ_z , etc.) of a normal distribution is the value of the random variable corresponding to the 50th percentile of the cumulative function. The standard deviation (σ_x, σ_z , etc.) of a normal population is the difference between the value of the random variable

[†] An excellent discussion of the use of logarithmic probability paper and the plotting of logarithmic normal distribution data is found in A. Hald, Statistical Theory with Engineering Applications, New York, Wiley and Sons, Chap. 7 (1952).

corresponding to the 84.13 percentile and the mean. These relationships may be seen in Figures 7 and 8.

Summary

It is the purpose of this article to accomplish four main objectives:

1. Suggest a standard notation for use in the field of small particle statistics.
2. Set a standard technique for the comparison of log normal representations of small particle data obtained by different experimental methods.
3. Present the transformations required for proper data comparisons.
4. Emphasize differences among histograms, populations, functional relationships, and probability relationships.

It has been emphasized that in comparing particle size data represented by a log normal function, the parameters obtained should be transformed to μ_x and σ_x^2 and these values used as data comparison criteria.

The transformations may be summarized as follows:

Table I

Linear Transforms [ordinate = $p(z_k)$]

	Abcissa Value			
	$x = \ln d$	$z_2 = \ln a = \ln C_2 d^2$	$z_3 = \ln V = \ln C_3 d^3$	$z_k = \ln C_k d^k$
Parameters	μ_x	$\mu_{z_2} = \ln C_2 + 2\mu_x$	$\mu_{z_3} = \ln C_3 + 3\mu_x$	$\mu_{z_k} = \ln C_k + k\mu_x$
	σ^2	$\sigma_{z_2}^2 = 4\sigma_x^2$	$\sigma_{z_3}^2 = 9\sigma_x^2$	$k^2\sigma_x^2$

Table II

Moment Transforms (abscissa = x)

Ordinate Value

	$p(x)$	$f_2(x)$	$f_3(x)$	$f_k(x)$
Parameters	μ_x	$\mu_x + 2\sigma_x^2 = \mu_{x2}$	$\mu_x + 3\sigma_x^2 = \mu_{x3}$	$\mu_x + k\sigma_x^2 = \mu_{xk}$
	σ_x^2	σ_x^2	σ_x^2	σ_x^2

It is the intent of the authors to discuss, in subsequent articles, the relations among parameters of the log normal distribution of x and the distribution of d, moment relationships, geometric properties of particulate solids, such as specific surface of air classified fractions, theory of mixing of distributions, and practical applications.

Acknowledgments

The authors would like to acknowledge the encouragement of Dr. R. K. Zeigler, Dr. R. B. Gibney, and M. C. Smith in the preparation of this article. Helpful correspondence with Dr. G. Herdan is also gratefully acknowledged.

APPENDIX I

Theorem: If t is distributed normally with mean, μ_t , and variance σ_t^2 ; then $s = a + bt$ is also normal with mean, $\mu_s = a + b \mu_t$, and variance $\sigma_s^2 = b^2 \sigma_t^2$; where a and b are constants.

Proof: The distribution of t is given by

$$p(t) = \frac{1}{\sqrt{2\pi} \sigma_t} e^{-\frac{(t-\mu_t)^2}{2 \sigma_t^2}}; \quad -\infty < t < \infty.$$

By a change of variable

$$s = a + bt,$$

then

$$p(s) = \frac{1}{\sqrt{2\pi} b \sigma_t} e^{-\frac{[s - (a+b \mu_t)]^2}{2 b^2 \sigma_t^2}}.$$

Thus

$$\mu_s = a + b \mu_t,$$

$$\sigma_s^2 = b^2 \sigma_t^2,$$

and the theorem is proved.

Example: As a specific example of this theorem, if $x = \ln d$ is distributed normally with mean, μ_x , and variance σ_x^2 ; then $z_k = \ln C_k d^k = \ln C_k + kx$ is also normal with mean, $\mu_{zk} = \ln C_k + k\mu_x$, and variance, $\sigma_{zk}^2 = k^2 \sigma_x^2$.

APPENDIX II

Theorem: If $x = \ln d$ is distributed normally with mean, μ_x , and variance, σ_x^2 ; then $f(x) = g_k d^k p(x)$ is a function with moments $\mu_{xk} = \mu_x + k\sigma_x^2$ and $\sigma_{xk}^2 = \sigma_x^2$; where

$$- [k\mu_x + \frac{k^2}{2} \sigma_x^2]$$

$$g_k = e$$

Proof: By definition,

$$p(x) = \frac{1}{\sqrt{2\pi} \sigma_x} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}}$$

In order for $f(x)$ to have the properties of a density function there must exist a g_k such that

$$g_k \int_{-\infty}^{\infty} d^k p(x) dx = 1;$$

where

$$x = \ln d,$$

and

$$d^k = e^{kx}.$$

Thus,

$$\frac{g_k}{\sqrt{2\pi} \sigma_x} \int_{-\infty}^{\infty} e^{kx} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} dx = 1,$$

implies that

$$g_k = \exp \left[- \left(k\mu_x + \frac{k^2}{2} \sigma_x^2 \right) \right];$$

consequently,

$$f(x) = \frac{1}{\sqrt{2\pi} \sigma_x} e^{-\frac{[x - (\mu_x + k\sigma_x^2)]^2}{2\sigma_x^2}}$$

and the theorem is proved. The converse of the theorem, if the variance of $f(x)$ equals the variance of $p(y)$ then x and y are both normally distributed, has been shown in a paper submitted for publication by Dr. R. K. Zeigler.

APPENDIX III

Example of Appendix 2: Assume spherically shaped particles such that d = diameter, and $k = 3$. The problem is to find a weight-diameter functional relationship that approximates $\frac{w_j}{W}$ depicted in Figure 2 of the text when $x = \ln d$ is distributed normally, $p(x)$.

The relative weight in the j^{th} interval of the weight relationship would be found to be

$$\frac{w_j}{W} = \frac{C_3 n_j \delta_j^3}{C_3 \sum_{i=1}^I n_i \delta_i^3} = \frac{n_j \delta_j^3}{\sum_{i=1}^I n_i \delta_i^3};$$

where C_3 is the shape factor, and δ_i is a value of d in the interval $(d_{i(1)}^*, d_{i+1(1)}^*)$. The meaning of δ_i will be defined more precisely later.

Dividing numerator and denominator by N

$$\frac{w_j}{W} = \frac{n_j \delta_j^3}{N} \bigg/ \sum_{i=1}^I \frac{n_i \delta_i^3}{N}$$

Since

$$\frac{n_j}{N} \approx \int_{x_j}^{x_{j+1}} p(x) dx,$$

and

$$x = \ln d;$$

then

$$\frac{w_j}{W} \approx \int_{x_j}^{x_{j+1}} e^{3\xi_j} p(x) dx$$

where ξ_j is a value of x in the interval (x_j, x_{j+1}) .

By applying the mean value theorem

$$\int_{x_j}^{x_{j+1}} e^{3\xi_j} p(x) dx = e^{3\xi_j} (x_{j+1} - x_j) p(X)$$

where

$$x_j < X < x_{j+1}.$$

Since ξ_j is as yet unspecified, let $\xi_j = X$.

Then

$$\frac{w_j}{W} \approx \frac{(x_{j+1} - x_j) e^{3X} p(X)}{W}$$

By applying the mean value theorem again

$$\begin{aligned} \frac{w_j}{W} &\approx \frac{\int_{x_j}^{x_{j+1}} e^{3x} p(x) dx}{\int_{-\infty}^{\infty} e^{3x} p(x) dx} = g_3 \int_{x_j}^{x_{j+1}} d^3 p(x) dx \\ &= g_3 \int_{x_j}^{x_{j+1}} e^{3x} p(x) dx \end{aligned}$$

and the theorem is applicable to this problem.

Consequently, by applying the theorem

$$f(x) = \frac{1}{\sqrt{2\pi} \sigma_x} e^{-\frac{[x - (\mu_x + 3\sigma_x^2)]^2}{2\sigma_x^2}}$$

REFERENCES

1. R. P. Loveland and A. P. H. Trivelli, "Mathematical Methods of Frequency Analysis of Size of Particles," Jour. Franklin Inst., Aug. and Sept. 1927.
2. T. Hatch and S. Choate, "Statistical Description of the Size Properties of Non-Uniform Particulate Substances," Jour. Franklin Inst., pp. 369-387, March 1929.
3. M. J. Thornton, "The Application of the Coulter Counter to Problems in the Size Analysis of Insoluble Drugs," Jour. of Pharmacy and Pharmacology, Vol. 15, pp 742-752, Nov. 1963.
4. T. Hatch, "Determination of 'Average Particle Size' from the Screen-Analysis of Non-Uniform Particulate Substances." Jour. Franklin Inst., Jan. 1933, pp. 27-37.
5. G. Herdan, "Small Particle Statistics," 2nd Ed., p. 81, New York, Academic Press Inc., 1960.
6. R. Irani and C. Callis, "Particle Size: Measurement, Interpretation, and Application," p. 44, New York, Wiley and Sons, Inc., 1963.
7. R. Cadle, "Particle Size Determination," p. 41, New York, Interscience Publishers, Inc., 1960.
8. C. Orr, Jr., and J. M. DallaValle, "Fine Particle Measurement," p. 29, New York, The MacMillan Company, 1959.

9. DalleValle, "Micromeritics," 2nd Edition, p. 54, New York, Pitman Publishing, 1948.
10. R. Irani and C. Callis, "Particle Size: Measurement Interpretation, and Application," Chap. 6, New York, John Wiley and Sons, Inc., 1963.
11. R. Irani and C. Callis, *ibid*, Chap. 7.
12. F. S. Eadie and R. E. Rayne, "Particle Size Distribution Analyzed Quickly, Accurately," pp. 99-102, *The Iron Age*, Sept. 2, 1954.
13. M. A. VanDilla, Personal Communication, Los Alamos Scientific Laboratory.
14. F. Kottler, "The Distribution of Particle Sizes," *Jour. Franklin Inst.*, Vol. 250, pp 339,419 (1950).
15. A. Hald, "Statistical Theory with Engineering Applications," Chapters 6 and 7, New York, John Wiley and Sons, Inc., 1952.
16. J. C. Kapteyn, "Skew Frequency Curves in Biology and Statistics," Groningen, 1903.
17. J. Smith and M. Jordan, "Mathematical and Graphical Interpretation of the Log-Normal Law for Particle Size Distribution Analysis," pp 549-559, *Jour. Colloid Science*, (19), 1964.
18. Atchison and Brown, "The Lognormal Distribution" p. 12, Cambridge, Cambridge University Press, 1957.

Figure Captions

- Fig. 1: Histogram: Relative Frequency of Occurrence of Sample Particles of Dimension d_1 .
- Fig. 2: Weight Fraction Occurrence of Sample Particles of Dimension d_1 .
- Fig. 3: Histogram: Relative Frequency of Occurrence of Sample Particles of Volume v_1 .
- Fig. 4: Relative Frequency Distribution of $x = \ln d$.
- Fig. 5: Weight Fraction vs. $\ln d$ Functional Relationship.
- Fig. 6: Relative Frequency Distribution of $z_3 = \ln v$.
- Fig. 7: Comparison of log probability plots $P(x)$ and $F_3(x)$.
- Fig. 8: Comparison of log probability plots $P(x)$ and $P(z_3)$.

25