211
7-3-73

LA-5254-MS

INFORMAL REPORT

33398

th. 208

Small Particle Statistics

The Analysis of Particle "Size" Data

los alamos
scientific laboratory
of the University of California
LOS ALAMOS, NEW MEXICO 87544

UNITED STATES
ATOMIC ENERGY COMMISSION
CONTRACT W-7405-ENG. 36

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

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.

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Atomic Energy Commission, nor any of their employees, nor any of their contractors, subcontractors, or 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.

In the interest of prompt distribution, this LAMS report was not edited by the Technical Information staff.

Printed in the United States of America. Available from
National Technical Information Service
U. S. Department of Commerce
5285 Port Royal Road
Springfield, Virginia 22151

Price: Printed Copy ~~\$3.00~~; Microfiche \$0.95

4.00

los alamos
scientific laboratory
of the University of California
LOS ALAMOS, NEW MEXICO 87544

LA-5254-MS
Informal Report
UC-32

ISSUED: May 1973

Small Particle Statistics

The Analysis of Particle "Size" Data

by

H. D. Lewis

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Atomic Energy Commission, nor any of their employees, nor any of their contractors, subcontractors, or 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.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

leg

SMALL PARTICLE STATISTICS
THE ANALYSIS OF PARTICLE "SIZE" DATA

by

H. D. Lewis

ABSTRACT

Fundamentals of "small particle statistics" are developed from basic statistical concepts applied to examples of analysis of idealized particulate systems. The development of methods for "particle size" data analysis is based on the form of the data generated by various methods for measurement of "particle size distributions". It is shown that data can be classified into two general types. Methods for computing and comparing sample statistics are developed on the basis of Type I and II classifications for both functional and interval representations of the sample data. The compatibility of recognized statistical notation and definition with the problem of particle "size" analysis is demonstrated.

I. INTRODUCTION

It is impossible to present a comprehensive course in small particle statistics in these few pages. However, we will attempt to build a framework of simple statistical concepts, and from these, develop the basic "small particle statistics." It will be necessary to omit detailed discussion of the very important aspects of statistics, sampling,^{1,2} statistical inference and tests of hypotheses,³ in order to concentrate on the basic problems requiring the use of small particle statistics. This problem is almost universally stated in disarming simplicity as "characterization of particulate systems (or powders)."

The characterization of a particulate system is a fundamental problem in all branches of industry and science concerned with bulk properties of matter composed of small particles, particulate raw materials or physical, chemical, mechanical, or physiological phenomena related to some property or properties of a particulate system.^{4,6} The complete characterization of a particulate system should include definitions of "particle" and "particulate

system", a description of particle "size", particle "shape", particle density, surface characteristics, and mechanical, physical, and chemical properties of the particles. The purpose of our discussion is to develop aspects of particulate system characterization concerned with the analytical concepts of particle "size" and related concepts of particle "shape" and surface characteristics. We will use the following definitions of particle and particulate system. A particle is a unit of matter having a defined boundary, whose "size" and "shape" depend on interatomic or intermolecular bonding forces and its process of formation. A particulate system is a set of particles in some environment. (The importance of the particle-environment concept is illustrated by the lubricant property of graphite particles in air, and the abrasive properties of graphite particles in vacuum.) Although the following development of data analysis techniques is generally applicable to particles in any "size" range, we shall assume our discussion applies to particles in the generally accepted "small particle range", 10^{-1} to 10^4 or 10^5 microns ($1 \text{ micron} = 10^{-3} \text{ mm}$).

Figures 1 and 2 illustrate some of the problems involved in characterization of a particulate. Assume Fig. 1 represents a monosized particulate system containing an unknown number of perfect spheres of pure copper having exactly the same diameter. The problem of characterizing this system with respect to particle "size" and "shape" is quite simple. A sample consisting of one particle can be removed from the system and its diameter measured. It can then be stated that the "shape" and diameter ("size") of every particle in the system is known with precision limited only by the device used to measure the diameter of the sample. The specific surface (surface area per unit volume or unit weight of particles) could be calculated. Consider the microscopic sample of a real particulate shown in Fig. 2. This is a scanning electronmicroscope photograph of a graphite powder at 2000x magnification. Notice that "size", "shape", and surface character, and perhaps density, cannot be defined uniquely and are all variable in this sample of many particles. Therefore the characterization of the system from which this sample was taken must involve some type of statistical analysis of the sample. Such analysis should be related to at least one of the following three objectives:

1. Sample characterization from which characteristics of the particulate system (population) can be inferred.
2. Sample comparisons from which comparative conclusions can be drawn concerning characteristics of different particulate systems, or different measurement methods.
3. Explanation or prediction of the behavior of the particulate in a process.

Figure 1. Schematic system of copper spheres

II. FUNDAMENTAL STATISTICAL CONCEPTS

We will now develop some basic statistical concepts which will permit a generalized development of the methods of "Small Particle Statistics". The diagrams and examples presented in this section are not "real" particle "size" data. The synthetic "sample data" are used in order to permit a simplified presentation of calculation methods.

The basic foundation of statistical analysis is the mathematics of probability.^{7,10} Methods of statistical analysis are techniques used to aid interpretation in the analysis of data obtained by repetitive observation. In characterizing a particulate system (powder) measurements or observations would in general be made on a sample of particles removed from the system. The system is the population or total mass of powder whose characteristics are in question. The sample is the portion of the population on which measurements are made in order to obtain estimates of the characteristics of the population. In order that the sample reliably represent the population, it should be a random sample; that is, be chosen so that it has the same probability of being chosen as any other sample. The statistics which characterize the population are the variates (numerical quantities) or functions of the variates obtained from the observations made on the sample.

Consider a hypothetical system (population) of perfect sodium chloride crystals containing $N = 10$ crystals or particles. Assume we know the exact length, (linear dimension, d) of the major cube diagonal of each particle as given in Table I.

TABLE I
HYPOTHETICAL "SIZE" DISTRIBUTION
OF $N = 10$ NaCl CRYSTALS

Number of Particles	"Size" = Cube Diagonal
n	d_1 (microns)
2	4
6	6
2	8

"Size" is defined for this particular system as the dimension d = cube diagonal.

Figure 2. Graphite powder, magnification 2000X.

The question is, how can we describe this particle "size" distribution. One method would be to draw a line chart as shown in Figure 3. It is readily seen that a chart would not provide a convenient quantitative description of a population consisting of many particles of varied "size". The population distribution of the 10 NaCl crystals could be adequately described by three parameters, the mean μ_d , variance σ_d^2 , and skewness γ_d . These parameters are defined as:

$$\text{Mean: } \mu_d = \frac{1}{N} \sum_{i=1}^N d_i \quad (1)$$

$$\text{Variance: } \sigma_d^2 = \frac{1}{N} \sum_{i=1}^N (d_i - \mu_d)^2 \quad (2)$$

$$\text{Skewness: } \gamma_d = \frac{1}{N} \sum_{i=1}^N (d_i - \mu_d)^3 \quad (3)$$

Figure 3. Description of particle "size" distribution by line chart.

For the system of NaCl crystals,

$$\mu_d = (4+4+6+6+6+6+6+6+8+8)/10 = 6.00$$

$$\sigma_d^2 = [(4-6)^2 + (4-6)^2 + (8-6)^2 + (8-6)^2]/10 = 1.60$$

$$\gamma_d = [(4-6)^3 + (4-6)^3 + (8-6)^3 + (8-6)^3]/10 = 0.0$$

These parameters are also respectively the first moment about the origin, second moment about the mean, and third moment about the mean of the distribution of d .

These parameters are considered respectively as measures of central value, dispersion or variability, and asymmetry about the mean for the population distribution of d . Although, for our purpose, these parameters adequately describe a distribution, it can be argued in general that a distribution is not quantitatively defined unless its moment generating function (i.e. moments of all order) is known.¹¹ Notice that $\gamma_d = 0$ for the symmetric distribution.

In practice, the population distribution is never known. Therefore consider a more "realistic" example of a sample of $N = 1000$ particles taken from a population (particulate system) of 10^9 perfect NaCl crystals. The N particles in the sample yield many different measured values of "size" $d =$ cube diagonal as listed in Table II.

Because the sample contains a large number of particles having varied values of d , presentation and interpretation of the sample data is simplified

TABLE II

SCHEMATIC PARTICLE "SIZE" DATA FOR SAMPLE OF NaCl CRYSTALS

Interval End Point Value	Midpoint Value	n_i	n_i/N
d_i^*	d_i		
1	2	6	0.006
3	4	224	0.224
5	6	540	0.540
7	8	220	0.220
9	10	10	0.010
11			

$$N = \sum_{i=1}^I n_i = 1000$$

by grouping or classification into I "size" intervals ($d_i^* < d \leq d_{i+1}^*$) where d_i values are the midpoint values of each of the i "size" intervals, $i = 1, 2, 3, \dots, I$. The midpoint values are defined as

$$d_i = \frac{d_i^* + d_{i+1}^*}{2} \quad (4)$$

The values of n_i are the frequency or number of particles observed in the i^{th} interval (d_i^*, d_{i+1}^*) and the values $\frac{n_i}{N}$, where $N = \sum_{i=1}^I n_i$, the relative frequency of occurrence of particles in the i^{th} interval. The observed n_i/N values give an estimate of the probability of occurrence of particles in the population having "size" d in the range $d_i^* < d \leq d_{i+1}^*$. In a real analysis of $N = 1000$ particles (or observations), the number of "size" intervals would be greater than 5. Cochran¹² provides some useful rules for selecting a proper value of I (number of intervals) with respect to N . These are summarized in Table III.

TABLE III

N	200	400	600	800	1000	1500	2000
I	12	20	24	27	30	35	39

Methods for graphical display of the data of Table II are illustrated in Figures 4 and 5. A plot of n_i/N vs d_i^* values as shown by the solid line is a histogram or sample distribution of d . From this

Figure 4. Graphical display of Table II data. Histogram and relative frequency polygon.

type of graph the sample M_d , an estimator of the population mode or most probable value of d , can be obtained. M_d is the value of d having the greatest observed value of n_i or can be approximated as d_i of the "size" interval having the greatest value of n_i/N . A plot of n_i/N vs d_i values, as shown by the dashed line of Figure 4, is called the relative frequency polygon.

Figure 5. Cumulative relative frequency polygon plot of Table II data.

A plot of the values $\sum_{i=1}^r n_i/N$ vs d_i^* would produce the cumulative relative frequency polygon of Figure 5. The values $\sum_{i=1}^r n_i/N$ are the total fraction of occurrence of particles through the r^{th} interval where $i = 1, 2, 3 \dots r \dots I$. From this graph estimates can be obtained of the probability of occurrence of particles of "size" less than any particular "size" d in the population distribution. A statistic called the median of d , another measure of central value, can also be obtained. The median is the .5 quantile value of d or $d_{.5} = 6.0^+$ for the data of Figure 5.

Values of mean, variance, and skewness of the sample distribution of d cannot be obtained directly from either of these graphs. These sample statistics are defined as follows.

$$\text{Mean: } \bar{d} = \frac{1}{N} \sum_{i=1}^I n_i d_i \quad (5)$$

$$\text{Variance: } s_d^2 = \frac{1}{N-1} \sum_{i=1}^I n_i (d_i - \bar{d})^2 \quad (6)$$

$$\text{Skewness: } g_d = \frac{1}{N} \sum_{i=1}^I n_i (d_i - \bar{d})^3 \quad (7)$$

These statistical formulae are related to the population parameters in the following way: $\bar{d} \approx \mu_d$, $s_d^2 \approx \sigma_d^2$, and $g_d \approx \gamma_d$, where $\approx$ is read as "is an estimator of." The term $N - 1$ is used in the denominator of (6) to make s_d^2 an unbiased estimator of σ_d^2 .¹³ However, because the sample size N should be large i.e. > 500 in "small particle statistics," s_d^2 can be defined as $s_d^2 \approx \sum_{i=1}^I (n_i/N) (d_i - \bar{d})^2$. The values of the sample statistics, for the data of Table II are listed in Figure 4. Notice the + value of g_d . This positive skewness indicates the sample distribution is asymmetric with $M_d < d_{0.5} < \bar{d}$; however, it is not proof that $\gamma_d \neq 0$. Negative skewness would indicate $\bar{d} < d_{.5} < M_d$.

III. DISTRIBUTION FUNCTION AND FUNCTIONAL MODEL

If the measurements of d_i^* in Table II and Figure 4 could be made with great accuracy and precision, it would be possible to construct a histogram having many more "size" intervals, and in the limit, the values of d would be continuous. If some function, $h(d)$ could be found that approximated the histogram, $h(d)$ would be a mathematical model for the distribution of, d . Such a function is called the

probability density, or density function. The density function must have the following properties:

$$h(d) \geq 0 \quad (8)$$

$$\int_{-\infty}^{\infty} h(d) dd = 1 \quad (9)$$

$$\Pr[d_i^* < d < d_{i+1}^*] = \int_{d_i^*}^{d_{i+1}^*} h(d) dd \quad (10)$$

Expression (10) indicates that integration of $h(d)$ over any particular interval gives the probability that the variable d will have values in that interval. The use of a functional model to represent statistical data similar to the data we have discussed implies

$$\frac{n_i}{N}(d_i^*, d_{i+1}^*) \approx \int_{d_i^*}^{d_{i+1}^*} h(d) dd \quad (11)$$

That is, integration of the function over the interval (d_i^*, d_{i+1}^*) gives a fraction approximately equal to n_i/N observed in the same sample data interval as indicated in Figs. 6(a) and 6(b).

Figure 6. Functional representation of sample data.

Further, as illustrated in Figs. 6(c) and 6(d),

$$\frac{r}{S} \frac{n_i}{N} \approx H(d) = \int_{-\infty}^{d_{i+1}^*} h(d) dd \quad (12)$$

for cumulative data. The histogram and $h(d)$ are drawn to indicate the model need not be symmetric about some "central" value. This will be discussed in detail in the next section.

In order to facilitate the discussion of "small particle statistics" in the next section as well as illustrate some basic concepts here, assume it can be demonstrated that the normal or Gaussian density function,

$$p(d) = \frac{1}{\sqrt{2\pi}\sigma_d} \exp \left[-\frac{(d - \mu_d)^2}{2\sigma_d^2} \right], \quad (13)$$

having mean μ_d , variance σ_d^2 ,¹⁴ and $\gamma_d = 0$, represents the data of Fig. 4. A particular advantage of using the normal function as the model is that the function is completely defined by μ_d and σ_d^2 , hence the sample statistics $\hat{\mu}_d$ and $\hat{\sigma}_d^2$ are the only estimators required to characterize the "size" distribution of the population. A plot of the cumulative form of the sample data, Fig. 5, on normal probability paper¹⁵ would give a straight line, from which first approximations of the sample statistics $\hat{\mu}_d$ and $\hat{\sigma}_d^2$ can be obtained as shown in Fig. 7. Probability paper is ruled such that a plot of

$$P(d) = \int_{-\infty}^{d_{i+1}^*} p(d) dd$$

on the ordinate vs d on the abscissa gives a straight line. Because $p(d)$ is normal, symmetric about μ_d ; the mean, median, and mode are identical, and μ_d is defined such that $\int_{-\infty}^{\mu_d} p(d) dd = 0.5$. Notice that because the ordinate values, Fig. 7, are usually labelled as $P(d) \times 100$, the value of $\hat{\mu}_d = d_{.5} = M_d$ is the 50th percentile value of d , or d_{50} . The variance $\hat{\sigma}_d^2$ is computed as $(d_{84.13} - d_{50})^2$ since

Figure 7. Probability plot of Table II data.

$P(d)$ in the interval $[-\infty, (\mu_d + \sigma_d)]$ is,

$$P(d) [-\infty, (\mu_d + \sigma_d)] = \int_{-\infty}^{(\mu_d + \sigma_d)} p(d) dd = .8413$$

It is emphasized that the estimate of σ_d^2 , $(d_{.8413} - d_{.50})^2 = \hat{\sigma}_d^2$ is a property of $p(d)$. This estimate is only valid for the normal function and can be made for our data only because in plotting on normal probability paper we assumed our sample data were normally distributed.

To illustrate the origin of the estimate $\hat{\sigma}_d = (d_{.8413} - d_{.50})$, consider calculating the value of $P(d)$ in the interval $[(\mu_d - \sigma_d), (\mu_d + \sigma_d)]$. By change of variable

$$t = \frac{d - \mu_d}{\sigma_d}, \quad \sigma_d dt = dd; \quad \text{then } P(t) = \frac{1}{\sqrt{2\pi}} \int_{-1}^{+1} e^{-\frac{t^2}{2}} dt.$$

$P(t)$ can be expanded as a power series, and it is easily shown that $P(t)(-1, +1) = P(d) [(\mu_d - \sigma_d), (\mu_d + \sigma_d)] = .68249$. Therefore, in the interval (μ_d, σ_d) , $P(d) = .34125$, and

$$\int_{-\infty}^{\mu_d + \sigma_d} p(d) dd = .8413.$$

This is, in general, the method by which tables of the normal function are devised.^{16,17}

It is emphasized that the values of $\hat{\mu}_d$, and $\hat{\sigma}_d^2$ obtained from the model $p(d)$, or values of statistics obtained from any functional model, should be obtained from an objective fitting procedure such as Probit¹⁸ for cumulative data, or Minimum Chi Squared¹⁹ for interval data.

IV. TEST OF STATISTICAL HYPOTHESES. (STATISTICAL INFERENCE)

Statistical inference and tests of statistical hypotheses, are aspects of statistics much too complicated to develop here in detail.²⁰ However, it is important that the fundamental meanings of these terms be understood. A statistical hypothesis is the statement of a theory concerning some population. For example, in the case of the "sample data" plotted in Fig. 7 the hypothesis is that the data are from a normal population having certain values of μ_d and σ_d^2 . A complete test of the hypothesis requires first comparison of the sample data to theoretical values obtained from a hypothetical normally distributed population, then comparison of the sample mean and variance to the hypothetical mean and variance.

To illustrate the basic idea of test of statistical hypotheses, we shall test the hypothesis that the "sample data" of Table II were obtained from a normal population, i.e. the population is a normal distribution. The basic procedure for test of hypothesis is:²⁰

1. State the hypothesis.
2. Select a level of significance for the test.
3. Compute the appropriate test statistic from experimental data.
4. Determine from tabulated theoretical values of the distribution of the test statistic the range of values which will cause acceptance or rejection of the hypothesis.
5. Accept or reject the hypothesis on the basis of the computed value of the test statistic.

Test of the hypothesis that the data of Table II are from a normal population is a chi squared^{19,21} "goodness of fit" test.

Let us choose a 5% level of significance, α , for the test. The choice of $\alpha = 5\%$ implies²³ we are willing to accept a .05 probability of rejecting the hypothesis even if it is true.

The test statistic $\hat{\chi}^2$ has a χ^2 distribution²² with $(k-1)-p$ degrees of freedom. The test consists of comparing the observed frequencies, n_i , in the intervals (d_i^*, d_{i+1}^*) of Table II or Fig. 4 with the theoretical frequencies expected in the same interval. The test statistic is calculated as

$$\hat{\chi}^2 = \sum_{i=1}^{I=k} \frac{(n_i - F_i)^2}{F_i}, \quad (14)$$

where F_i = theoretical values of n_i , and number of degrees of freedom = $(k-1)-p$ where p = number of parameters estimated for the population. For this example $p = 2$ and number of degrees of freedom = $(5-1)-2 = 2$.

The normal curve which was fitted to the data has the equation

$$N p(d) \approx \frac{N}{\sqrt{2\pi}\sigma_d} \exp \left[-\frac{(d - \hat{\mu}_d)^2}{2\hat{\sigma}_d^2} \right].$$

F_i values are calculated as

$$F_i = \int_{d_i^*}^{d_{i+1}^*} p(d) dd$$

actually obtained from normal distribution tables

$$\text{of } z = \frac{d - \mu_d}{\sigma_d}. \quad \text{The computations for } \hat{\chi}^2 \text{ are illus-}$$

trated in Table IV. Since we chose $\alpha = .05$, the value $\hat{\chi}^2$ is compared with the theoretical 95th percentile value²² of χ^2 . For a chi squared test, if $\hat{\chi}^2 < \chi^2$, the hypothesis is not rejected. The value of $\chi^2_{.95}(2) = 5.99$. Since $\hat{\chi}^2 = 4.00 < 5.99$ the conclusion drawn from the test is; there is no reason to reject the hypothesis that the population distribution is normal. Basic procedures for testing hypotheses concerning the mean and variance are the same as those outlined earlier, however, different test statistics must be used.²⁵

SUMMARY

The basic concepts and techniques of statistical analysis have been illustrated using a hypothetical small particle sample. The basic procedures for analysis of small particle data are

Table IV

d_i^*	$\frac{d_i^* - \hat{\mu}_d}{\hat{\sigma}_d}$	d_i	$P(d)$	F_i	n_i	$\frac{(n_i - F_i)^2}{F_i}$
1	-3.82		0			
		2		11	6	2.273
3	-2.30		0.0107			
		4		210	224	0.933
5	-0.77		0.2207			
		6		555	540	0.405
7	+0.76		0.7763			
		8		212	220	0.302
9	+2.28		0.9887			
		10		11	10	0.091
11	+3.81		0.9999			

$$N = 1000$$

$$\hat{\mu}_d = 6.01$$

$$\hat{\sigma}_d = \sqrt{\hat{\sigma}_d^2} = 1.31$$

$$\sum_{i=1}^5 \frac{(n_i - F_i)^2}{F_i} = \hat{\chi}^2 = 4.00$$

exactly the same. It will be seen in the next section that some methods of particle "size" measurement produce data which do not permit calculation of estimators of parameters of the population distribution in the technical sense. These data types also make very difficult, if not impossible, use of the procedures for test of hypothesis.

V. SMALL PARTICLE STATISTICS

Dalla Valle stated²⁶ "The characteristics and behavior of small particles can be understood satisfactorily only when the methods of determining size-distribution are placed on a firm basis." A very important aspect of this "firm basis" is the understanding of what is meant by 'size' and the understanding of the differences among various forms of particle 'size' data.

What is "Size"? The definition depends upon the measuring device used to obtain the so-called "size-distribution." This is illustrated by Fig. 8 which shows the definition of "size" imposed by three typical measurement methods. Notice that sedimentation methods measure occurrence of particles as weight, w_i rather than frequency, n_i . The difficulty in unique definition of "size" has an attendant problem in the definition of particle shape. A development of a generalized definition of shape factor will aid in the following development of a completely general treatment of particle "size" data.

As illustrated in Table V, particle "size" can be defined as a linear dimension, an area, volume, or weight, and for a given shape, the concepts of "size" are related by a shape factor. That is a general shape factor C_k relates the k^{th} power of the linear dimension d to some "geometric" particle

TABLE V

"SIZE" DEFINITIONS	PARTICLE SHAPES
	SPHERE	IRREGULAR	REGULAR CUBIC	REGULAR PARALLELEPIPED
$d =$	DIAMETER	MAXIMUM DIMENSION	EDGE	EDGE OR LENGTH
$v = C_3 d^3 =$	$\frac{\pi}{6} d^3$	$C_3 d^3$	d^3	$\frac{L}{d} d^3$
$a = C_2 d^2 =$	πd^2	$C_2 d^2$	$6d^2$	$(2 + \frac{4L}{d}) d^2$
$w =$	ρv	ρv	ρv	ρv

property. For example, $v = C_3 d^3$ and C_3 is a volume shape factor, or $A = C_2 d^2$ and C_2 is the surface area shape factor. Notice the volume and surface shape factors for the idealized cubic particle. Recall the example discussed previously of measurement of cube diagonal as "size" d . In this case, the volume shape factor $C_3 = \frac{1}{3\sqrt{3}}$ and the surface area shape factor is $C_2 = 2$. Suppose the elongated "particle" in Table V had the measure d . An "acicularity ratio" could be defined as L/d . The volume shape factor would also be L/d since $v = (L/d)d^3$.

This discussion indicates that shape factor can be defined in many ways, therefore, careful definition of the meaning of shape factor must be made with respect to the application of the term. For the remainder of the discussion of small particle statistics, the assumption is made that an "average" or equivalent shape, hence C_3 and C_2 , can be defined for the particles in a powder, no matter how irregular the particles.

Figure 8. Examples of the relationship between definitions of "size" and the measuring device.

Figure 9. The two general types of sample data.

VI. GENERAL TYPING OF PARTICLE "SIZE" DATA

Particle "size" data obtained from all possible methods of "size" measurement can be classified into two general Types²⁷ as illustrated in Fig. 9. Type I data is sample distribution data, and diagrams of the type shown are histograms. Type II data are sample distribution data only if $k = 0$. Diagrams of this type are called moment relationships. In the case of Coulter counter data, "size" is defined as $C_3 d^3 = v$, where v is measured, and fraction of occurrence in the various "size" intervals is measured as n_i/N . This would be called type (3,I) data. Sedimentation methods measure "size" as d = equivalent Stokes diameter, and fraction of occurrence as w_i'/W or weight fraction. The weight of particles in an interval (d_i^*, d_{i+1}^*) can be estimated as

$$w_i' \approx \rho C_3 n_i d_i^3, \quad (15)$$

and weight fraction as,

$$\frac{w_i'}{W} \approx \frac{\rho C_3 n_i d_i^3}{\sum \rho C_3 n_i d_i^3}; \quad (16)$$

hence sedimentation data is of the general Type II or data type (3,II). Data obtained from microscopic count can be considered either Type I or II.

VII. CALCULATION OF SAMPLE STATISTICS AND DATA COMPARISON FOR THE FINITE INTERVAL MODEL

If a functional model representing particle "size" sample data is not recognized, sample statistics can be computed in a manner analogous to the methods previously described for calculating mean, variance and skewness. The definitions of these statistics are given in Table VI. The statistics in the first and last columns are estimators of the three parameters of a population distribution, however, those in the middle are related to the distribution parameters only through the approximation $w_i' \approx \rho C_3 n_i d_i^3$. However, they are sample statistics and do define the sample data. If an instrument measuring weight fraction data were consistently used, it would certainly be legitimate to use these sample statistics for both sample characterization and sample comparison. As a matter of fact, the statistics computed directly from the sample, give the better basis for sample comparison than, for example, a $\bar{d}_3$ value computed from Coulter Counter data.

Table VI

	Optical	Sediment	Coulter C.
Mean:	$\bar{d} = \sum \frac{n_i}{N} d_i$	$\bar{d}_3 = \sum \frac{w_i'}{W} d_i$	$\bar{v} = \sum \frac{n_i}{N} v_i$
Variance:	$s_d^2 = \sum \frac{n_i (d_i - \bar{d})^2}{N - 1}$	$s_{d_3}^2 = \sum \frac{w_i'}{W} (d_i - \bar{d})^2$	$s_v^2 = \sum \frac{n_i (v_i - \bar{v})^2}{N - 1}$
Skewness:	$g_d = \sum \frac{n_i}{N} (d_i - \bar{d})^3$	$g_{d_3} = \sum \frac{w_i'}{W} (d_i - \bar{d})^3$	$g_v = \sum \frac{n_i}{N} (v_i - \bar{v})^3$

It is possible that there may be a special requirement to compare a $\bar{d}$ value from optical count data with a $\bar{d}$ value obtained from weight fraction or Coulter Counter data. These calculations should be based on the proper definition of midpoint values, endpoint values, and the approximation $w'_i \approx \rho C_3 n_i d_i^3$.

Figure 10 illustrates the effect of transforming (0,I) data to (3,I) data. The variable endpoint values are transformed using the definition $v_i^* = C_3 d_i^3$. Notice the distortion of the interval width on transforming the variable to v .

Figure 11 illustrates the effect of converting relative frequency to weight fraction. Notice that w'_i/W in a particular interval is greater than n_i/N for values of $d > 1$.

The general conversions and transformations for Types I and II data have been developed in detail in reference 27. To illustrate the method, consider computing $\bar{d}$ from weight fraction and Coulter Counter data.

In computing $\bar{d}$ from Coulter Counter data, we use the endpoint transformations,

$$d_i^* = \left(\frac{v_i^*}{C_3} \right)^{1/3}$$

and the midpoint definition

$$d_i = \left[\left(\frac{v_i^*}{C_3} \right)^{1/3} + \left(\frac{v_{i+1}^*}{C_3} \right)^{1/3} \right] / 2, \quad (17)$$

and $\bar{d}$ is calculated as,

$$\bar{d} \approx \frac{1}{N} \sum_{i=1}^I \frac{n_i \left[\left(\frac{v_i^*}{C_3} \right)^{1/3} + \left(\frac{v_{i+1}^*}{C_3} \right)^{1/3} \right]}{2} \quad (18)$$

Figure 10. The effect of Type I transforms.

Figure 11. The effect of Type II transforms.

The calculation of $\bar{d}$ from weight fraction data is slightly more complex. The fraction of occurrence in a given "size" interval is converted using the approximations,

$$w'_i \approx \rho C_3 n_i d_i^3, \text{ and}$$

$$\frac{w'_i}{W} \approx \frac{(\rho C_3 n_i d_i^3)}{\sum_{i=1}^I \rho C_3 n_i d_i^3},$$

assuming C_3 and ρ constant for all intervals.

$$\text{Since } \frac{w'_i}{W} \approx \frac{n_i d_i^3}{\sum_{i=1}^I n_i d_i^3}; \text{ then } \frac{w'_i}{W} \frac{1}{d_i^3} \approx \frac{n_i}{N}$$

and $\bar{d}$ may be written,

$$\bar{d} \approx \frac{\sum_{i=1}^I \frac{w'_i}{W} \frac{d_i}{d_i^3}}{\sum_{i=1}^I \frac{w'_i}{W} \frac{1}{d_i^3}} \quad (19)$$

In expression (19) the assumption is made that the midpoint value, d_i , represents the diameter of a particle of average weight in each of the I intervals. This is not necessarily true. In 1910 J. Mellor suggested the following definitions to estimate the average weight $\bar{w}_i$, of a particle in the i^{th} interval:

$$w'_i \approx \int_{d_i^*}^{d_{i+1}^*} \rho C_3 y^3 dy; \quad (20)$$

then,

$$\bar{w}_i = \rho C_3 \int_{d_i^*}^{d_{i+1}^*} y^3 dy / \int_{d_i^*}^{d_{i+1}^*} dy ; \quad (21)$$

and

$$\bar{w}_i = \frac{\rho C_3 [d_{i+1}^{*4} - d_i^{*4}]}{4 [d_{i+1}^* - d_i^*]} \quad (22)$$

A value of d , related to $\bar{w}_i$ can be defined such that,

$$d_{Mi}^3 = \left(\frac{\bar{w}_i}{\rho C_3} \right) \quad (23)$$

Then

$$d_{Mi} = \left[\frac{(d_{i+1}^{*2} + d_i^{*2})(d_{i+1}^* + d_i^*)}{4} \right]^{1/3} \quad (24)$$

and we can rewrite the expression for $\bar{d}$ (from weight fraction data) as

$$\bar{d} = \left[\frac{\sum_1 \frac{w_i}{W} \frac{d_i}{d_{Mi}^3}}{\sum_1 \frac{w_i}{W} \frac{1}{d_{Mi}^3}} \right] \quad (25)$$

By expanding the expressions in Table VI for variance and skewness in d_3 and v , we can make the same transformations and conversions to get s_d^2 and g_d from these data.

To summarize the methods for finite interval data:
It should now be clear that data or statistics computed from data from different types of measuring devices cannot be directly compared. Sedimentation data or C.C. can be compared with microscope count data by using proper conversions of fraction of occurrence or transforms of "size" intervals. However, for data comparisons we should, where possible, compare those statistics computed directly from sample data. One should be especially careful of converting n_i/N data to w_i/W . Computer experiments have indicated²⁸ a sample of $N = 2500$ may not be adequate.

VIII. CALCULATION OF SAMPLE STATISTICS & DATA COMPARISON, LOG NORMAL MODEL

Many different types of functions have been used to represent particle "size" data. Perhaps the most useful is the log normal function. Development of the use of the log normal function began with the work of Galton and Macallister²⁹ in 1879. Their work lead to the use of the log normal function by Hazen³⁰ in 1914, Kapteyn³¹ in 1916, Wightman,³² 1921-1924, Loveland and Trevelli,³³ 1927, Hatch and Choate³⁴ and many others.

Histograms of particle "size" data type (0, II) or (1, I) are in general positively skewed. If the sample distribution variable is defined as

$$x = \ln d, \quad (26)$$

where $\ln = \log_e$

i.e. "size" is defined as $x = \ln d$, the normal density function,

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma_x} \exp \left[-\frac{(x - \mu_x)^2}{2\sigma_x^2} \right] \quad (27)$$

is often a good approximation of the sample data. The use of $p(x)$ as a model implies the same approximation method as expression (11). That is

$$\frac{n_i}{N} (x_i^*, x_{i+1}^*) \approx \int_{x_i^*}^{x_{i+1}^*} p(x) dx \quad (28)$$

and the distribution variable is $x = \ln d$. The effect of transforming the variable is illustrated by Fig. (12). Notice $p(x)$ is normal having identical values of mean, median and mode. The transformation implies the population is log normal in d , and normal in $x = \ln d$.

The use of the log normal function as a model has three basic advantages,

1. The sample data are defined by only two statistics.
2. Various data forms are related through functional relationships among their sample statistics.
3. "Geometric properties" such as particle surface area and volume can be estimated as simple exponential functions.

Figure 12. Illustrating the relation between a log normal density function and its normal transform.

IX. THE GENERAL TYPES OF LOG NORMAL FUNCTIONS

The various forms of the log normal function can be classified as one of two types analogous to Types I and II data forms, as shown in Fig. 13. Consider Type II first and define $f_k(x)$ in terms of $p(x)$ using the approximation

$$\frac{\frac{1}{N} \sum_{i=1}^N \frac{n_i d_i^k}{d_i}}{\frac{1}{N} \sum_{i=1}^N \frac{n_i}{d_i}} \approx \int_{x_i^*}^{x_{i+1}^*} f_k(x) dx. \quad (29)$$

Since $x = \ln d$, $d^k = e^{kx}$, and

$$\frac{\frac{1}{N} \sum_{i=1}^N \frac{n_i d_i^k}{d_i}}{\frac{1}{N} \sum_{i=1}^N \frac{n_i}{d_i}} \approx \frac{\int_{x_i^*}^{x_{i+1}^*} e^{kx} p(x) dx}{\int_{-\infty}^{\infty} e^{kx} p(x) dx}. \quad (30)$$

Figure 13. Schematic representation of the general form of Type I and Type II log normal density functions.

After multiplying the exponential terms, completing the square in the integrands, and rearranging constant terms, expression (30) is written,

$$\frac{\frac{1}{N} \sum_{i=1}^N \frac{n_i d_i^k}{d_i}}{\frac{1}{N} \sum_{i=1}^N \frac{n_i}{d_i}} \approx \frac{\int_{x_i^*}^{x_{i+1}^*} f_k(x) dx}{\int_{-\infty}^{\infty} f_k(x) dx} = \frac{\frac{\exp[k\mu_x + \frac{k^2}{2}\sigma_x^2]}{\sqrt{2\pi}\sigma_x}}{\frac{\exp[k\mu_x + \frac{k^2}{2}\sigma_x^2]}{\sqrt{2\pi}\sigma_x}} \frac{\int_{x_i^*}^{x_{i+1}^*} \exp\left[-\frac{\{x - (\mu_x + k\sigma_x^2)\}^2}{2\sigma_x^2}\right] dx}{\int_{-\infty}^{\infty} \exp\left[-\frac{\{x - (\mu_x + k\sigma_x^2)\}^2}{2\sigma_x^2}\right] dx}$$

This demonstrates the following properties of the functions representing type II data.

1. $\int_{-\infty}^{\infty} f_k(x) dx = 1$, this is a necessary condition if the integral is to have properties of a density function.

2. $\int_{-\infty}^{\infty} d^k p(x) dx = \exp[k\mu_x + \frac{k^2}{2}\sigma_x^2]$, (32)
is the value of the k^{th} moment of the log normal function, $p(x)$.

$$f_k(x) = \frac{e^{kx} p(x)}{\exp[k\mu_x + \frac{k^2}{2}\sigma_x^2]}, \quad (33)$$

is the general definition of a k^{th} moment function analogous to the normal density function.

4. The relationships among the parameters of Type II functions are,

$$\mu_{xk} = \mu_x + k\sigma_x^2, \quad (34)$$

$$\sigma_{xk}^2 = \sigma_x^2. \quad (35)$$

Now let us develop in brief the relationships among the parameters of the Type I functions. If $z_k = \ln(C_k d^k)$, the variable x in expression (27) can be transformed using the expressions, $z_k = \ln C_k + kx$, and $dz_k = kdx$ to obtain the approximation,

$$\frac{n_i}{N} \approx \frac{1}{\sqrt{2\pi k} \sigma_x} \int_{z_{k,i}}^{z_{k,i+1}^*} \exp \left\{ -\frac{[z_k - (\ln C_k + k \mu_x)]^2}{2 k^2 \sigma_x^2} \right\} dz_k$$

$$= \int_{z_{k,i}}^{z_{k,i+1}^*} p(z_k) dz_k \quad (36)$$

The form of expression (36) indicates $p(z_k)$ is a normal function of z_k . The relationships among the parameters of the Type I functions are,

$$\mu_{z_k} = \ln C_k + k \mu_x, \quad (37)$$

and,

$$\sigma_{z_k}^2 = k^2 \sigma_x^2 \quad (38)$$

X. COMPARISON OF SAMPLE STATISTICS - LOG NORMAL MODEL

The utility of these relationships is illustrated by Figures 14 and 15 which schematically demonstrate the method for obtaining sample statistics from log probability plot of three typical forms of sample data. These plots are analogous to Fig. 7. It is emphasized that the assumption has been made that $\log_e$ "size" is the normally distributed variable. The values of the variances are properties of the appropriate density function just as discussed for Fig. 7. For example $\hat{\sigma}_x^2 = (x_{.8413} - x_{.50})^2$ has no direct relationship to the variance of the sample distribution of d , in fact $\hat{\sigma}_d^2$ is an exponential function of $\hat{\mu}_x$ and $\hat{\sigma}_x^2$ as we shall see.

Figure 14. Log probability plot - Type I.

It is also emphasized that values of mean and variance should be obtained from a fitting routine such as Probit¹⁸ or Chi Squared Minimum.¹⁹

The sample statistics $\hat{\mu}_x$, $\hat{\sigma}_x^2$, $\hat{\mu}_{x3}$, $\hat{\sigma}_{x3}^2$, $\hat{\mu}_{z_3}$ and $\hat{\sigma}_{z_3}^2$ are the logical statistics to use for sample characterization comparison. The relationships among the distribution parameters given in expressions (34), (35), (37) and (38) are the basis for comparing the sample statistics obtained from different data forms i.e. measurement methods. The sample statistics are related as follows:

$$\hat{\mu}_{x3} \approx \hat{\mu}_x + 3 \hat{\sigma}_x^2, \quad (39)$$

$$\hat{\sigma}_{x3}^2 \approx \hat{\sigma}_x^2, \quad (40)$$

Figure 15. Log probability plot - Type II.

$$\hat{\mu}_{z_3} \approx \ln C_3 + 3\hat{\mu}_x, \text{ and} \quad (41)$$

$$\hat{\sigma}_{z_3}^2 \approx 9\hat{\sigma}_x^2. \quad (42)$$

If for some reason it is necessary to examine the statistics of the distribution of "size" rather than $\ln$ "size", the expressions of Table VI are used as the basic definitions with the proper function approximating fraction of occurrence. For example, to compute the values of $\bar{d}$, s_d^2 and g_d , the approximation

$$\frac{n_i}{N} \approx \int_{x_i}^{x_{i+1}^*} p(x) dx \text{ is used, the expressions of}$$

Table VI are expanded where necessary, and the following estimators are obtained.

$$\bar{d} \approx \exp [\hat{\mu}_x + .5\hat{\sigma}_x^2] \quad (43)$$

$$s_d^2 \approx \exp [2\hat{\mu}_x + 2\hat{\sigma}_x^2] - \bar{d}^2, \quad (44)$$

and

$$g_d \approx \exp [3\hat{\mu}_x + 4.5\hat{\sigma}_x^2] - 3 \exp [3\hat{\mu}_x + 2.5\hat{\sigma}_x^2] + 2\bar{d}^3 \quad (45)$$

The mode and median are obtained as,

$$M_d \approx \exp [\hat{\mu}_x - \hat{\sigma}_x^2] \quad (46)$$

$$d_{.50} \approx e^{\hat{\mu}_x} \quad (47)$$

By using the relations (39), (40), (41) and (42), these statistics can be obtained from any data form. For example, an estimate of $\hat{\mu}_d \approx \bar{d}$ from Coulter Counter data is obtained as

$$\bar{d} \approx \left[\frac{1}{C_3} \exp [\hat{\mu}_{z_3} + \frac{.5}{3} \hat{\sigma}_{z_3}^2] \right]^{\frac{1}{3}},$$

and from weight fraction data,

$$\bar{d} \approx \exp [\hat{\mu}_{x3} - 2.5 \hat{\sigma}_{x3}^2].$$

Expressions for $\bar{d}_3$, $s_{d_3}^2$, g_{d_3} , $\bar{v}$, s_v^2 and g_v are obtained in the same manner as expressions (43), (44), and (45). For example, $\bar{d}_3 \approx \exp [\hat{\mu}_{x3} + .5\hat{\sigma}_{x3}^2]$ and $\bar{v} \approx \exp [\hat{\mu}_{z_3} + .5\hat{\sigma}_{z_3}^2]$. From this discussion, inequalities of the following type should be obvious,

$$\bar{d} \neq e^{\hat{\mu}_x}, s_d^2 \neq e^{\hat{\sigma}_x^2}, \bar{d}_3 \neq \bar{d}, s_{d_3}^2 \neq s_d^2 \text{ and } \bar{d} \neq \left(\frac{\bar{v}}{C_3} \right)^{\frac{1}{3}}.$$

XI. SURFACE AREA & SPECIFIC SURFACE CALCULATIONS

Specific surface is one of the characteristics of a powder often considered in attempts at process behavior correlation. S_W can be measured by adsorption techniques (BET)³⁵ or estimated from the particle "size" distribution. Such estimates assume non-adsorbative "smooth shell particles". Comparison of calculated values of S_W with BET values can give some insight into the character of the particle surface.

To illustrate the computation of S_W as a sample statistic,²⁷ consider calculating S_W from weight fraction data for the finite interval model and from microscopic count data for log normal model. For the finite interval model the approximation $a'_i \approx C_2 n_i d_i^2$ is used to define S_W as,

$$S_W \approx \frac{\sum a'_i}{W} = \frac{C_2}{\rho C_3} \frac{\sum \frac{w_i d_i^2}{W d_i^3}}{\sum \frac{1}{Mi}} \quad (48)$$

Notice this expression implies a knowledge of C_2 and C_3 . We believe a better estimate of S_W is obtained if a value representing the diameter of a particle having "average" surface area in the interval is used rather than d_i in the numerator of (48). This value of d^{27} is defined as,

$$d_{Li} = \left[\frac{(d_{i+1}^*)^2 + (d_{i+1}^*)(d_i^*) + (d_i^*)^2}{3} \right]^{\frac{1}{2}}. \quad (49)$$

Calculation of S_W by this method is laborious if I is large enough to adequately define the data. If appropriate, the use of the log normal model greatly facilitates the calculation. For example, from

microscopic count, (0,II) data and the definition of the general moment function (32), the approximation for S_W can be written

$$S_W \approx \frac{NC_2}{\rho NC_3} \frac{\int_{-\infty}^{\infty} d^2 p(x) dx}{\int_{-\infty}^{\infty} d^3 p(x) dx}$$

$$= \frac{C_2}{\rho C_3} \exp \left[-(\hat{\mu}_x + 2.5 \hat{\sigma}_x^2) \right]. \quad (50)$$

Notice that expression (50) shows that $S_W \neq \frac{C_2}{\rho C_3}$.

From the relationships among sample statistics for various data forms, S_W can be easily estimated from any type of data as illustrated by Table VII.²⁷

SUMMARY

An attempt has been made to develop the small particle statistics as a natural evolution from basic statistical concepts. It was indicated that tests of statistical hypothesis were possible for "goodness of fit" and comparisons of mean and variance respectively for particle "size" data of the type n_i/N vs "size". However, the methods of statistical inference do not apply to weight fraction or general moment data in a neat, clearcut manner. This is because the test statistic distributions have not been defined for Type II or k^{th} order moment data. Sample data comparisons for moment data must in general be based on subjective analysis tempered by a good portion of practical experience.

We have shown that all forms of particle 'size' data can be classified as one of two general types, and that such classification directly relates data forms generated by a particular measurement method with the proper mathematical treatment of the data. We believe this development provides a complete basis for the understanding of the analytical relationships among the various forms of particle "size" data, and for the development of a theory for mixtures.²⁷

REFERENCES

1. J. M. DallaValle, *Micromeritics*, Pitman Publishing Corp., New York, 2nd Ed. Chapt. 23.
2. T. Allen, *Particle Size Measurement*, Chapman Hall Ltd., London, 1968, Chapt. 1.
3. W. Dixon & F. Massey Jr., *Introduction to Statistical Analysis*, McGraw Hill, 2nd Ed., Chapt. 7.
4. R. D. Cadle, *Particle Size, Theory & Industrial Application*, Reinhold, New York, 1965.
5. N. A. Fuchs, *The Mechanics of Aerosols*, MacMillan Company, New York, 1964.
6. R. Irani & C. Callis, *Particle Size Measurement, Interpretation & Application*, J. Wiley & Sons, N.Y., 1963.
7. P. G. Hoel, *Introduction to Mathematical Statistics*, J. Wiley & Sons, New York, 2nd Ed., Chaps. 1 and 2.
8. K. Brownlee, *Statistical Theory & Methodology in Science & Engineering*, Wiley & Sons, N.Y., 1960.
9. N. Johnson & F. Leone, *Statistics and Experimental Design in Engineering and Physical Sciences*, Vol. I, J. Wiley & Sons, New York, 1964.
10. Dixon & Massey, *Particle Size Measurement*, Chapman Hall Ltd., London, 1968, Chapt. 20.
11. P. G. Hoel, *Introduction to Mathematical Statistics*, J. Wiley & Sons, New York, 2nd Ed., Chapt. 5.
12. W. G. Cochran, "The χ^2 Test of Goodness of Fit", *Annals of Math. Stat.*, Sept. 1952, Vol. 23, No. 3.
13. P. C. Hoel, *Introduction to Mathematical Statistics*, J. Wiley & Sons, New York, 2nd Ed., pp. 196-199.
14. A. Hald, *Statistical Theory with Engineering Applications*, Wiley & Sons, New York, 1952, pp. 120-121.
15. A. Hald, *Ibid*, p. 130.
16. *Handbook of Chemistry & Physics*, Chemical Rubber Company.
17. A. Hald, *Statistical Tables and Formulas*, J. Wiley and Sons, New York, 1952.
18. D. J. Finney, *Probit Analysis*, Cambridge Press, London, 1947.
19. Harald Cramer, *Mathematical Methods of Statistics*, Princeton Univ. Press, 1951, pp. 424-452.
20. Dixon & Massey, *Introduction to Statistical Analysis*, McGraw Hill, 2nd Ed. Chaps. 5-14.
21. *Ibid*, Chapt. 13.

TABLE VII
FUNCTIONAL ESTIMATES OF PARTICLE SURFACE AREA, VOLUME, AND WEIGHT

Property Estimated	Measuring Devices		
	Microscope (relative frequency of d)	Micromerograph, Sédibal, etc. (weight fraction of d)	Coulter Counter (relative frequency of v)
N	Measured	$\frac{W}{\rho C_3} \exp \left[-(\hat{\mu}_{x3} - 4.5 \hat{\sigma}_x^2) \right] 10^{-12}$	Measured
$A(\text{cm}^2)$	$NC_2 \exp \left[2\hat{\mu}_x + 2\hat{\sigma}_x^2 \right] 10^{-8}$	$\frac{WC_2}{\rho C_3} \exp \left[-(\hat{\mu}_{x3} - .5 \hat{\sigma}_x^2) \right] 10^4$	$\frac{NC_2}{C_3^{2/3}} \exp \left[\frac{2}{9} (3\hat{\mu}_{z3} + \hat{\sigma}_{z3}^2) \right] 10^{-8}$
$V(\text{cm}^3)$	$NC_3 \exp \left[3\hat{\mu}_x + 4.5 \hat{\sigma}_x^2 \right] 10^{-12}$	$\frac{W}{\rho}$	$N \exp \left[\hat{\mu}_{z3} + .5 \hat{\sigma}_{z3}^2 \right] 10^{-12}$
$W(\text{g})$	$\rho NC_3 \exp \left[3\hat{\mu}_x + 4.5 \hat{\sigma}_x^2 \right] 10^{-12}$	Measured	$\rho N \exp \left[\hat{\mu}_{z3} + .5 \hat{\sigma}_{z3}^2 \right] 10^{-12}$
$S_v(\text{cm}^{-1})$	$\frac{C_2}{C_3} \exp \left[-(\hat{\mu}_x + 2.5 \hat{\sigma}_x^2) \right] 10^4$	$\frac{C_2}{C_3} \exp \left[-(\hat{\mu}_{x3} - .5 \hat{\sigma}_x^2) \right] 10^4$	$\frac{C_2}{C_3^{2/3}} \exp \left[-\frac{1}{18} (6\hat{\mu}_{z3} + 5\hat{\sigma}_{z3}^2) \right] 10^4$
$S_v(\text{cm}^2/\text{g})$	$\frac{C_2}{\rho C_3} \exp \left[-(\hat{\mu}_x + 2.5 \hat{\sigma}_x^2) \right] 10^4$	$\frac{C_2}{\rho C_3} \exp \left[-(\hat{\mu}_{x3} - .5 \hat{\sigma}_x^2) \right] 10^4$	$\frac{C_2}{\rho C_3^{2/3}} \exp \left[-\frac{1}{18} (6\hat{\mu}_{z3} + 5\hat{\sigma}_{z3}^2) \right] 10^4$

Units: d in microns, ρ in g/cm^3 .

22. Ibid, Table A-6a.
 23. Ibid, Chapt. 7.
 24. Ibid, Chapt. 5, Sec. 3.
 25. Ibid, Chapt. 7-14.
 26. J. M. DallaValle, Micromeritics, Pitman Publishing Corp., New York, 2nd Ed., Chapt. 23, p. 6.
 27. H. D. Lewis and A. Goldman, Theoretical Small Particle Statistics: A Summary of Methods for Data Analysis with Recent Developments in Data Comparison, Notation, and Mixture Theory, Los Alamos Scientific Laboratory report LA-3656, 1968.
 28. Ibid, Appendix III.
 29. F. Galton, D. MacAllister, The Law of the Geometric Mean, Proc. Roy. Soc., Vol. 29, 1879, pp. 365-367.
 30. A. Hazen, Trans. ASCE, vol 77, p. 1549-1550, 1914.
 31. Kapteyn & VanUhlen, Skew Frequency Curves in Biology and Statistics, Groniugen, 1916.
 32. Wightman et al. I. J. Phys. Chem. Vol. 25, p. 181, 1921, Vol. 25, p. 561, 1921, Papers III, IV, V, J. Phys. Chem., Vol. 27, p. 1, p. 141, p. 466, 1923, also Vol. 28, p. 529, 1924.
 33. R. Loveland and A. Trivelli, J. FI. Aug.-Sept. 1927.
 34. T. Hatch & S. Choate, J. F. I., March 1929.
 35. S. Gregg and K. Sing, Adsorption, Surface Area, and Porosity, Academic Press, New York, 1967.
- XII. GENERAL REFERENCE TEXT LIST
- A. Statistics (Introductory)
- Kendall & Buckland, A Dictionary of Statistical Terms, Hafner Pub. Co., N.Y., 1957.
- W. Dixon & F. Massey Jr., Introduction to Statistical Analysis, McGraw Hill, 2nd Ed.
- A. Hald, Statistical Theory With Engineering Applications, Wiley & Sons, New York, 1952.
- B. Statistics (Elementary Mathematical)
- P. G. Hoel, Introduction to Mathematical Statistics, J. Wiley & Sons, New York, 2nd Ed.
- K. Brownlee, Statistical Theory & Methodology in Science & Engineering, Wiley & Sons, N.Y., 1960.
- C. "Small Particle Statistics"
- H. D. Lewis and A. Goldman, Theoretical Small Particle Statistics:, etc., Los Alamos Scientific

Laboratory LA-3656, 1968.

G. Herdan, Small Particle Statistics, Elsevier Pub. Co., N.Y., 1953.

D. "Particle Size" Measurement Methods

Symposium on Particle Size Analysis, Feb. 4, 1947, Supplement to Transactions, Institution of Chemical Engineers, Vol. 25, 1947.

Symposium on Particle Size Measurement, ASTM Special Technical Publication No. 234, American Society for Testing Materials, 1959.

C. Orr & J. M. DallaValle, Fine Particle Measurement, MacMillan Co., N.Y., 1959.

T. Allen, Particle Size Measurement, Chapman & Hall Ltd., London, 1968.

R. Irani & C. Callis, Particle Size, Measurement, Interpretation, and Application, Wiley & Sons, N.Y., 1963.

APPENDIX I

NOTATION AND DEFINITION

A List of Symbols and Definitions in Approximate

Order of Introduction

Powder: (Particulate Material)	A set of particles of matter in some environment.
Particle:	A unit of matter having a defined boundary, whose "size" and "shape" depend on interatomic or intermolecular bonding forces and its process of formation.
Population:	A population of powder particles is the main or total mass of powder whose properties or characteristics are in question.
Population Parameters:	Constants which define the population distribution.
Sample:	The sample is a portion of the population upon which measurements are performed in order to obtain estimates of the characteristics of the population.
Statistics:	The variates (numerical quantities), or functions of the variates obtained from measurements made on a sample.
Random Sample:	A sample chosen by a method so that it has the same probability of being selected as any other sample.
d:	A linear particle dimension characteristically measurable for all particles in the population of particles.
k:	An integer = 0, 1, 2, 3, ... determined by a particular defined geometric property of the particle.

C_k : Shape Factor; a coefficient constant relating the k^{th} power of d to a geometric property of the particle (e.g. for $k = 2$, and spherical particles, $C_2 = \pi$ and $C_2 d^2$ = surface area of the particle).

Weight Factor: ρC_3 = volume shape factor multiplied by particle density.

a : surface area of an individual particle

v : volume of an individual particle

w : weight of an individual particle

"Size": a measured particle characteristic that may be 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" intervals; $i = 1, 2, 3, \dots I$, where I = total number of intervals.

d_i, a_i, v_i, w_i : The "midpoint" values of the respective i^{th} "size" intervals.

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

n_i : The frequency (i.e. number) of particles occurring in the i^{th} "size" interval.

N : $\sum_{i=1}^I n_i = n_1 + n_2 + n_3 \dots n_I$: Total number of particles in I size intervals.

$\frac{n_i}{N}$: Relative frequency of occurrence of particles in the i^{th} "size" interval.

(Note: $\sum \frac{n_i}{N} = 1$)

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

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

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

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

ρ : particle density, weight/unit volume

$W = \sum_{i=1}^I w_i'$: The total weight of all the particles in I "size" intervals.

$W \approx \sum_{i=1}^I \rho n_i C_3 d_i^3$

$$\frac{w_i}{W} :$$

Relative weight or weight fraction of particles in the i th "size" interval.

$$\frac{r}{S} = \frac{\sum_{i=1}^r C_{k,i} n_i d_i^k}{\sum_{i=1}^r C_{k,i} n_i d_i^k} :$$

The general expression for the cumulative fraction less than "size" for a fractile diagram, or cumulative frequency polygon, where r refers to the r th "size" interval through which the partial sum is computed. For example, if $k = 0$, then,

$$\frac{r}{S} = \frac{n_i}{N}$$

is the partial sum of all values of n_i/N of a relative frequency histogram for $i = 1, 2, 3, \dots, r$, and r may take any value of $i = 1, 2, \dots, I$. The series of partial sums,

$$\frac{n_1}{N}, \frac{n_1 + n_2}{N},$$

$$\frac{n_1 + n_2 + n_3}{N}, \dots, \frac{\sum_{i=1}^I n_i}{N} = 1,$$

gives the fraction less than "size" for the fractile diagram.

$$m_k' = \frac{\sum_{i=1}^I n_i d_i^k}{N} : \text{The } k^{\text{th}} \text{ moment about the origin of the frequency histogram of } d.$$

$$m_1' = \bar{d} = \frac{\sum_{i=1}^I n_i d_i}{N} : \text{The first moment about the origin of the frequency histogram of } d, \text{ the arithmetic mean of } d.$$

$$\bar{d}_k = \frac{\sum_{i=1}^I n_i d_i^{(k+1)}}{\sum_{i=1}^I n_i d_i^k} : \text{An average dimension of sample particles based on a } k^{\text{th}} \text{ moment relationship. (When } k = 0,$$

$$\bar{d}_0 = \bar{d} = \frac{\sum_{i=1}^I n_i d_i}{N} \text{ and } \bar{d} \text{ is the arithmetic mean which is identically equal to } m_1'.$$

$$m_k = \frac{\sum_{i=1}^I n_i (d_i - \bar{d})^k}{N} : \text{The } k^{\text{th}} \text{ moment about the mean of the frequency histogram of } d.$$

$$m_2 = s^2 = \frac{\sum_{i=1}^I n_i (d_i - \bar{d})^2}{N} : \text{The second moment about the mean of the sample distribution (frequency histogram) of } d, \text{ the sample variance of } d. \text{ For } s^2 \text{ to be}$$

an unbiased estimate of the population variance, the denominator must be $N - 1$.

$$s_{d,k}^2 = \frac{\sum_{i=1}^I n_i d_i^k (d_i - \bar{d}_k)^2}{\sum_{i=1}^I n_i d_i^k} : \text{The measure of sample dispersion for the } k^{\text{th}} \text{ moment relationship. } s_{d,k}^2 \text{ will be}$$

called the sample variance of d for the k^{th} moment relationship.

$$\bar{a} = \frac{\sum_{i=1}^I n_i a_i}{N} \approx \frac{\sum_{i=1}^I n_i C_{2,i} d_i^2}{N} : \text{Average individual surface area of particles in the sample.}$$

$$\bar{v} = \frac{\sum_{i=1}^I n_i v_i}{N} \approx \frac{\sum_{i=1}^I n_i C_{3,i} d_i^3}{N} : \text{Average individual volume of particles in the sample.}$$

$$\bar{w} = \rho \bar{v} \approx \frac{\sum_{i=1}^I n_i \rho C_{3,i} d_i^3}{N} : \text{Average individual weight of particles in the sample.}$$

$$s_a^2 = \frac{\sum_{i=1}^I n_i (a_i - \bar{a})^2}{N - 1} : \text{Sample variance of particle surface areas.}$$

$$s_v^2 = \frac{\sum_{i=1}^I n_i (v_i - \bar{v})^2}{N - 1} : \text{Sample variance of particle volumes.}$$

$$s_w^2 = \rho^2 s_v^2 \approx \frac{\sum_{i=1}^I n_i (\rho C_{3,i} d_i^3 - \bar{w})^2}{N - 1} : \text{Sample variance of particle weights.}$$

$$g_{d,k} = \frac{\sum_{i=1}^I n_i d_i^k (d_i - \bar{d}_k)^3}{\sum_{i=1}^I n_i d_i^k} : \text{The basic definition of sample skewness.}$$

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

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

$x = \ln d$: 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 .

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.

$\approx$: Approximately equal to.

$\hat{\approx}$: Estimates or is estimated by.

$\wedge$: Designator of a sample statistic for samples represented by a functional model.

(Example: $\bar{d} \approx \hat{\mu}_d \hat{\approx} \mu_d$; read as $\bar{d}$ is approximately equal to $\hat{\mu}_d$ which estimates μ_d , the mean of the population of d .)

$$f_k(x) = \frac{e^{kx} e^{-\frac{(x - \mu_x)^2}{2\sigma_x^2}}}{\sqrt{2\pi} \sigma_x \exp \left[k\mu_x + \frac{k^2 \sigma_x^2}{2} \right]} : \text{The definition of the } k^{\text{th}} \text{ moment relationship for the log normal function.}$$