FOLK (1968) GRAPHIC PRESENTATION OF SIZE ANALYSIS DATA Grabe

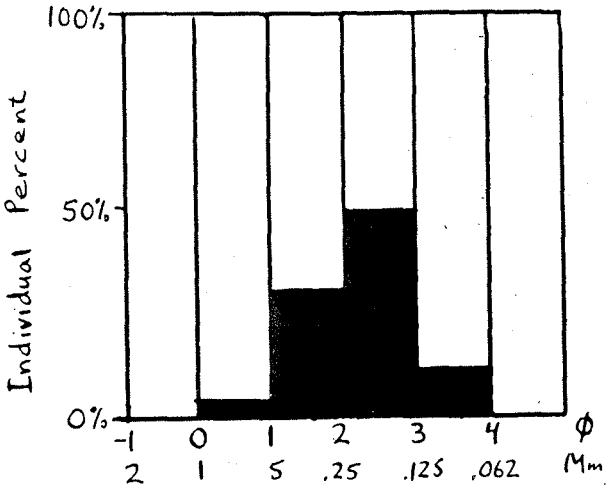
The data obtained from grain-size analysis may be plotted in many different ways. For some purposes, one method is best; for other purposes, other methods are best. Get familiar with all methods, so that you do not become so blindly used to using one method that you ignore other ways which may be more suitable for a given problem.

All the methods use grain size as the abscissa (horizontal scale) and some measure of percentage frequency as the ordinate (vertical scale). Grain size analyses may either be plotted directly in millimeters, using a logarithmic-base paper; or they may be plotted in phi units (ϕ), in which case arithmeticbase paper is used. The latter is much more convenient and accurate to read.

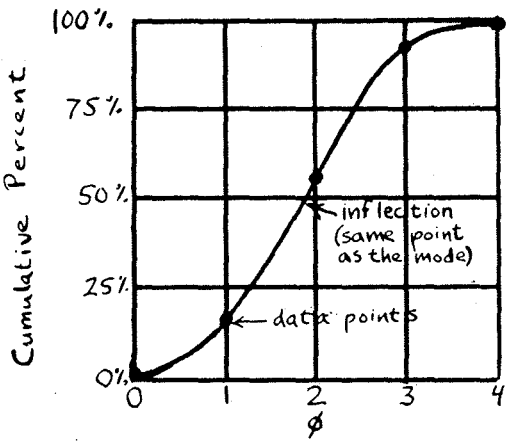
1. Histogram. A histogram is essentially a bar graph in which the percentages for each grade size are plotted as a column. It is very easy to prepare and one can easily interpret general features of the sediment. However, it is a pictorial method, no more, and cannot be used for determination of any statistical parameters such as median, sorting etc. Furthermore its shape is greatly affected by the sieve interval chosen; the same sample may look entirely different if it is analyzed on a different set of screens and another histogram prepared. Nevertheless it proves of value in plotting distribution of sediments on a map or stratigraphic section, as the heights of the columns may be more easily compared by eye than if the data were plotted as cumulative curves. For pictures, yes: otherwise no.

2. Cumulative Curve, Arithmetic Ordinate. This is the most commonly used method. As the abscissa, one may use either millimeters (in which case he must use semi-log paper) or phi units (ordinary "squared" arithmetic paper). The ordinate is an arithmetic scale running from 0 to 100%; grain size is plotted on the abscissa with coarser particles to the left (this is customary in all size analysis plots). Cumulative percentages of the sediment are plotted on this graph; for example, if 32% of the material is coarser than 2ϕ (caught on the 2ϕ screen) then 32 is plotted as the ordinate against 2.0 as abscissa. Draw a curve through all the resulting points. YOUR CURVE MUST PASS THROUGH ALL THE PLOTTED POINTS -- NEVER USE A FRENCH CURVE TO "SMOOTH" OUT THE GRAPH. The sample analysis normally forms an S-shaped curve. The advantage of this curve is that all statistical parameters may be read from it exactly, thus one can compare samples quantitatively as to median, skewness, etc. The shape of the curve is independent of the sieves used. Its only disadvantage is that it is difficult for the untrained eye to look at the curve and interpret it at a glance; it is not "pictorial". Also if the sieve interval is wide, sketching the curve between data points is subject to considerable error.

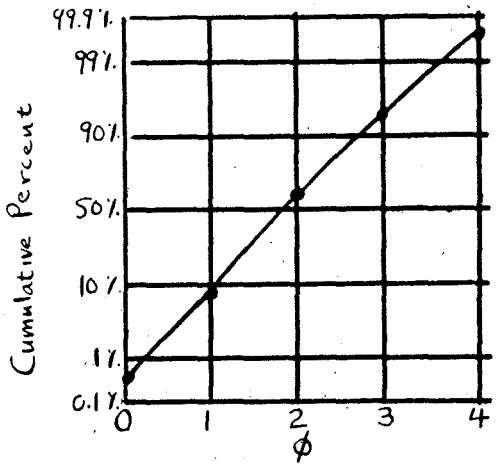
3. Cumulative Curve, Probability Ordinate. Most sediments tend to approach the "normal probability curve" in their size frequency distribution--in other words, most of the particles are clustered about a given size, with less and less material on each side of this size. If the cumulative curve of a sediment following the normal, symmetrical probability distribution is plotted on probability paper, the result is a perfectly straight line whose position depends on the average particle size and whose slope depends on the sorting. This happens because the probability scale is very condensed in the middle of the scale (30 to 70%) and very much expanded at the ends (under 10 or over 90%), thereby straightening out the S-shaped curve which would result if arithmetic ordinates were used. Thus it is very valuable for studying the departure of sediments from the normal probability law. Moreover, since the "tails"



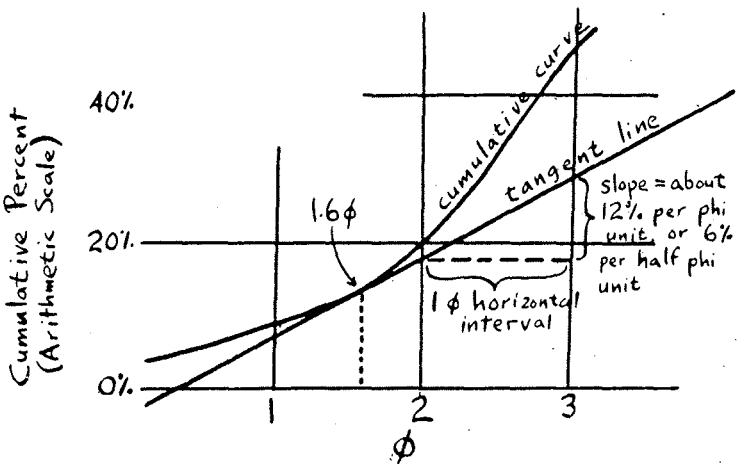
Histogram



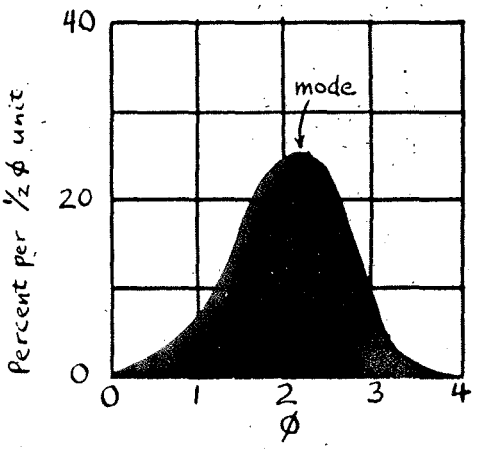
Cumulative Curve, Arithmetic Ordinate



Cumulative Curve, Probability Ordinate



Method of construction. Here we are obtaining the value of the frequency curve at a diameter of 1.6ϕ. It is about 6% per half phi unit.



Thefinal curve (not the same sample as the one at left).

F r e q u e n c y C u r v e

are straightened out and the sample tends to plot as a straight line, it is possible to read the statistical parameters with much greater accuracy because of the ease of interpolation and extrapolation. HENCE, THIS IS THE CURVE THAT MUST BE USED FOR ALL DETERMINATION OF PARAMETERS. The only disadvantage is that it is even less pictorial than the arithmetic cumulative curve, and is not commonly used.

4. Frequency Curve. The frequency curve represents in essence a smoothed histogram in which a continuous bell-shaped curve takes the place of the discontinuous bar graph. Again, it is chiefly of pictorial value because no statistical parameters can be read for it. Although strictly pictorial, it gives a much better picture than the histogram because it is independent of the sieve interval used. Mathematically, it is the first derivative of the cumulative curve, and is thus obtained by measuring the slopes of tangents to the cumulative curve. To construct it, one plots a cumulative curve with arithmetic (not probability) ordinates. Now one measures the slopes of tangents to this curve at various grain-size values. For example, it you want to find the frequency at a diameter of 2.83ϕ, lay a straight-edge tangent to the curve at the point where the 2.83ϕ line intersects it. Measure the slope of this tangent by noting how much the tangent rises over a horizontal distance of 1/2 phi unit. This value then is plotted at 2.83ϕ on the frequency curve, whose vertical scale is stated as "so many percent per half phi unit" (or half Wentworth grade--both being identical). As many points as plotted as needed, spaced anywhere along the curve (not necessarily at the analysis points). Be sure to get all points of inflection (steeper places on cumulative curve) and all "sags" (flatter places); thus you will accurately determine all the modes and minimums on the frequency curve. The mode may be fixed accurately by repeated approximation (page 44).

For very accurate work, a cumulative probability curve should be plotted first, then from this probability curve, data points may be taken to construct a much more accurate cumulative curve on the arithmetic graph paper.

STATISTICAL PARAMETERS OF GRAIN SIZE

For evaluation of sets of samples it is probably best to compare the curves directly by eye as only in that way can the entire character of the sediment curves be revealed. But this is inconvenient and furthermore not very quantitative: it is often difficult to decide whether curve A represents a better sorted or finer sample than curve B, or how great the differences are. To solve this problem one resorts to various statistical measures which describe quantitatively certain features of the curves; these values can then be tabulated and, it is hoped, certain combinations of values may be indicative of different sedimentary environments.

There are two basic methods of obtaining statistical parameters. The most commonly used method involves plotting the cumulative curve of the sample and reading the diameter represented by various cumulative percentages (as, what grain size value corresponds to the 25% mark of the sediment--meaning 25% of the material is coarser than that diameter). In this method, much more accurate results are obtained if one plots the cumulative curve on probability paper, because of the superior accuracy of extrapolation and interpolation. The second method, called the method of moments, is far more complicated and probably of no greater value. It is explained at the end of this section.

The significance of most of the measures here described is shown graphically on page 51. Refer to it continually.

MEASURES OF AVERAGE SIZE

It is desirable to have a measure which will say, "sample A is so much coarser than sample B." This is not nearly as easy as it looks, though, because there are many different measures of average size. Using one measure, sample A might be "coarser"; using another measure, sample B might be coarser. There is no consensus yet as to which is best. Until then, become familiar with all of them.

Mode (mo) is the most frequently-occurring particle diameter. It is the diameter corresponding to the steepest point (point of inflection) on the cumulative curve (only if the curve has an arithmetic frequency scale). It corresponds to the highest point on the frequency curve. Several formulas have been developed for determination of the mode, but none of them are satisfactory. The only way the mode can be determined is by successive trials. Using the graph of the sample plotted on probability paper, one selects a point where the mode ought to be, and measures the percentage of the sample that occurs within the diameter range from 1/4ϕ coarser than that point to 1/4ϕ finer than that point (i.e., within 1/2ϕ interval centered on the presumed modal point). Then he moves over a small distance (say 0.1 or 0.2ϕ) to a new presumed mode and measures the percentage occurring in the 1/2ϕ interval centered on that new point. This is done repeatedly until the highest value is obtained which then corresponds to the modal diameter. It is often difficult to fix the mode more accurately than 0.1 or 0.2ϕ. Sediments not uncommonly have two or more modes, located by finding other points of inflection on the cumulative curve or other peaks on the frequency curve. Advantages: the mode is quite valuable in sediment genesis and transport studies, especially when two or more sources are contributing. The modal diameter often stays fairly constant in an area while the other, more "synthetic" measures tend to vary more erratically. It deserves more common use. The disadvantages are its lack of common usage, and in the fact that it is difficult to determine. Also, it is independent of the grain size of the rest of the sediment, therefore is not a good

measure of overall average size.

Median (Md). Half of the particles by weight are coarser than the median, and half are finer. It is the diameter corresponding to the 50% mark on the cumulative curve and may be expressed either in ϕ or mm. (Md_{ϕ} or Md_{mm}). The advantage is that it is by far the most commonly used measure and the easiest to determine. The disadvantage is that it is not affected by the extremes of the curve, therefore does not reflect the overall size of sediments (especially skewed ones) well. For bimodal sediments it is almost worthless. Its use is not recommended.

Graphic Mean (M_z) (Folk). The best graphic measure for determining overall size is the Graphic Mean, given by the formula $M_z = (\phi 16 + \phi 50 + \phi 84)/3$. It corresponds very closely to the mean as computed by the method of moments, yet is much easier to find. It is much superior to the median because it is based on three points and gives a better overall picture. This will be the standard measure of size used. Inman has used $(\phi 16 + \phi 84)/2$ as a measure of mean size but this is not satisfactory in skewed curves.

MEASURES OF UNIFORMITY

Several measures are available for measuring the uniformity or sorting of sediments. As a general rule, the more of the curve that enters into the sorting coefficient, the better the measure.

Trask's Sorting Coefficient (S_o) is used only with millimeter values, and is given by $\sqrt{Mm25/Mm75}$. Most beach sands have S_o = 1.3-1.5. In the past, it has been almost the only measure of sorting used, but its use is declining because it measures only the sorting in the central part of the curve. It should be abandoned.

Phi Quartile Deviation ($QD\phi$) is the exact analogue of S_o but adapted for the ϕ scale; it is given by $(\phi 75 - \phi 25)/2$. This measure, like S_o, fails to give a good indication of sorting because they indicate only the sorting in the middle of the curve and ignore the ends, where the differences between samples are most marked. For example, a beach sand consisting of nothing else but fine and medium sand might have the same $QD\phi$ and S_o as a sediment consisting of sand with 15% pebbles and 10% clay! Therefore these measures should be no longer used.

Graphic Standard Deviation (σ_G) is $(\phi 84 - \phi 16)/2$. It is very close to the standard deviation of the statistician (see the method of moments) but is obtained by reading two values on the cumulative curve instead of by lengthy computation. This sorting measure embraces the central 68% of the distribution, thus is better than $QD\phi$ but not as good as σ_I If a sediment has σ_G of 0.5ϕ, it means that two thirds (68%) of the grains fall within 1ϕ unit or 1 Wentworth grade centered on the mean - i.e., the mean $\pm$ one standard deviation.

Inclusive Graphic Standard Deviation (σ_I) (Folk) The Graphic Standard Deviation, σ_G, is a good measure of sorting and is computed as ($\phi 84 - \phi 16)/2$. However, this takes in only the central two-thirds of the curve and a better measure is the Inclusive Graphic Standard Deviation, σ_I, given by the formula $\frac{\phi 84 - \phi 16}{4} + \frac{\phi 95 - \phi 5}{6.6}$. This formula includes 90% of the distribution and is the best overall measure of sorting. It is simply the average of (1) the standard deviation computed from $\phi 16$ and $\phi 84$, and (2) the standard deviation as computed from $\phi 5$ and $\phi 95$--since this interval (from 5 to 95%) embraces 3.30σ, the standard deviation is found as $(\phi 95 - \phi 5)/3.30$. The two are simply averaged together (which explains why the denominators are both multiplied by 2).

Note that the standard deviation here is a measure of the spread in phi units of the sample, therefore the symbol ϕ must always be attached to the value for σ_I

Measurement of sorting values for a large number of sediments has suggested the following verbal classification scale for sorting:

σ_I		
under .35ϕ	very well sorted	1.0-2.0ϕ, poorly sorted
.35-.50ϕ	well sorted	2.0-4.0ϕ, very poorly sorted
.50-.71ϕ	moderately well sorted	over 4.0ϕ, extremely poorly sorted
.71-1.0ϕ	moderately sorted	

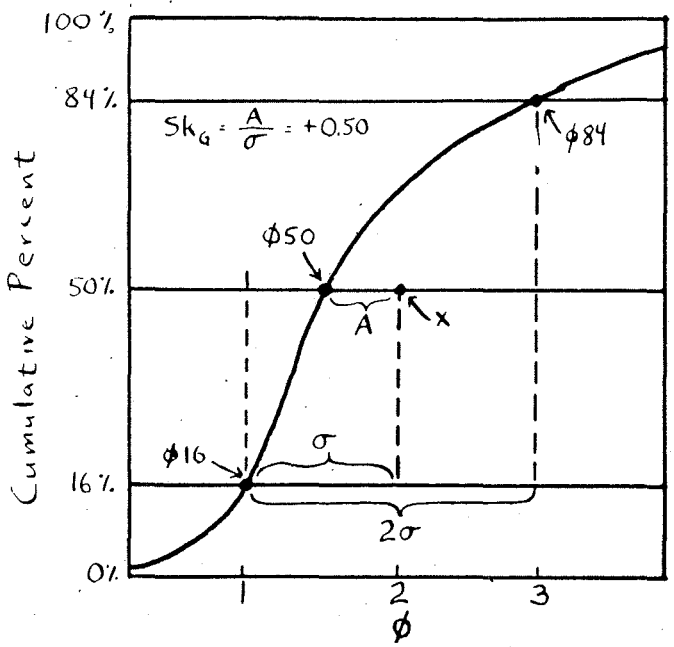
The best sorting attained by natural sediments is about .20-25ϕ, and Texas dune and beach sands run about .25-.35ϕ. Texas river sediments so far measured range between .40-2.5ϕ, and pipetted flood plain or neritic silts and clays average about 2.0-3.5ϕ. The poorest sorted sediments, such as glacial tills, mudflows, etc., have σ_I values in the neighborhood of 5ϕ to 8 or even 10ϕ.

MEASURES OF SKEWNESS OR ASYMMETRY

Curves may be similar in average size and in sorting but one may be symmetrical, the other asymmetrical. Skewness measures the degree of asymmetry as well as the "sign"--i.e., whether a curve has an asymmetrical tail on the left or right.

Phi Quartile Skewness (Skqϕ). This is found by $(\phi 25 + \phi 75 - 2(Md\phi))/2$ A (+) value indicates that the sediment has an excess amount of fines (the frequency curve shows a tail on the right) and a (-) value indicates a tail in the coarse (left). The disadvantage of this measure is that it measures only the skewness in the central part of the curve, thus is very insensitive; also, it is greatly affected by sorting so is not a "pure" measure of skewness. In two curves with the same amount of asymmetry, one with poor sorting will have a much higher quartile skewness than a well-sorted sample.

Graphic Skewness. As a measure of skewness, the Graphic Skewness (Sk_G) given by the formula $\frac{\phi 16 + \phi 84 - 2\phi 50}{(\phi 84 - \phi 16)}$ may be used (Inman). This measures the displacement of the median from the average of the $\phi 16$ and $\phi 84$ points (see figure below), expressed as a fraction of the standard deviation, thus the measure is geometrically independent of sorting. The derivation follows:



Let "x" be the midpoint of the ϕ16 and ϕ84 values, found by (ϕ16+ϕ84) /2--in this case (1+3) /2 or 2.0ϕ. Then the distance "A" is the displacement of the Median (ϕ50) from the x midpoint. The skewness measure is then $\frac{A}{\sigma}$

But $A = \frac{\phi16+\phi84}{2} - \phi50$, and

$\sigma = \frac{\phi84-\phi16}{2}$, so clearing fractions gives $\frac{\phi16+\phi84-2\phi50}{(\phi84-\phi16)}$

In this case, $\frac{1 + 3 - 2(1.5)}{(3-1)}$

or $Sk_G = +0.50$. Note that the median is displaced 0.50 of the way from the "x" midpoint to the ϕ16 or standard deviation mark.

Inclusive Graphic Skewness (Sk_I) (Folk). The skewness measure discussed above covers only the central 68% of the curve. Inasmuch as most skewness occurs in the "tails" of the curve, this is not a sensitive enough measure. A much better statistic, one that includes 90% of the curve, is the Inclusive Graphic Skewness given by the formula $\frac{\phi16 + \phi84 - 2\phi50}{2(\phi84 - \phi16)} + \frac{\phi5 + \phi95 - 2\phi50}{2(\phi95 - \phi5)}$. This formula simply averages the skewness obtained using the ϕ16 and ϕ84 points with the skewness obtained by using the ϕ5 and ϕ95 points, both determined by exactly the same principle. This is the best skewness measure to use because it determines the skewness of the "tails" of the curve, not just the central portion, and the "tails" are just where the most critical differences between samples lie. Furthermore it is geometrically independent of the sorting of the sample. Because in the skewness formula a measure of phi spread occurs both in numerator and denominator, the Sk_I value is a pure number and should not be written with ϕ attached. Skewness values should always be recorded with a + or - sign to avoid possible confusion.

Symmetrical curves have $Sk_I = .00$; those with excess fine material (a tail to the right) have positive skewness and those with excess coarse material (a tail to the left) have negative skewness. The more the skewness value departs from .00, the greater the degree of asymmetry. The following verbal limits on skewness are suggested: Sk_I from +1.00 to +.30, strongly fine-skewed; +.30 to +.10, fine-skewed; +.10 to -.10, near-symmetrical; -.10 to -.30, coarse-skewed; and -.30 to -1.00, strongly coarse-skewed. The absolute mathematical limits of

the measure are +1.00 to -1.00, and few curves have Sk_I values beyond +.80 and -.80.

MEASURES OF KURTOSIS OR PEAKEDNESS

In the normal probability curve, defined by the Gaussian formula, the phi diameter interval between the ϕ5 and ϕ95 points should be exactly 2.44 times the phi diameter interval between the ϕ25 and ϕ75 points. If the sample curve plots as a stright line on probability paper (i.e., if it follows the normal curve), this ratio will be obeyed and we say it has normal kurtosis (1.00). Departure from a straight line will alter this ratio, and kurtosis is the quantitative measure used to describe this departure from normality. It measures the ratio between the sorting in the "tails" of the curve and the sorting in the central portion. If the central portion is better sorted than the tails, the curve is said to be excessively peaked or leptokurtic; if the tails are better sorted than the central portion, the curve is deficiently or flat-peaked and platykurtic. Strongly platykurtic curves are often bimodal with subequal amounts of the two modes; these plot out as a two-peaked frequency curve, with the sag in the middle of the two peaks accounting for its platykurtic character. The kurtosis measure used here is the Graphic Kurtosis, K_G, (Folk) given by the formula $K_G = \frac{\phi 95 - \phi 5}{2.44\ (\phi 75 - \phi 25)}$. This value answers the question, "for a given spread between the ϕ25 and ϕ75 points, how much is the ϕ5 to ϕ95 spread deficient (or in excess)?" For normal curves, $K_G = 1.00$; leptokurtic curves have K_G over 1.00 (for example a curve with $K_G = 2.00$ has exactly twice as large a spread in the tails as it should have for its ϕ25-ϕ75 spread, hence is much poorer sorted in the tails than in the central portion); and platykurtic curves have K_G under 1.00 (in a curve with $K_G = 0.70$, the tails have only 0.7 the spread they should have with a given ϕ25-ϕ75 spread). Kurtosis, like skewness, involves a ratio of spreads hence is a pure number and should not be written with ϕ attached.

The following verbal limits are suggested: K_G under 0.67, very platykurtic; 0.67-0.90, platykurtic; 0.90-1.11, mesokurtic; 1.11-1.50, leptokurtic; K_G over 1.50-3.00, very leptokurtic; K_G over 3.00, extremely leptokurtic. The absolute mathematical limits of the measure are from 0.41 to virtually infinity; few analyzed samples fall beyond the range from 0.60 to 5.0, however.

The distribution of K_G values in natural sediments is itself strongly skewed, since most sediments are around .85 to 1.4, yet some values as high as 3 or 4 are not uncommon. Thus for all graphic and statistical analysis (computation of mean or standard deviation of kurtosis, running of t tests, etc.) the kurtosis distribution must be normalized by using the transformation $K_G/(1 + K_G)$. Using transformed kurtosis (written K_G') a normal curve has a value of .50, and most sediments fall between .40-.65.

CHARACTERIZATION OF FREQUENCY DISTRIBUTIONS

For characterization of the size frequency distributions of sediments, the limits suggested above should be followed. For size terms use the grain-size triangle (page 27). Here are some examples:

Fine sand, well-sorted, fine-skewed mesokurtic.

Sandy pebble gravel, moderately sorted, strongly fine skewed leptokurtic.

Granular medium sand, very poorly sorted, coarse-skewed very platykurtic.

Very fine sandy mud, very poorly sorted, near-symmetrical platykurtic.

These various statistical measures may be plotted against each other to see how, for example, skewness values may be related to mean size (although geometrically independent, in any given set of samples the two values may show some correlation). They may be plotted on recent sediment maps and contoured to show the regional variation of the measures, and provide a clue to identification of ancient environments.

Frequency distributions of other sets of data may be statistically analyzed in exactly the same way as grain size. The measures of mean size, standard deviation, skewness and kurtosis used here may be used for any type of data at all, in any field of science. The verbal limits for skewness and kurtosis suggested here may also be used for data in any other field, but the verbal limits on size and standard deviation are of course inapplicable.

THE METHOD OF MOMENTS

The second method of obtaining statistical parameters is called the method of moments. It is a computational (not graphical) method of obtaining values, in which every grain in the sediment affects the measure. Thus it probably gives a truer picture than the graphic methods, which rely on only a few selected percentage lines. For example, the median is obtained graphically by merely reading the diameter at the 50% mark of the cumulative curve, and is not at all affected by the character of the rest of the curve; but the mean, computed by the method of moments, is affected by the distribution over every part of the curve (see sheet at the end giving graphic significance of measures). Details on the computations involved are given in Krumbein and Pettijohn. It is possible to obtain skewness and kurtosis also by the method of moments, but we will confine ourselves to determination of the mean and standard deviation. To begin with, one sets up the following form, using the ϕ scale. A midpoint of each ϕ class is selected (usually 2.5, 3.5ϕ) but in some cases (as in the Pan fraction) a different midpoint must be selected. In these "open-ended" distribution, where there is a lot of material in the pan fraction of unknown grain size, the method of moments is severely handicapped and probably its use is not justified.

ϕ class interval	ϕ mid-point	Weight, grams	Product	Midpoint Deviation	Mid. Dev. squared	Product
	(D)	(W)	(D. W)	(Mϕ-D)	$(M\phi-D)^2$	$W(M\phi-D)^2$
0.0-1.0	0.5	5.0	2.5	2.18	4.74	23.7
1.0-2.0	1.5	10.0	15.0	1.18	1.39	13.9
2.0-3.0	2.5	30.0	75.0	0.18	0.03	0.9
3.0-4.0	3.5	20.0	70.0	0.82	0.67	13.4
4.0-pan	5*	5.0	25.0	2.32	5.39	27.0
sum(Σ)	--	70.0	187.5			78.9

* arbitrary assumption.

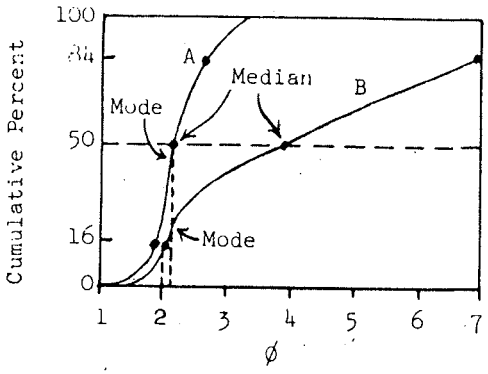
The phi arithmetic mean of the sample (Mϕ) is then $\frac{\Sigma DW}{\Sigma W} = \frac{187.5}{70.0}$ or 2.68ϕ.

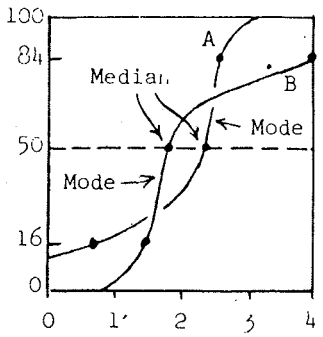
The "midpoint deviation" column is then obtained by subtracting this mean (2.68) from each of the phi midpoints of each of the classes. These deviations are then squared, multiplied by the weights, and the grand total obtained. Then the standard deviation (σ_ϕ) is obtained by the following formula:

$$\sigma_\phi = \sqrt{\frac{\Sigma [W(M\phi-D)^2]}{\Sigma W}} = \sqrt{\frac{78.9}{70}} = \sqrt{1.127} = 1.06\ \phi$$

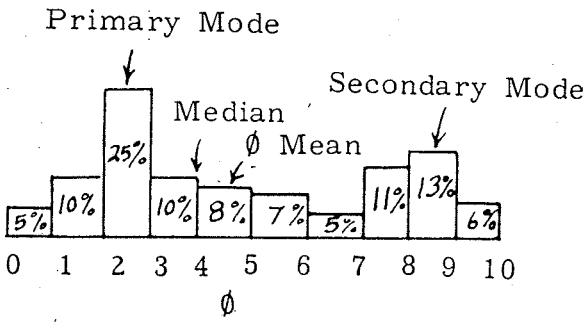
SPECIAL MEASURES

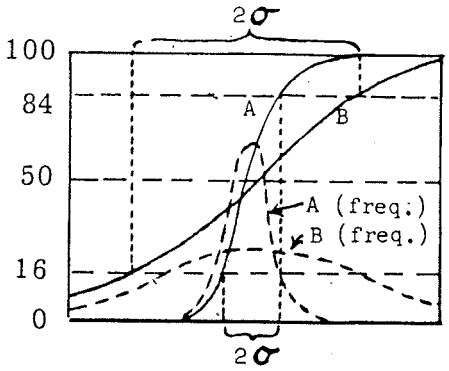
This whole problem may be approached fruitfully in another way. Instead of asking, "what diameter corresponds to the 50% mark of a sample" we can ask "what percent of the sample is coarser than a given diameter." This is an especially valuable method of analysis if one is plotting contour maps of sediment distribution in terms of percent mud, percent gravel, percent of material between 3ϕ and 4ϕ, etc. It works better than any other method if the sediment is bimodal. Or, one may plot as a contour map the diameter of the largest particle in the sample. Try new methods all the time.





Mode, Median and Mean. In the left-hand figure, both curves have a mode at about 2ø, but B has a much finer median. The Graphic Mean (which is the average of the three heavy dots at the 16, 50, and 84 percentiles) is about 2.2ø for A and about 4.2ø for B. In the right-hand figure, B. has a much coarser median and mode, but actually is finer if you consider the entire sediment; the Graphic Mean for A would be about (0.7 + 2.4 + 2.6)/3, or 1.9ø, and for curve B would be about (1.4 + 1.7 + 4.1)/3, or 2.4ø. The Mean gives the best overall indication of average size; the mode is of real value, but the median is not a good measure.

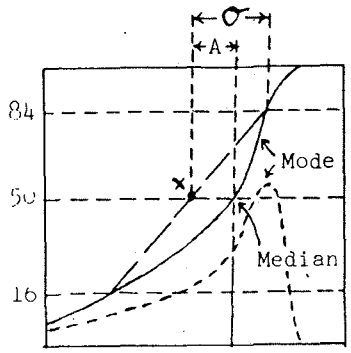




This histogram illustrates the difference between the several measures of average size. Note there are two modes, at about 2.5ø and 8.3ø. The median has half the sample finer and half the sample coarser than itself; the left four columns total 50% of the sample so the median is 4.0ø. The ø Mean, computed by the method of moments, is based on the entire curve (5 × 0.5ø + 10 × 1.5ø + 25 × 2.5ø..... 13 × 8.5ø + 6 × 9.5ø)/100; the ø Mean is thus 4.7ø. The Graphic Mean, which is a rapid and close approximation to the ø Mean, is also 4.7ø but this must be found on the cumulative curve Changing the 6-7ø grade to 16% and the 8-9ø grade to 2%, will affect the means but will not at all affect the primary mode or median (prove this for yourself).

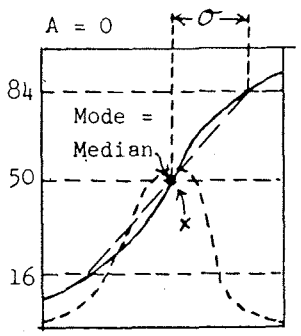
Cumulative and Frequency (dashed) curves for well-sorted sediment (A) and poorly sorted sediment (B). In the well-sorted sediment, the cumulative curve is much steeper, the frequency curve is much higher and much less spread out, and the range between 16 and 84 percentiles (which equals 2σ) is much smaller, hence the standard deviation is much smaller.

Measures of Skewness and Kurtosis



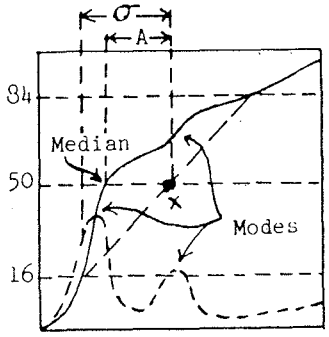
Skewed with excess coarse material (tail in the coarse, or left side). X is the midpoint of the 16 and 84 marks;

$Sk_G = \frac{A}{\sigma} = -.60$



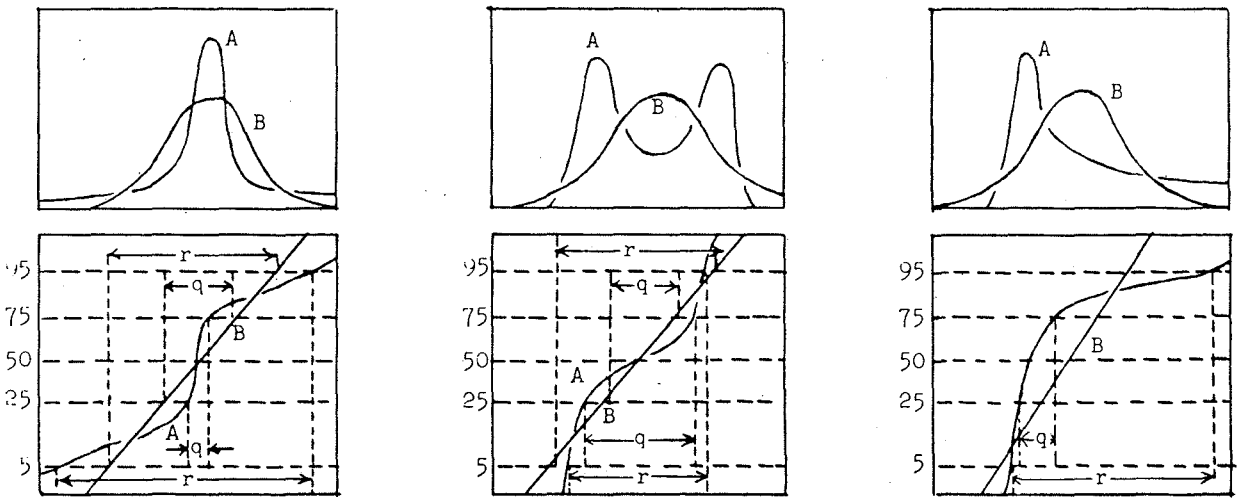
Symmetrical. Mode equals median.

Sk = .00



Skewed with excess fine material (tail in the fine, or right side); also bimodal.

$Sk_G = \frac{A}{\sigma} = +.75$



Kurtosis. All the cumulative curves are drawn on probability paper. The corresponding frequency curves are shown above. Each graph has a straight line (B) which represents a normal probability curve with the same mean and standard deviation as the sample curve (A). In a normal curve, the ratio r/q (r is the 5-95 spread, q the 25-75 spread) must be exactly 2.44. The left-hand figure is extremely leptokurtic, because r/q is much greater than 2.44. Note the frequency curve which is excessively peaked at the middle, is deficient in the "shoulders" of the curve, and the tails are extremely long (to get the same overall spread as the corresponding normal curve, the tails must be spread wider to compensate for the condensation of the center). The middle figure is extremely platykurtic, with r/q much less than 2.44; it is also bimodal, although not all platykurtic curves are bimodal. The right-hand figure is a curve that is highly leptokurtic and also has strong positive skewness; r/q is much greater than 2.44.

A FEW STATISTICAL MEASURES FOR USE IN SEDIMENTARY PETROLOGY

Introduction. A good knowledge of statistics is becoming essential for anyone who wishes to work in any of the sciences, because the whole of scientific work from laying out the experiment to interpretation of data is based on statistics. Trying to use numerical data without a knowledge of statistics is like trying to drive without a brake. You never know where you will end up and the odds are you will end up in the wrong place and get the wrong conclusion.

In sedimentary petrography, statistics is used in laying out the sampling program; in determining the best experimental technique for analysis; in collecting the analytical data; and in drawing correct geological conclusions, such as: what is the content of feldspar in X formation? Within what limits am I certain this value is correct? What is the spread of values to be expected? Does X formation have more feldspar than Y formation and how confident am I of this? Does its heavy mineral content differ significantly from that of formation Y? What is the relation between grain size and zircon content, expressed mathematically?

This outline is not intended to make you an expert in statistics. It merely shows examples of the use of statistics in petrography, with the hope that it will stimulate you to take several courses or read up on your own. It is super-simplified and condensed, therefore omits a lot of material that should really be covered. For further information refer to any standard textbook.

The Normal Probability Curve. In order to understand some of the assumption and underlying principles, it is essential to study the statisticians' most fundamental concept, that of the normal probability curve. This is the basis for study of experimental data of all kinds.

As a first step in the analysis of data from any field of science, one usually constructs a frequency distribution. For example, if one is studying the batting averages of baseball players, he would select convenient class intervals to divide the entire range of data into about 10 to 20 classes and proceed to find how many batters had averages between .200 and .210, how many between .210 and .220, and so on; here the class interval would be .010. Or if an anthropologist were studying the lengths of human thigh bones, he would first ascertain the spread between the largest and smallest bone (say for example 12" to 31"), and divide this into a convenient number of classes. Here a convenient class interval would be 1", and he would proceed to find how many thighbones were between 12" and 13", how many between 13" and 14", and so on. When data of this type is plotted up in the form of a histogram of frequency curve, it is usually found that most of the items are clustered around the central part of the distribution with a rapid "tailing off" in the extremes. For example, far more baseball players hit between .260 and .280 than hit between .320 and .340, or between .180 and .200. Even less hit between .100 and .120, or between .380 and .400. A great many types of data follow this distribution, and the type of frequency curve resulting is called the "probability curve," or the "normal curve," or often a "Gaussian curve" after Gauss who was a pioneer in the field. The curve is defined as the kind of distribution resulting if one had 100 well-balanced coins and tossed them all repeatedly to count the number of heads appearing. Naturally, the most frequent occurrence would be 50 heads and 50 tails; 40 heads or 60 heads would be less likely, and so on down to occurrences of 10 heads and 90 tails, which would be very few; a throw of 1 head and 99 tails would be exceedingly rare. By tedious computations, one could figure

the chances of throwing any combination of heads and tails and this is the useful feature of the normal probability curve: The probabilities fall off at a definite, predictable rate which is fixed by a mathematical equation. Furthermore, the curve is symmetrical about the mean--a throw of 34 heads (50-16) is exactly as likely as a throw of 66 heads (50 + 16); and a throw of 18 heads is just as likely as 82 heads.

Many types of data follow closely this curve which is defined by coin-tossing experiments and rigidly fixed by an equation. Baseball batting averages; weights of bolts turned out by a factory; life spans of electric light bulbs; mean daily temperatures for any month if records are kept over some years; heights of people; densities of granite samples; widths of brachiopod valves; slope angles of geomorphic features and many others often follow this normal probability curve, providing enough data is collected. For example if one chose ten people and weighed them his curve would be rather irregular; by the time he weighed 1000 people it would be much smoother and if he weighed 10,000 the distribution would hew very closely to the normal curve.

Many distributions do not follow the theoretical normal distribution, however. One of the most common ways that a distribution departs from normality is in its lack of symmetry. The graph of a normal distribution is a perfectly symmetrical bell-shaped curve, with equal frequencies on both sides of the most common value (i.e. in the 100-coin toss, 25 heads are just as common as 75 heads). Many kinds of data are asymmetrical, though. Consider the prices of houses in an average American city. The most common price might be somewhere around $10,000. In this city there might be many $25,000 homes; in order to have a symmetrical frequency distribution, this would demand that there be many homes that cost minus $5000! This curve, then would be highly asymmetrical with the lowest value being perhaps $3000, the peak frequency at about $10,000, and a long "tail" in the high values going out to perhaps $100,000 or even more. The distribution of the length of time of long-distance telephone calls is also a distribution of this type, since most calls last between two and three minutes, very few are less than one minute but there are some long distance calls lasting as long as 15 minutes or even an hour. The frequency distribution of percentage of insoluble materials in limestone samples is also a highly asymmetrical or skewed distribution, with most limestones in one formation having for example between 5 and 10 percent insoluble, but some samples having as much as 50 or 75% insoluble, with 0% as the obvious minimum percentage.

A further way in which distributions depart from normality is that they may have two or more peak frequencies (termed modes). If one took a large college building and obtained the frequency distribution of ages of all people in the building at a given time, he would find a curve that had two peaks (bimodal) instead of one. The highest peak would be between 19 and 20 (the average age of the students who would make up most of the population), and another peak might occur at around 40 (the average age of the professors), with a minimum at perhaps 30 (too young for most professors, and too old for most students). This distribution would be distinctly non-normal; technically, it would be said to have deficient kurtosis. In geology we would obtain a similarly non-normal, bimodal distribution if we measured the sizes of crystals in a porphyritic granite, or if we measured the percentage of quartz in a sand-shale sequence (sand beds might be almost pure quartz, while the shale beds might have 20% or less).

In analyzing frequency distributions, the most common measures used are the arithmetic mean (or average) and the standard deviation (or degree of scatter about the mean). Skewness and kurtosis are also used for special purposes (see pages 48, 49).

Arithmetic Mean: The most common and useful type of average for ordinary purposes is the arithmetic mean, symbolized $\bar{x}$. To obtain this value, simply add up the values and divide the result by the number (n) of values. This is symbolized as $\bar{x} = \frac{\sum X}{n}$. There are a number of other averages (median, mode, geometric mean, etc.) which are useful in special cases--especially when the data is skewed, as when you are taking an average consisting of a lot of small values and a few very large values, which would tend to distort the simple arithmetic mean.

Standard Deviation: For most data we are interested in determining the spread or scatter of the values. For example, a set of readings 4, 6, 7, 4, 5, 6, obviously is less scattered than a set such as 2, 5, 14, 1, 9, 3, although both have the same mean. The standard deviation (s or σ) is a precise measure of this scatter and, next to the mean, is the most useful value in statistics. It is the only measure of spread that has wide use. The standard deviation is computed such that, on an average, 68% of the samples will fall within plus or minus one standard deviation ($\pm \sigma$) from the mean; 95% will fall within plus or minus two standard deviations ($\pm 2\sigma$); and 99% of the samples will fall within $\pm 3\sigma$. For example, if we have 100 porosity values on a sandstone formation and the mean porosity is 15.0% with a standard deviation of 3.0%, we know that approximately 68 of these 100 samples will nave porosities ranging between 12 and 18% ($\bar{x} \pm \sigma$), that 95 samples will range between 9 and 21% ($\bar{x} \pm 2\sigma$) and that 99 samples will range between 6 and 24% ($\bar{x} \pm 3\sigma$). Of course this assumes that we have a sufficient number of values, and that the values follow the normal bell-shaped probability curve (in which most of the values cluster about the mean, and "tail off" symmetrically to each side). This is tacitly assumed for most statistical work. A complete table of the percent of values included in a number of standard deviations is given on page 64.

To obtain the standard deviation there are two different methods which give identical answers. In the long method, one finds the arithmetic mean of his values; then finds the deviation of each value from this mean, and squares these deviations; then adds up these squared values, divides by the number of values and takes the square root of this quotient. In the short method, which we will use, one substitutes values in the following formula:

$$s = \sqrt{\frac{\sum (x^2) - \frac{(\sum x)^2}{n}}{n-1}}$$

For example, let us say that we have five thin sections of a formation and wanted to know the feldspar content. By point counter we obtain the following percentages on each slide: 8, 11, 6, 15, 10. What is the mean and standard deviation? We set up the following table:

	x	x^2
	8	64
	11	121
	6	36
	15	225
	10	100
Σ =	50	546

n, the number of samples, is 5; Σx (the sum of the x values--in this case the percentages of feldspar) equals 50, therefore the arithmetic mean, $\bar{x}$, is:

$$\frac{\Sigma x}{n} = \frac{50}{5} \quad \text{or } 10.0\%$$

To find the standard deviation, we substitute values in the equation above.

(x^2), the sum of the squared values of x, equals 546; (Σ x)^2/n is $50^2/5$ or 500; and n-1 is 5-1 or 4. Substituting,

$$s = \sqrt{\frac{546-500}{4}} = \sqrt{\frac{46}{4}} = \sqrt{11.5} = 3.4\%$$

If our samples are representative and the feldspar content follows the normal bell-shaped probability curve, then if we collected 100 samples of the same formation we would expect 68 of them to have feldspar contents between 6.6 and 13.4%. If someone asks you, "what are the odds of finding a sample with over 17% feldspar, you can immediately answer "only 2 or 3 out of 100." Why? Well, 17% is about equal to $\bar{x}$ + 2s, or 10.0 + 2(3.4); we know that 95% of the samples will fall in the range of $\bar{x} \pm 2s$, therefore only 5% of the samples will fall outside this range; of these half will be higher and half lower, so 2.5% of the samples will have less than 3% feldspar, and 2.5% will have over 17%. Again, this assumes a normal distribution symmetrical about the mean.

Confidence Limits on a single mean. In the example above, we got a mean of 10.00 percent feldspar based on five samples of the formation. Thisis our estimate, based on five samples, of the feldspar content of the entire formation. Now the true feldspar content could not be determined unless we analyzed every one of the millions of sand grains in the entire formation. This we cannot do, so we have to estimate the mean by taking a small number of representative samples. But by the test to be described, we can tell how close our estimated mean is likely to be to the true mean of the formation--in otherwords, we can assign confidence limits to the mean. We can say, "I am 95% sure that the true formation mean lies between 9.0 and 11.0%," which means we will be wrong only 5% of the time--or one time in twenty.

We realize immediately that if we take 100 samples we will be more confident of our mean, then if we took only five samples. Also, we will be more confident of the mean if the values show a small spread (standard deviation) than if they show a large spread (Example: if we have values of 4.2, 3.6, 4.0, 4.8, 3.4 we are more confident that our mean is nearly 4.0 than if we have values 1.2, 4.8, 7.6, 3.1, 3.3). The formula for computing the confidence limits appropriately then takes these two factors into account.

$$\pm L = \frac{ts}{\sqrt{n}}$$

where n is the number of observations or values, s is the standard deviation, and t is explained below.

The factor t is put into the equation so that we can choose our "confidence level." Assume a formation with a mean porosity of 12.0%, standard deviation 3.0%. If we want to be only 50% sure of our mean, we look up in a table the value of t at the 50% level, and insert this in the formula. Let us say this makes "L" come out to $\pm$ 1.5. This means that we would be only 50% sure that our true formation mean lay between 10.5 and 13.5%, and there is one chance in two that the true mean lies either above or below these limits. However, if we chose the 95% level, we insert this value of t in the formula. Say that L then comes out to 3.0%; then we would be 95% sure that the true formation mean lay between 9.0 and 15.0%, and there would be only one chance in twenty of our being wrong. Most statisticians use the 95% level all the time. (P = 0.05 column, page 62).

The value t is found in a table with the levels to be selected (5%, 10%, 20%, etc.) plotted against the "degrees of freedom" (which in this case simply means

the number of values minus 1); the value of t to be used lies at the intersection of our chosen confidence level and the degrees of freedom. For example, with 21 values, at the 5% (.05) level, t = 2.09, as shown in t table on page 62

POPULATIONS AND PROBABILITY

In the tests that follow it is essential that we understand the two concepts of "populations" and "probability."

Population is the term used for the data we are sampling; if we are measuring porosities, the "population" that we are sampling is the vast number of porosities present in every cubic inch of that formation, of which we test what we assume are a representative few; if we are measuring the mica content of a bed, the "population" is the percent of mica in each minute part of that bed; if we fill a jar with 500 black and 1000 red beans and proceed blindfolded to pull out 50 beans, then we hope that our sample of 50 is representative of the true "population"--i. e., the 1500 total beans. The population is the vast amount of numerical data available, of which we take only a small sample.

In many statistical tests, we try to answer the question, "what are the odds that we could have obtained as great a difference or greater by chance sampling of the same population?" This gives us an insight whether or not there is a real difference in the properties of the two formations we have sampled, or whether we could have gotten just as large a difference in, say chert content, by chance sampling of a single formation (a "homogenous" population). Let's think about it this way. We have a square jar and a round jar, each filled with a certain mixture of 1000 black and white beans. The round jar has a population with a certain number of blacks, the square jar has a different proportion of blacks. You are now blindfolded and asked to pull 50 beans from one of the jars; let's say you got 30 blacks and 20 whites. Now, still blindfolded, you are asked to reach again into a jar (you still don't know which one you are reaching into) and pick 50 more beans. This time you get 22 blacks and 28 whites. Statistics enables us to answer the question, "on the second drawing, did I take from the same jar as I did the first time; what is the probability that I drew both sets of beans from the same jar?" Or phrased differently, "what are the odds that I drew both samples of 50 from the same population (i.e., the same jar?)" This is essentially what we do when we obtain numerical data from rocks. We take only a small sample of the numerical data available in the formation, ' en sample another formation and see if there is any difference between the two fo ations;--any difference in the population of beans in the jar, so to speak.

Probability. Many statistical tables have a critical value called "P" as an important part. P stands for probability--in most cases P, represented as a percent, stands for the probability of a certain event happening. In the jar experiment above, after running through the computations of the statistical test, we enter a table and come out with a certain value for P. In this case let us say P came out to be .10 (10%). This means that if we had repeatedly sampled the same jar, only once in every ten times would we have gotten as large, or larger, a difference as we did on the two draws. Or, there is only one chance in ten that we have sampled the same population. Although it is not technically correct to say it this way, we may state the corollary that there is a 90% chance that we have sampled two different jars. Mose statisticians do not accept an experiment as significant unless P reaches the 5% level (.05 on the table, pages 62 and 63).

Comparison of two means. We have studied the heavy mineral suites of two formations. We examine 50 samples of the Abner formation and find that within the heavy minerals,

the garnet content averages 15.0% with a standard deviation of 5%; and 30 samples of the Benjamin formation show a garnet content of 10.0 with a standard deviation of 4.0%. The question arises, does the Abner really have more garnet than the Benjamin, or could such differences have arisen by chance sampling of a homogeneous population? The so-called "t test" has been devised to answer this question and is one of the most useful statistical devices. We substitute our values in the following equation, where $\bar{x}$ is our mean garnet content, and n equals the number of samples (subscript $_a$ refers to values from the Abner formation, $_b$ to the Benjamin formation).

$$t = \frac{\bar{x}_a - \bar{x}_b}{s} \sqrt{\frac{n_a \cdot n_b}{n_a + n_b}}$$

where s approximately equals the average standard deviation of the two sets of values, and is found by the following equation (which is merely an expansion of our regular formula for the standard deviation):

$$s = \sqrt{\frac{\Sigma (x_a{}^2) - \frac{(\Sigma x_a)^2}{n_a} + \Sigma (x_b)^2 - \frac{(\Sigma x_b)^2}{n_b}}{n_a + n_b - 2}}$$

We come out with a certain value for t. From here we enter a table (p. 62) showing t values as a function of the number of "degrees of freedom" and of P (probability). To find the degrees of freedom in this case, we add up the number of values in both sets of data and subtract 2 from the result (in the example, 50+30-2 = 78 d.f.) Entering the table with 78 degrees of freedom, we read across the horizontal row at this value until we encounter the correct t value. From this t value we read straight up to see what P corresponds to our t. If P lies between .05 and .10, then we know that there is something between a 5% and 10% chance of our obtaining such differences (or larger) by chance sampling of a homogeneous population. This may be colloquially stated as saying "our experiment has shown that there are only 5 to 10 chances in 100 that the garnet content of the two formations is the same," or the reverse, "there is a 90 to 95% chance that the Abner really does contain more garnet than the Benjamin." Both these statements are technically not exactly precise, but may be considered as pretty close to the truth.

As stated before, statisticians ordinarily consider that in any experiment that fails to reach the 5% level, the data do not warrant making a conclusion. In other words if your P comes out to the 10% level, it means that you have failed to find a really significant difference between the two formations, either because the difference in means is too small, or the standard deviation is too large, or you took too few samples. The only way to remedy this situation is to take enough samples to push the results beyond the 5% level.

The X^2 (Chi square) test. The t test is used when you are comparing means of measurements (like grain size, percentages, porosities, densities, etc.) between, two formations and can generally be done on only one property at a time.

The X^2 test is used when you are comparing counts of discrete objects between formations (like the number of grains larger than a given size in a formation; or the number of grains of a certain heavy mineral; or in ordinary life, the number of accidents in a

given city each month; or the number of home runs hit by a team each week.) In some cases it is possible to use both the X^2 test or the t test, as in heavy mineral counts (which may be converted to percentages); but the X^2 test has this advantage, that more than one item or property can be compared at a time. For instance, if we count the number of tourmaline, rutile and zircon grains in two formations, we can use this whole stack of data at once with the X^2 test and see if the formations are significantly different. If we used the t test we'd have to make a separate test for the tourmaline, another one for the zircon, another one for the rutile and there would be no simple way to combine them.

To use the X^2 test we must in some way arrive at a predicted or expected frequency or occurrence, based on a prior reasoning or else on long experience. We then compare our observed frequency with the given frequency. For example, if we roll a dice 600 times, we would expect that if the dice is well balanced each number (1 through 6) would appear equally and our expected frequency would therefore be 100. When we actually do roll the dice, we may find that we got only 95 1's, 123 2's, 82 3's, etc. The purpose of the X^2 test is to compare the observed with the expected frequency and see how likely it is that the deviations are due to chance, or whether the dice really is loaded. As another example, say that data collected over a period of 20 years shows that American League teams hit a grand total of 1035 homers per season, on the average; thus we would expect them to hit about 1035 homers this season. Well, this season they actually slammed 1230 round trippers; is this "deviation" just due to chance, or to a new "lively" ball? The X^2 test helps us to answer this question.

There are a number of different ways to set up the X^2 test. The one we'll use most is called a "2 by n" table--usually we are comparing <u>two</u> formations based on counts of "n" types of heavy minerals, where "n" may be anything from 1 to 20 or more. For example, let us say that we count varieties of tourmaline in the Eli and Nineteenten formations and want to know if a significant difference exists between them in this respect. We tabulate the <u>number</u> of grains thus:

	Eli fm.	Nineteenten fm.	Total	(This is a 2 by n table where n, the number of different mineral types equals 3)
Green tourmalines	15	25	40	
Brown tourmalines	10	12	22	
Other colors	20	14	34	
Total	45	51	96	

Does this data indicate that there is a real difference in the types of tourmaline between the two formations? Or could we have gotten counts as different as this by chance sampling of formations that were identical in tourmaline varieties?

Please remember, to use the X^2 test, data must be stated in terms of actual numbers of grains counted--you cannot convert to percentages and then use the X^2 test.

We start out with the assumption that both have the same heavy mineral content and proceed to compute our expected frequencies on this basis. To figure the number of green tourmalines expected in a count of 45 grains of the Eli formation, we decide (from looking at the totals) that 40/96 of the grains should be green; thus we expect 45 timer 40/96 green grains in the Eli, or 18.7. Since there were a total of 40 green

grains counted, then we'd expect 40-18.7 or 21.3 green grains to occur in the Nineteenten.

The expected number of browns is 45 times 22/96 or 10.3 for Eli, and 22-10.3 or 11.7 for Nineteenten. For "others" we expect 45 times 34/96 or 15.9 for Eli, and 34-15.9 or 18.1 for Nineteenten. Now we set up the following table, subtracting 0.5 from each of the differences, giving our corrected differences (the reason we subtract 0.5 from each difference is a "correction for continuity" because obviously you cannot count fractional grains, and 10 grains is as close as you can possibly get to an expected frequency of 10.3 grains).

		Eli	Nineteenten
	Observed	15	25
Green	Expected	18.7	21.3
	Difference	3.7	3.7
	Corr. Diff.	3.2	3.2
	Observed	10	12
Brown	Expected	10.3	11.7
	Difference	0.3	0.3
	Corr. Diff.	0.3	0
	Observed	20	14
Other	Expected	15.9	18.1
	Difference	4.1	4.1
	Corr. Diff.	3.6	3.6

For each of the six "cells" (three tourmaline varieties in each of two formations) we square the corrected difference and divide it by the expected frequency for that one cell. Then we add up these values for the entire six cells and the total gives us the value of X^2, in this example 2.56. Symbolically, this operation is

$$X^2 = \sum \left(\frac{D^2}{E} \right)$$

$$X^2 = \frac{10.24}{18.7} + \frac{10.24}{21.3} + \frac{0}{10.3} + \frac{0}{11.7} + \frac{12.95}{5.9} + \frac{12.95}{18.1} = 2.56$$

We enter a X^2 table (p. 63) and must know two things: our value for X^2 and the degrees of freedom (d.f.). This time the d.f. is obtained by multiplying these two quantities: (number of horizontal data columns minus one) times (number of vertical data columns minus one); in this case (3-1)(2-1), or two degrees of freedom. Again, we read across the horizontal row corresponding with the proper number of degrees of freedom until we find our value of X^2 in the body of the table. Then we read directly up to the top of the table to find the corresponding "P". This will answer the question, what are the chances (out of 100) that such differences--or larger--would be obtained in random sampling of two uniform formations (or in sampling a homogeneous population?) Our result came out P = approximately .20, in other words there is one chance in five that we have sampled a homogeneous population, or conversely, four chances in five that the formations have a differing tourmaline content--though this last statement is not strictly true.

For another example, consider a tire company that for years has averaged 31 blowouts per million tire miles. After switching to a new type rubber, they find 44 blowouts occur in the next million miles. Is the new rubber inferior, or is this merely an expectable chance fluctuation? Using the X^2 test, the expected frequency would have been 31 blowouts for the new tires. (O-E) is 44-31 or 13, and correcting for discontinuity gives 12.5. $X^2 = (12.5)^2/31 = 156/31 = 5.0$. In the table, for $X^2 = 5.0$ and d.f. = 1, P = .03; hence it is 97% certain that the new rubber is inferior.

There is one serious caution about using the X^2 test. In case the expected frequency in any cell is less than 5, this cell must be combined with another to bring the total expected frequency for the combined cells over 5. In the sample above, let's say I also counted 4 orange tourmalines and 3 yellow ones; in order that no expected

cell frequency be under 5, I would have to lump these with other rare types in the cell labeled "others."

Other Techniques. Any statistical text will list many other valuable tests and techniques. Some of these, of more interest to geologists, are simply mentioned here; for details, go to the texts, e.g. Miller and Kahn, Snedecor, etc.

Much geologic data can be presented in the form of scatter plots, wherein we wish to see how one property is related to another. Examples are plots of roundness vs. distance; mean grain size versus sorting; feldspar percentage versus stratigraphic position; for a collection of dinosaur bones, length of thigh bone versus thickness of the bone; zircon/tourmaline ratio versus grain size; percent carbonate mud versus roundness of shell fragments, etc. To analyze such associations, two main procedures are applied: (1) the perfection of the association is tested, and (2) the equation of the relationship is determined.

If the two properties are very closely related, they give a long narrow "train" of points on a scatter diagram. If the two properties are not associated, a random "buckshot" pattern emerges. The correlation coefficient, r, computes the perfection of correlation. For perfect correlation r = 1.00, which means that, knowing one property, we can predict the other property exactly, and that both increase together. An r of -1.00 means perfect negative correlation, a correlation just as exact except that as one property increases the other decreases. If the two properties are not correlated, r may be .00; weak correlation would be +.25 or -.15 etc. Coefficients beyond $\pm$.50 are considered "good" for most geological work. The normal correlation coefficient is valid only for straightline trends. Other methods must be used for hyperbolic, parabolic, sinusoidal etc., trends.

If a small number of data points are available, it is possible for "good-looking" correlations to arise purely by chance. Thus one should always refer to tables which show whether the given value of r shows a significant correlation; this depends of course on the number of samples and the value of r, thus is similar in principle to the t test.

Squaring the correlation coefficient, r, gives the coefficient of causation r^2; this tells one how much of the variation in one property is explained by the variation in the other property. For example, if we find that in a series of pebbles, the roundness shows a correlation coefficient of r = +.60 with grain size, then r^2 = .36, and we can say that 36% of the variation in roundness is caused by changes in grain size (thus 64% of the roundness variation would be due to other causes: differences in lithology, distances of travel, "chance", etc.). Further analyses may be carried out, such as partial or multiple correlations, analysis of variance etc.--see standard texts.

A trend line may be fitted to a scatter diagram, and an equation may be fitted to this line so that, given a value of one property, the other property may be predicted. Trend lines can be drawn in by eye, but this process is usually sneered at; a mathematical way of doing it is the "least squares" method. It is important to realize that, unless the two properties are almost perfectly correlated, two trend lines can be computed; let's say we have a graph of feldspar content versus roundness, with a correlation coefficient of +.45 (good, but far from perfect correlation). We can either (1) compute the line to predict most accurately the roundness, given the feldspar content; or (2) a line to predict most accurately the feldspar content given the roundness. The two lines may easily form a cross with as much as 30° or more difference in angle. Again, these work only for straight-line trends; non-linear trends require more complicated arithmetic.

To the trend line is usually attached a "standard error of estimate" band, essentially equal to the standard deviation. This band runs parallel to the computed trend and includes two-thirds of all the points in the scatter diagram. Its purpose is to show the accuracy of the relationship. For example, in a certain brachiopod the length and width of the shell are related by the equation $L = 2W - 1 \pm 0.5$. The last figure is the standard error of estimate; if a given specimen has a width of 3.5 cm, the length is most likely 8 cm, but we can expect two-thirds of the specimens to range between 7.5 and 8.5 cm. (one-sixth of them will be over 8.5 cm, and one-sixth under 7.5 cm).

Abridged Table of t

This table can be used only when the data is in the form of means of continuous variables

Degrees of Freedom	P						
	0.50	0.20	0.10	0.05	0.02	0.01	0.001
1	1.0	3.0	6.3	12.7	31.8	63.7	636.6
2	0.84	1.89	2.92	4.30	6.97	9.93	31.60
3	0.79	1.62	2.35	3.18	4.54	5.84	12.94
4	0.78	1.52	2.13	2.78	3.75	4.60	8.61
7	0.73	1.42	1.90	2.37	3.00	3.50	5.41
10	0.70	1.36	1.81	2.23	2.76	3.17	4.59
20	0.69	1.30	1.73	2.09	2.53	2.85	3.85
30 - ∞	0.68	1.28	1.65	1.96	2.33	2.58	3.29

Enter the table with the proper degrees of freedom and read right until you reach the (interpolated) value of t you obtained by calculation. Then read up to the top of the table the corresponding P. Example: for 10 d.f., $t = 3.0$; therefore P - about .013, i.e., there is a little more than 1 chance in 100 that the differences are due to chance. As a general rule, if P is .05 or less the differences are considered as real; if P is between .05 and .20, there may be real differences present and further investigations are warranted with the collection of more samples if possible; if P is over .20 differences are insignificant.

$$t = \frac{\bar{x}_a - \bar{x}_b}{s} \sqrt{\frac{n_a \cdot n_b}{n_a + n_b}}$$

Abridged Table of X^2

This table can be used only when the data is in the form of discrete counts of individuals

Degrees of Freedom	P						
	0.50	0.20	0.10	0.05	0.02	0.01	0.001
1	0.45	1.65	2.71	3.84	5.41	6.64	10.83
2	1.39	3.26	4.61	5.99	7.82	9.21	13.82
3	2.37	4.65	6.25	7.82	9.84	11.3	16.2
4	3.36	5.91	7.7	9.4	11.6	13.2	18.4
6	5.35	8.60	10.6	12.5	15.0'	16.8	22.4
8	7.34	11.0	13.3	15.5	18.1	20.0	26.1
10	9.34	13.4	15.9	18.3	21.1	23.2	29.5
15	14.3	19.3	22.3	25.0	28.2	30.5	37.7
20	19.3	25.1	28.4	31.4	35.0	37.5	45.3
25	24.3	30.7	34.3	37.6	41.5	44.3	52.6
30	29.3	36.2	40.2	43.7	47.9	50.8	59.7

Enter the table with the proper degrees of freedom and read right until you reach the (interpolated) value of X^2 you obtained by calculation. Then read up to the top of the table the corresponding P. Example: for 8 d.f., $X^2 = 14.1$; therefore P = about .08, i.e., there are 8 chances in 100 that the differences are to chance. As a general rule, if P is .05 or less the differences are considered as real; if P is between .05 and .20, there may be real differences present and further investigations are warranted with the collection of more samples, if possible; if P is over .20 differences are insignificant.

$$X^2 = \sum \left(\frac{D^2}{E} \right)$$

Table of Areas of the Normal Probability Curve

The following table gives the percentage of values included within a range of the mean plus or minus the number of standard deviations listed in the left-hand column. More complete tables can be found in textbooks.

Number of Standard Deviations	Percentage of Samples Included in Range	Percentages of Samples Outside the Range
d	$\bar{x} \pm d$	$\bar{x} \pm d$
0.1	8%	92%
0.2	16%	84%
0.3	24%	76%
0.4	31%	69%
0.5	38%	62%
0.6	45%	55%
0.68	50%	50%
0.7	52%	48%
0.8	58%	42%
0.9	63%	37%
1.0	68.3%	31.7%
1.2	77%	23%
1.4	84%	16%
1.6	89%	11%
1.8	93%	7%
2.0	95.5%	4.5%
2.2	97.22%	2.78%
2.4	98.36%	1.64%
2.6	99.06%	0.94%
2.8	99.48%	0.52%
3.0	99.73%	0.27%
3.5	99.954%	0.046%
4.0	99.9947%	0.0063%
5.0	99.999943%	0.000057%
6.0	99.9999998%	0.00000020%

The table is used in solving the following types of problems:
(1) Chert pebbles on a beach have a mean sphericity of 0.71, standard deviation of 0.08. What percentage of the pebbles will have sphericities between 0.68 and 0.74, assuming the distribution is normal? The stated limits are 0.03 on either side of the mean; .03 is in this example 0.375 standard deviations, and approximately 29% of the pebbles will fall within this range. (b) Out of 500 pebbles, how many will be expected to have sphericities higher than .90? This value, .90, is .19 higher than the mean, i.e., 2.375 standard deviations; 2% of samples fall outside the range of $\bar{x} \pm 2.375\ \sigma$, and of these half will be above, half below; therefore the answer is 1%, or 5 pebbles.

(2) A coin weighs an average of 15.0 gm as minted. Due to irregularities in the minting process, the standard deviation is 0.6 gm. Ten percent of the coins will be heavier than what weight? The table is constructed to give us the percentage of samples falling outside the given range, both above and below; we need only the ones higher, therefore we look in column 2 for twice 10%, or 20%; the corresponding value is 1.3 standard deviations. In other words, 20% of the samples will fall outside the range of $\bar{x} - 1.3, \sigma$, 10% being below and 10% above. 1.3 standard deviations in this case is about 0.8 gm; therefore the answer is 15.8 gm.