Chapter 9 – Spatial Variability of Soil Physical Properties

- Expressing Variability
- Probability Distributions
- Measured Variabilities

Soil varies naturally.
They are a product of the factors of formation and continuously change over the earth's surface.

Expressing Variability:

Usually one can determine variability using only a sample of the total population.

One can say something about the mean and variation (standard deviation) of a specific soil physical property of a population, by measuring that physical property for a limited number of soil samples, which together make up the **sample of the population**.

Collect N soil samples with the physical property z_1, z_2, z_3, z_4, ………. z_N

(e.g. let z_1 be the bulk density of soil sample number 1)

Then, the mean (m) is defined as

$$m = \frac{\sum_{j=1}^{N} z_j}{N}$$

and the standard deviation is $$s = \sqrt{\frac{\sum_{j=1}^{N} (z_j - m)^2}{N-1}}$$

and the variance is defined as s^2.

The coefficient of variation (CV) is defined as the ratio of the standard deviation over the mean, or

$$CV = \left(\frac{s}{m}\right)100(\%)$$

What do these statistical properties tell us:

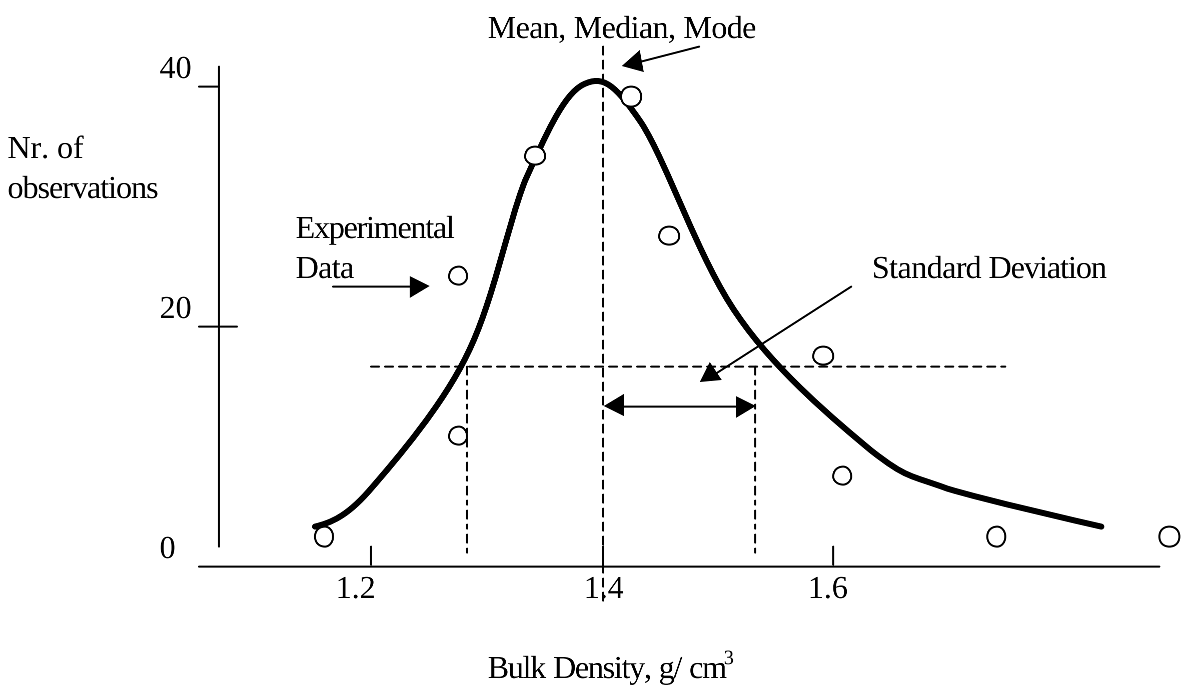
1. The Mean tells us about the average value of the soil property
2. The Standard Deviation gives information about the scatter around the Mean
3. The Coefficient of Variation normalizes that variation, so that variations of properties with different mean magnitudes can be easily compared

Other statistical properties include the mode (most frequently measured value), kurtosis (skewness of distribution), median (1/2 of z-values are smaller and ½ of z-values are larger than median).

None of these properties either by themselves or combined say anything about the type of distribution from which the samples are derived (frequency or probability density function).

Probability distributions

The population is more completely defined by its frequency distribution, which includes the mean and variation values:



For example, the frequency distribution of bulk density might be normally distributed, or

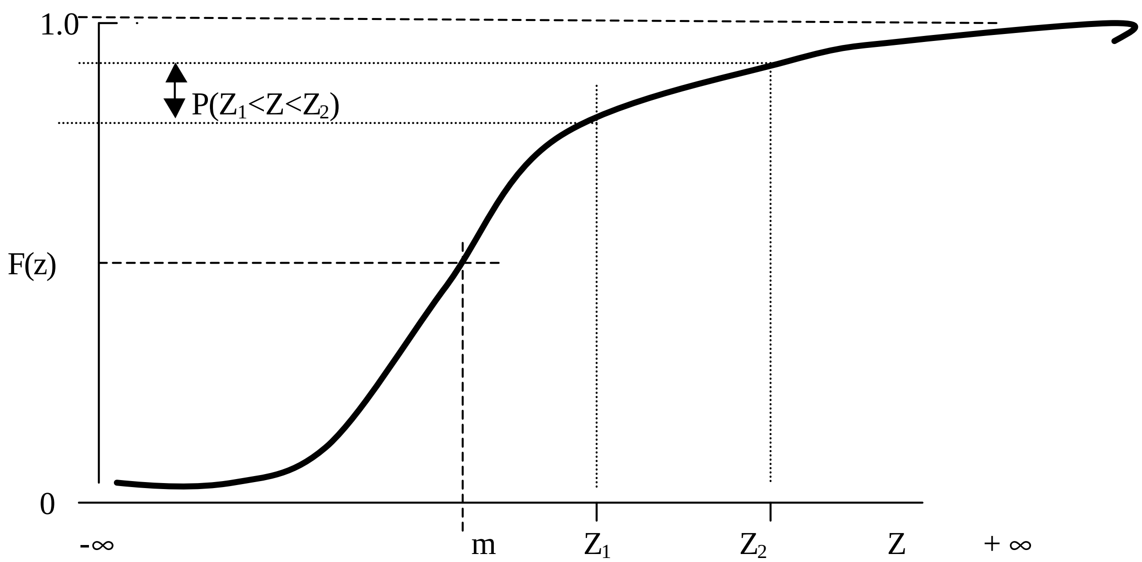
$$f(Z) = \frac{1}{s\sqrt{2\pi}} \exp\left(\frac{-(z-m)^2}{2s^2} \right)$$

where f(Z) is defined as the probability density function, which describes the frequency distribution as measured (see figure).

The cumulative distribution function, F(Z) is defined as:

$$F(Z) = P(z < Z) = \int_{-\infty}^{z} f(z)dz \qquad \text{and } F(\infty) = 1$$

Alternatively, we can say that the area under the curve of f(x=∞) is always equal to 1.0



Also, use the figure of above to confirm that:

$$P(Z_1 < Z < Z_2) = F(Z_2) - F(Z_1) = \int_{Z_1}^{Z_2} f(z)dz$$

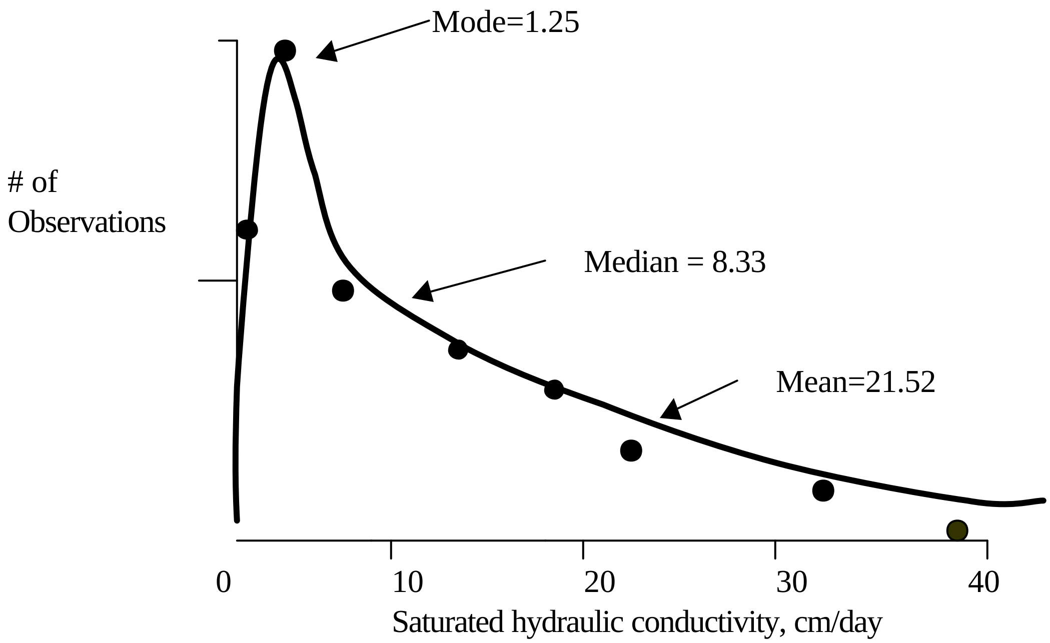
The probability Table (will be used in the laboratory) is based on normality of the distribution. However, rather than using the variable z, the standard normal deviate (μ) is defined:

$\mu = \dfrac{z - m}{s}$, so that these tables can be used for any normally distributed variable, independent of its mean or standard deviation

The standard normal distribution is a normalized distribution, which has a mean (m) of zero and standard deviation (s) of one, or μ = N[0,1].

However, normality is only one of the many type of distributions that have been described analytically. Others are Poisson Distribution, Beta and Gamma Distribution, Gumbel Distribution, and Lognormal Distribution:

<u>Log-normal distribution:</u>
(of saturated hydraulic conductivity)



If a soil property (z) is log-normally distributed than the logarithm of this property (lnz) will be normally distributed. In that case, the mean, median and mode are different:

Let m_{lnx} be the mean of the lnz values and s_{lnx} be the standard deviation of the log-transformed z-values, then:

$$\text{Mode} = \exp(m_{lnx} - s^2_{lnx})$$
$$\text{Median} = \exp(m_{lnx})$$
$$\text{Mean} = = \exp(m_{lnx} + 0.5s^2_{lnx})$$
$$\text{Variance} = \text{Mean}^2[\exp(s^2_{lnx}) - 1]$$

How do you determine that a set of soil physical observations is normally or lognormally distributed ??

This is done by constructing a fractile diagram:

1. Rank the values of the variable from low to high in ascending order,

2. Compute the cumulative probability distribution of z as approximated by
$P_i = \frac{i - 0.5}{N}, i = 1,.., N$, N = total number of observations.

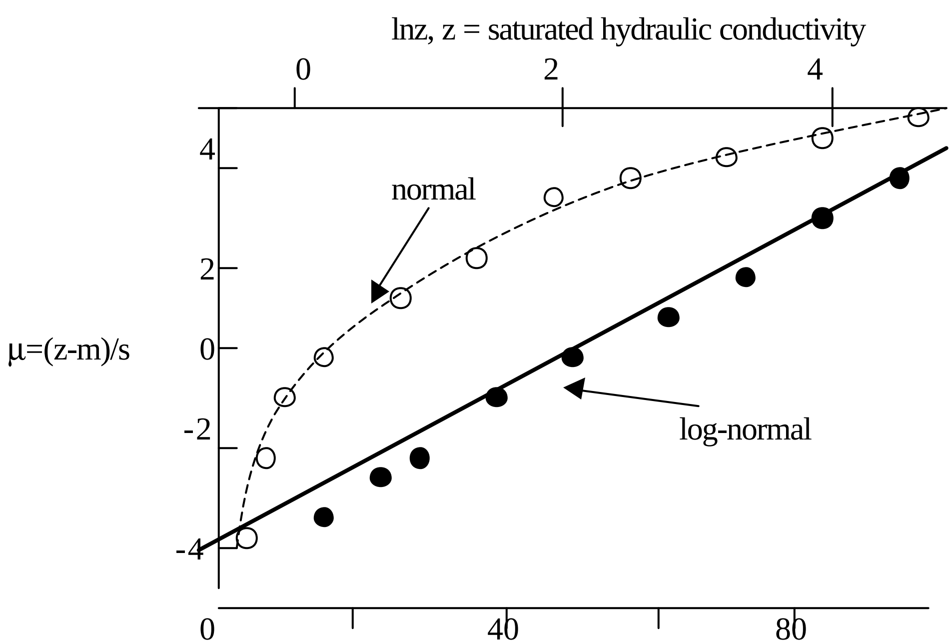
3. Determine the standard normal deviate (μ_i) for the distributed variable from the standard normal probability table, i.e. P[$\mu < \mu_i$], for each value computed in 2 above.

4. Plot the standard normal deviate (μ_i) determine in 3 versus the variable z,

5. Repeat, but in instead of z, do the ranking and probability calculations for lnz,

6. Repeat step 4,

7. The plot with the straight line relationship is normally distributed.



z, saturated hydraulic conductivity

An example of how to do these calculations for the fractile diagram are given in the following table:

Table 13.1
OBSERVED VALUES (N = 20) OF PORE WATER VELOCITY υ (CM DAY^{-1}), SOIL WATER CONTENT θ (CM^3 CM^{-3}), AND CORRESPONDING STATISTICAL PARAMETERS[a]

i	$(\upsilon)_i$	$(\ln \upsilon)_i$	θ_i	(i-0.5)/N[b]	(z-m)/s[c]
1	0.66	-0.42	0.297	0.025	-1.96
2	1.16	0.15	0.310	0.075	-1.44
3	1.52	0.42	0.318	0.125	-1.15
4	2.29	0.83	0.326	0.175	-0.93
5	2.94	1.08	0.330	0.225	-0.76
6	3.46	1.24	0.333	0.275	-0.60
7	4.35	1.47	0.339	0.325	-0.45
8	4.95	1.60	0.341	0.375	-0.32
9	6.11	1.81	0.345	0.425	-0.19
10	7.39	2.00	0.350	0.475	-0.06
11	8.41	2.13	0.351	0.525	0.06
12	10.38	2.34	0.356	0.575	0.19
13	12.81	2.55	0.359	0.625	0.32
14	15.64	2.75	0.363	0.675	0.45
15	18.64	2.92	0.366	0.725	0.60
16	24.53	3.20	0.372	0.775	0.76
17	31.14	3.47	0.375	0.825	0.93
18	47.47	3.86	0.382	0.875	1.15
19	88.23	4.48	0.390	0.925	1.44
20	90.75	4.51	0.404	0.975	1.96
μ_E	19.19	2.12[d]	0.350	--	--
σ_E	26.800	1.38[e]	0.027	--	--

[a] Used to construct the frequency distribution and the fractile diagram given in Fig. 2.

[b] (i - 0.5)/N approximates the value of the cumulative probability function P{u} where

$$P\{u\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} exp(-x^2/2)\,dx$$

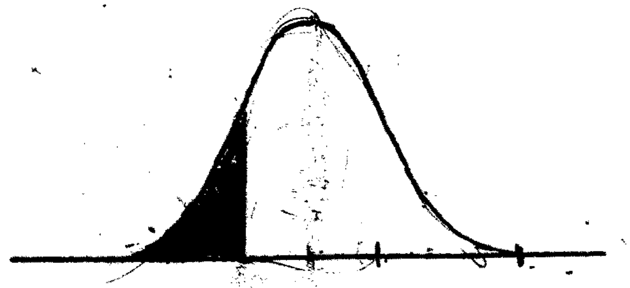
[c] Values of $[z-m]s^{-1} = u$ are obtained from tables of P{u} for each value of (i - 0.5)/N.

[d] For $g(\upsilon) = \ln \upsilon$, $m_E = \sum_{i=1}^{20} [(ln \ \upsilon)_i / 20] = 2.12$.

[e] For $g(\upsilon) = \ln \upsilon$,

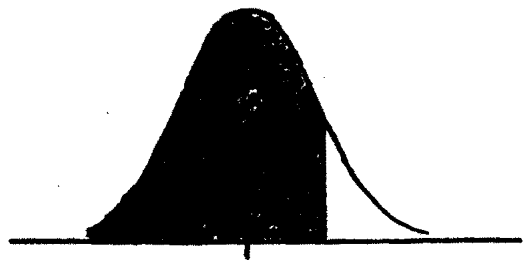
$$s_E = \left\{ \sum_{i=1}^{20} [(\ln \ \upsilon)_i - m_E]^2 / (N-1) \right\}^{1/2} \quad 1.38$$

THE STANDARD NORMAL DISTRIBUTION



z_0	0.000	0.010	0.020	0.030	0.040	0.050	0.060	0.070	0.080	0.090
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0004
-3.2	.0007	.0007	.0007	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0009	.0008	.0008	.0008	.0007	.0007
-3.0	.0014	.0013	.0013	.0012	.0012	.0012	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014
-2.8	.0026	.0025	.0024	.0024	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0027
-2.6	.0047	.0046	.0044	.0043	.0042	.0041	.0039	.0038	.0037	.0036
-2.5	.0063	.0061	.0059	.0057	.0056	.0054	.0053	.0051	.0050	.0048
-2.4	.0082	.0080	.0078	.0076	.0074	.0072	.0070	.0068	.0066	.0064
-2.3	.0108	.0105	.0102	.0100	.0097	.0094	.0092	.0089	.0087	.0085
-2.2	.0140	.0136	.0133	.0129	.0126	.0123	.0120	.0117	.0114	.0111
-2.1	.0179	.0175	.0171	.0166	.0162	.0158	.0154	.0151	.0147	.0143
-2.0	.0228	.0223	.0218	.0212	.0207	.0202	.0198	.0193	.0188	.0184
-1.9	.0288	.0281	.0275	.0269	.0263	.0257	.0251	.0245	.0239	.0234
-1.8	.0360	.0352	.0345	.0337	.0330	.0322	.0315	.0308	.0301	.0295
-1.7	.0446	.0437	.0428	.0419	.0410	.0401	.0393	.0384	.0376	.0368
-1.6	.0549	.0538	.0527	.0516	.0506	.0496	.0485	.0475	.0466	.0456
-1.5	.0669	.0656	.0643	.0631	.0619	.0606	.0595	.0583	.0571	.0560
-1.4	.0808	.0793	.0779	.0764	.0750	.0736	.0722	.0709	.0695	.0682
-1.3	.0969	.0952	.0935	.0918	.0902	.0886	.0870	.0854	.0839	.0823
-1.2	.1151	.1132	.1113	.1094	.1076	.1057	.1039	.1021	.1003	.0986
-1.1	.1357	.1336	.1314	.1293	.1272	.1251	.1231	.1211	.1191	.1171
-1.0	.1587	.1563	.1539	.1516	.1492	.1469	.1446	.1424	.1401	.1379
-.9	.1841	.1815	.1788	.1762	.1737	.1711	.1686	.1661	.1636	.1611
-.8	.2119	.2090	.2062	.2033	.2005	.1977	.1949	.1922	.1895	.1868
-.7	.2420	.2389	.2358	.2327	.2297	.2267	.2237	.2207	.2177	.2148
-.6	.2743	.2710	.2677	.2644	.2611	.2579	.2547	.2515	.2483	.2451
-.5	.3086	.3051	.3016	.2981	.2946	.2912	.2878	.2844	.2810	.2776
-.4	.3446	.3409	.3373	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-.2	.4208	.4168	.4129	.4091	.4052	.4013	.3974	.3936	.3898	.3859
-.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

THE STANDARD NORMAL DISTRIBUTION



z	0.000	0.010	0.020	0.030	0.040	0.050	0.060	0.070	0.080	0.090
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5792	.5832	.5871	.5909	.5948	.5987	.6026	.6064	.6102	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6627	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6914	.6949	.6984	.7019	.7054	.7088	.7122	.7156	.7190	.7224
.6	.7257	.7290	.7323	.7356	.7389	.7421	.7453	.7485	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7703	.7733	.7763	.7793	.7823	.7852
.8	.7881	.7910	.7938	.7967	.7995	.8023	.8051	.8078	.8105	.8132
.9	.8159	.8185	.8212	.8238	.8263	.8289	.8314	.8339	.8364	.8389
1.0	.8413	.8437	.8461	.8484	.8508	.8531	.8554	.8576	.8599	.8621
1.1	.8643	.8664	.8686	.8707	.8728	.8749	.8769	.8789	.8809	.8829
1.2	.8849	.8868	.8887	.8906	.8924	.8943	.8961	.8979	.8997	.9014
1.3	.9031	.9048	.9065	.9082	.9098	.9114	.9130	.9146	.9161	.9177
1.4	.9192	.9207	.9221	.9236	.9250	.9264	.9278	.9291	.9305	.9318
1.5	.9331	.9344	.9357	.9369	.9381	.9394	.9405	.9417	.9429	.9440
1.6	.9451	.9462	.9473	.9484	.9494	.9504	.9515	.9525	.9534	.9544
1.7	.9554	.9563	.9572	.9581	.9590	.9599	.9607	.9616	.9624	.9632
1.8	.9640	.9648	.9655	.9663	.9670	.9678	.9685	.9692	.9699	.9705
1.9	.9712	.9719	.9725	.9731	.9737	.9743	.9749	.9755	.9761	.9766
2.0	.9772	.9777	.9782	.9788	.9793	.9798	.9802	.9807	.9812	.9816
2.1	.9821	.9825	.9829	.9834	.9838	.9842	.9846	.9849	.9853	.9857
2.2	.9860	.9864	.9867	.9871	.9874	.9877	.9880	.9883	.9886	.9889
2.3	.9892	.9895	.9898	.9900	.9903	.9906	.9908	.9911	.9913	.9915
2.4	.9918	.9920	.9922	.9924	.9926	.9928	.9930	.9932	.9934	.9936
2.5	.9938	.9939	.9941	.9943	.9944	.9946	.9947	.9949	.9950	.9952
2.6	.9953	.9954	.9956	.9957	.9958	.9959	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9973
2.8	.9974	.9975	.9976	.9976	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9983	.9984	.9984	.9985	.9985	.9986
3.0	.9986	.9987	.9987	.9988	.9988	.9988	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9991	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9993	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9996

Measured Variabilities:

Bulk density and porosity: CV = 7 – 10 %

Particle size and θ at 0.1 – 15 bar: CV = 10 – 100 %

K_{sat}, K(h), pore water velocity and EC (salinity): CV = 100 – 1000 %

> In general: Static soil properties are normally distributed and dynamic soil properties are log-normally distributed and have larger CV's.

If the distribution is normal, the minimum number of samples (N) required to estimate the mean of the soil property within a pre-defined percentage of the true value (d = m in %) at a certain confidence level (α) can be determined from the following expression:

$$N = \frac{X_\alpha^2 s^2}{d^2}$$

where X_α is the normal deviate corresponding with confidence interval or probability of making an error:

α	X_α
0.05	1.96
0.10	1.65
0.50	0.84

Example:

m = 9.5
s = 3.1
d = (0.15)(9.5) = 1.4 (15% of the mean)
Confidence level (α) = 0.05

Then: $N = (1.96)^2 (3.1)^2 / (1.4)^2 = 19$

Consequently, the soil properties with the lowest CV requires the least number of samples to be taken to estimate their mean value within a certain percentage at any significance level.

If 10 % uncertainty of the mean is allowed, and for a 0.05 significance level (experimental data):

Property	N
Bulk density	2
0.1bar water content	61
15 bar water content	98
Clay (%)	110
Silt (%)	150
Sand (%)	53
Hydraulic conductivity	1,300