6.0 SAMPLING INTERVAL AND NUMBER OF SAMPLES

6A.1 INTRODUCTION TO THE DETERMINATION OF WATER SAMPLE NUMBERS

Selecting how often to sample and how many samples to take are two of the most important, yet difficult, tasks in the implementation of the spill remediation program. The methodologies discussed in this section are intended to be used in the spill response program for ground water. Statistical methods for determining the required number of samples may not be necessary until the operation is near the end of remediation.

Ground water sampling has traditionally been done quarterly or semi-annually and, in cases of hazardous waste sites, as often as monthly. Where knowledge exists about the dynamics of the ground water flow, sampling frequency can be established by considering prevailing ground water flow velocity, degree of potential environmental and/or economical impact, the objective of the sampling, and statistics. For example, where the ground water flow velocity is one foot per day or 30 feet per month, monthly sampling amounts to collection of ground water every 30 feet along the line of flow. In a rural, sparsely populated area, this may be too excessive regardless of the nature or toxicity of the pollutant being monitored because, in most situations, the fact that the plume has moved 30 feet before it was detected would not jeopardize the effectiveness of potential remedial measures. In an urban area, or where there is a possible threat to public water supply, monthly sampling may be inadequate.

The purposes of sampling are: (A) to detect the plume at an early stage in its development so that remedial measures can be implemented with relative ease; (B) to determine if and to what extent a plume has been formed; or (C) to monitor how the plume has been progressing and how effective the remedial action has been. In the case of B, a more technically sound monitoring program might consist of a larger number of wells and more comprehensive, but not more frequent, sampling and analysis. For A and C, greater frequency of sampling at fewer wells might meet the need if the monitoring system is properly designed. A method for determining sampling frequency is discussed below in the following categories:

- Sampling Interval for Spill Monitoring Well (Section 6A.2)
- Sampling Interval for Residential Well After Report of Contamination (Section 6A.3)
- Sampling Frequency for Water Treatment Device or Apparatus (Section 6A.4) (For Residential Water Treatment and Ground Water Treatment)
- Sampling Frequency of Plume Monitoring (Section 6A.5)
- Sampling Schedule for Groundwater After Remediation Has Stopped or In Other Normally Distributed, Steady Conditions (Section 6A.6)

6A.1.1 General Considerations

The following procedures and factors should be followed:

- Specific groundwater sampling procedures described in the applicable EPA or DOH protocol shall be followed. Samples shall be handled and data shall be collected according to these procedures.
- Local ordinances or State and Federal regulations which require specific sampling frequencies or intervals should be followed.
- In some situations, the statistical methods recommended below may become impractical. Judgements and appropriate modification may have to be made.

6A.2 SAMPLING INTERVAL FOR SPILL MONITORING WELL

6A.2.1 At Populated Areas or At Areas With A Threat To Public Water(Well)

When groundwater contamination has been discovered, the rate of the spread of pollutant and its possible impacts should be considered. The recommended sampling interval can be determined as follows:

(A) At the site where the spill source has not been identified, or the spill source has been identified but no remediation has been started:

(A1)(i) if the groundwater flow velocity is greater than three tenths of one foot a day, the sample interval can be determined by dividing half of the distance (feet) between the suspected source, such as an underground storage tank and the possible affected party or well which is directly downgradient, by the average groundwater flow velocity (feet/day).

(ii) if the groundwater flow velocity is less than three tenths of one foot a day, the sample interval can be determined by dividing two thirds of the distance (feet) between the suspected source, such as a storage tank and the first possible affected party or well, by the average groundwater flow velocity (feet/day).

The results provide a recommended sampling interval which should be used until the source is found and remediation has been started.

(A2) Where the nearest affected party or well is not directly downgradient in the flow path:

use the closest potentially affected party or well located in a direction downgradient of the spill source and calculate the sampling interval as it is recommended above in Section 6A.2.1.A1.

(B) At the site where remediation has been started, the sampling interval can be determined by:

(i) using the sampling interval recommended in Section 6A.2.1.A. for the first three samples and plotting the concentrations of the first three samples against the time (days) lapsed in a graph.

(ii) estimating the reduction rate of the pollutant from the graph. Subsequent sampling should take place at the time the pollutant would be reduced to half the concentration of the immediate prior sample. The reduction rate of the pollutant may need an adjustment as additional information concerning concentrations reduction over time is obtained. (see 6B.6.2 for example)

In addition, samples should also be taken at the time the water table is at its highest and lowest levels.

6A.2.2 At Non-Populated Rural Areas and Where No Groundwater Is Being Used

In unpopulated rural areas consideration must be given to the migration of the pollutant, but there is less urgency in addressing the impact. The following recommended sampling schedules are based on an assumption that the average ground water flow rate is 150 feet per year*.

(A) At the site where the spill source has not been confirmed, the sampling period should be determined by dividing 150 feet by the average ground water flow velocity (feet/day) at the site.

(B) At the site where remediation has been started, the sampling schedule should be determined by using the same method as that described in Section 6A.2.1.B. However, use Section 6A.2.2.A. to obtain the first three sample points. In addition, samples should also be taken at the time the water table is at its highest and lowest levels.

6A.3 SAMPLING INTERVAL FOR RESIDENTIAL WELL AFTER REPORT OF CONTAMINATION

Once the contamination is reported and confirmed, the sampling schedule is recommended as follows:

* Nacht, Steve J., "Monitoring Sampling Protocol Considerations", Groundwater Monitoring Review Summer 1983, vol. 3, No. 3, 23-29.

(A) At sites where the spill source has not been identified:

(A1) When there is a nearby neighbor directly downgradient along the flow path and where the average groundwater flow velocity is greater than three tenths of a foot a day, the sampling interval should be determined by dividing half of the distance (ft) between the contaminated resident and the next possibly affected resident by the average ground water flow velocity(ft/day) at the site until the source is found and remediation has been started;

(A2) At the area where the average ground water flow velocity is below three-tenths of a foot a day, the sampling interval should be determined by dividing two thirds of the distance (ft) between the contaminated resident and the nearest neighbor by the average ground water flow velocity (ft/day) until the source is found and remediation has been started.

(B) At the site where spill source has been identified and remediation has been started, the sampling schedule should be determined by:

First, using the sampling interval recommended in Section 6.3.A. for the first three samples and plotting the concentrations of the first three samples against the time (days) lapsed in a graph;

Second, estimating the reduction rate of the pollutant from the graph and subsequent sampling at the time the pollutant is reduced to half the concentration from the immediate prior sample. The reduction rate of the pollutant may need an adjustment upon additional information concerning concentration reduction over time.

In addition, samples should also be taken at the time the water table is at its highest and lowest levels.

6A.4 SAMPLING FREQUENCY FOR WATER TREATMENT DEVICE OR APPARATUS (For Residential Water Treatment and Groundwater Treatment During Remediation)

When selecting a water treatment device or apparatus, obtain and evaluate all technical data from the manufacturer, including targeted pollutants, treatment capability, operational parameters, and its manufacturing quality control and quality assurance. Operational parameters such as pH value, temperature, water hardness, micro organism, suspended matters, treatment rate, treatment capacity, etc. are important factors to be considered for application.

Sampling in this category includes taking influent and effluent simultaneously and is recommended at the beginning of service and at 90% of the estimated useful life of the device. If the interval between the beginning of service and 90% of the useful life of the device is more than three months, it is recommended that the sampling should be done at the beginning of service, every quarter, and then at 90% of the useful life of the device.

6A.5 SAMPLING FREQUENCY OF PLUME MONITORING FOR GROUNDWATER

In addition to the determination of the sampling interval after remediation has been started as discussed earlier, a sampling schedule and the number of samples for an appropriate monitoring can be determined based upon other objectives related to the spill.

(A) In populated areas or at areas with a threat to public water;

(A1) When the objective is to define the plume concentrations in a statistically meaningful way, the number of samples should be determined by using Section 6A.6.

(A2) When the objective is to monitor the movement or the remedial effectiveness of the plume, the sampling schedule is the same as Section 6A.2.1. Additional downgradient wells in both the longitudinal and the transverse directions may be required.

(B) In non-populated rural areas and where no groundwater is being used;

(B1) When the objective is to define the plume concentrations in a statistically meaningful way, the sampling schedule and the number of samples required can be determined by using Section 6A.6; and

(B2) When the objective is to monitor the movement or the remedial effectiveness of the plume, the sampling schedule is the same as Section 6A.2.2. Additional downgradient wells in both the longitudinal and the transverse direction may be required.

6A.6 SAMPLING SCHEDULE FOR GROUNDWATER AFTER REMEDIATION HAS STOPPED OR IN OTHER NORMALLY DISTRIBUTED, STEADY (NOT A SIGNIFICANTLY VARIANT) CONDITIONS (Sampling Number for Contaminant Plume in Transitory Conditions)

The following recommended procedures are based on statistical concepts and can be utilized to determine the number of samples required for any specified confidence over a period of time (generally one year) or at a site where a statistically reliable definition or concentration of a plume is needed, (e.g. when determining that no further cleanup is being accomplished). A period of time can be a short period of time, such as one or two days to collect enough samples for finding a statistically meaningful concentration over a plume, or longer, such as a year for a complete cycle of monitoring of overall water quality in an investigated area. The confidence (interval) or the probability of detection can be started from 80 percent and up depending on the amount of resources you have and the degree of reliability you desire. Two techniques are presented here, but the user has to determine whether single or combined techniques are most appropriate for his or her application. The sampling interval can be determined by dividing the number of samples by the period of time available (up to a year). For a definition of statistical terms, see Table 6-1.

6A.6.1 Determination of Sample Number or Sampling Schedule by Using Standard Normal Value (Example)

STEP 1. Set the probability of detection at the expected percent, say 85%, and the standard deviation (measure of the spread) at any desired number, say 0.25.

2. Subtract the expected probability of detection in decimal form, 85% = 0.85, from 1.0 (100% detection probability).

3. The remainder, 1.00 - 0.85 = 0.15, from above Step 2 will be for a two-tailed z score. Divide the remainder, 0.15, by 2 to obtain the one-tailed value, 0.075.

4. Use this one-tailed value, 0.075, and look up Table 6-3 for a two-tailed z score at the probability of 0.075. The z score is 1.44 in this example.

5. Put the z value, 1.44, and the desired standard deviation, 0.25, into Equation (12) in Table 6-1 to obtain the number of samples required to meet the expected probability of detection and the desired standard deviation, $(1.44/0.25)^2 = 33$ samples in this example.

Using a period of one year (365 days) to collect these samples, a sampling period of once every 11 days (365 days/33 samples) is obtained.

6A.6.2 Determination of Sample Number or Sample Schedule By Using Simple Random Sampling

The steps described below also lead further to the determination of whether the contaminant in the groundwater is at a hazardous level or not, in addition to the determination of number of samples.

TABLE 6-1. BASIC STATISTICAL TERMINOLOGY APPLICABLE TO SAMPLING PLANS FOR GROUNDWATER

Terminology	Symbol	Mathematical equation	(Equation)
★ Variable (e.g., benzene, toluene or TOC)	x	---	
★ Individual Measurement of variable	x_i	---	
★ Mean of all possible measurements of variable (population mean)	μ	$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$, where N = number of possible measurements	(1)
★ Mean of measurements generated by sample (sample mean)	$\bar{x}$	For Simple Random or Systematic Random Sampling: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$, where n = number of sample measurements	(2a)
		For Stratified Random Sampling: $\bar{x} = \sum_{k=1}^{r} W_k \bar{x}_k$, where $\bar{x}_k$ = stratum mean and W_k = fraction of population represented by Stratum k (number of Strata [k] range from 1 to r)	(2b)
★ Variance of sample	s^2	For Simple Random or Systematic Random Sampling: $s^2 = \frac{\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2/n}{n-1}$, or $s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$	(3a)
		For Stratified Random Sampling: $s^2 = \sum_{k=1}^{r} W_k s_k^2$, where s_k^2 = stratum variance and W_k = fraction of population represented by Stratum k (number of strata [k] range from 1 to r)	(3b)
★ Standard deviation of sample	s	$s = \sqrt{s^2}$	(4)
★ Standard error (also standard error of mean and standard deviation of mean) of sample.	$s_{\bar{x}}$	$s_{\bar{x}} = \frac{s}{\sqrt{n}}$	(5)
★ Confidence interval for μ*	CI	$CI = \bar{x} \pm t \times s_{\bar{x}}$, where t obtained from Table 6-2 for appropriate degree of freedom	(6)
★ Regulatory threshold*	RT	Defined by DEC or EPA, or DOH (e.g. 100 ppb for TOC)	(7)

TABLE 6-1. **BASIC STATISTICAL TERMINOLOGY APPLICABLE TO SAMPLING PLANS FOR GROUNDWATER** (continued)

	Terminology	Symbol	Mathematical equation	(Equation)
★	Appropriate number of samples to collect from investigated site (by statistics)	n	$n = \frac{t^2 x s^2}{\Delta^2}$, where $\Delta = RT - \bar{x}$	(8)
★	Degree of freedom	df	$df = n - 1$	
★	Square root transformation	-	$\sqrt{x_i + 1/2}$	(10)
★	Arcsin transformation	-	Arcsin p: if necessary, refer to any text on basic statistics; measurements must be converted to percentage (p)	(11)
★	Standard normal variable z	z	$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$, with σ = population deviation or $\sqrt{n} = \frac{z}{x'}$ or $n = (\frac{z}{x'})^2$ if $x' = \frac{\bar{x}-\mu}{\sigma}$ standard deviation (measure of the spread)	(12)

* The upper limit of the CI for is compared with the applicable regulatory threshold (RT) to determine if a sample contains the variable (pollutant) of concern at a hazardous level. The pollutant of concern is not considered to be present in the sample at a hazardous level if the upper limit of the CI is less than the applicable RT. Otherwise, the opposite conclusion is reached.

TABLE 6-2

PERCENTILES OF THE t DISTRIBUTION

Degrees of Freedom(df) (n-1)	Confidence Level (%): $1-\alpha/_2$ for two-tailed test							
	20	30	60	80	90	95	98	99
	Confidence Level (%): $1-\alpha$ for one-tailed test							
	60	70	80	90	95	97.5	99	99.5
1	.325	.727	1.376	3.078	6.314	12.706	31.821	63.657
2	.289	.617	1.061	1.886	2.920	4.303	6.965	9.925
3	.277	.584	.978	1.638	2.353	3.182	4.541	5.641
4	.271	.569	.941	1.533	2.132	2.776	3.747	4.604
5	.267	.559	.920	1.476	2.015	2.571	3.365	4.032
6	.265	.553	.906	1.440	1.943	2.447	3.143	3.707
7	.263	.549	.896	1.415	1.895	2.365	2.998	3.499
8	.262	.546	.889	1.397	1.860	2.306	2.896	3.355
9	.261	.543	.883	1.383	1.833	2.262	2.821	3.250
10	.260	.542	.879	1.372	1.812	2.228	2.764	3.169
11	.260	.540	.876	1.363	1.796	2.201	2.718	3.106
12	.259	.539	.873	1.356	1.782	2.179	2.681	3.055
13	.259	.538	.870	1.350	1.771	2.160	2.650	3.012
14	.258	.537	.868	1.345	1.761	2.145	2.624	2.977
15	.258	.536	.866	1.341	1.753	2.131	2.602	2.947
16	.258	.535	.865	1.337	1.746	2.120	2.583	2.921
17	.257	.534	.863	1.333	1.740	2.110	2.567	2.898
18	.257	.534	.862	1.330	1.734	2.101	2.552	2.878
19	.257	.533	.861	1.328	1.729	2.093	2.539	2.861
20	.257	.533	.860	1.325	1.725	2.386	2.528	2.845
21	.257	.532	.859	1.323	1.721	2.080	2.518	2.831
22	.256	.532	.858	1.321	1.717	2.074	2.508	2.819
23	.256	.532	.858	1.319	1.714	2.069	2.500	2.807
24	.256	.531	.857	1.318	1.711	2.064	2.492	2.797
25	.256	.531	.856	1.316	1.708	2.060	2.485	2.787
26	.256	.531	.856	1.315	1.706	2.056	2.479	2.779
27	.256	.531	.855	1.314	1.703	2.052	2.473	2.771
28	.256	.530	.855	1.313	1.701	2.048	2.467	2.763
29	.256	.530	.854	1.311	1.699	2.045	2.462	2.756
30	.256	.530	.854	1.310	1.697	2.042	2.457	2.750
40	.255	.529	.851	1.303	1.684	2.021	2.423	2.704
60	.254	.527	.848	1.296	1.671	2.000	2.390	2.660
120	.254	.526	.845	1.289	1.658	1.980	2.358	2.617
∞	.253	.524	.842	1.282	1.645	1.960	2.326	2.576

STEP 1. Obtain preliminary estimates of $\bar{x}$ and s^2 (from preliminary investigation and sampling) for each contaminant of a groundwater that is of concern. The two identified statistics above are calculated using Equations 2a and 3a (Table 6-1), respectively.

2. Estimate the appropriate number of samples (n_1) to be collected from the groundwater through use of Equation 8 (Table 6-1) and Table 6-2. Derive individual values of n_1 for each contaminant of concern. The appropriate number of samples to be taken from the groundwater is the greatest of the individual n_1 values.

3. Randomly collect at least n_1 (or $n_2 - n_1$, $n_3 - n_2$, etc., as will be indicated later in this procedure) samples from the groundwater (collection of a few extra samples will provide protection against poor preliminary estimates of $\bar{x}$ and s^2).

4. Analyze the n_1 (or $n_2 - n_1$, $n_3 - n_2$, etc.) samples for each contaminant of concern. Superficially (graphically) examine each set of analytical data for obvious difference from what might be expected.

5. Calculate $\bar{x}$, s^2, the standard deviation (s), and $s_{\bar{x}}$ for each set of analytical data using, respectively, Equations 2a, 3a, 4, and 5 (Table 6-1).

6. If $\bar{x}$ for a contaminant is equal to or greater than the applicable regulatory threshold (RT) (Equation 7, Table 6-1) and is believed to be an accurate estimator of μ, the contaminant is considered to be present in the groundwater at a hazardous concentration, and the study is completed. Otherwise, continue the study. In the case of a set of analytical data that does not exhibit obvious abnormality and for which $\bar{x}$ is greater than s^2, perform the following calculations with non-transformed data. Otherwise, consider transforming the data by the square root transformation (if $\bar{x}$ is about equal to s^2) or the arcsine transformation (if $\bar{x}$ is less than s^2) and performing all subsequent calculations with transformed data. Square root and arcsine transformation are defined by Equations 10 and 11 (Table 6-1) respectively.

7. Determine the CI for each contaminant of concern using Equation 6 (Table 6-1) and Table 6-2. If the upper limit of the CI is less than the applicable RT (Equations 6 and 7, Table 6-1), the contaminant is not considered to be present in the water at a hazardous concentration and the study is completed. Otherwise, the opposite conclusion is tentatively reached.

8. If a tentative conclusion of a hazardous concentration is reached, re-estimate the total number of samples (n_2) to be collected from the groundwater by use of Equation 8 (Table 6-1) and Table 6-2. When deriving n_2, employ the newly calculated (not preliminary) values of $\bar{x}$ and s^2. If additional $n_2 - n_1$ samples of groundwater cannot reasonably be collected, the study is completed, and a definitive conclusion of hazard is reached. Otherwise, collect extra $n_2 - n_1$ samples of groundwater.

9. Repeat the basic operations described in Step 3 through 8 until the water is judged to be nonhazardous or, if the opposite conclusion continues to be reached, until increased sampling effort is impractical.

EXAMPLE (HYPOTHETICAL)

STEP 1. Assume that four samples were taken from four wells of a plume in a preliminary study for total organic contaminants (TOC). Their analytical results are: 86, 90, 98, 104 ppb. The preliminary estimates of $\bar{x}$ and s^2 are calculated as follows:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{86+90+98+104}{4} = 94.5$$ (by Equation 2a) and

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2/n}{n-1} = \frac{35916.00-35721.00}{3} = 65.00$$ (by Equation 3a)

2. Based on the preliminary estimates of $\bar{x}$ and s^2, as well as on the knowledge that the RT for TOC is 100 ppb,

$$n_1 = \frac{t^2 x s^2}{\Delta^2} = \frac{(1.638^2)(65.00)}{(100-94.5)^2} = 5.77$$ by Equation 8(Table 6-1) and $t=1.638$ by choosing 80% CI (Table 6-2)

3. As indicated in Step 2, the appropriate number of samples to be collected from the plume is six. As mentioned before, collection of a few more samples will provide protection against poor preliminary estimates of $\bar{x}$ and s^2.

4. The six samples of groundwater designated for immediate analysis generate the following results: 89, 90, 87, 96, 93, and 113 ppb. Although the value of 113 ppb appears unusual as compared with other data, there is no obvious indication that the data is not normally distributed.

5. New value for $\bar{x}$ and s^2 and associated values for the standard deviation (s) and $s_{\bar{x}}$ are calculated as:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{89+90+87+96+93+113}{6} = 94.67$$ (Equation 2a)

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2/n}{n-1} = \frac{54224.00-53770.67}{5} = 90.67$$ (Equation 3a)

$$s = \sqrt{s^2} = \sqrt{90.67} = 9.52$$ (Equation 4)

$$s_{\bar{x}} = s/\sqrt{n} = 9.52/\sqrt{6} = 3.89$$ (Equation 5)

6. The new value for $\bar{x}$ (94.67) is less than the RT (100). In addition, $\bar{x}$ is greater (only slightly) than s^2 (90.67), and, as previously indicated, the raw data are not characterized by obvious abnormality. Consequently, the study is continued, with the following calculations performed with nontransformed data.

7. $$CI = \bar{x} \pm t x s_{\bar{x}} = 94.67 \pm (1.476)(3.89) = 94.67 \pm 5.74$$ (Equation 6)

   Because the upper limit of CI (100.41) is greater than the applicable RT (100), it is tentatively concluded that the TOC is present in the water at a hazardous concentration.

8. n is now reestimated as:

$$n_2 = \frac{t^2 x s^2}{\Delta^2} = \frac{(1.476^2)(90.67)}{(100-94.67)^2} = 6.95$$

by Equation 8, and t = 1.476 and 80% CI, Table 6-2.

The value for n_2 (approximately 7) indicates that an additional (n_2 - n_1 = 1) sample should be collected.

9. It happened that three extra samples were collected along with the six samples. The analytical results for these three samples are 93, 90 and 91 ppb. Consequently, $\bar{x}$, s^2, the standard deviation (s), and $s_{\bar{x}}$ are calculated as follows:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{86+90+....+91}{9} = 93.56$$ (Equation 2a)

$$s^2 = \frac{\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2/n}{n-1} = \frac{79254.00-78773.78}{8} = 60.03$$ (Equation 3a)

$$s = \sqrt{s^2} = \sqrt{60.03} = 7.75$$ (Equation 4)

$$s_{\bar{x}} = s/\sqrt{n} = 7.75/\sqrt{9} = 2.58$$ (Equation 5)

The value for $\bar{x}$ (93.56) is again less than the RT (100), and there is no indication that the nine data points, considered collectively, are abnormally distributed (in particular, $\bar{x}$ is now substantially greater than s^2). Consequently, CI, calculated with non-transformed data, is determined to be:

$$CI = \bar{x} \pm t x s_{\bar{x}} = 93.56 \pm (1.397)(2.58) = 93.56 \pm 3.60$$

Equation 6 and Table 6-2 at 80% CI

The upper limit of the CI (97.16) is now less than the RT (100). Consequently, it is definitively concluded that TOC is not present in the water at a hazardous level.

6B.1 INTRODUCTION TO DETERMINING THE NUMBER OF SAMPLES FOR SOIL OR SOLID MEDIUM

Soil sampling is an important adjunct to groundwater monitoring. Sampling of soil above the groundwater table can detect contaminants before they migrate into the water table and can help establish the amount of contamination sorbed on vadose or aquifer soils that have the potential for contributing to groundwater contamination.

As indicated in Section 1.3, soils are very complex and variable. As a consequence, a good sampling plan and a multitude of sampling methods are necessary to complete a proper evaluation of the soil. The methodologies discussed below are intended to be used in the spill response program. High heterogeneity of soil means the methods of acquiring reliable information on soil contamination should be pursued with statistical or possibly geostatistical techniques along with specific soil handling methods. Associated with determining the number of soil sample, methods for determining the appropriate sample locations are discussed in the Appendix 6X.B at the end of this section.

Some of the major statistical terms are explained in the GLOSSARY in Appendix 6X.A. The statistical formulas used in arriving at the required sample numbers are based on these terms. For further introduction to the concepts the reader can refer to statistics textbooks.

The number of samples derived from theoretical statistical concepts may often be large to achieve reasonable confidence in the data, but resources may often not support this requirement. Important considerations such as number of samples, reliability of data, allowable error, resources, and possible impact to environment demand an appropriate compromise. An analysis or comparison of cost for different sampling schemes and the cost of an expected remedy from a worst case should be considered. Statistically less reliable sampling schemes might be used on less significant areas and projects. A sampling scheme of higher reliability should be used for more populated areas or for areas having a greater potential of causing immediate damage to resources. A method for determining soil sample number is discussed in the following categories:

- Soil Sample Number When Identifying Pollutants
- Soil Sample Number When Estimating The Average Concentration or The Amount of Contamination on Soil In-Situ
- Soil Sample Number When Determining The Average Concentration of Contamination On Soils Which are Excavated And Treated
- Soil Sample Number When Estimating The Average Concentration of Contamination On Soil Immediate After Incineration
- Frequency of Soil Sampling

6B.1.1 <u>GENERAL CONSIDERATIONS</u>

The following factors should be considered when designing a soil sampling scheme:

- Most contaminated soil sampling projects require multiple phases of investigation.
- Preliminary field investigation data such as site history, hydrogeology, geology...etc. are as important as other information obtained afterwards.
- The statistical design of the sample size (number) should incorporate an adequate and verifiable quality assurance/quality control (QA/QC) program for the overall investigation.
- Control or background sample should be taken under similar conditions, such as the same depth to correspond with other collected samples.

- Samples shall be handled and data shall be collected according to the established procedures.
- Local ordinances or state and federal regulations which stipulate any requirements should be followed.
- In some situations, the statistical methods recommended below may become impractical. Judgements and appropriate modifications may have to be made.

6B.2 SOIL SAMPLE NUMBER WHEN IDENTIFYING POLLUTANTS

There are times in which the pollutants need to be identified and cross-matched with the substances from a suspected source. In such situations, the soils mostly saturated with the pollutants or free floating product (preferred if available) should be taken. The samples should be at least in duplicate. Triplicates may be needed sometimes if more analytical work or further evaluation of statistical parameters is to be done.

If the visible evidence of physical characteristics is not clear, an appropriate instrument should be used to determine the most contaminated soil. Note that soils which are predominantly contaminated by migrating petroleum hydrocarbon vapors, rather than liquid, can show higher contaminant concentrations in the field vapor analysis than that which will be found by laboratory analysis.

6B.3 SOIL SAMPLE NUMBER WHEN ESTIMATING THE AVERAGE CONCENTRATION OR THE AMOUNT OF CONTAMINATION ON SOIL IN-SITU

The minimum number (n) of samples required to achieve a specified precision and confidence level at a defined minimum detectable relative difference, $[100(\mu_s - \mu_B)/\mu_B]$ can be estimated by the use of Table 6-5 or one of the following equations (if other conditions such as different precision, confidence level, or coefficient of variation, etc. happen to occur).

$$n \geq [(Z_\alpha + Z_\beta)/D]^2 + 0.5Z^2_\alpha \qquad \text{(equation 13)*}$$

for a one-sided, one-sample t-test, and

$$n \geq [(Z_\alpha + Z_\beta)/D]^2 + 0.25Z^2_\alpha \qquad \text{(equation 14)*}$$

for a one-sided, two-sample test

where: Z_α = A percentile of the standard normal distribution such that $P(Z \geq Z_\alpha) = \alpha$;

Z_β = A percentile of the standard normal distribution such that $P(Z \geq Z_\beta) = \beta$;

D = (minimum relative detectable difference)/CV;

CV (coefficient of variation) = the ratio of the standard deviation (s) of sample to the mean ($\bar{x}$) of sample = $s / \bar{x}$;

Relative detectable difference = $100(\mu_s - \mu_B)/\mu_B$;

μ_s = Mean of pollutant concentration of the site after contamination;

μ_B = Mean of pollutant concentration of the site before contamination or the non-contaminated area (background).

For a two-sided t-test, the values for Z_α should be changed to $Z_{\alpha/2}$. Since the determination of whether a medium has been contaminated or not is based on the concentration of pollutant in the medium, as compared to the allowable contamination limit, only a one-sided test needs to be used for all the cases.

Regarding the confidence level, power, and relative increase over background or an action level to be

* - These formulas are recommended in the Soil Sampling Quality Assurance User's Guide of U.S. Environmental Protection Agency.

detectable with a probability, USEPA has suggested the following guidelines for three different investigative stages.

Investigative Stage	Confidence Level $(1-\alpha)$	Power $(1-\beta)$	Relative Increase over Background or an Action Level to be Detectable with Probability $(1-\beta)$
Preliminary Investigation	70-80%	90-95%	10-30% $[100\ (\mu_r - \mu_B/\mu_B)]$
Emergency Cleanup	80-90%	90-95%	10-20%
Planned Removal & Remedial Response Studies	90-95%	90-95%	10-20%

EXAMPLE: APPLICATION of EQUATION 13

How to use equation 13 to find the minimum number of samples required is as follows:

Assume: CV was found to be 30% (from preliminary investigation)(see Glossary, Appendix 6X.A)
Confidence Level $(1-\alpha)$ is to be 80%
Power $(1-\beta)$ is to be 95%
Minimum detectable relative difference is 20%

Find: D = minimum detectable relative difference/CV = 20%/30%
From the Percentiles of the t Distribution in Table 6-2 and for infinite degree of freedom (when t distribution becomes a normal one), $Z_\alpha = 0.842$, and $Z_\beta = 1.645$

Therefore, the minimum number of samples, (n) is:

$$n \geq [(Z_\alpha + Z_\beta/D]^2 + 0.5Z_\alpha^2$$
$$\geq [(0.842 + 1.645)/(20/30)]^2 + 0.5\ (0.842)^2$$
$$\geq 13.92 + 0.35 = 14.27 \rightarrow 15$$

This number, 15, agrees with the number in Table 6-5

6B.4 SOIL SAMPLE NUMBER WHEN DETERMINING THE AVERAGE CONCENTRATION OF CONTAMINATION ON SOILS WHICH ARE EXCAVATED AND TREATED

The same method described in Section 6B.3 can be applied to determine the required number of samples here. Excavated soil is often known to be contaminated. So sampling would be done to confirm the concentration at any point in time to provide a basis for a decision for next actions.

The soils excavated and treated with micro-organisms and/or by a venting process may be spread onto a fixed area of land and in two physical forms, plain flat and ridges-and-furrows (like a saw tooth). When the plain flat form is used, the sampling procedure including the required number, location, and sample taking is not different from the one in Section 6B.3. With the form of the ridges-and-furrows, the sample taking or core drilling should aim perpendicularly at the side line and near or slightly below the mid-point. Half of the samples should be taken from one side of the ridge and the other half from the other side.

6B.5 SOIL SAMPLE NUMBER WHEN ESTIMATING THE AVERAGE CONCENTRATION OF CONTAMINATION ON SOIL IMMEDIATELY AFTER INCINERATION

The effectiveness of incinerating the pollutant on soil depends on the temperature, duration at that temperature, characteristics of the pollutant and soil, and the proper equipment design and operation of the incinerating process. The characteristics of the soil includes the composition of material in it, physical form and size, thermal capacity and conductivity etc.

To use the relatively economical sample number procedure (systematic and composite sampling) recommended below, the operating parameters of incineration process for petroleum products listed below should be followed. All conditions of incineration should be at least the equivalent in effectiveness. Before it is certain for the equivalent effectiveness for any different incineration, the method in determining a sample number described in section 6B.3 should be used.

The operating requirements of the incineration process are: 1) the physical size of soil aggregate should not be larger than 4 inches at the entry of the incinerator and at least 90% of soil should be less than 2 mm in size at the exit of the incinerator; 2) the burning temperature should be no less than 1,000°F; 3) the duration of burning at 1,000°F or higher should be no less than 5 minutes; and 4) the temperature of soil at the exit of incinerator should be no less than 600°F. Furthermore, the temperature should be chart-recorded.

With the above operating conditions, the sample number can be determined as follows:

"For each project per day of operation, at least 5 properly arranged samples should be analyzed; 2 composites, 2 duplicates, and 1 coarse soil sample.

Step 1. 15 minutes after the start of the incineration, the first sample should be taken in duplicate, one of the duplicates to be combined with the next three consecutive samples to form a composite. Both samples, the composite and one of the duplicates, are to be sent to the laboratory for analysis. Be sure that the soil which did not meet the burning requirements set in the last paragraph is returned to the incinerator for further processing. A second duplicate sample should be performed in the same fashion as the first during the last third of the entire daily incineration for each project. The coarse soil sample is the soil with particle size larger than 3mm and should be taken from the first composite sample.

2. <u>When incineration is less than four hours</u>, collect a sample every 30 minutes thereafter, and combine 4 consecutive samples, including the one taken after 15 minutes, to form a composite for laboratory analysis if the daily incineration is 4 hours or shorter. Or

3. <u>When incineration is longer than four hours</u>, collect a sample every 30 minutes thereafter, and combine 5-6 consecutive samples, including the one taken after 15 minutes, to form a composite for laboratory analysis if the daily incineration is longer than 4 hours.

4. One additional duplicate and one additional coarse sample at a random time should be added to make at least a total of 7 samples for laboratory analysis when the incineration is 4 hours or longer in the daily operation."

The number of samples used to form each composite and the time of sampling should be properly recorded.

6B.6 FREQUENCY OF SOIL SAMPLING

Frequency of soil sampling depends on program objectives, sources and effects of pollution, pollutants of interest, rates of movement and degradation of the pollutant. The rates of movement and degradation of pollutants in soil are by nature associated with precipitation. Precipitation both on the site and upgradient of the site influences the transport of pollutants and aids in decomposition. Therefore, soil sampling frequency may be related to change over time, season, or precipitation, and evaluation on remedial process to provide information for decision-making.

6B.6.1 SOIL SAMPLING FREQUENCY FOR MONITORING CHANGES OVER TIME WITHOUT OR PRIOR TO TREATMENT

Since the major concerns about pollutants in soil are the potentiality that the pollutants will reach groundwater and when it will occur, the important factors are the amount of precipitation, the depth to the groundwater table, and the characteristics of the soil. Although there are different characteristics of water movement in different soils, the shortest time the pollutants may reach the groundwater is what we want to determine in the monitoring. Therefore, a soil sampling frequency can be based on environmental factors defined as:

$$\text{Environmental(E) Factor} = \left[\frac{\text{Depth } to \text{ Groundwater Level, feet}}{\text{Precipitation, inches/month}}\right] \left[\frac{\log(2\times10^{3}\times N)}{\log(HC_{inv.}\times N)}\right]^{1/2}$$

Where " 2×10^3 ", gal/day/ft^2, is the average hydraulic conductivity for sand;
$HC_{inv.}$ is the hydraulic conductivity, gal./day/ft^2, from the investigated area;
N is any numerical number which will make the term " $HC_{inv.}\times N$ " greater than 1, but less than 5, before taking logarithm if the HC happens to be less than 1. Example, saying $HC_{inv.}$ happens to be 1.2×10^{-1}, $N=2\times10^{1}$ can be used to make $HC_{inv.}\times N = 2.4$, which is between 1 and 5. Let $N = 1$ if the $HC_{inv.}$ happens to be greater than 1.

1. For the first year, if the E factor falls below one or a major rainfall makes the factor smaller than one, the sampling of soil should be done monthly. If the factor happens to be larger than one, such as 1.5 or 2.1, for example, the sampling schedule can be performed at the interval of whatever the number is, such as one and a half months for 1.5 or two months for 2.1.

2. For the following years, the sampling may be done quarterly for the second year; semi-annually for the next two to three years; then annually thereafter. In any case, the depth to groundwater, precipitation, and possible impact on resources should be thoroughly evaluated and acted upon accordingly.

6B.6.2 SOIL SAMPLING FREQUENCY FOR MONITORING THE EFFECTS OF REMEDIATION

The following techniques can be applied to any soil treatment, most commonly ex-situ bioremediation:

1. At the beginning, perform monthly sampling, or perform the sampling at a reasonable period until at least three consecutive samples are taken. The decision of the interval on the initial sampling should be based on and determined by the data and progress from past pilot experiment or field work. The analytical data from any two consecutive samples at a proper period should be statistically differentiable.

2. When the first three consecutive samples at a proper time interval and their analytical results are obtained, establish a concentration-reduction chart with the concentration as the Y-axis and time

(months or weeks) as the X-axis.

3. Plot the concentration of the first three consecutive samples against its corresponding time (months or weeks) of sampling on the chart, and draw a best-fit line through the three points.

4. Extend the best fit line of the three points to a point, called the 4th point, where its concentration is one-half of the concentration of its immediate preceding sample point. From the 4th point, trace vertically down the line and find the time on the X-axis for the next sampling.

5. Once the additional information of the concentration from the 4th point is obtained, plot it with the first 3 data points into a new chart and draw a best fit line through all four points.

6. Extend the best fit line of the 4 data points to a point, the 5th point, where its concentration is one-half of the concentration of its immediate preceding sample, the 4th sample point this case. From the 5th point, find the time interval for the next sampling schedule as in Step 4.

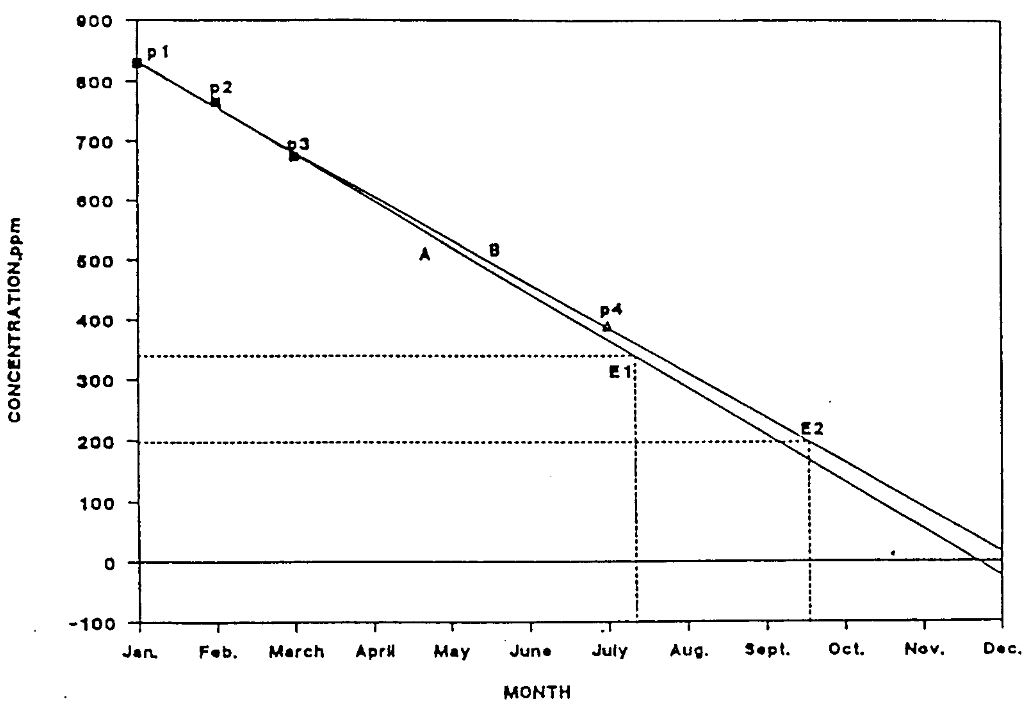
7. Repeat the above procedures to determine the next sampling schedules until the cleanup criteria are met.

If the sample points become a curve instead of a straight line, use a semilog chart.

<u>EXAMPLE</u> (HYPOTHETICAL)

STEP 1. Assume that five samples each time were taken from a site for 3 consecutive months in a bioremediation project. The samples were taken on the 5th day of each month. Their average concentration of analytical results each time were: 830 ppm for January, 765 ppm for Februaιy, and 674 ppm for March for the total organic contaminants.

Figure 6-1 Sampling Schedule



2. Establish a concentration (ppm) versus time (month) chart, as in Figure 6-1, with the concentration as the Y-axis and the time as the X-axis.

3. Plot the concentration of the above 3 data points into the chart as follows: Point P1 for 830 ppm at the location of January, P2 for 765 ppm at the location of February, P3 for 674 ppm at the location of March. Draw a best fit line, A, through the three points, P1, P2 and P3.

4. Extend the best fit line A to a point, E1, where its concentration is one-half of the concentration of its immediate preceding sample. P3 is its immediate preceding sample in this case, and one-half of the concentration at P3 is (674/2) 337 ppm. From the point E1, trace vertically down to the X-axis and find the 4th sampling time to be in July. Therefore, the fourth sampling event can be done in July. Assume it was done on July 5 and its analytical result was 390 ppm.

5. Plot the result, 390 ppm, into a chart as the 4th point P4 accordingly and along with the previous sample results, P1, P2, P3. Draw a best-fit line, B, through all 4 points, P1, P2, P3 and P4.

6. Similar to Step 4, extend the best fit line B to a point, E2, where its concentration is one-half of the concentration of the last sample, 390 ppm. The concentration at E2 is 195 ppm (390/2). Trace vertically down to the X-axis from the point E2 and find the next sampling schedule to be in September.

7. Repeat the above procedures to determine the next sampling schedule until the cleaning criteria are met.

APPENDIX 6X.A
GLOSSARY OF STATISTICAL TERMS

1	**Alpha (α)**	In the context of a statistical test, a greek letter is used to denote the significance level or probability of a Type I error.
2	**Alpha-error**	Sometimes used for Type I error.
3	**Alternative hypothesis**	An alternative hypothesis specifies that the underlying distribution differs from the null hypothesis. The alternative hypothesis usually specifies the value of a parameter (for example, the mean concentration) that one is trying to detect.
4	**Beta (β)**	In the context of a statistical test, β is the probability of a Type II error.
5	**Binomial Distribution**	A probability distribution used to describe the number of occurrences of a specified event in n independent trials. The binomial distribution can be used to develop statistical tests concerned with testing the proportion of any sample units in a simple random sampling having excessive concentrations of a contaminant.
6	**Coefficient of Variation** ($CV = s/\bar{x}$)	The ratio of the standard deviation to the mean for a set of data or distribution, abbreviated CV. For data that can only have positive values, such as concentration measurements, the coefficient of variation provides a crude measure of skewness.
7	**Confidence Coefficient**	The confidence coefficient of a confidence interval for a parameter is the probability that the random interval constructed from the sample data contains the true value of the parameter. The confidence coefficient is related to the significance level of an associated hypothesis test by the fact that the significance level (in percent) is one hundred minus the confidence coefficient (in percent).
8	**Confidence Interval**	A confidence interval, not a single value, for a parameter is a random interval constructed from sample data in such a way that the probability that the interval will contain the true value of the parameter is a specified value.
9	**Confidence Level**	The degree of confidence associated with an interval estimate. For example, with a 95% confidence interval, we would be 95% certain that the interval contains the true value being estimated. The confidence level is equal to 1 minus the Type I error ($1-\alpha$) (false Positive rate).
10	**Conservative Test**	A statistical test for which the Type I error rate (false positive rate) is actually less than that specified for the test. For a conservative test there will be a greater tendency to accept the null hypothesis when it is not true than for a non-conservative test.
11	**Cumulative Distribution Function**	Distribution function.
12	**Distribution**	The frequencies (either relative or absolute) with which measurements in a data set fall within specified classes. A graphical display of a

distribution is referred to as a histogram.

13	**Distribution-free**	This is sometimes used as a synonym for nonparametric. A statistic is distribution-free if its distribution does not depend upon which specific distribution function (in a large class) the observations follow.
14	**Distribution Function**	The distribution function for a random variable, x, is a function that specifies the probability that x is less than or equal to t, for all real values of t.
15	**Estimate**	Any numerical quantity computed from a sample of data. For example, a sample mean is an estimate of the corresponding population mean.
16	**Experimentwise Error Rate**	This term refers to multiple comparisons. If a total of n decisions are made about comparisons (for example of compliance wells to background wells) and x of the decisions are wrong, then the experimentwise error rate is x/n.
17	**False Positive Rate**	The probability of mistakenly concluding that the sample area is clean when it is dirty. It is the probability of making a Type I error (α).
18	**False Negative Rate**	The probability of mistakenly concluding that the sample area is dirty when it is clean. It is the probability of making a Type II error (β).
19	**Geostatistics**	A methodology for the analysis of spatially correlated data. The characteristic feature is the use of variograms or related techniques to quantify and model the spatial correlation structure. Also includes the various techniques such as kriging, which utilize spatial correlation models.
20	**Histogram**	A graphical display of a frequency distribution.
21	**Hot Spot**	Localized elliptical areas with concentrations in excess of the cleanup standard, either a volume defined by the projection of the surface area through the soil zone that will be sampled or a discrete horizon within the soil zone that will be sampled.
22	**Hypothesis**	An assumption about a property or characteristic of a population under study. The goal of statistical inference is to decide which of two complementary hypothesis is likely to be true. In the context of this guidance document, the null hypothesis is that the sample area is "dirty" and the alternative hypothesis is that the sample area is "clean".
23	**Independence**	A set of events are independent if the probability of the joint occurrence of any subset of the events factors into the product of the probabilities of the events. A set of observations is independent if the joint distribution function of the random errors associated with the observations factors into the product of the distribution functions.
24	**Inference**	The process of generalizing (extrapolating) results from a sample to a larger population.
25	**Judgment Sample**	A sample of data selected according to non-probabilistic methods.

26 **Kriging** A weighted-moving-average interpolation method where the set of weights assigned to samples minimizes the estimation variance, which is computed as a function of the variogram model and locations of the samples relative to each other, and to the point or block being estimated. This technique is used to model the contours of contamination levels at a polluted site.

27 **Less-Than-Detection-Limit** A concentration value that is below the detection limit. It is generally recommended that these values be included in the analysis as values at the detection limit.

28 **Lognormal Distribution** A family of positive-values, skewed distributions commonly used in environmental work.

29 **Mean** The arithmetic average of a set of data values. Specifically, the mean of a data set, $x_1, x_2, ..., x_n$, is defined by:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{1}$$

30 **Median** This is the middle value of a sample when the observations have been ordered from least to greatest. If the number of observations is odd, it is the middle observation. If the number of observations is even, it is customary to take the midpoint between the two middle observations. For a distribution, the median is a value such that the probability is one-half of that an observation will fall above or below the median.

31 **Multiple Comparison Procedure** This is a statistical procedure that makes a large number of decisions or comparisons on one set of data. For example, at a sampling period, several compliance well concentrations may be compared to the background well concentration.

32 **Nonparametric Test** A test based on relatively few assumptions about the underlying process generating the data. In particular, few assumptions are made about the exact form of the underlying probability distributions. As a consequence, nonparametric tests are valid for a fairly broad class of distributions.

33 **Normal Distribution** A family of "bell-shaped" distributions described by the mean and variance, μ and σ^2. Refer to a statistical text [(e.g., Sokal and Rohlf (1973)] for a formal definition. The standard normal distribution has $\mu = 0$ and $\sigma^2 = 1$.

34 **Null Hypothesis** A null hypothesis specifies the underlying distribution of the data completely. Often the null distribution specifies that, for example, there is no difference between the mean concentration in background well water samples and compliance well water samples.

35 **One-sided Test** A one-sided test is appropriate if concentrations higher than those specified by the null hypothesis are of concern. A one-sided test only rejects for differences that are large and in a prespecified direction.

36 **One-sided Tolerance Limit** This is an upper limit on observations from a specified distribution.

37 **One-sided Confidence Limit** This is an upper limit on a parameter of a distribution.

38	**Order Statistics**	The sample values observed after they have been arranged in increasing order.
39	**Ordinary Kriging**	A variety of kriging which assumes that local means are not necessarily closely related to the population mean, and which therefore uses only the samples in the local neighborhood for the estimate. Ordinary kriging is the most commonly used method for environmental situations.
40	**Outlier**	A measurement that is extremely large or small relative to the rest of the data gathered and that is suspected of misrepresenting the true concentration at the sample location.
41	**Parameter**	A statistical property of characteristic of a population of values. Statistical quantities such as means, standard deviations, percentiles, etc. are parameters if they refer to a population of values rather than to a sample of values.
42	**Parametric Test**	A test based on relatively strong assumptions about the underlying process generating the data. For example, most paremetric tests assume that the underlying data are normally distributed. As a consequence, parametric tests are not valid unless the underlying assumptions are met. See robust test.
43	**Percentile**	The specific value of a distribution that divides the set of measurements in such a way that P percent of the measurements fall below (or are equal to) this value, and 1-P percent of the measurements exceed this value. For specificity, a percentile is described by the value of P (expressed as a percentage). For example, the 95th percentile ($P=0.95$) is that value X such that 95 percent of the data have values less than X, and 5 percent have values exceeding X. By definition, the median is the 50th percentile.
44	**Point Estimate**	See estimate.
45	**Population**	The totality of the units of environmental media such as soil, water and air at a polluted site for which inferences regarding attainment of cleanup standards are to be made.
46	**Power**	The probability that a statistical test will result in rejecting the null hypothesis when the null hypothesis is false. Power $=1 - \beta$, where β is the Type II error rate associated with the test. The term "power function" is more accurate because it reflects the fact that power is a function of a particular value of the parameter of interest under the alternative hypothesis.
47	**Precision**	See standard error.
48	**Proportion**	The number of units of environmental media such as soil and water in a set of units of environmental media that have a specified characteristic, divided by the total number of units of environmental media in the set. This may also be expressed as a proportion of area or proportion of volume that has a specified characteristic.
49	**Random Sample**	A sample of units of environmental media such as soil and water selected using the simple random sampling procedures.

50 **Robust Test** — A statistical test that is approximately valid under a wide range of conditions.

51 **Sample Standard Deviation** — This is the square root of the sample variance.

52 **Sample Variance** — This is a statistic (computed on a sample of observations rather than on the whole population) that measures the variability or spread of the observations about the sample mean. It is the sum of the squared differences from the sample mean, divided by the number of observations less one.

53 **Serial Correlation** — This is the correlation of observations spaced a constant interval apart in a series. For example, the first order serial correlation is the correlation between adjacent observations. The first order serial correlation is found by correlating the pairs consisting of the first and second, second and third, third and fourth, etc., observations.

54 **Sequential Test** — A statistical test in which the decision to accept or reject the null hypothesis is made in a sequential fashion.

55 **Semi-variogram** — Identical to the term "variogram". There is a disagreement in the geostatistical literature as to which term should be used.

56 **Significance Level** — The probability of a Type I error associated with a statistical test. In the context of the statistical tests presented in this document, it is the probability that the sample area is declared to be clean when it is dirty. The significance level is often denoted by the symbol α (Greek letter alpha).

57 **Skewed Distribution** — Any nonsymmetric distribution.

58 **Standard Deviation** — A measure of dispersion of a set of data. Specifically, given a set of measurements, $x_1, x_2, \ldots, x_n$, the standard deviation is defined to be the quantity,

$$S = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

where $\bar{x}$ is the sample mean.

59 **Standard Error** — A measure of the variability (or precision) of a sample estimate. Standard errors are often used to construct confidence intervals.

60 **Statistical Test** — A formal statistical procedure and decision rule for deciding whether a sample area attains the specified cleanup standard.

61 **Stratified Sample** — A sample comprised of a number of separate samples from different strata.

62 **Stratum** — A subset of a sample area within which a random or systematic sample is selected. The primary purpose of creating strata for sampling is to improve the precision of the sample design.

63	**Symmetric Distribution**	A distribution of measurements for which the two sides of its overall shape are mirror images of each other about a center line.
64	**Systematic Sample**	A "grid" sample with a random start position.
65	**t-Test**	A statistical test to accept or reject a hypothesis by computing and comparing t value with the value of the hypothesis.
66	**Tolerance Interval**	A confidence interval around a percentile of a distribution of concentrations.
67	**Type I Error**	The error made when the sample area is declared to be clean when it is contaminated. This is also referred to as a false positive.
68	**Type II Error**	The error made when the sample area is declared to be dirty when it is clean. This is also referred to as a false negative.
69	**Variance**	The square of the standard deviation.
70	**Variogram**	A plot of the variance (one-half the mean squared difference) of paired sample measurements as a function of the distance (and optionally of the direction) between samples. Typically, all possible sample pairs are examined, distance and direction. Variograms provide a means of quantifying the commonly observed relationship that samples close together will tend to have more similar values than samples far apart.
71	**Z Value**	Percentage of a standard normal distribution.

APPENDIX 6X.B - DETERMINING THE SAMPLING LOCATIONS

This section discusses the techniques of how to determine the sampling locations after the sample number (n) is determined. In most cases, these procedures have to be repeated to obtain finer definition of the boundaries of the contaminated site or the different contaminated strata. The techniques can be modified and used under different situations other than the example given in this section.

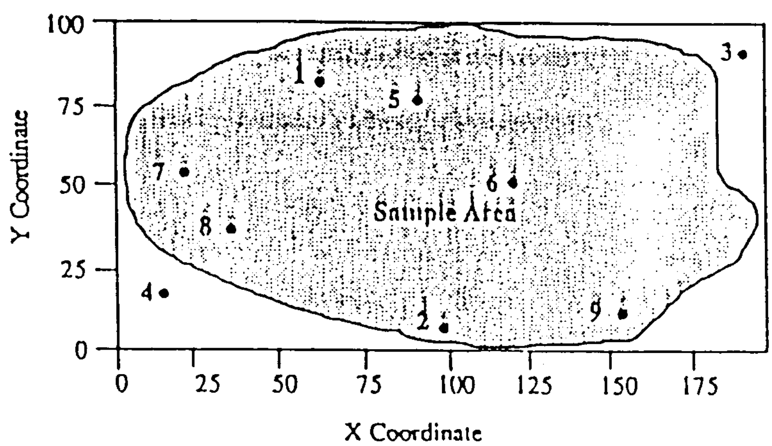
6X.1 SIMPLE RANDOM SAMPLING LOCATION

6X.1.1 Use X and Y (Z for Depth If Necessary) Coordinates

Once the number of samples is determined, their locations can be identified by X and Y coordinates within a grid system as shown in Figure 6-2. It is not necessary to draw a grid on the entire investigated area, but the actual coordinates selected need to be identified. One way of doing this is to overlay a map of the investigated area with a grid of an appropriate scale. The origin of the coordinate system can be any bench mark or landmark on the ground. The sampling should be randomly selected rather than located for convenience. Steps for generating random coordinates that define sampling locations are as follows:

Figure 6-2 Map of a Sample Area Showing Random Sampling Locations

Locations of the random samples are indicated by a "•". The numbers reference the XY pairs in Table 6-4.



1) Generate a set of coordinates (X, Y, Z) using the following equations:

$X_i = X_{min} + (X_{max} - X_{min}) \times RND$ (Equation 15)
$Y_i = Y_{min} + (Y_{max} - Y_{min}) \times RND$ (Equation 16)
$Z_i = Z_{min} + (Z_{max} - Z_{min}) \times RND$ (Equation 17)

Max = Maximum; Min = Minimum; i = 1, 2, 3 --- n
RND is the next unused random number between 0 and 1 in an sequence of random numbers.
Z can be the depth. More consideration on depth will be discussed in Appendix 6X.4. Random numbers can be obtained from calculators, computer software, or the random number table at the end of this section.

2) If (X, Y, Z) is outside the sample area, repeat step 1 to generate another random coordinate; otherwise go to Step 3.

3) Round all X_i, Y_i, Z_i to the nearest unit, in the same fashion, that can be implemented in the field.

4) Repeat steps 1 through 3 to generate the next random coordinate until the determined sample numbers, n, is reached.

EXAMPLE OF GENERATING RANDOM SAMPLING LOCATIONS

By following the steps explained above, sets of random numbers, (X_i, Y_i, Z_i), can be generated and used to identify each sample point. In this example seven sample points were selected without considering the depth (Z_i = 0), as shown in Table 6.4 and Figure 6-2. X will be measured on the map's coordinate system in the horizontal direction and Y in the vertical direction. The coordinate can use units of feet, meter or any appropriate unit. The X_i must be between 0 and 190 (i.e. X_{min} = 0 and X_{max} = 190) and the Y_i between 0 and 100 (i.e., Y_{min} = 0 and Y_{max} = 100) for this example. If the X, Y and Z coordinates for any set identify a point outside the area of interest, they are ignored and the process is continued until the sample size n has been achieved. It took nine attempts to secure seven coordinates that fall within the sample area.

Table 6-4 Example of Generating Random Sampling Locations

(X,Y,Z) pair	Random X Coordinate	Random Y Coordinate	Random Z Coordinate
1	67	80	0
2	97	4	0
3	190	88 (outside of sample area)	0
4	17	15 (outside of sample area)	0
5	94	76	0
6	123	49	0
7	25	52	0
8	35	39	0
9	152	14	0

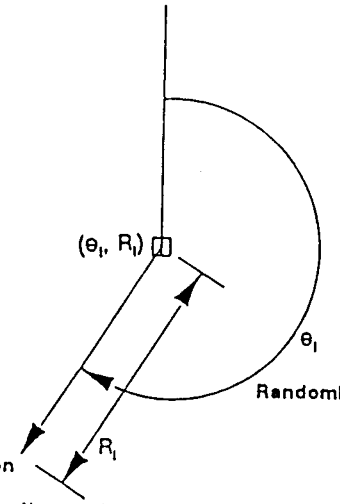
6X.1.2 USE CIRCULAR ANGLE, Θ, AND RADIUS R, (Z FOR DEPTH IF NECESSARY)

The approach of this technique is essentially similar to the one in Appendix 6X.1.1, but it uses a different coordinate system. The angle, Θ, and radius, R, of the geometric circle are used to identify the sample point instead of X and Y coordinates as indicated in Figure 6-3. The point or reference point for starting the process can be any landmark or benchmark. When it happens to be in the sample area, the angle, Θ, can be a randomly chosen angle between 0° and 360°, and the radius can be a randomly selected distance within the sampling location. When the reference or pivot point is located outside of the sampling area, the angle, Θ, can be any set range of between 0° - 90°, 0° - 180°, or 0° - 270°, depending on the position of the reference point and the shape of the sampling area.

The steps of securing the pair point are the same as that explained in Appendix 6X.1.1, but the coordinates of X and Y are changed to Θ and R. No direct example would be needed here.

Another similar approach is to assign a number to the block located in the sampling area as the one

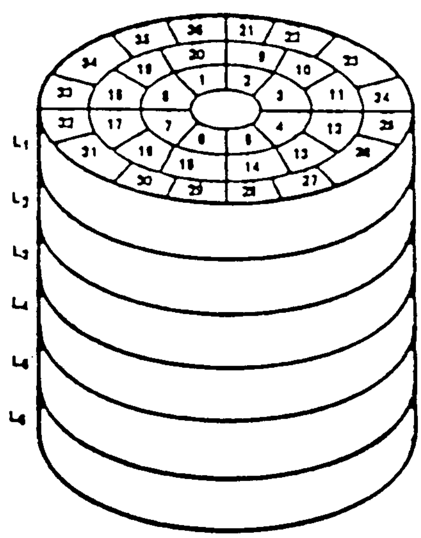
Figure 6-3 An example of how to choose a field sampling locations



The random numbers can be generated in the field using a hand-held calculator or by generating the random numbers prior to sampling. The sample should be collected as close to this sampling location as possible.

shown in Figure 6-4, then find the sample locations by using a random number table or computer. The problem with this approach is that the location may not be as exact as the other techniques; therefore, bias may occur.

Figure 6-4 A column like a drum or tank divided into a three-dimensional grid



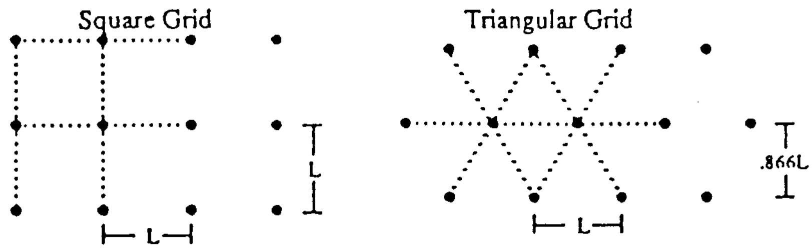
6X.2 SYSTEMATIC RANDOM SAMPLING LOCATION

6X.2.1 USE SQUARE OR TRIANGULAR GRID PATTERN METHOD

Two common patterns used in the systematic sampling location are square and triangular grids as shown in Figure 6-5. It is noted that the rows of points in the triangular grid are closer (Lxsin 60° = 0.866L) than the distance between points in a row (L) and that the points in every other row are offset by half a grid width.

The distance, L, between the sampling locations in the systematic grid is calculated from the size of the sampled or investigated area, A, and the number of samples, n, which is determined in Sections 6B.1. For the

Figure 6-5 Examples of a Square and a Triangular grid for Systematic Sampling



square grid,

$$L = \sqrt{A/n} \quad \text{(equation 18)}$$

for the triangular grid

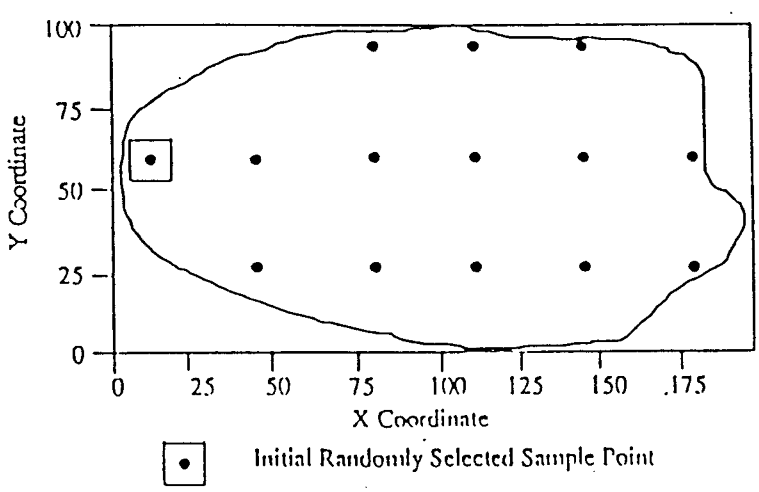
$$L = \sqrt{A/0.866n} \quad \text{(equation 19)}$$

The sampled area, A, can be measured on a map by using a planimeter. The units of the area are square feet, square meters or square yards, whichever is convenient and practical in the field.

After computing L, the actual location of the first point in the grid should be chosen by a random procedure as explained in section Appendix 6X.1.1. Using this location as one intersection of two gridlines, construct gridlines running parallel to the coordinate axes and separated by a distance L. The sampling locations are the points at the intersections of the gridlines that are within the samples area boundaries. The grid intersections that lie outside the area are not used.

There will be some variation in sample size, depending on the location of the initial randomly drawn point. However, the relative variation in number of sample points becomes small as the number of desired or calculated sample points increases. For unusually shaped sample areas or strata, the number of sample points can vary considerably from the desired number. In that case, there may be a chance to reduce the sample number by repeating the process from the beginning. Or the number can be accepted as is. Figure 6-6 is an example of this technique. The same process can also apply to the use of the triangular grid.

Figure 6-6 Map of a Sample Site Showing Systematic Sampling Location

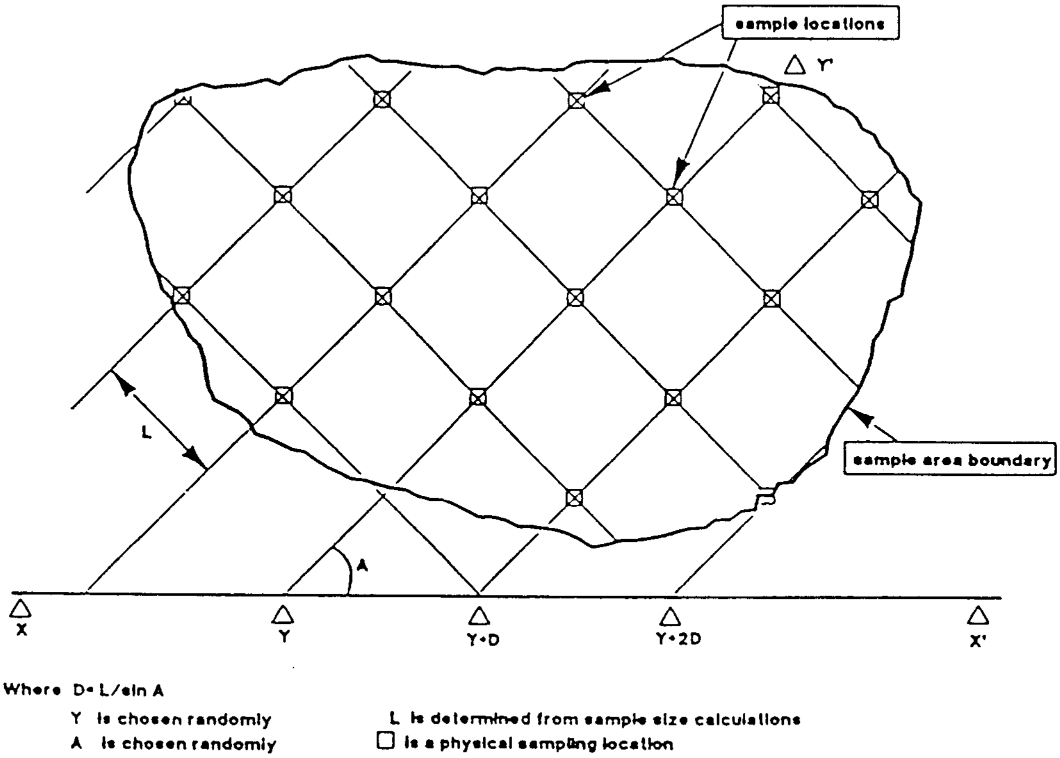


6X2.2 AN ALTERNATIVE METHOD FOR LOCATING THE RANDOM START POSITION FOR A SYSTEMATIC SAMPLING

This approach is useful under circumstances where a transit and stadia rod are available for turning angles, measuring distances, and establishing transects. The method is essentially equivalent to the method described above. The steps are:

1) Establish the main transect with endpoints X and X' using any convenient reference line (e.g. established boundary). The transect X - X' must be large enough, such as the line indicated in Figure 6-7, in order to site all of the transects that intersect the sample area.

Figure 6-7 Method for positioning systematic sample locations in the field



2) Randomly select a point Y on the transect between X and X'.
3) Randomly select an angle A between 0° and 90°.
4) Locate and mark transect with endpoints Y and Y', (A) degrees from transect X and X' which intersects the boundary of the sampled area.
5) Move away from point Y on transect X - X' a distance D, where D = L/sin (A). L is the desired interval between sampling points along the grid pattern.
6) At the point D units away from Y on transect X - X', establish two more transects: One (A) degrees from transect X - X' and parallel to transect Y - Y', and the other [90 + (A)] degrees from X - X' also beginning at the point D units from point Y.
7) Continue to move intervals of distance D along the transect X - X' until two transects intersect within the boundary of the sampled area. Establish the sample location at that point. Repeat these steps until all sample points are established.

6X3 STRATIFIED RANDOM SAMPLING LOCATION

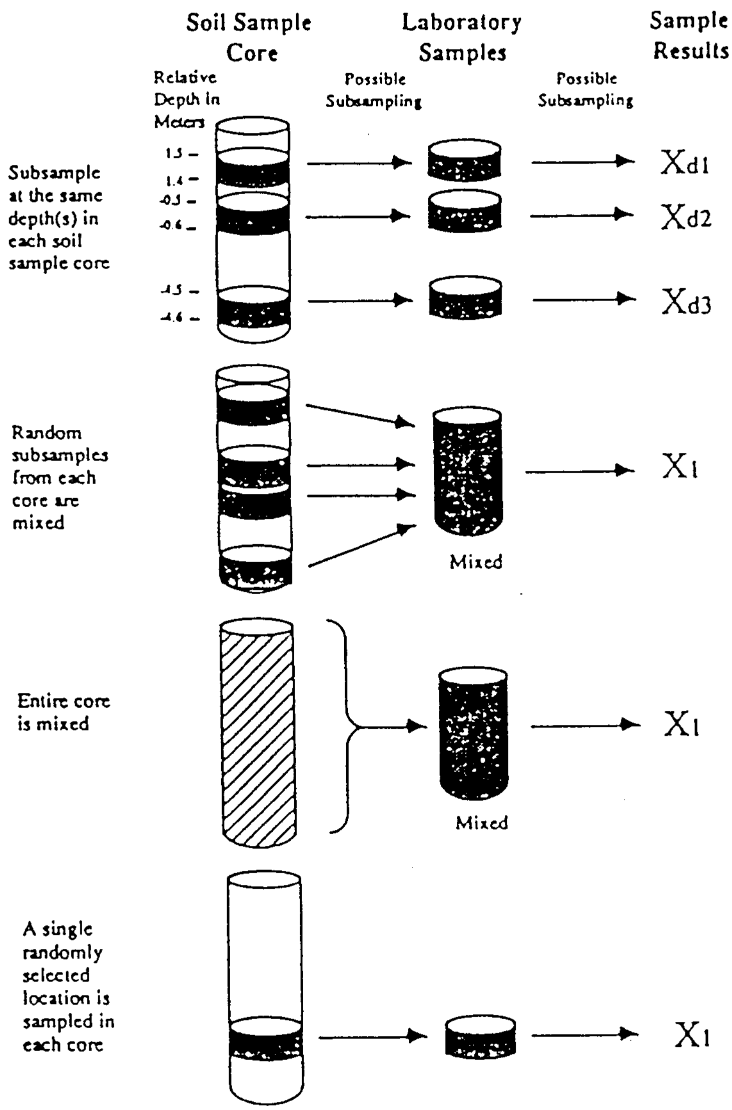
The above procedures can be extended to stratified sampling, and each stratum is sampled separately.

Different random numbers for locating the grids should be used in each stratum within the sampled area. The sampling approach chosen for one stratum does not have to be used in another stratum. For example, if a sample area is made up of a small waste pile and a large 200 acre hillside, then it would be possible to use systematic sampling for the hillside and random sampling for the waste pile.

6X.4 SAMPLING or SUB-SAMPLING/GRAB or COMPOSITE SAMPLING ACROSS DEPTH

Deciding how and where to perform grab or composite sampling of a soil core is an important step. The methods chosen should be executed consistently throughout the site, and they will depend on many things, including the soil sampling device, the quantity of material needed for analysis, the contaminants that are present, the consistency of the solid or soil media that is being sampled, and the objectives. There are several methods explained in Figure 6-8 for soil acquisition across depth once an exact augering or coring position has been determined. Be reminded that some of these methods, such as composite sampling, will not be used for volatile pollutants when a true representation of contamination is needed for the sample point.

Figure 6-8 Subsampling and Sampling Across Depth

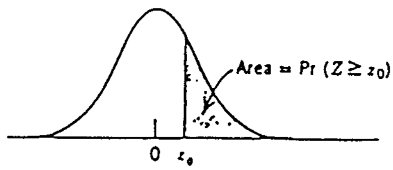


The first approach is to decide before sampling on an exact depth or depths that will be retained for analysis. The size of the interval would depend on the purpose and the volume required by the laboratory. The next step is to determine whether grab or depth discrete sampling and the composite sampling will be used. In the grab or depth discrete sampling, various depths or elevations, relative to a geodetic or site standard elevation, can be selected. Advantages of this approach are that each depth or elevation can be considered a different sample area and conclusions regarding the attainment of cleanup standards can be made independently for each soil horizon. This is also a preferred method when the presence of volatiles in soil media prevents the use of compositing methods.

Composite sampling can cover more area of soil, which would result in lowering sampling cost. Composite sampling should generally be used with caution unless the statistical parameter of interest is the mean concentration, because the variance of the mean from composite samples will be lower than the same variance associated with the mean from non-composite samples. However, composite sampling will restrict the evaluation of the proportion of soil above an established cleanup standard due to the physical mixing that takes place in the process.

Like the different sampling techniques, further selection of the segment or segments of a soil core to make up an ultimate sample can be done randomly, systematically or judgmentally.

Table 6-3
Standard Normal, Cumulative Probability in Right-Hand Tail
(For Negative Values of z, Areas are Found by Symmetry)



	Second Decimal Place of z_0									
$\downarrow z_0$ $\rightarrow$	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.00135									
3.5	.000 233									
4.0	.000 031 7	To interpolate carefully, see Table X.								
4.5	.000 003 40									
5.0	.000 000 287									

TABLE 6-5 NUMBER OF SAMPLES REQUIRED IN A ONE-SIDED ONE-SAMPLE t-TEST TO ACHIEVE A MINIMUM DETECTABLE RELATIVE DIFFERENCE AT CONFIDENCE LEVEL (1-α) AND POWER OF (1-ß)

Coefficient of Variation (%) CV=100 s/x̄	Power (%) 1-ß	Confidence Level (%) 1-α	Minimum Detectable Relative Difference (%) 5	10	20	30	40
10	95	99	66	19	7	5	*
		95	45	13	5	*	*
		90	36	10	*	*	*
		80	26	7	*	*	*
	90	99	55	16	6	5	*
		95	36	10	*	*	*
		90	28	8	*	*	*
		80	19	5	*	*	*
	80	99	43	13	6	*	*
		95	27	8	*	*	*
		90	19	6	*	*	*
		80	12	*	*	*	*
15	95	99	145	39	12	7	5
		95	99	26	8	5	*
		90	78	21	6	*	*
		80	57	15	*	*	*
	90	99	120	32	11	6	5
		95	79	21	7	*	*
		90	60	16	5	*	*
		80	41	11	*	*	*
	80	99	94	26	9	6	5
		95	58	16	5	*	*
		90	42	11	*	*	*
		80	26	7	*	*	*
20	95	99	256	66	19	10	7
		95	175	45	13	9	5
		90	138	36	10	5	*
		80	100	26	7	*	*
	90	99	211	55	16	9	6
		95	139	36	10	6	*
		90	107	28	8	*	*
		80	73	19	5	*	*
	80	99	164	43	13	8	6
		95	101	27	8	5	*
		90	73	19	6	*	*
		80	46	12	*	*	*

TABLE 6-5 (Continued)

Coefficient of Variation (%) CV=100 s/ $\bar{x}$	Power (%) 1-β	Confidence Level (%) 1-α	Minimum Detectable Relative Difference (%) 5	10	20	30	40
25	95	99	397	102	28	14	9
		95	272	69	19	9	6
		90	216	55	15	7	5
		80	155	40	11	5	*
	90	99	329	85	24	12	8
		95	272	70	19	9	6
		90	166	42	12	6	*
		80	114	29	8	*	*
	80	99	254	66	19	10	7
		95	156	41	12	6	*
		90	114	30	8	*	*
		80	72	19	5	*	*
30	95	99	571	145	39	19	12
		95	391	99	26	13	8
		90	310	78	21	10	6
		80	223	57	15	7	*
	90	99	472	120	32	16	11
		95	310	79	21	10	7
		90	238	61	16	8	5
		80	163	41	11	5	*
	80	99	364	84	26	13	9
		95	224	58	16	8	5
		90	164	42	11	6	*
		80	103	26	7	*	*
35	95	99	775	196	42	25	15
		95	532	134	35	17	10
		90	421	106	28	13	8
		80	304	77	20	9	6
	90	99	641	163	43	21	13
		95	421	107	28	14	8
		90	323	82	21	10	6
		80	222	56	15	7	*
	80	99	495	126	34	17	11
		95	305	78	21	10	7
		90	222	57	15	7	5
		80	140	36	10	5	*

* - At least 5 samples when performing a one-sample t-test and 3 samples when using a two-sample t-test.

Table of Random Units

RANDOM UNITS

Use of Table. If anyone wishes to select a random sample of N items from a universe of M items, the following procedure may be applied. (M > N)

1) Decide upon some arbitrary scheme of selecting entries from the table. For example, one may decide to use the entries in the first line, second column; second line, third column; third line, fourth column; etc..

2) Assign numbers to each of the items in the universe from 1 to M. Thus, If M = 500, the items would be numbered from 001 to 500, and therefore, each designated item is associated with a three digit number.

3) Decide upon some arbitrary scheme of selecting positional digits from each entry chosen according to Step 1. Thus, if M = 500, one may decide to use the first, third, and fourth digit of each entry selected, and as consequence a three digit number is created for each entry choice.

4) If the number formed is ≤ M, the correspondingly designated item in the unverse is chosen for the random sample of N items. If a number formed is > M or is a repeated number of one already chosen, it is passed over and the next desirable number is taken. This process is continued until the random of N items is selected.

Random numbers

Line	Column 1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	10480	15011	01536	02011	81647	91646	69179	14194	62590	36207	20969	99570	91291	90700
2	22368	46573	25595	85393	30995	89198	27982	53402	93965	34095	52666	19174	39615	99505
3	24130	48360	22527	97265	76393	64809	15179	24830	49340	32081	30680	19655	63348	58629
4	42167	93093	06243	61680	07856	16376	39440	53537	71341	57004	00849	74917	97758	16379
5	37570	39975	81837	16656	06121	91782	60468	81305	49684	60672	14110	06927	01263	54613
6	77921	06907	11008	42751	27756	53498	18602	70659	90655	15053	21916	81825	44394	42880
7	99562	72905	56420	69994	98872	31016	71194	18738	44013	48840	63213	21069	10634	12952
8	96301	91977	05463	07972	18876	20922	94595	56869	69014	60045	18425	84903	42508	32307
9	89579	14342	63661	10281	17453	18103	57740	84378	25331	12566	58678	44947	05585	56941
10	85475	36857	53342	53988	53060	59533	38867	62300	08158	17983	16439	11458	18593	64952
11	28918	69578	88231	33276	70997	79936	56865	05859	90106	31595	01547	85590	91610	78188
12	63553	40961	48235	03427	49626	69445	18663	72695	52180	20847	12234	90511	33703	90322
13	09429	93969	52636	92737	88974	33488	36320	17617	30015	08272	84115	27156	30613	74952
14	10365	61129	87529	85689	48237	52267	67689	93394	01511	26358	85104	20285	29975	89868
15	07119	97336	71048	08178	77233	13916	47564	81056	97735	85977	29372	74461	28551	90707
16	51085	12765	51821	51259	77452	16308	60756	92144	49442	53900	70960	63990	75601	40719
17	02368	21382	52404	60268	89368	19885	55322	44819	01188	65255	64835	44919	05944	55157
18	01011	54092	33362	94904	31273	04146	18594	29852	71585	85030	51132	01915	92747	64951
19	52162	53916	46369	58586	23216	14513	83149	98736	23495	64350	94738	17752	35156	35749
20	07056	97628	33787	09998	42698	06691	76988	13602	51851	46104	88916	19509	25625	58104
21	48663	91245	85828	14346	09172	30168	90229	04734	59193	22178	30421	61666	99904	32812
22	54164	58492	22421	74103	47070	25306	76468	26384	58151	06646	21524	15227	96909	44592
23	32639	32363	05597	24200	13363	38005	94342	28728	35806	06912	17012	64161	18296	22851
24	29334	27001	87637	87308	58731	00256	45834	15398	46557	41135	10367	07684	36188	18510
25	02488	33062	28834	07351	19731	92420	60952	61280	50001	67658	32586	86679	50720	94953
26	81525	72295	04839	96423	24878	82651	66566	14778	76797	14780	13300	87074	79666	95725
27	29676	20591	68086	26432	46901	20849	89768	81536	86645	12659	92259	57102	80428	25280
28	00742	57392	39064	66432	84673	40027	32832	61362	98947	96067	64760	64584	96096	98253
29	05366	04213	25669	26422	44407	44048	37937	63904	45766	66134	75470	66520	34693	90449
30	91921	26418	64117	94305	26766	25940	39972	22209	71500	64568	91402	42416	07844	69618
31	00582	04711	87917	77341	42206	35126	74087	99547	81817	42607	43808	76655	62028	76630
32	00725	69884	62797	56170	86324	88072	76222	36086	84637	93161	76038	65855	77919	88006
34	69011	65795	95876	55293	18988	27354	26575	08625	40801	59920	29841	80150	12777	48501
34	25976	57948	29888	88604	67917	48708	18912	82271	65424	69774	33611	54262	85963	03547
35	09763	83473	73577	12908	30883	18317	28290	35797	05998	41688	34952	37888	38917	88050
36	91567	42595	27958	30134	04024	86385	29880	99730	55536	84855	29080	09250	79656	73211
37	17955	56349	90999	49127	20044	59931	06115	20542	18059	02008	73708	83517	36103	42791
38	46503	18584	18845	49618	02304	51038	20655	58727	28168	15475	56942	53389	20562	87338
39	92157	89634	94824	78171	84610	82834	09922	25417	44137	48413	25555	21246	35509	20468
40	14577	62765	35605	81263	39667	47358	56873	56307	61607	49518	89656	20103	77490	18062
41	98427	07523	33362	64270	01638	92477	66969	98420	04880	45585	46565	04102	46880	45709
42	34914	63976	88720	82765	34476	17032	87589	40836	32427	70002	70663	88863	77775	69348
43	70060	28277	39475	46473	23219	53416	94970	25832	69975	94884	19661	72828	00102	66794
44	53976	54914	06990	67245	68350	82948	11398	42878	80287	88267	47363	46634	06541	97809
45	76072	29515	40980	07391	58745	25774	22987	80059	39911	96189	41151	14222	60697	59583
46	90725	52210	83974	29992	65831	38857	50490	83765	55657	14361	31720	57375	56228	41546
47	64364	67412	33339	31926	14883	24413	59744	92351	97473	89286	35931	04110	23726	51900
48	08962	00358	31662	25388	61642	34072	81249	35648	56891	69352	48373	45578	78547	81788
49	95012	68379	93526	70765	10592	04542	76463	54328	02349	17247	28865	14777	62730	92277
50	15664	10493	20492	38391	91132	21999	59516	81652	27195	48223	46751	22923	32261	85653
51	16408	81899	04153	53381	79401	21438	83035	92350	36693	31238	59649	91754	72772	02338

Random numbers (*Concluded*)

Line	1	2	3	4	5	6	7	8	9	10	11	12	13	14
52	18629	81953	05520	91962	04739	13092	97662	24822	94730	06496	35090	04822	86774	98289
53	73115	35101	47498	87637	99016	71060	88824	71013	18735	20286	23153	72924	35165	43040
54	57491	16703	23167	49323	45021	33132	12544	41035	80780	45393	44812	12515	98931	91202
55	30405	83946	23792	14422	15059	45799	22716	19792	09983	74353	68668	30429	70735	25499
56	16631	35006	85900	98275	32388	52390	16815	69298	82732	38480	73817	32523	41961	44437
57	96773	20206	42559	78985	05300	22164	24369	54224	35083	19687	11052	91491	60383	19746
58	38935	64202	14349	82674	66523	44133	00697	35552	35970	19124	63318	29686	03387	59846
59	31624	76384	17403	53363	44167	64486	64758	75366	76554	31601	12614	33072	60332	92325
60	78919	19474	23632	27889	47914	02584	37680	20801	72152	39339	34806	08930	85001	87820
61	03931	33309	57047	74211	63445	17361	62825	39908	05607	91284	68833	25570	38818	46920
62	74426	33278	43972	10119	89917	15665	52872	73823	73144	88662	88970	74492	51805	99378
63	09066	00903	20795	95452	92648	45454	09552	88815	16553	51125	79375	97596	16296	66092
64	42238	12426	87025	14267	20979	04508	64535	31355	86064	29472	47689	05974	52468	16834
65	16153	08002	26504	41744	81959	65642	74240	56302	00033	67107	77510	70625	28725	34191
66	21457	40742	29820	96783	29400	21840	15035	34537	33310	06116	95240	15957	16572	06004
67	21581	57802	02050	89728	17937	37621	47075	42080	97403	48626	68995	43805	33386	21597
68	55612	78095	83197	33732	05810	24813	86902	60397	16489	03264	88525	42786	05269	92532
69	44657	66999	99324	51281	84463	60563	79312	93454	68876	25471	93911	25650	12682	73572
70	91340	84979	46949	81973	37949	61023	43997	15263	80644	43942	89203	71795	99533	50501
71	91227	21199	31935	27022	84067	05462	35216	14486	29891	68607	41867	14951	91696	85065
72	50001	38140	66321	19924	72163	09538	12151	06878	91903	18749	34405	56087	82790	70925
73	65390	05224	72958	28609	81406	39147	25549	48542	42627	45233	57202	94617	23772	07896
74	27504	96131	83944	41575	10573	08619	64482	73923	36152	05184	94142	25299	84387	34925
75	37169	94851	39117	89632	00959	16487	65536	49071	39782	17095	02330	74301	00275	48280
76	11508	70225	51111	38351	19444	66499	71945	05422	13442	78675	84081	66938	93654	59894
77	37449	30362	06694	54690	04052	53115	62757	95348	78662	11163	81651	50245	34971	52924
78	46515	70331	85922	38329	57015	15765	97161	17869	45349	61796	66345	81073	49106	79860
79	30986	81223	42416	58353	21532	30502	32305	86482	05174	07901	54339	58861	74818	46942
80	63798	64995	46583	09785	44160	78128	83991	42865	92520	83531	80377	35909	81250	54238
81	82486	84846	99254	67632	43218	50076	21361	64816	51202	88124	41870	52689	51275	83556
82	21885	32906	92431	09060	64297	51674	64126	62570	26123	05155	59194	52799	28225	85762
83	60336	98782	07408	53458	13564	59089	26445	29789	85205	41001	12535	12133	14645	23541
84	43937	46891	24010	25560	86355	33941	25786	54990	71899	15475	95434	98227	21824	19585
85	97656	63175	89303	16275	07100	92063	21942	18611	47348	20203	18534	03862	78095	50136
86	03299	01221	05418	38982	55758	92237	26759	86367	21216	98442	08303	56613	91511	75928
87	79626	06486	03574	17668	07785	76020	79924	25651	83325	88428	85076	72811	22717	50585
88	85636	68335	47539	03129	65651	11977	02510	26113	99447	68645	34327	15152	55230	93448
89	18039	14367	61337	06177	12143	46609	32989	74014	64708	00533	35398	58408	13261	47908
90	08362	15656	60627	36478	65648	16764	53412	09013	07832	41574	17639	82163	60859	75567
91	79556	29068	04142	16268	15387	12856	66227	38358	22478	73373	88732	09443	82558	05250
92	92608	82674	27072	32534	17075	27698	98204	63863	11951	34648	88022	56148	34925	57031
93	23982	25835	40055	67006	12293	02753	14827	23235	35071	99704	37543	11601	35503	85171
94	09915	96306	05908	97901	28395	14186	00821	80703	70426	75647	76310	88717	37890	40129
95	59037	33300	26695	62247	69927	76123	50842	43834	86654	70959	79725	93872	28117	19233
96	42488	78077	69882	61657	34136	79180	97526	43092	04098	73571	80799	76536	71255	64239
97	46764	86273	63003	93017	31204	36692	40202	35275	57306	55543	53203	18098	47625	88684
98	03237	45430	55417	63282	90816	17349	88298	90183	36600	78406	06216	95787	42579	90730
99	86591	81482	52667	61582	14972	90053	89534	76036	49199	43716	97548	04379	46370	28672
100	38534	01715	94964	87288	65680	43772	39560	12918	86537	62738	19636	51132	25739	56947

Abridged from *Handbook of Tables for Probability and Statistics*, 2nd ed. Edited by William H. Beyer (Cleveland: The Chemical Rubber Company, 1968). Reproduced by permission of CRC Press, Inc.