Uncertainty

Despite the importance of the concept of uncertainty in analytical chemistry, questions and controversies still remain over the ease of its interpretation by legal and statutory bodies and non experts, and about the best methods of calculating it.

Two quantities are used to express uncertainty:

1) Standard uncertainty (u) expresses the concept as a standard deviation.
2) Expanded uncertainty (U) defines an interval containing a high proportion of the distribution of values that could reasonably be attributed to the measurand and is obtained by multiplying u by a coverage factor, k, chosen according to the degree of confidence required for the interval, i.e., $U = u \times k$.

Since u is analogous to a standard deviation, if $k = 2$ (this is generally taken as the default value if no other information is given) U gives approximately one-half of the 95% confidence interval.

In principle, two basic approaches are available to estimate uncertainty:

1) top-down
2) bottom-up

The bottom-up approach identifies each separate stage of an analysis, including sampling steps, whenever possible, assigns appropriate random and systematic errors to each, and then combines these components to give an overall u value.

This approach is not as simple as it would seem at a first glance.
The first problem is that even simple analytical processes may involve many individual experimental steps and possible sources of error.

Examples of error sources that should be considered but are easily overlooked include:

1) operator bias
2) instrument bias, including sample carry-over
3) reagent purity and stability
4) use of volumetric apparatus at a temperature different from the one at which it was calibrated
5) changes in the composition of the sample during the analysis, either because of contamination or because of inherent instability;
6) use of computers with inadequate capabilities or with the wrong statistical model applied.

All these factors may arise in addition to the random errors that inevitably occur in repeated measurements and may have to be estimated using experience, or equipment manufacturers' information, such as calibration certificates or instrument specifications.

Unfortunately, although they can be minimized by the use of standards and reference materials, systematic errors can be present and their combination with random errors to give an overall u value is not easy.

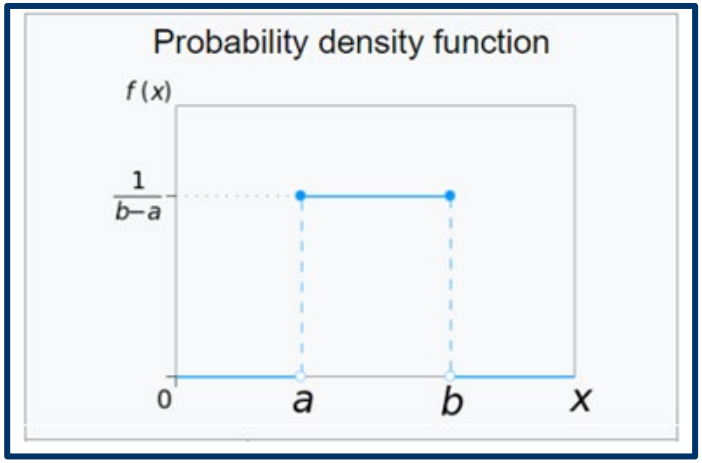
The usual method of tackling systematic errors is to treat them as coming from a rectangular distribution.

Suppose, for example, that a manufacturer quotes the purity of a reagent as 99.9 ± 0.1%. This means that the purity of the reagent in a single bottle is comprised between 99.8% and 100.0%. Consequently, there is no reason to suppose that the actual purity is closer to 99.9% than to any other value in the range 99.8–100.0%. In other words, the purity has a uniform distribution over this range.

In such cases, the contribution to the standard uncertainty is obtained by dividing the error by $3^{1/2}$, thus giving a value of 0.0577.

Indeed, as shown at the beginning of this course, the variance for a uniform distribution like the one shown in the figure on the right is $(b-a)^2/12$.

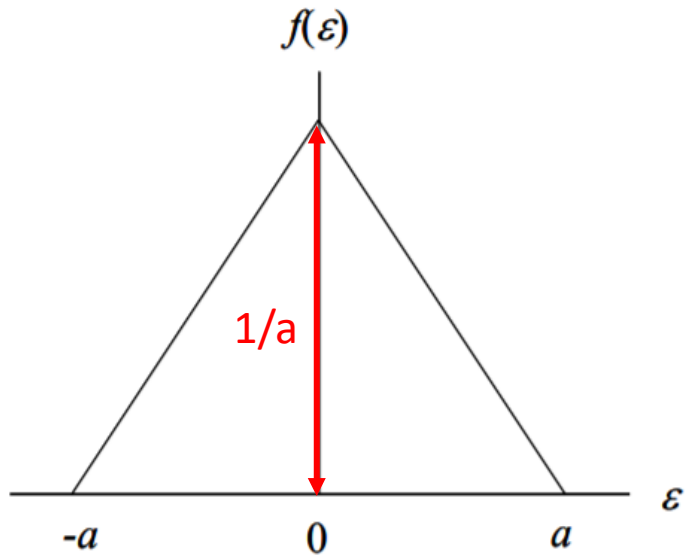
In the specific example, b-a = 0.2 %, thus the variance is $(0.2)^2/12$, i.e., $(\text{error} \times 2)^2/12$, and the standard deviation is: $(\text{error} \times 2)/(4 * 3)^{1/2} = \text{error}/3^{1/2}$.



Uncertainties of this kind, derived from the uncertainty stated by the manufacturer of a reagent, are referred to as type B uncertainties in the ISO Guide to the Expression of Uncertainty in Measurement.

Another example of type B uncertainty is the one depending on triangular distributions. The latter are adopted to describe the volume measured using a volumetric flask, when a trained person is involved. In fact, in this case the probability of filling the flask so that the liquid level is very near to the line, and thus the volume is close to the nominal one, is higher.

The general expression of the triangular probability density function, referred to the case in which it is centered on 0 and the error is $\pm$ a, is reported below:



$$f(\varepsilon)=\begin{cases}(a+\varepsilon)/a^2, & -a\le\varepsilon\le 0\\ (a-\varepsilon)/a^2, & 0\le\varepsilon\le a\\ 0, & \text{otherwise}\end{cases}$$

Note that if a <1, as it is usual in the present case, 1/a > a (e.g., a = 0.2 implies that 1/a = 5).

The variance can be calculated as follows, considering that the distribution mean is 0:

$$\sigma^2 = \int_{-a}^{a} f(\varepsilon)\varepsilon^2 d\varepsilon = \frac{1}{a^2}\int_{-a}^{0}(a+\varepsilon)\varepsilon^2 d\varepsilon + \frac{1}{a^2}\int_{0}^{a}(a-\varepsilon)\varepsilon^2 d\varepsilon =$$

$$= \frac{1}{a^2}\left(a\frac{\varepsilon^3}{3}+\frac{\varepsilon^4}{4}\right)_{-a}^{0} + \frac{1}{a^2}\left(a\frac{\varepsilon^3}{3}-\frac{\varepsilon^4}{4}\right)_{0}^{a} =$$

$$= \frac{1}{a^2}\left(\frac{a^4}{3}-\frac{a^4}{4}\right) + \frac{1}{a^2}\left(\frac{a^4}{3}-\frac{a^4}{4}\right) =$$

$$= 2a^2(1/3-1/4) = \frac{a^2}{6}$$

Since a is the error indicated on the flask, the standard deviation on the flask volume if a triangular distribution is considered is evaluated as the ratio between that error and $6^{1/2}$.

It is worth noting that random errors that can be calculated using the usual methods are called type A contributions, that have to be combined with type B contributions, in order to have an estimate of total uncertainty based on a bottom-up approach.

The following simplified example of a bottom-up uncertainty calculation shows some of these principles in operation.

Suppose we wish to determine the concentration of a solution of sodium hydroxide by titration with a standard acid such as potassium hydrogen phthalate (KHP).

The molar concentration of NaOH given by this experiment will depend on the volume of the NaOH solution used in the titration, and the mass, purity and molecular weight of the KHP.

The uncertainty in the molecular weight of the acid can be computed from the atomic weights table of the International Union of Pure and Applied Chemistry. It would be treated as a type B uncertainty, but it is so small that it is negligible for most practical purposes.

The mass of the KHP used would almost certainly be determined by difference, i.e., by weighing a container with the KHP in it, then weighing the container after the KHP has been removed for dissolution.

Each of these weightings would have an uncertainty derived as a type B estimate from the calibration certificate of the balance used. If the certificate indicated a balance error of ± 0.2 mg the uncertainty in each weighing would be $0.2/3^{1/2} = 0.1155$ mg.
The overall uncertainty in the weighing stage is then determined using the usual approach of error propagation: $[(0.1155)^2 + (0.1155)^2]^{1/2} = 0.1633$ mg.

The contribution to the overall uncertainty due to the uncertainty in the purity of the KHP is another type B estimate, again obtained by dividing the fractional impurity level by $3^{1/2}$.

The uncertainty contribution from the volume of NaOH used will have several sources, including a temperature effect (i.e., is the glassware used at the temperature at which it was calibrated?) and the calibration uncertainty of the burette (often assumed to derive from a triangular distribution), if it was used to measure the volume.

Finally, replicate titrations will show the extent of random errors during the analysis.
Although most attention will in practice be given to the major contributors to the uncertainty, it is clear that even in a simple titrimetric analysis a full uncertainty estimate requires much care.

A further problem is that the rules usually adopted for combining errors assume that they are independent of each other.

In reality, it seems quite possible that this is not always true. For example, laboratory temperature fluctuations might have several effects, such as altering the capacity of volumetric apparatus, causing sample losses through volatility, affecting the sensitivity of optical or electrochemical detectors, and so on.

Since all these errors would arise from a single source, they would be correlated, and could not be combined using the simple formula for independent errors.

In particular, the covariance between different sources should be considered in this case, as generally expressed by the following formula:

$$u\left[y\left(x_i, x_{j,\ldots}, x_n\right)\right] = \sqrt{\sum_{i=1,n} {c_i}^2 u(x_i)^2 + \sum_{\substack{i,k=1,n \\ i \neq k}} c_i c_k u(x_i, x_k)}$$

where x_i are the different sources of uncertainty, $u(x_i, x_k)$ corresponds to the covariance between sources i and k and each coefficient c_i is given by:

$$c_i = \partial y / \partial x_i$$

Numerical examples

1. A method requires 100 mg of an internal standard to be weighed out on a four-figure balance. The balance calibration certificate gives a standard uncertainty of ±0.0002 g, excluding weighing-to-weighing precision.

   Replicate weighings of a 100 mg check weight on the four-figure balance had a standard deviation of 0.000071 g, which can be used directly as an estimate of the standard uncertainty associated with weighing precision. There are thus two uncertainties, both directly affecting the observed weight and arising from additive effects.
   The standard uncertainty in the result of the weighing is therefore:

$$\sqrt{0.0002^2 + 0.000071^2} = 0.00021 \text{ g}$$

2. The concentration c of a (nominally) 1000 mg L^{-1} solution of an organic analyte is calculated from the mass m of the analyte and the volume v (expressed in mL) of the volumetric flask used to prepare the standard solution.

The concentration is, then: c = 1000 m/v. The mass was 100.4 mg with standard uncertainty 0.21 mg; the volumetric flask has a volume of 100 mL with standard uncertainty 0.16 mL.

The equation for c is entirely multiplicative, so, based on the general rule on error propagation, the uncertainties can be easily combined as relative uncertainties, giving:

$$\frac{u(c)}{c} = \sqrt{\left(\frac{0.21}{100.4}\right)^2 + \left(\frac{0.16}{100}\right)^2} = 0.0026$$

Since c = 1000 x 100.4/100 = 1004 mg L^{-1}, the standard uncertainty is 1004 × 0.0026 = 2.6 mg L^{-1} .

Although it can be very time-consuming, the bottom-up approach is the most common approach to uncertainty estimate, especially when legal or regulatory issues arise, and an estimate must be provided for each disputed sample.

A completely different approach is the top-down method, which makes use of the results of proficiency testing (PT) schemes in a number of laboratories to give estimates of the overall uncertainties.

The method is only applicable in areas where data from properly run proficiency schemes are available, though such schemes are rapidly expanding in number and may thus provide a real alternative to bottom-up methods in many fields.

It is worth noting that PT schemes use a variety of analytical methods, so it might reasonably be claimed that the uncertainty of results from a laboratory that has long experience on a single method might be better (smaller) than the one which PT results from many laboratories would suggest.

Moreover, PT schemes utilize single sample materials prepared with great care and tested for their homogeneity.
Some sampling errors that would occur in a genuine analysis might thus be overlooked.

In order to overcome problems related to the described approach, regulatory bodies have proposed simpler methods, explicitly designed to minimize the workload in laboratories that use a range of analytical procedures.

The basic principles of these methods are:

1. Systematic errors are not included in the uncertainty estimates, but are assessed using reference materials, and thus corrected or eliminated.

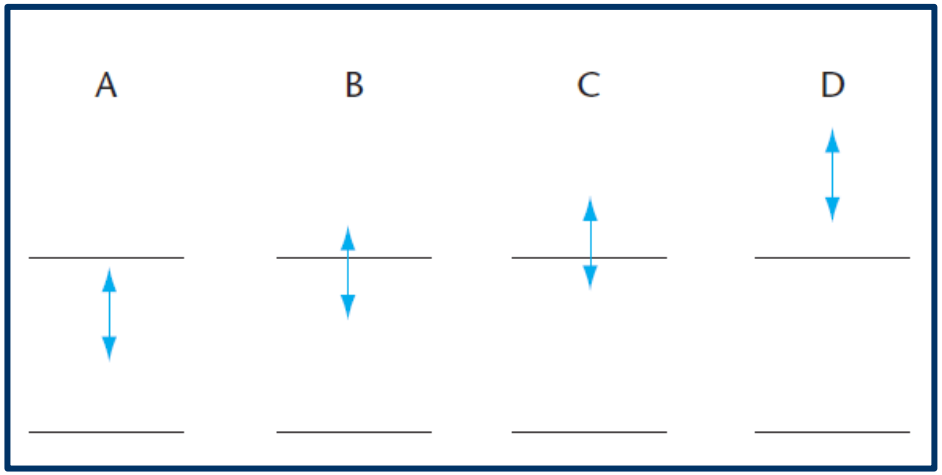
2. At least 10 replicate measurements are made on stable and well-characterized authentic samples or reference materials.

3. Uncertainties are calculated from the standard deviations of measurements made in internal reproducibility conditions, i.e., with different analysts, using different concentrations (including all those relevant to legal requirements), and in all relevant matrices.

Attention is paid, in this case, to special, intrinsically unstable samples, for which reproducibility conditions might be difficult to achieve.

Uncertainty estimates are important in demonstrating that a laboratory has the capacity to perform analyses of legal or statutory significance.
Once an uncertainty value for a particular analysis in a given laboratory is known, it is easy to interpret the results in relation to such statutory limits.

Four possible situations are shown in the following figure, where it is assumed that a coverage factor of 2 has been used to determine U at the 95% level (the 95% interval is shown by the vertical double arrows), and where both upper and lower limits for the concentration of the analyte are indicated by the horizontal lines:



In case A the uncertainty interval lies completely between the upper and lower specified limits, so compliance with the specification has been achieved. In case B the 95% interval extends just beyond the upper limit, so although compliance is more likely than not, it cannot be fully verified at the 95% level. In case C compliance is very unlikely, though not impossible, and in case D there is a clear failure to comply.

Sampling

In some occasions even a carefully designed analytical method may give poor results because errors associated with the sample have not been accounted for properly.

Indeed, a critical step is often represented by the need to obtain individual samples that accurately represent the whole population from which they are drawn, i.e., the target population. If this does not occur, even a careful analysis may yield an inaccurate result.

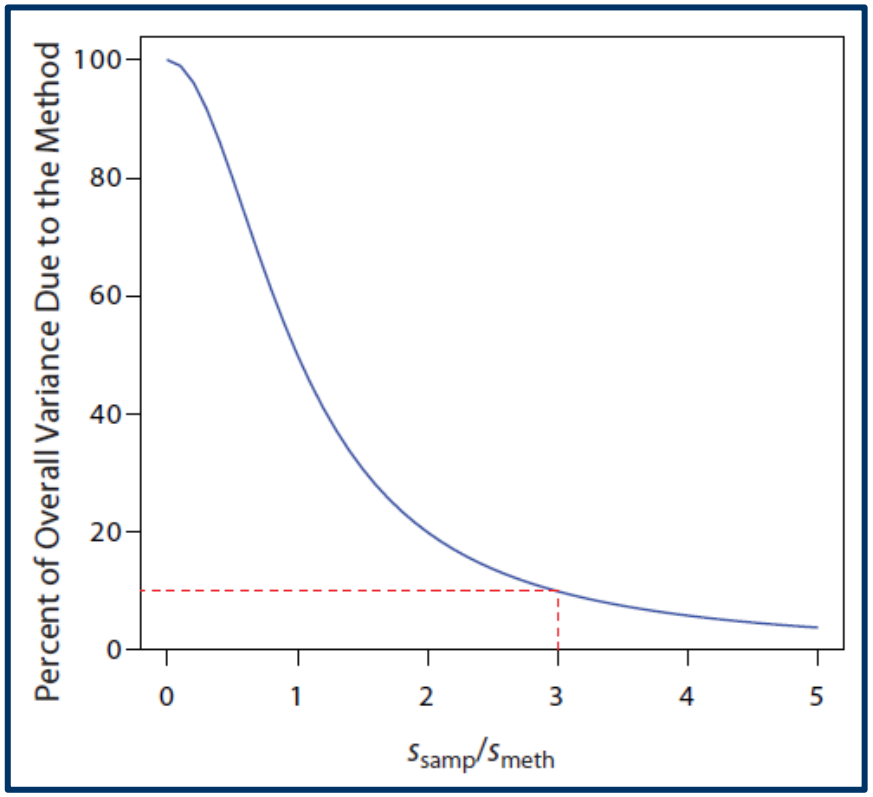
Extrapolating the result obtained for a sample to its target population introduces a sampling error. Using the rule of propagation of uncertainty, the relationship between the overall variance, s^2, and the variances due to sampling, s_{samp}^2, and to the analytical method, s_{meth}^2, can be easily obtained:

$$s^2 = s_{samp}^2 + s_{meth}^2$$

Unfortunately, analysts often try to minimize the overall variance by improving only the method's precision. This is a futile effort, however, if the standard deviation for sampling is more than three times greater than that for the method.

The figure reported on the right shows how the ratio s_{samp}/s_{meth} affects the method's contribution to the overall variance.

As shown by the dashed line, if the sample's standard deviation is more than three times the method's standard deviation, the percentage of overall variance due to the method becomes less than 10%.



In order to separate the contribution due to sampling an estimate of variance is made on samples obtained under conditions where both s_{samp} and s_{meth} contribute to the overall variance and on samples obtained under conditions where s_{samp} is known to be insignificant.

As an example, the following data, obtained after the analysis of a residual drug in an animal feed formulation, can be considered:

% Drug (w/w)			% Drug (w/w)		
0.0114	0.0099	0.0105	0.0105	0.0109	0.0107
0.0102	0.0106	0.0087	0.0103	0.0103	0.0104
0.0100	0.0095	0.0098	0.0101	0.0101	0.0103
0.0105	0.0095	0.0097			

Data on the left were obtained under conditions where both s_{samp} and s_{meth} contributed to the overall variance; those on the right were obtained under conditions where s_{samp} was known to be insignificant, i.e., by replicating the analysis of the same sample.

The corresponding variances were 4.71×10^{-7} and 7.00×10^{-8}, thus the variance due specifically to sampling was estimated to be equal to 4.01×10^{-7}.

Since the sampling standard deviation is more than twice as large as the standard deviation related to the method, the variance due to the sampling process accounts for almost 80% of the overall variance, thus improving the precision of that process has the greatest impact on the overall precision.

Sampling plan

A sampling plan must support the goals of an analysis.

For example, a material scientist interested in characterizing a metal's surface chemistry is more likely to choose a freshly exposed surface, created by cleaving the sample under vacuum, than a surface previously exposed to the atmosphere.

For a quantitative analysis, the sample composition must accurately represent the target population, a requirement implying a careful sampling plan. The following five aspects need to be considered in this case:

1. Where/when samples should be collected from the target population
2. What types of samples should be collected
3. The minimum amount of sample for each analysis
4. How many samples should be analyzed
5. The minimization of the overall variance for the analysis.

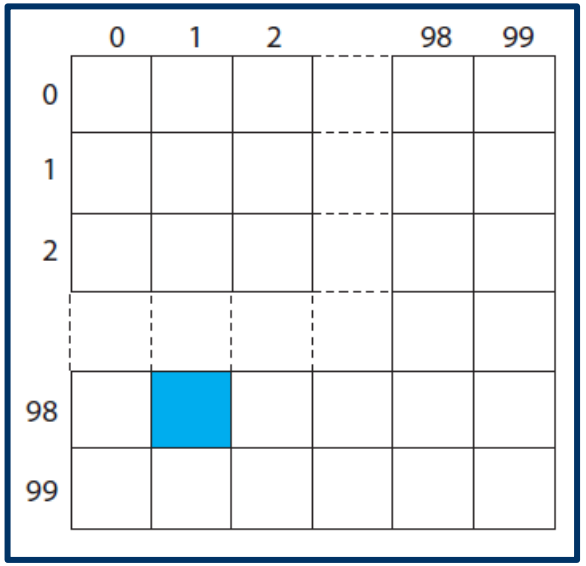
1. Where/when to sample the target population

If the target population is homogeneous in space/time, there is no need to collect individual samples in a specific position/moment.
Unfortunately, the target population is usually heterogeneous and only occasionally its homogenization can be a solution, since homogenization destroys the information about the analyte's spatial/temporal distribution, that can be important.

The ideal sampling plan provides an unbiased estimate of the target population's properties, and a Random Sampling is the easiest way to satisfy this requirement.
A simple method for ensuring the collection of random samples consists in dividing the target population into equal units and assigning a unique number to each unit.
Then, a random number table (or a random number generator) is adopted to select the units to sample.

As an example, let us suppose that a polymer's tensile strength has to be analyzed and ten 1 cm × 1 cm samples from a 100 cm × 100 cm polymer sheet need to be tested.

The sheet can be divided into 10000 1 cm × 1 cm squares, each identified by its row number and its column number, with numbers running from 0 to 99.



In order to select 10 squares at random, a table of random numbers, like the one shown on the right, can be used.

The last four digits of each random number could represent a row and a column (two digits each) in the table shown before, thus indicating a specific sample to be collected.
As an example, row 65 and column 18 would be inferred from the first number of the table.

The ten random numbers to be used can be selected following a predetermined path on the random number table.

Table of Random Numbers

36518	36777	89116	05542	29705	83775	21564	81639	27973	62413	85652	62817	57881
46132	81380	75635	19428	88048	08747	20092	12615	35046	67753	69630	10883	13683
31841	77367	40791	97402	27569	90184	02338	39318	54936	34641	95525	86316	87384
84180	93793	64953	51472	65358	23701	75230	47200	78176	85248	90589	74567	22633
78435	37586	07015	98729	76703	16224	97661	79907	06611	26501	93389	92725	68158
41859	94198	37182	61345	88857	53204	86721	59613	67494	17292	94457	89520	77771
13019	07274	51068	93129	40386	51731	44254	66685	72835	01270	42523	45323	63481
82448	72430	29041	59208	95266	33978	70958	60017	39723	00606	17956	19024	15819
25432	96593	83112	96997	55340	80312	78839	09815	16887	22228	06206	54272	83516
69226	38655	03811	08342	47863	02743	11547	38250	58140	98470	24364	99797	73498
25837	68821	66426	20496	84843	18360	91252	99134	48931	99538	21160	09411	44659
38914	82707	24769	72026	56813	49336	71767	04474	32909	74162	50404	68562	14088
04070	60681	64290	26905	65617	76039	91657	71362	32246	49595	50663	47459	57072
01674	14751	28637	86980	11951	10479	41454	48527	53868	37846	85912	15156	00865
70294	35450	39982	79503	34382	43186	69890	63222	30110	56004	04879	05138	57476
73903	98066	52136	89925	50000	96334	30773	80571	31178	52799	41050	76298	43995
87789	56408	77107	88452	80975	03406	36114	64549	79244	82044	00202	45727	35709
92320	95929	58545	70699	07679	23296	03002	63885	54677	55745	52540	62154	33314
46391	60276	92061	43591	42118	73094	53608	58949	42927	90993	46795	05947	01934
67090	45063	84584	66022	48268	74971	94861	61749	61085	81758	89640	39437	90044
11666	99916	35165	29420	73213	15275	62532	47319	39842	62273	94980	23415	64668
40910	59068	04594	94576	51187	54796	17411	56123	66545	82163	61868	22752	40101
41169	37965	47578	92180	05257	19143	77486	02457	00985	31960	39033	44374	28352
76418												

Random sampling often requires a relevant number of samples to ensure that the target population is adequately sampled.

From a statistic point of view, the mean, $\bar{x}$, of values x_i measured for each test sample resulting from random sampling provides an unbiased estimate of the bulk material composition.

The uncertainty in $\bar{x}$ due to sampling variation depends on the number of test samples n and the number of items N in the population:

$$s(\bar{x}) = s_{\text{sam}} \sqrt{\frac{1-f}{n}}$$

where $f = n/N$ and s_{sam} is the standard deviation found for the sampling process. If the analytical uncertainty is small, s_{sam} is simply the standard deviation s of the observations x_i. Otherwise, if the test samples are measured with replication, s_{sam} can be obtained from the between-group component of variance from a one-way ANOVA.

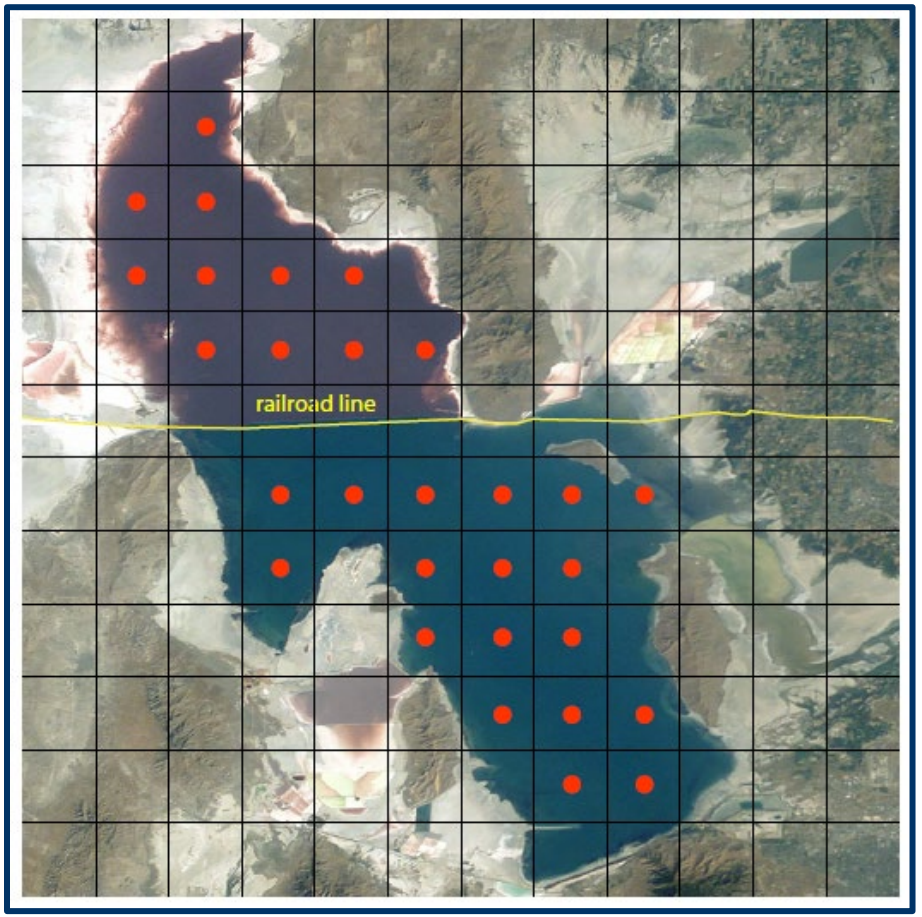
Notably, if N is much larger than n, the standard deviation converges to $s_{sam}/(n)^{1/2}$. As n approaches N, the uncertainty associated with sampling variability approaches zero; in fact, if the entire population is sampled, there is no remaining possibility of variation due a different choice of test samples.

In Systematic Sampling the target population is sampled at regular intervals in space or time.

In the figure reported on the right, the two-dimensional grid adopted for a systematic sampling of Great Salt Lake in Utah is shown, superimposed to an aerial photo of the lake.

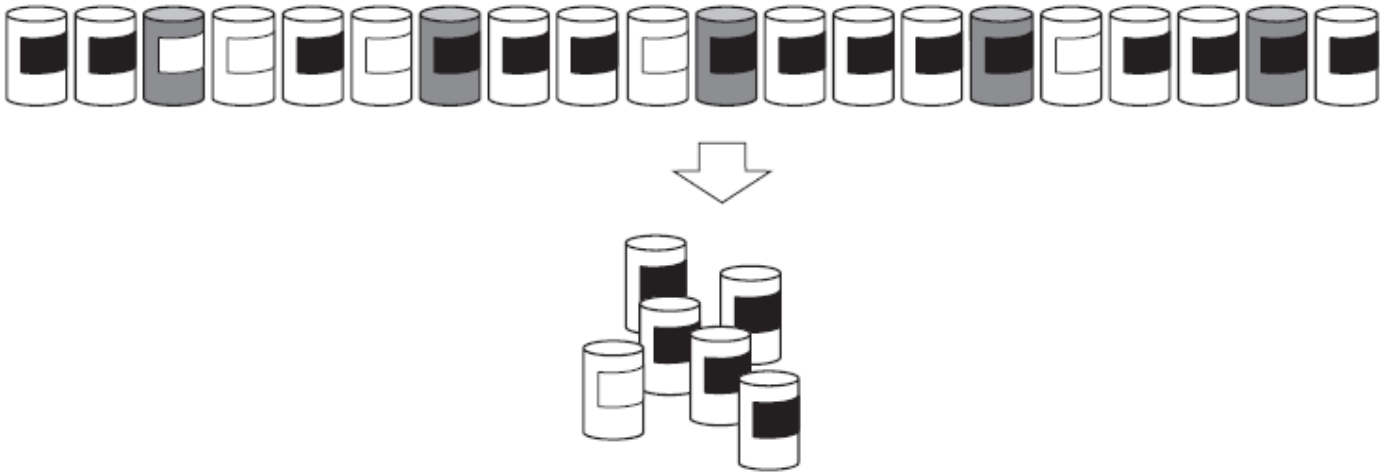
Notably, the lake is divided in two sections by a long stretch of land on which a railroad line was built. The different color of water in the two sections is due to the different growth of algae, in turn related to the different salinity, which is higher in the Northern part.

In this case the sampling aimed at providing samples to compare the two sections in terms of chemical composition.



Systematic Sampling can be performed also on discrete items, like cans containing a specific foodstuff obtained from a production line.

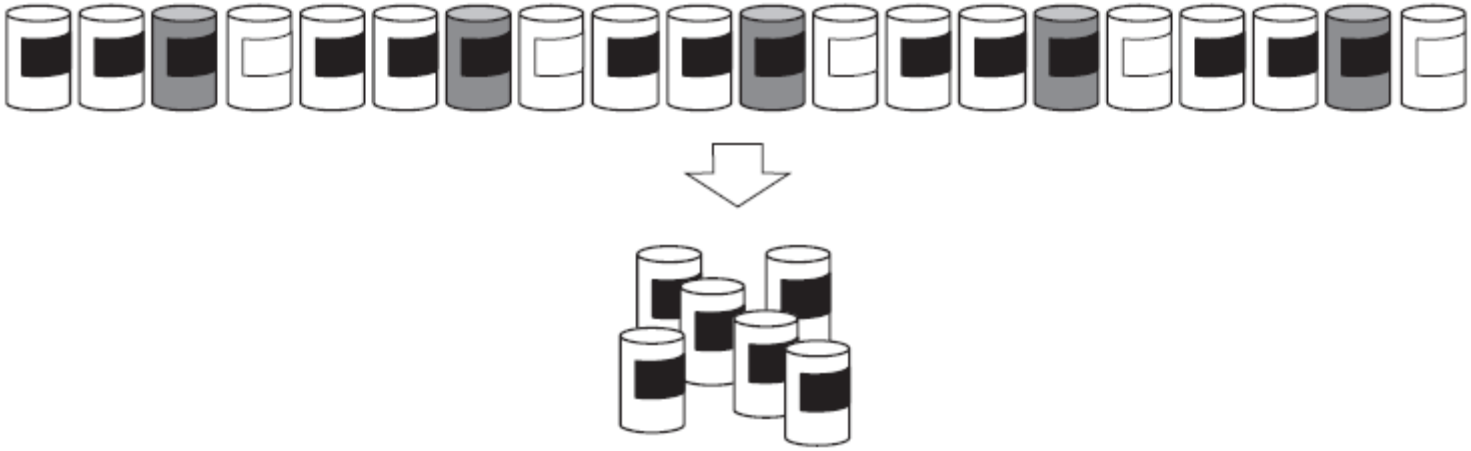
In this case the first test sample is selected at random but subsequent samples are then selected at a fixed interval (one out of four cans):



In this example, dark grey cans are those sampled from the line and the color of their labels emphasizes the existence of inhomogeneity between cans (which is, however, unknown to the sampler).

Systematic Sampling is usually simpler to implement than random, but it may imply a risk of bias if a systematic variation correlated with the sampling points occurs.

As shown in the following picture, a systematic variation is present in the items obtained from a production line, emphasized by the periodical occurrence of cans with white labels (one every four cans):



In this case, the choice of a systematic sampling scheme starting from one of the cans with black-label and implying the same frequency (one every four) leads to a very biased sampling, since no can with white label is selected. This is an extreme case but, in principle, it is not impossible.

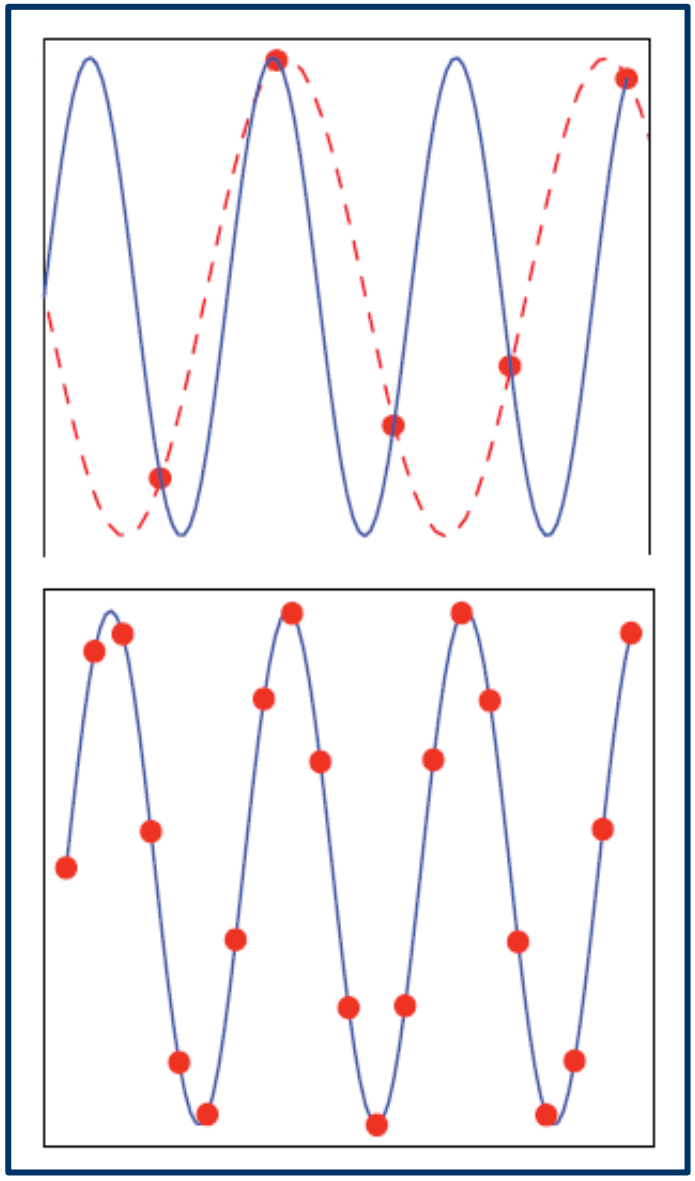
When a population's heterogeneity is time-dependent, as it is common in clinical studies, systematic sampling occurs at regular intervals in time.

If the target population's properties have a periodic trend, a systematic sampling will lead to a significant bias if the sampling frequency is too small.

This is a common problem when sampling electronic signals, leading to the phenomenon known as aliasing.

Let us consider, for example, a signal consisting of a simple sine wave.

In the figure shown on the right it is apparent how an insufficient sampling frequency underestimates the signal's true frequency. The apparent signal, shown by the dashed red line passing through the five data points, is significantly different from the true signal shown by the solid blue line.



According to the Nyquist theorem, in order to accurately determine a periodic signal's true frequency, the signal must be sampled at least twice during each cycle or period.

If samples are collected at an interval Δt, the highest frequency we can accurately monitor is $(2\ \Delta t)^{-1}$.

For example, if the sampling rate is 1 sample/hr, the highest frequency that can be monitored is 0.5 hr^{-1}, corresponding to a period of 2 hours.

If the signal's period is less than 2 hours (i.e., the highest frequency is higher than 0.5 hr^{-1}), a faster sampling rate must be used.

Ideally, the sampling rate should be at least 3-4 times greater than the highest frequency signal of interest. In other words, if the signal has a period of one hour, a new sample should be collected every 15-20 minutes.

Judgmental (also known as purposive) Sampling is adopted when prior information about the target population, or simply the researcher's judgment, is exploited to select samples for testing.

This kind of sampling was applied to assess the total mercury contamination in soil located in and around the dumping ground of Deonar, near Mumbay (India).

In the figure on the right the yellow border indicates the boundary of the dumping ground and the area highlighted in green indicates the site where active dumping is still ongoing. The white dots represent the sampling sites.

Based on judgmental sampling, 5 sampling points were established to cover the entire area and the remaining 4 were distributed at locations of the previous dumping.

P.P. Bhave, K. Sadhwani, Env. For., 23, 2022, 75-92

Judgmental sampling is clearly more biased than random sampling but requires fewer samples. Moreover, it can be useful if the analyst wishes to limit the number of independent variables that can potentially influence the results.

Since judgmental sampling is typically used to identify test samples of particular individual interest, it is not usually appropriate to average the results or consider the sampling variance other than to summarize the range of results found.

Instead, the individual results for particular test samples should be reported.

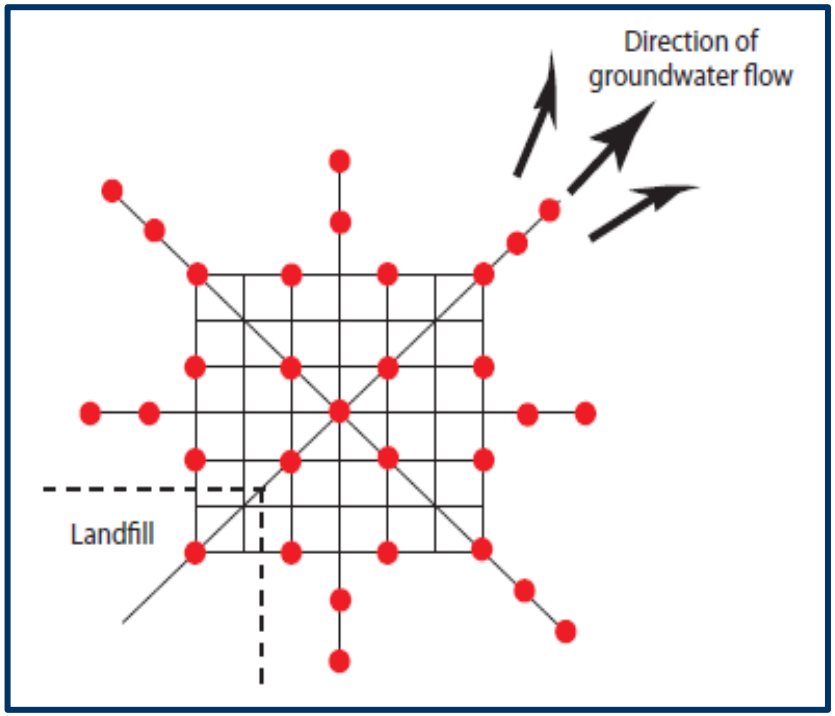
Combinations of the three primary approaches to sampling described so far are also possible.

One such combination is systematic–judgmental sampling, in which prior knowledge about a system is used to guide a systematic sampling plan.

For example, when monitoring waste leaching from a landfill, the plume is expected to move in the same direction as the flow of groundwater. This helps in focusing the sampling, thus saving money and time.

In the figure shown on the right, the corresponding systematic–judgment sampling plan is shown:

The plan includes a rectangular grid for most of the samples and additional samples along diagonal, vertical and horizontal paths, to explore the plume's limits.



2. What types of samples to collect

After determining where/when to collect samples, the next step in designing a sampling plan is to decide what type of samples to collect.

There are three common methods for obtaining samples:

1) grab sampling, 2) composite sampling, 3) *in situ* sampling.

In grab sampling a portion of the target population is collected at a specific time and/or location, providing a "snapshot" of the target population.

If the target population is homogeneous, a series of random grab samples allows to establish its properties. For a heterogeneous target population, systematic grab sampling allows to characterize how its properties change over space and/or time.

In composite sampling, a set of grab samples is combined into a single sample before analysis. It can be adopted if only an average composition of the target analyte is searched for.

For example, wastewater treatment plants must monitor and report the average daily composition of the treated water they release to the environment. The analyst can collect and analyze individual grab samples using a systematic sampling plan, reporting the average result, or can combine them into a single composite sample, thus saving time and money.

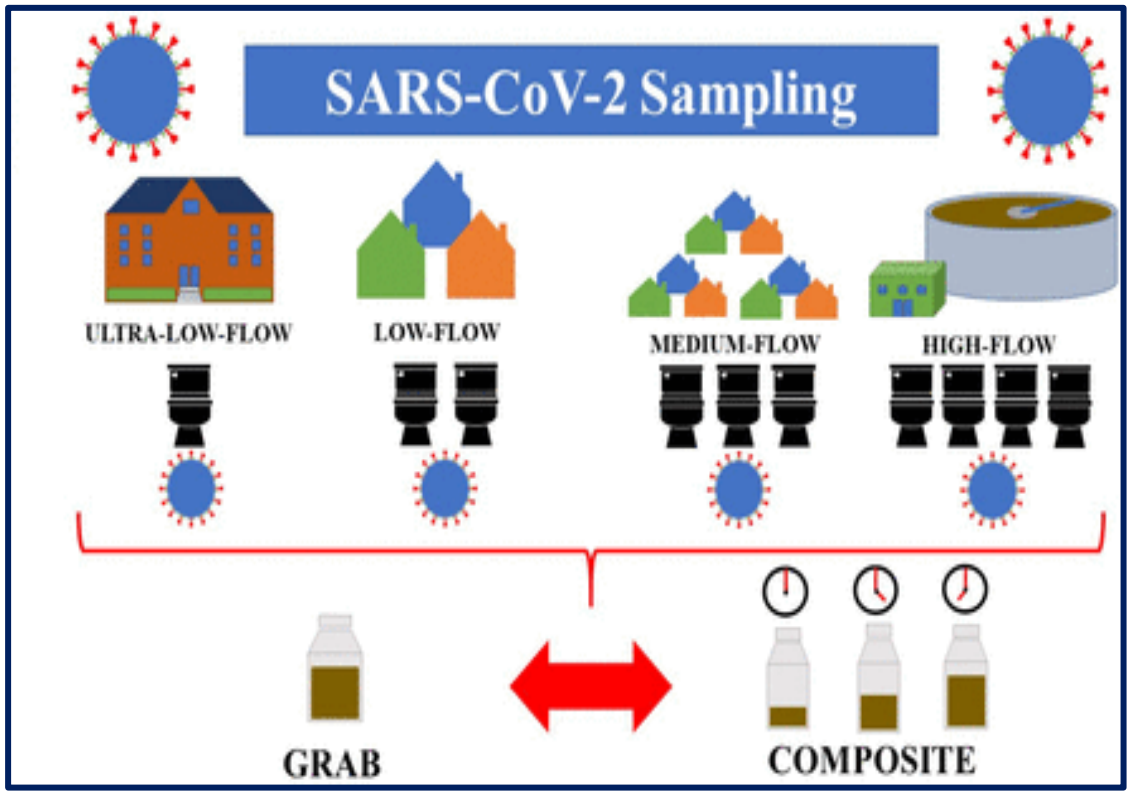
An interesting comparison between the outcomes of grab and composite sampling was reported for the detection of SARS-CoV-2 virus in wastewaters.
Specifically, four sites, characterized by ultra-low-, low, medium, and high flow rates in the Forest Grove, Oregon, sewershed were selected for 24 h sampling and subsequent detection of the virus:

Ultra-low-flow site:
four college dormitory buildings, one of which used to house COVID-19 infected students for convalescence;

Low-flow site (0.42 m^3/min):
small residential community (400 persons)

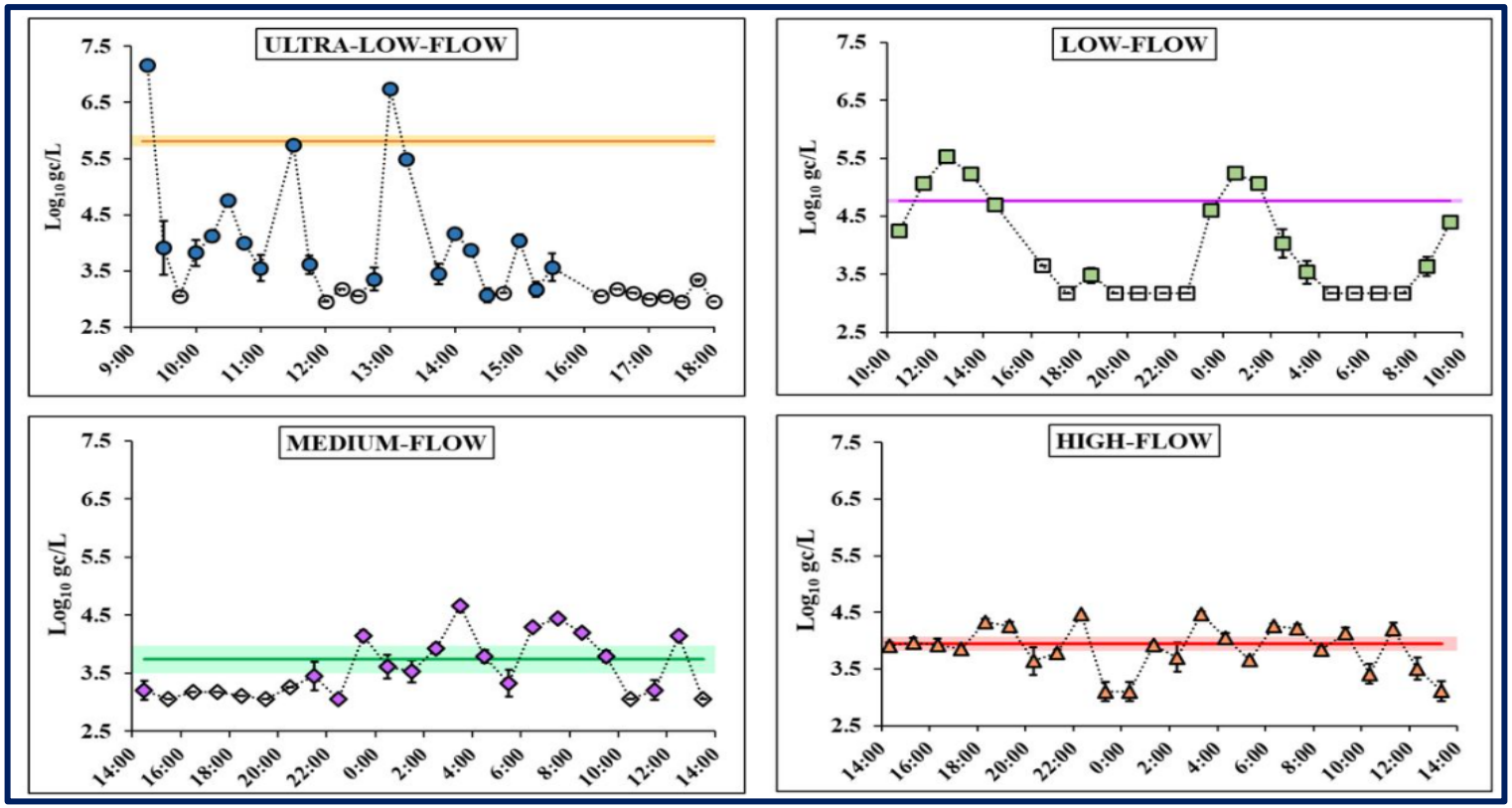
Medium-flow site (2.65 m^3/min):
medium residential community (2200 persons)

High-flow site (9.20 m^3/min):
Forest Grove community (40000 persons) plus a mixture of industrial wastewater.



A.D. George et al., Environ. Sci. Technol. Lett., 9, 2022, 160–165

The following graphs show SARS-CoV-2 concentrations measured over time from grab samples collected from (a) ultra-low-flow (15-min sampling frequency), (b) low-flow, (c) medium-flow, and (d) high-flow sites. The solid line denotes the composite value for each time series. The error bars on the grab samples and the shaded range on the composite lines denote standard error; non-detects are represented by open markers.



The results (express as $\log_{10}$ of genome copies/mL) indicate that while grab samples may provide fairly representative SARS-CoV-2 concentrations at high flow sites, they fail to provide representative SARS-CoV-2 concentrations at lower flow sites (e.g., buildings) and may lead to over- or under-estimates of the viral burden.

Composite sampling is also useful when a single sample can not supply sufficient material for the analysis.

For example, analytical methods for determining PolyChloroBiphenyls (PCB's) in fish often require as much as 50 g of tissue, an amount that may be difficult to obtain from a single fish. By combining and homogenizing tissue samples from several fish, it is easy to obtain the necessary 50-g sample. This can be easier to do if fish are obtained from an aquaculture plant, whereas it can be obviously more difficult if wild-type fish has to be considered.

A significant disadvantage of grab samples and composite samples is that they cannot be used to continuously monitor a time-dependent change in the target population.

In situ sampling, in which an analytical sensor is inserted into the target population, allows the continuous monitoring of the target population without withdrawing individual grab samples.

For example, the pH of a solution moving through an industrial production line can be monitored by immersing a pH electrode in the solution's flow.

3. Minimum amount of sample to collect

To minimize sampling errors, samples must be of an appropriate size.

If a sample is too small, its composition may differ substantially from that of the target population, introducing a sampling error.
Samples that are too large, however, require more time and money to be collected and analyzed, without providing a significant improvement in the sampling error.

Let us assume, as a simple model, that the target population is a homogeneous mixture of two types of particles. Particles of type A contain a fixed concentration of analyte, whereas particles of type B do not contain the analyte.

Samples from this type of target population follow a binomial distribution, whose probability function is:

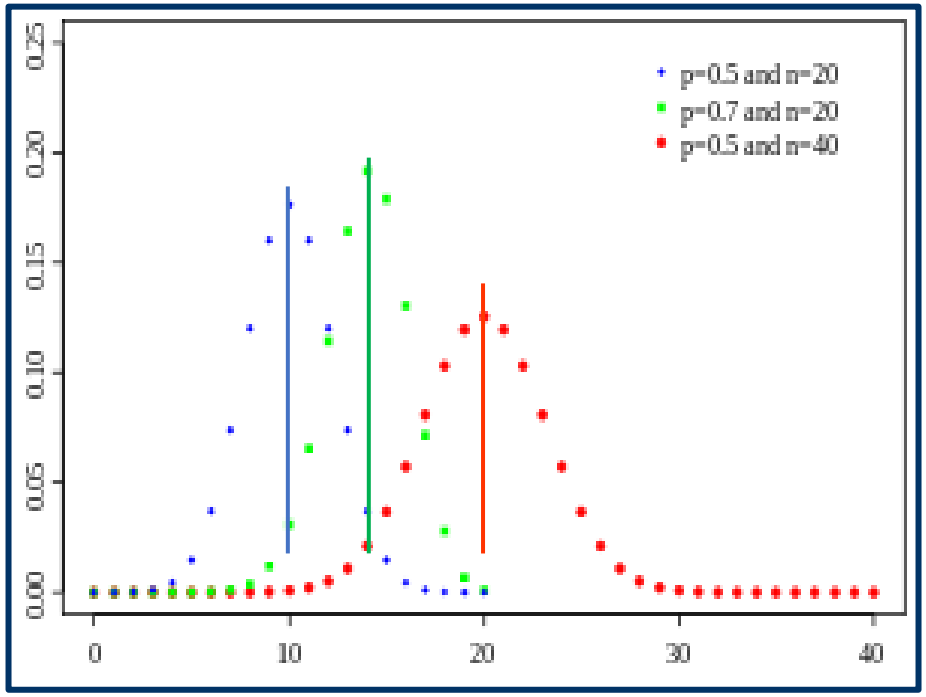
$$b(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}$$

where: $\binom{n}{x} = \frac{n!}{x!(n-x)!}$

represents the simple combination of x objects taken from a population of n objects.

The equation for b(x; n, p) indicates the probability of having x successes in a series of n independent trials when the probability of success in any one of the trials is p.

For example, if a success is accounted for as 1 (like the head of a coin) and a failure (like the tail of a coin) as 0, the curves shown in the figure on the right express the distribution of probability as a function of the total count.



If n = 20 and p = 0.5, the most probable total count is 20*0.5 = 10 (i.e., having 10 heads and 10 tails); if n = 20 and p = 0.7 it corresponds to 20*0.7 = 14 and if n = 40 and p = 0.5 it corresponds to 40*0.5 = 20.

In general terms, it can be demonstrated that the expectation of a binomial distribution is np, whereas the variance is np(1-p).

Returning to the case of a sample containing n particles, that can be of type A or B, the expected number of particles containing analyte, n_A, is:

$$n_A = np$$

where p is the probability of selecting a particle of type A.

The standard deviation for sampling is:

$$s_{samp} = \sqrt{np(1-p)}$$

The relative standard deviation for sampling is:

$$s_{samp}^{rel} = \frac{\sqrt{np(1-p)}}{np}$$

It this equation is re-arranged, the number of particles to be collected to provide the desired relative sampling variance is:

$$n = \frac{1-p}{p} \times \frac{1}{(s_{\text{samp}}^{\text{rel}})^2}$$

As an example, if a soil in which particles containing the analyte represent only 1×10^{-7} % of the population, 10^{13} particles should be collected to obtain a relative standard deviation for sampling of 1 %.

As a general concept, it is clear that if a very low percentage of particles contain the analyte, only sampling a very high number of particles can lead to a reasonable relative standard deviation.

Treating a population as if it contains only two types of particles is a useful exercise because it shows that the relative standard deviation for sampling can be improved by collecting more particles.

Of course, a real population may contain more than two types of particles, with the analyte present at several levels of concentration, according to the type of particle.

Nevertheless, many well-mixed populations, in which the population's composition is homogeneous on the scale at which the sampling occurs, approximate binomial sampling statistics. Under these conditions the following relationship between the mass of a random grab sample, m, and the percent relative standard deviation for sampling, R, is often valid:

$$mR^2 = K_s$$

Once the sampling constant K_s is determined, through appropriate experimental data, the mass of sample to be analyzed to obtain a certain relative standard deviation can be calculated.

As an example, the following data were obtained in a preliminary determination of the amount of inorganic ash in a breakfast cereal:

Mass of Cereal (g)	0.9956	0.9981	1.0036	0.9994	1.0067
% w/w Ash	1.34	1.29	1.32	1.26	1.28

In this case the average mass of the cereal samples is 1.0007 g, whereas the average % w/w ash and its absolute standard deviation are, respectively, 1.298% and 0.03194%.

The percent relative standard deviation R, therefore, is:

$$R = \frac{s_{samp}}{\bar{X}} \times 100 = \frac{0.03194\%\ \text{w/w}}{1.298\%\ \text{w/w}} \times 100 = 2.46\%$$

Solving for K_s gives its value as:

$$K_s = mR^2 = (1.0007\ \text{g})(2.46)^2 = 6.06\ \text{g}$$

Given this constant, in order to obtain a percent relative standard deviation of ±2%, samples need to have a mass of at least 1.5 g.

4. How many samples should be analyzed

If samples can be considered normally distributed, the confidence interval calculated using only the sampling error can be expressed as:

$$\mu = \bar{X} \pm \frac{ts_{samp}}{\sqrt{n_{samp}}}$$

where n_{samp} is the number of samples.

Rearranging the equation and substituting *e* for the quantity $\bar{X}-\mu$, gives the number of samples as:

$$n_{samp} = \frac{t^2 s_{samp}^2}{e^2}$$

Since the value of *t* depends on n_{samp}, the solution to this equation can be found iteratively.

As an example, if 2% is adopted as acceptable value for s_{samp} and a 95% confidence level is selected, the calculation to obtain *e* = 0.80% from data referred to inorganic ash in breakfast cereal is started considering an infinite number of degrees of freedom.

The Student's t coefficient in this case is equal to 1.960 (equivalent to the value referred to the standard Gaussian distribution), thus the first approximation for n_{samp} is:

$$n_{samp} = \frac{(1.960)^2(2.0)^2}{(0.80)^2} = 24.0 \approx 24$$

The $t_{0.975}$ coefficient for 24-1 = 23 degrees of freedom is equal to 2.073, thus the new calculation is:

$$n_{samp} = \frac{(2.073)^2(2.0)^2}{(0.80)^2} = 26.9 \approx 27$$

The t coefficient for 26 degrees of freedom is 2.060, thus the new n_{samp} value is:

$$n_{samp} = \frac{(2.060)^2(2.0)^2}{(0.80)^2} = 26.52 \approx 27$$

Since two successive calculations give the same value for n_{samp}, the iterative solution to the problem indicated that 27 samples are required to achieve a percent error of ±0.80% at a 95% confidence level.

5. Minimization of the overall variance for the analysis

A final consideration when developing a sampling plan is to minimize the overall variance for the analysis, which, as described before, corresponds to the sum between variances due to the method and to sampling.

If none of the two variances can be considered negligible, as it usually occurs, their combination has to be considered and the following equation can be obtained for the overall error:

$$e = t\sqrt{\frac{s^2_{samp}}{n_{samp}} + \frac{s^2_{meth}}{n_{samp}n_{rep}}}$$

where t is referred to n_{samp}* n_{rep} - 1 degrees of freedom.

This equation does not have a unique solution as different combinations of n_{samp} and n_{rep} may give the same overall error.

Consequently, the number of collected samples and of replicates performed for each of them is determined by other concerns, such as the cost of collecting and analyzing samples, and the amount of available sample.

As an example, let us consider an analytical method having a percent relative sampling variance of 0.4% and a percent relative method variance of 0.070%.

A comparison can be made for the error *e* (referred to a confidence level of 95%) obtained considering either 5 samples, each analyzed twice, or 2 samples, each analyzed five times.

In the first case (5 × 2) the value is (note that $t_{9,0.975} = 2.262$):

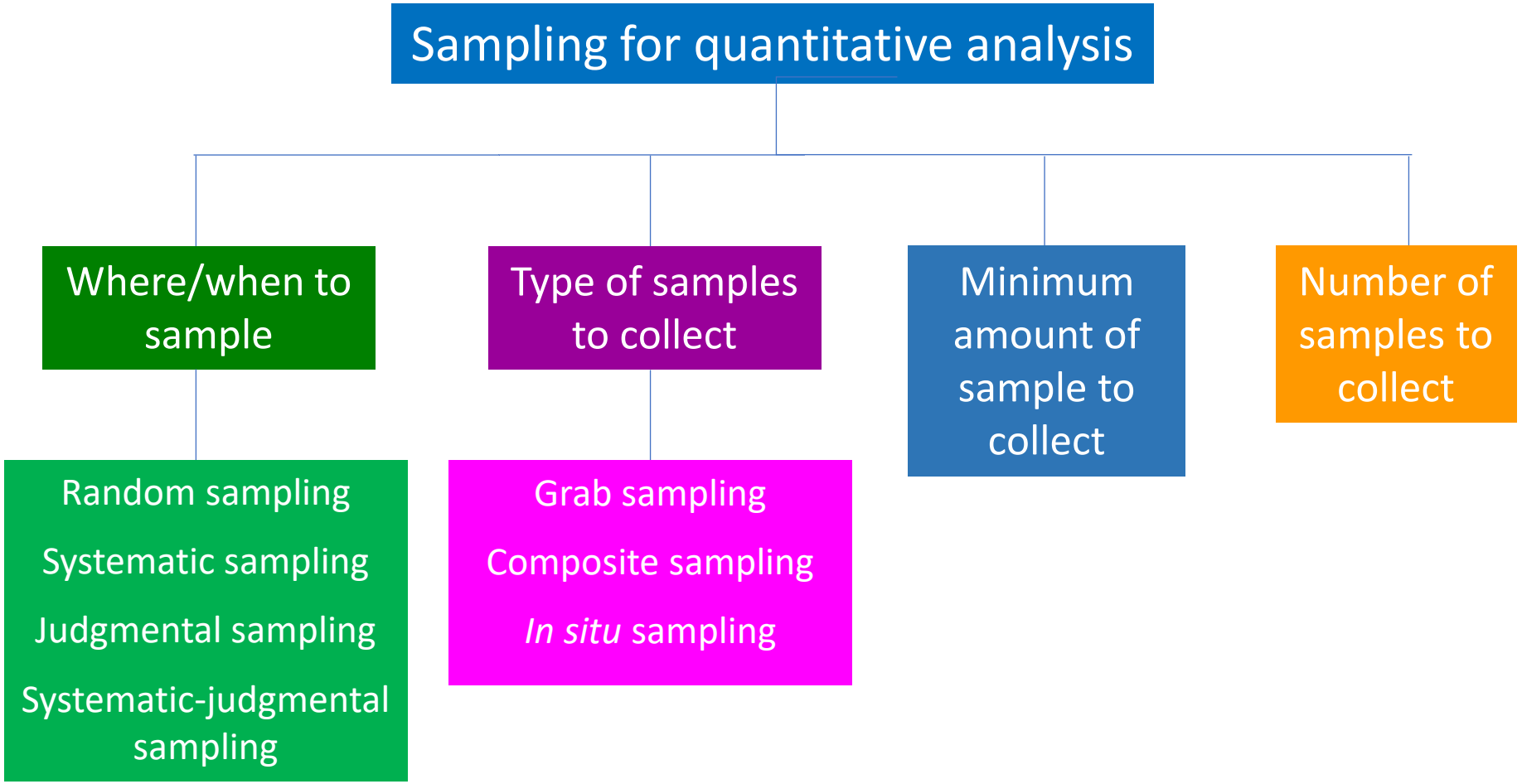
$$e = 2.262\sqrt{\frac{0.40}{5} + \frac{0.070}{5\times 2}} = 0.67\%$$

In the second case (2 × 5) it is:

$$e = 2.262\sqrt{\frac{0.40}{2} + \frac{0.070}{2\times 5}} = 1.0\%$$

Since the method variance is smaller than the sampling variance, a smaller relative error is obtained if more samples are collected, each analyzed fewer times.

Summary of relevant aspects of sampling for quantitative analysis



Acceptance sampling

Acceptance sampling represents a further important aspect involving both analysts and customers.

Specifically, it is used to determine whether to accept or reject a production lot of material and it is a common quality control technique used in industry.

It is usually done as products leave the factory, or in some cases even inside the factory. Usually, a producer supplies a customer with a specific number of items from a lot and the decision to accept or reject the items is made by determining the number of defective items among them. The lot is accepted if the number of defective items falls below the acceptance number, otherwise it is rejected.

In general, acceptance sampling is employed when one or more of the following circumstances occur:

1) testing is destructive
2) the cost of 100% inspection is very high
3) the time required for a 100% inspection would be too long

The approach known as acceptance sampling by variables is often adopted when the concentration of an analyte is involved.

Suppose that the manufacturer of a chemical is required to ensure that it does not contain more than a certain level of a particular impurity. This is called the acceptable quality level (AQL) of the product and is given the symbol μ_0.

The manufacturer's intention to ensure that this impurity level is not exceeded is monitored by testing batches of the product. Each test involves n test portions, whose mean impurity level is found to be $\overline{x}$.

The variation between portions, σ, is normally known from previous experience.

The practical problem that arises is that, even when a batch of manufactured material has an impurity level of μ_0, and is thus satisfactory, values of $\overline{x}$ greater than μ_0 will be found in 50% of the analyses, on a statistical basis.

Therefore, the manufacturer establishes a critical value for $\overline{x}$, indicated as $\overline{x}_0$.
If $\overline{x} > \overline{x}_0$ the batch is rejected.

Since the critical value is higher than μ_0, this approach leads to a decrease of the risk of rejecting a satisfactory batch for the manufacturer.

At the same time, the customer wishes to minimize the risk of accepting a batch with a mean impurity level greater than μ_0.

This can be achieved by setting an agreed tolerance quality level (TQL), μ_1, which has a smaller probability of acceptance.

The aim of acceptance sampling is that the critical value $\overline{x}_0$ should minimize the risk to the customer as well as to the manufacturer. At the same time, it should be ensured that n is not larger than necessary.
This can be achieved using the properties of the sampling distribution of the mean, since σ is known.

Suppose that the manufacturer accepts a 5% risk of rejecting a batch of the chemical that is in fact satisfactory, i.e., a batch for which $\overline{x} > \overline{x}_0$ even though $\mu = \mu_0$.

The following equation can be written:

$$(\overline{x}_0 - \mu_0)/(\sigma/\sqrt{n}) = 1.64$$

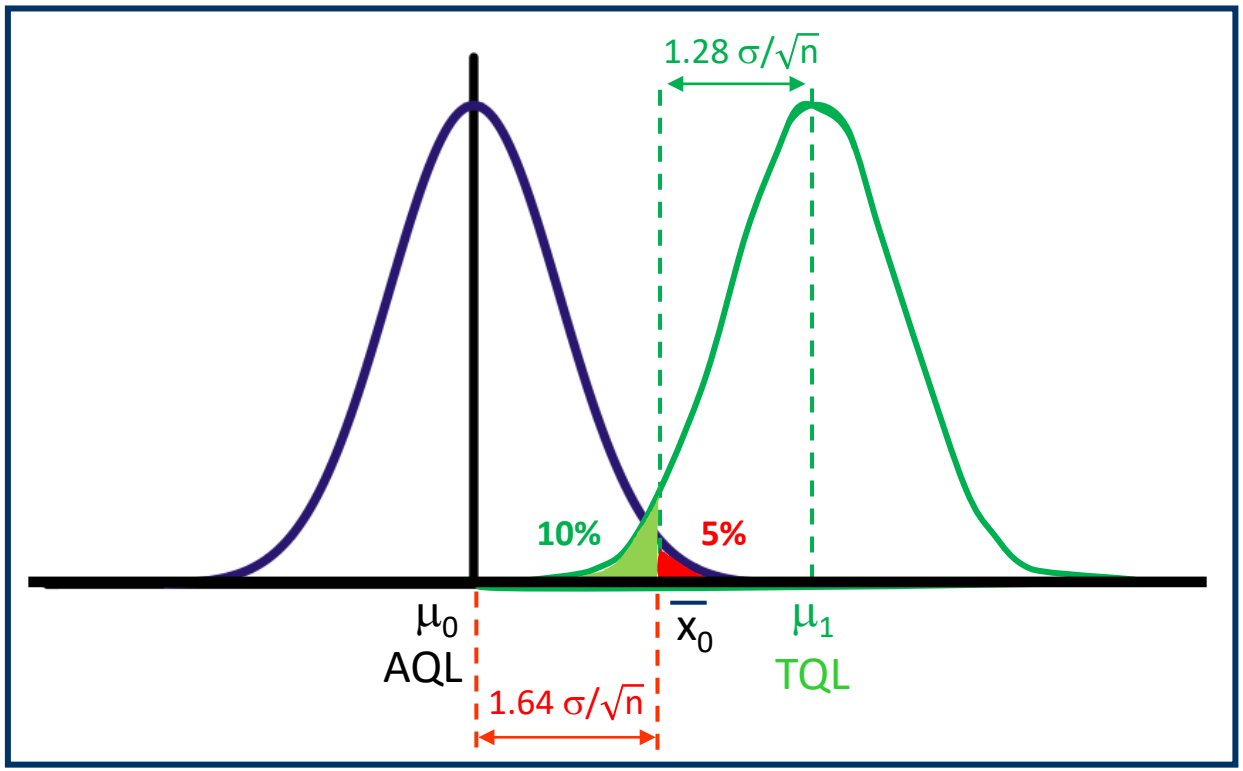
where 1.64 is the z value corresponding to an area under the standard Gaussian probability density function equal to 0.95.

Suppose also that the customer is prepared to accept a 10% risk of accepting a batch with the impurity at the TQL.

The following equation can be obtained:

$$(\bar{x}_0 - \mu_1)/(\sigma/\sqrt{n}) = -1.28$$

where -1.28 is the z value corresponding to an area under the standard Gaussian probability density function equal to 0.10.



A graphical representation of the different quantities expressed in the two equations is shown in the figure on the left.

Since, in practice, the values of μ_0 and μ_1 are fixed in advance, the two equations can be solved to find n and $\bar{x}_0$, if σ is known.

Suppose that the AQL and TQL for an impurity have been fixed as 1.00 g kg^{-1} and 1.05 g kg^{-1}, respectively, the manufacturer and customer risks are 5% and 10%, respectively, and σ is 0.05 g kg^{-1}.

By transformation of the equations shown before, the following equations are obtained:

$$n = [(1.64 + 1.28)0.05/(1.05 - 1.00)]^2$$

$$\bar{x}_0 = [(1.64 \times 1.05) + (1.28 \times 1.00)]/(1.64 + 1.28)$$

The resulting values are:

$n = 2.92^2 = 8.53$, which is rounded up to a sample size of 9;

$\bar{x}_0 = 3.002/2.92 = 1.028$.

Consequently, a critical value of 1.028 g kg^{-1} impurity and a sample size of 9 will provide both manufacturer and customer with the required assurances.