CODEX ALIMENTARIUS
INTERNATIONAL FOOD STANDARDS

Food and Agriculture Organization of the United Nations
World Health Organization

E-mail: codex@fao.org - www.codexalimentarius.org

GENERAL GUIDELINES ON SAMPLING

CXG 50-2004

Adopted in 2004. Revised in 2023.

1. REFERENCE GUIDELINES

1.1 Introduction

These guidelines are primarily intended for use by Codex commodity committees responsible for developing acceptance sampling plans for provisions in Codex standards, and by governments responsible for import or export inspection of foods to describe the design and evaluation of sampling plans for the international trade of food commodities.

Foods are frequently sampled throughout the food supply chain, from producers to consumers, for the purpose of checking their quality. Clear definition of sampling plans is an integral part of specifications for the sampling and testing of foods. Sampling plans are included in Codex standards and may be used by governments in standards for foods.

Codex sampling plans, in conjunction with methods of analysis, are intended as a means of verifying that foods comply with provisions relating to composition, chemical or microbiological contaminants or pesticide residues contained in Codex standards.

Sampling therefore plays an important role in achieving the Codex objectives of protecting consumers' health and ensuring fair practices in the food trade. Codex sampling plans also have an important role in harmonizing technical approaches to sampling and by results of analysis interpretation in relation to lots or consignments of foods, in the light of the relevant provision(s) of the applicable Codex standard.

It is important that sampling be undertaken in a way that contributes to these objectives.

Specification of these quality objectives, the quality level acceptable to the customer, and the rate of acceptance of compliant products, enables the development of sampling plans.

A Codex standard may set out a specific sampling plan for a particular context, or it may specify the outcome to be achieved by a sampling plan.

Although these guidelines provide a generic approach to the design of sampling plans, Codex sampling plans are intended primarily for inspection of foods upon receipt, for example by importing country regulatory agencies, and might not be suitable for use by producers. However, a clear definition of quality objectives in Codex standards will allow producers to devise appropriate control and inspection procedures to achieve them.

1.2 Scope

In these guidelines, the focus is on acceptance sampling plans for the inspection of isolated homogeneous lots, in which the risks to consumers and producers are controlled. Additionally, there are some guidelines for sampling inhomogeneous lots.

The term 'isolated' means that the inspection of each lot is done in isolation, without considering the outcome of the inspection of adjacent lots or, for example, other lots from the same producer. This does not mean that information from previous inspections cannot be used; in particular, there are cases where the lot standard deviation may be known from the inspection of previous lots.

The following situations are covered:

- acceptance sampling plans for the control of the percentage nonconforming for homogeneous lots by attributes or by variables, for goods in bulk or individual items;
- inspection by variables sampling plans for normally distributed characteristics;
- adjustment for measurement uncertainty in cases where it is non-negligible as compared to the lot standard deviation with a focus on cases where the measurement uncertainty is normally distributed;
- sampling plans for the control of the average content; and
- in addition, some information is provided on issues involved with the design of plans for bulk materials.

In Section 2, general concepts which are relevant for the sampling of foods are defined, Sections 3, 4 and 5 cover acceptance sampling plans for different situations of statistical food control. Section 6 covers other matters such as physical sampling and inhomogeneous lots.

Appendix I contains a step-by-step guide for the selection of sampling plans. Appendix II contains tables of ISO[i] attributes and variables plans indexed by producer's risk.

[i] The International Organization for Standardization (ISO).

These guidelines are not intended to be comprehensive; these guidelines do not provide information on all types of sampling plan options that may be available. Sampling plans from other sources, such as plans developed by other Codex committees, are still acceptable subject to their endorsement by the Codex Committee of Methods of Analysis and Sampling (CCMAS).

1.3 Definitions

For the terms commonly used in these guidelines, the following definitions are provided, in addition to those in the *Guidelines on Analytical Terminology* (CXG 72-2009).[1]

Note: In some of the definitions, reference is made to the process standard deviation or the process quality level. In these guidelines, the focus lies on lots rather than processes. For this reason, the relevant quantities in these guidelines are the lot standard deviation and the lot quality level.

Acceptance criterion

Acceptance criterion is used to cover terms such as acceptance and rejection numbers for attributes plans and acceptability constants for variables plans.

Note: In these guidelines, the term 'acceptance criterion' is used to describe the rule which is applied to the test results obtained during the lot inspection in the decision whether to accept the lot.

Acceptance sampling

Sampling after which decisions are made to accept a lot, or other grouping of products, materials or services, based on sample results.

Acceptance sampling by attributes

Acceptance sampling inspection whereby the presence or absence of one or more specified characteristics of each item in a sample is observed to establish statistically the acceptability of a lot or process.

Acceptance sampling by variables

Acceptance sampling inspection in which the acceptability of a process is determined statistically from measurements on specified quality characteristics of each item in a sample from a lot.

Acceptance sampling plan

Plan which states the sample size(s) to be used and the associated criteria for lot acceptance.

Conformity assessment

Activity to determine whether specified requirements relating to a product, process, system or person or body are fulfilled.

Consignment

A quantity of some commodity delivered at one time. It may consist of either a portion of a lot, or a set of several lots.

However, in the case of statistical inspection, the consignment shall be considered as a new lot for the interpretation of the results.

- If a consignment is a portion of a lot, the consignment shall be considered as a lot for the inspection.
- If the consignment is a set of several lots, before any inspection, care shall be given to the homogeneity of the consignment. If not homogeneous, a stratified sample may be used.

Consumer and producer

The terms 'consumer' and 'producer' are conventional and may apply to a range of different operators in the food supply chain, such as a grower, manufacturer, the manufacturer's own quality control system, supplier, exporting country, processor, on-seller, or importing country. In general, 'producer' refers to a supplier or seller of foodstuffs and 'consumer' to an importing country regulator, a purchaser, or an actual consumer of those foods.

Consumer's risk (CR)

Probability of acceptance when the quality level of the process has a value stated by the acceptance sampling plan as unsatisfactory.

Consumer's risk quality (CRQ)

Quality level of a lot or process which, in the acceptance sampling plan, corresponds to a specified CR.

Note: The CRQ corresponds to the limiting quality level (LQL) in the ISO 2859[ii] and ISO 3951[iii] standards.

Indifference quality level

Quality level which, in the acceptance sampling plan, corresponds to a probability of acceptance of 0.5, when a continuing series of lots is considered.

Laboratory sample

A sample as prepared (from the lot) for sending to the laboratory and intended for inspection or testing.

Lot

A quantity of product produced under conditions presumed uniform.

Operating characteristic curve

Curve showing the relationship between probability of acceptance of product and the incoming quality level for given acceptance sampling plan.

Plan

Refer acceptance sampling plan.

Producer's risk (PR)

Probability of non-acceptance when the quality level of the process has a value stated by the plan as acceptable.

Producer's risk quality (PRQ)

Quality level of a lot or process which, in the acceptance sampling plan, corresponds to a specified PR.

Note: The PRQ corresponds to the acceptance quality limit (AQL) in the ISO 2859[iv] and ISO 3951[v] standards.

Quality level

Quality expressed as a rate of nonconforming units or rate of number of nonconformities.

Note: In these guidelines, the quality level of a given lot is often expressed in terms of the percentage of nonconforming items.

Sample

One or more items taken from a population and intended to provide information on the population, and possibly serve as a basis for a decision on the population or on the process which had produced it.

Sampling plan

Refer acceptance sampling plan

2. ACCEPTANCE SAMPLING – GENERAL PRINCIPLES

2.1 Reasons for sampling

While various measures such as hazard analysis and critical control point systems (HACCP), good manufacturing practice (GMP), process control and sampling are available to producers to provide assurance about the quality of products they supply, consumers usually rely on acceptance sampling if they wish to verify the quality of incoming products.

[ii] ISO 2859: *Sampling procedures for inspection by attributes*. This ISO includes a series of standards (parts).
[iii] ISO 3951: *Sampling procedures for inspection by variables*. This ISO includes a series of standards (parts).
[iv] See note ii above.
[v] See note iii above.

Acceptance sampling procedures are used when goods are transferred between two parties. The purpose of these procedures is to provide unambiguous rules for releasing a product after inspection of only a limited sample. Both parties should be fully aware of the limitations and risks associated with using such procedures and therefore most acceptance sampling procedures should include provisions for dealing with disputes and non-conforming items found in lots that have been accepted by the sampling plan.

An acceptance sampling plan specifies the number of items to be taken and how they are to be taken, the acceptance criterion used to decide whether a lot should be accepted and how to take non-negligible measurement uncertainty into account.

In general acceptance sampling is used to:

- reduce costs;
- allow product assessment when tests are destructive; and
- enable faster decision-making.

2.2 Approaches to acceptance sampling

There are three possible approaches to acceptance sampling:

(a) 100 percent inspection, involving inspection of the entire (i.e. 100 percent) lot;

(b) sampling based on statistical principles; and

(c) ad hoc inspection, i.e. sampling plans without a statistical basis.

The risks and costs associated with each of these three options will be briefly discussed. Approach (a) is usually not feasible due to the prohibitive cost of testing and in addition, there might not be any product left to sell if the inspection method necessitates destructive testing.

Approach (b) has the disadvantage of higher risks as compared to approach (a), since a part of the lot is not inspected. However, by applying an approach based on statistical principles, the risks can be calculated, and a sampling plan can be chosen that ensures these risks are controlled to desired levels. It also has the advantage of practicability and lower costs.

In lot inspections, there are two types of risks:

- acceptance of a lot of unsatisfactory quality (CR); and
- rejection of a lot of acceptable quality (PR).

Sampling plans should be designed to control these risks to suitable levels, whereby suitable risk levels are determined based on fitness-for-purpose considerations.

Approach (c) is not recommended. It may be used for practical reasons, such as limited resources, or for simplicity. However, such plans might not provide the expected level of assurance of food quality and may inadvertently impose high costs, for instance through unwarranted acceptance of food that could lead to illness or unjustified rejection that, in turn, could lead to the imposition of fines, penalties or trade sanctions. The risks associated with such plans should be evaluated where possible. Decisions on acceptance or rejection should not be made solely based on these plans except by mutual agreement of the consumer and producer with an understanding of the risks involved.

In summary, approach (b) allows for practicability while ensuring that risks are controlled to levels considered appropriate based on fitness-for-purpose considerations.

2.2.1 *Acceptance sampling versus conformity assessment*

Acceptance sampling and conformity assessment do not have the same purpose. Conformity assessment is the use of a single measurement result to decide whether a single item conforms to a limit. Acceptance sampling is the process in which a sample[vi] is taken from a lot and involves the determination of acceptance criteria and sample size to decide whether a lot is accepted or rejected.

The broadest definition of conformity assessment may be considered to include acceptance sampling. However, in a narrower sense, conformity assessment can be understood to refer specifically to the situation where a one single measurement result is used to decide if one single item of interest conforms to a specified requirement. If conformity assessment is understood in this narrower sense, then it is important to distinguish conformity assessment and acceptance sampling. In this section, conformity assessment will be understood in the narrower sense.

Although acceptance sampling and conformity assessment involve similar procedures, and although consumer and producer risks are defined for both, they are performed in different contexts and follow different objectives.

Conformity assessment

In conformity assessment, conformity is assessed via the application of a decision rule which accounts for measurement uncertainty. Depending on the measurand, the measurement uncertainty may or may not include uncertainty from sampling. Depending on the decision rule, there may be cases where the assessment is inconclusive.

Acceptance sampling

In acceptance sampling, at least one measurement result (typically more than one) is used to decide whether to accept or reject a lot under inspection. The acceptance sampling plan consists in both requirements regarding the sampling procedure (e.g. the number of items to be taken from the lot) and an acceptance criterion. The acceptance sampling plan is determined in such a way as to ensure that producer and/or consumer risks are sufficiently low at a given quality level. The variation of the property of interest in the lot is always taken into consideration in acceptance sampling; however, analytical uncertainty is only taken into consideration if non-negligible. The context for lot inspection is typically a commercial agreement between two trading partners. In acceptance sampling, a lot is always either accepted or rejected; there are no cases of inconclusive lot inspections.

In the case that the quality level is expressed in terms of the percentage of nonconforming items, the distinction between acceptance sampling and conformity assessment is quite clear; the measurand is defined for the individual items, and thus the question of conformity to a specified requirement can only be framed in relation to the individual items. However, lot acceptance or rejection is not decided on the basis of the compliance or non-compliance of an individual item. Instead, the acceptance criterion is expressed in terms of the percentage of nonconforming items, in terms of the distribution of the property of interest among the items in the lot. The differences between acceptance sampling and conformity assessment are summarized in the following table.

[vi] Refer to the definition in Section 1.3.

Table 1. Differences between acceptance sampling and conformity assessment

	Conformity assessment	Acceptance sampling
Number of measurement results	Typically: one	Typically: several (For instance: if the lot consists of discrete items, several items are taken, and there is one measurement result per item)
Is analytical measurement uncertainty taken into account in the decision rule/acceptance criterion?	Always (if possible)	Only if the analytical measurement uncertainty is non-negligible (compared to the lot standard deviation)
Are any components of sampling uncertainty considered?	Depending on the measurand, it may or may not be necessary to include sampling uncertainty	The variation of the characteristic of interest within the lot is considered via the lot standard deviation
Context/Background	In many cases: conformity assessment is carried out against a legal limit	The context is often an agreement between trading partners
Inconclusive assessment	Depending on the decision rule, the assessment may be inconclusive	There are no inconclusive inspections: lots are either accepted or rejected.

Further clarifications regarding the term measurand and the distinction between sampling and analytical uncertainty are provided in Section 0.

Note 1: Figure 1 in the Guidelines on Measurement Uncertainty *(CXG 54-2004)[2] illustrates a procedure which can be applied in conformity assessment (this procedure may yield inconclusive results). This procedure should not be applied in acceptance sampling.*

Note 2: If the sample taken in a lot inspection consists of one single item, then producer/consumer risks may be poorly controlled. Nonetheless, there are special sampling plans for lot inspection based on a single item. These must not be confused with the procedure for conformity assessment illustrated in Figure 1 in the Guidelines on Measurement Uncertainty *(CXG 54-2004).[2]*

2.3 Acceptance sampling plan performance

Variation is present everywhere; raw materials vary in their composition, manufacturing processes vary and, consequently, the products manufactured by those processes will also vary. Therefore, when we take several samples from a lot, we do not expect those samples to be of the same composition. Furthermore, the presence of measurement uncertainty means that when those samples are tested, we will not get the same result, even if the same sample is retested. Similarly, we would not expect results from different sets of samples taken from the same lot or those taken from different lots (from the same process) to be the same; there will always be some variation.

Due to this variation, the incorrect acceptance or rejection of lots cannot be avoided. However, using a statistical description of the variation within a lot and of the uncertainty of the measurement process allows us to calculate the probability of correctly or incorrectly accepting a lot at any given quality level and for any given sampling plan.

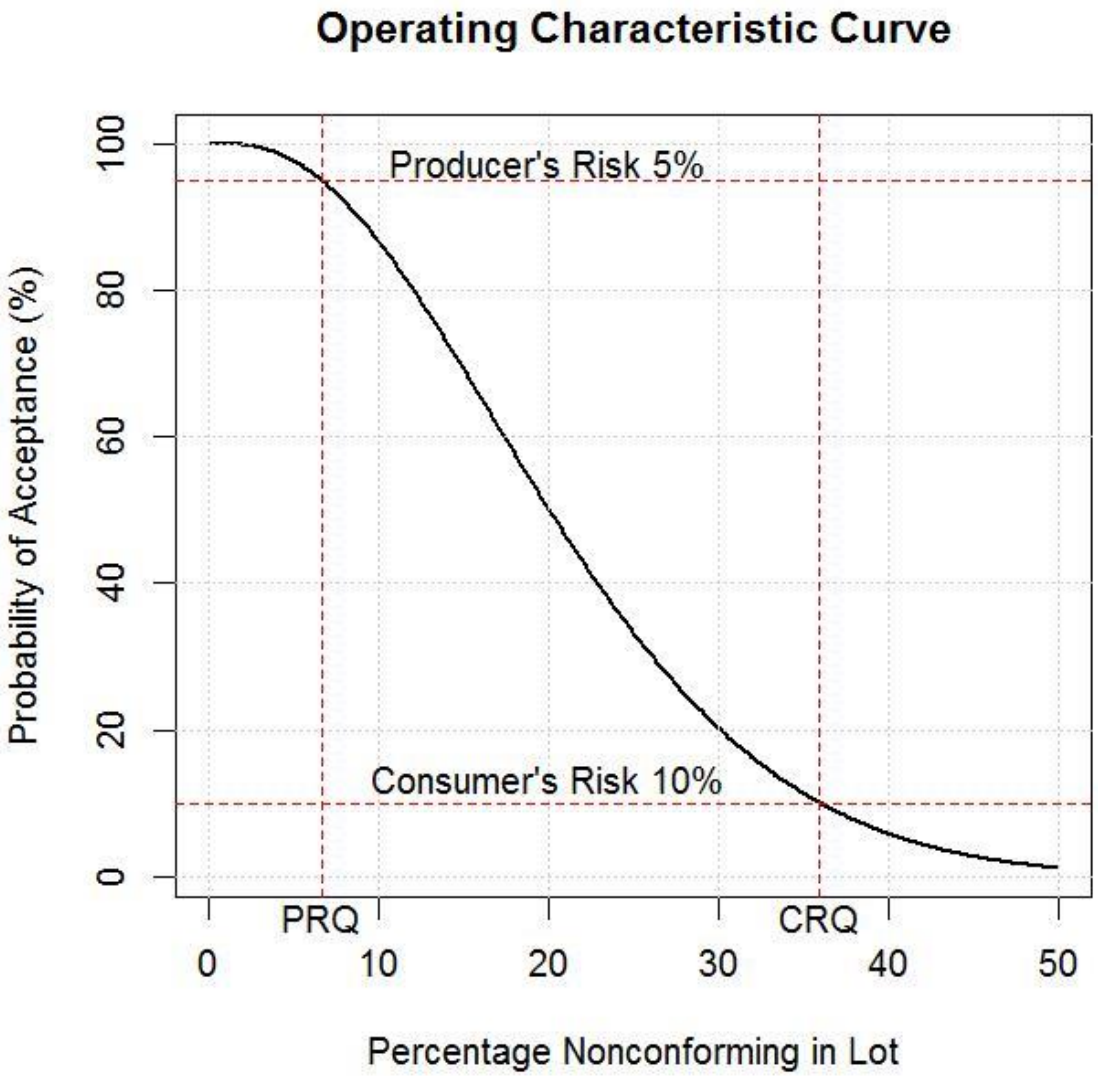
In acceptance sampling, the probability of acceptance depends on:

- the quality level (percent nonconforming) of the lot under inspection;
- the acceptance criterion (i.e. for the particular sampling plan);
- the variation of the characteristic within the lot; and
- the bias and variation inherent in the measurement process (in the case of non-negligible analytical uncertainty).

In practice, the quality level (percent nonconforming) of a lot is not known beforehand; however, for a particular acceptance sampling plan, it is possible to calculate the probability of acceptance at any quality level. The relationship between the probability of acceptance and the quality level for a particular sampling plan is described by the operating characteristic curve.

2.3.1 *Operating characteristic curve*

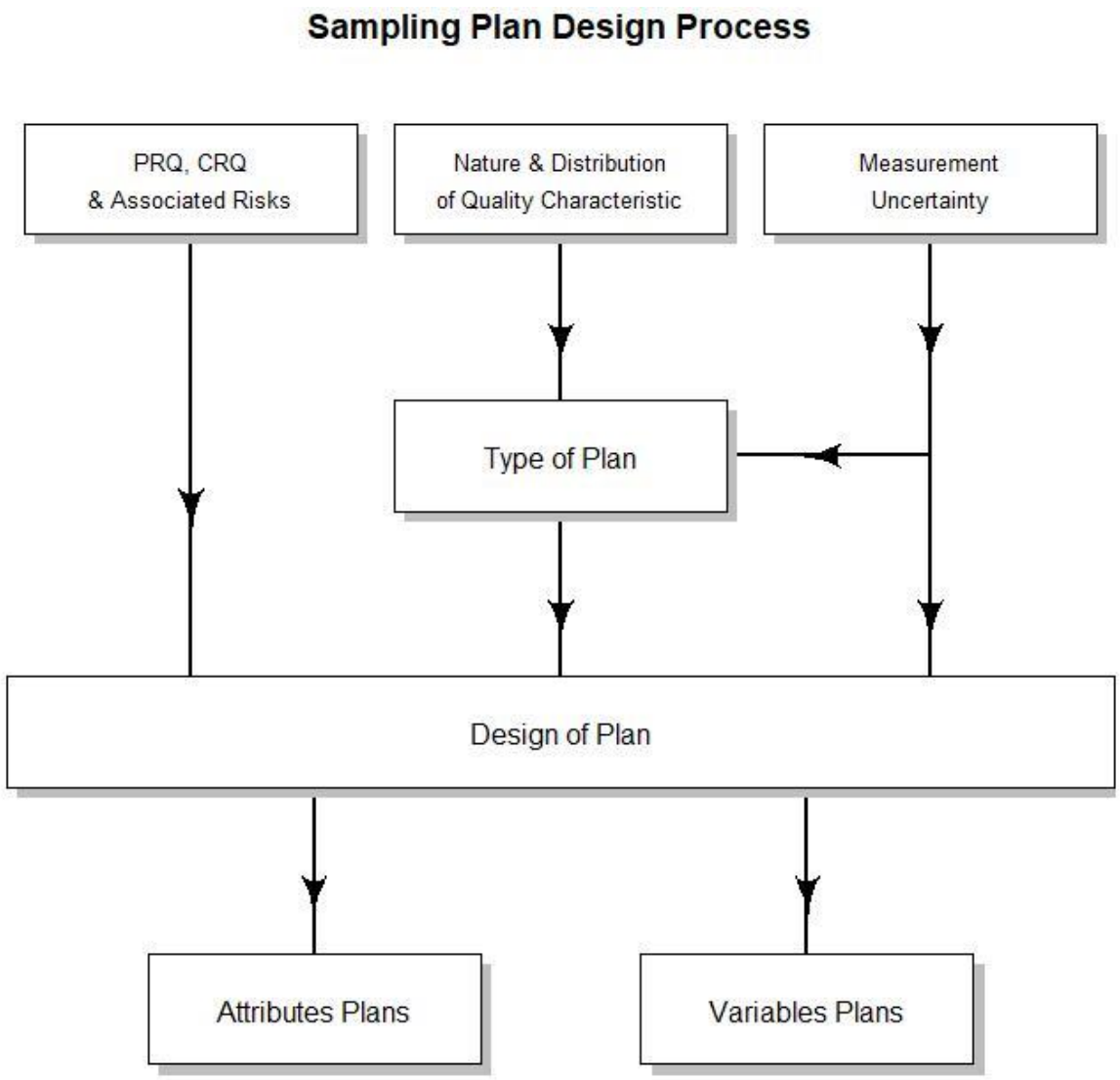
The following diagram is an example of an operating characteristic curve (OC curve) that shows the probability of accepting (or rejecting) a lot in terms of its quality level in the lot (expressed as percent nonconforming). This highlights that specification of the quality levels is fundamental to design of a sampling plan.



Note: The OC curve does not say anything about the quality of a given lot; it serves only to show the probability of accepting a lot with a particular quality level.

3. DESIGN OF SAMPLING PLANS

3.1 Sampling plan design process



3.2 Inputs to sampling plans

3.2.1 *Stringency*

As explained, the application of acceptance sampling plans does not eliminate the risk that a lot of poor quality will be incorrectly accepted nor that a lot of good quality will be incorrectly rejected.

However, designing such plans using statistical principles allows these risks to be controlled. This is achieved by specifying a particular producer's risk quality (PRQ) level, and a particular consumer's risk quality (CRQ) level, along with a corresponding producer's risk (PR) and a consumer's risk (CR) respectively. Once these four parameters (PRQ, CRQ, PR and CR) are specified, the probability of acceptance and therefore the producer's and consumer's risks at any quality level are uniquely determined.

The term stringency is used in these guidelines to refer to the ability of a sampling plan to control CRs and PRs, of incorrectly accepting or incorrectly rejecting a lot, at any specified quality level.

Often, the PR is specified as 5 percent, meaning that the probability of rejecting a lot with PRQ is at most five. Similarly, the consumer's risk is typically chosen as 10 percent, meaning that the probability of accepting a lot with CRQ is at most 10 percent. If any one of the four parameters is altered, the control of the producer's and consumer's risks will change.

In certain situations, such as characteristics relating to food safety where control of the CR is paramount, it might not be appropriate to take account of the PR in the design of sampling plans. This leads to two different options for the specification of risks.

Option 1: Plans that explicitly control both the CR and the PR:

- both the PRQ and CRQ, along with the respective allowable probabilities of incorrect rejection (PR) and incorrect acceptance (CR) are specified.

Option 2: Plans that explicitly control only the consumer's risk:

- plans for assessments of lots consisting of discrete items.

3.2.2 *Fitness for purpose*

Codex methods of sampling should be *'designed to ensure that fair and valid sampling procedures are used when food is being tested for compliance with a particular Codex commodity standard'.*[vii] When commodity committees have included sampling plans in a Codex commodity standard, these should be referred to CCMAS for endorsement along with relevant information relating to the sampling plan.

Sampling plans from other sources are still acceptable subject to their endorsement by CCMAS.

The *Principles for the Use of Sampling and Testing in International Food Trade* (CXG 83-2013)[3] states:

'Sampling and testing procedures are fit for purpose in a given product assessment, if, when used in conjunction with appropriate acceptance criteria, they have acceptable probabilities of wrongly accepting or wrongly rejecting a lot or consignment'.

Fairness

With regard to fairness, consideration of both the CR and the PR is necessary to avoid situations such as the following:

- sampling plans having inappropriate stringency, e.g. plans for the assessment of composition that are more stringent than for food safety;
- high producer or consumer risks that may arise due to the use of sampling plans not based on appropriate specifications of allowable risks; and
- sampling plans not based on statistically valid principles, e.g. ad hoc plans or plans that do not (properly) allow for measurement uncertainty.

In addition, in the interests of fairness, designers of plans should also take account of the measures that the producer may have to take to ensure compliance, given that it is usually not suitable for the producer to use the same sampling plan as that used by the consumer.

In selecting a sampling plan, it should be ensured that producers are not exposed to unreasonable costs in terms of sampling and testing, loss of yields, or excessive rejection of their products to achieve compliance.

Practicality

It is important to ensure that any sampling plan chosen will be practical to apply in terms of cost of sampling and testing and ease of use.

Other strategies could be used to develop sampling plans that are more economical in terms of sampling and testing, such as:

- managing average non-compliance rates over the medium to long term, rather than possibly paying a high premium in terms of testing costs for high levels of assurance on a lot-by-lot basis;
- the use of 'indifference' plans that are designed around the 'indifference quality level' (IQL), the level of defects at which there is 50 percent acceptance, rather than based on PRQ and CRQ. This leads to plans having more manageable sample sizes; and
- offsets, sometimes called guard-bands or buffers, between the limits used in the acceptance criteria and the actual specification limits for a provision can be used to reduce CR and to mitigate possibly unreasonably high sample numbers. However, offsets should be used with caution in the interest of fairness to producers.
-

[vii] Section 2: Elaboration of Codex Texts: Principles for the Establishment or Selection of Codex Sampling Procedures: Purpose of Codex Methods of Sampling (*Codex Procedural Manual*, latest edition).

3.2.3 *Specification limits*

For a given characteristic, a specification limit may be expressed as a minimum or a maximum limit (or both) applied either to each individual item in a lot, or to the average level.

Specification limits should apply to the 'true' values of the characteristics rather than to the measurements themselves. It follows that the assessments of lot compliance should also be in terms of the 'true' values of the characteristic within the lot (see Section 5.2.1).

Offsets

It is important to consider whether a given specification limit has an in-built offset (guard-band), and whether the offset reflects the measurement uncertainty associated with a particular sampling plan, that might include both analytical and sampling uncertainties.

Many provisions for chemical and microbiological contaminants have in-built offsets between the specification limits and the levels of contamination at which foods might become unsafe to consume. In such cases one may not need to design plans to provide high levels of protection against exceeding the limits as the CR is already well controlled by these offsets.

The use of offsets enables a reduction in sample size; for example, while large sample sizes are needed to show that a lot contains no more than say 1 percent nonconforming product, much smaller sample sizes are required to show that no more than 10 percent of the product in a lot exceeds a tightened limit.

3.2.4 *Lot homogeneity*

Acceptance sampling plans are usually based on the assumption that lots are homogeneous; indeed, the international definition of a lot is 'a quantity of product produced under conditions presumed uniform'.

In these guidelines, the term 'homogeneous' does not mean that the characteristic of interest does not vary within the lot. Rather, the term 'homogeneous' means that it is possible to characterize the variation of the characteristic of interest within the lot by means of a single standard deviation. Homogeneity applies only to variables plans.

In considering homogeneity, one needs to draw a distinction between:

- the type (shape) of the distribution, (e.g. *normal* distribution); and
- the *spatial distribution* of the characteristic within the lot.

If the lot consists of discrete items and if random sampling is used (as recommended for all plans in these guidelines), then the spatial distribution does not matter and the lot cannot always be considered homogeneous.

For this reason, if no prior information regarding the spatial distribution is available, then random sampling should be performed.

On the other hand, if prior knowledge indicates that the spatial distribution of the characteristic within the lot is random, then random sampling is not required. This case corresponds perhaps to the intuitive understanding of what homogeneity means in the context of acceptance sampling.

If random sampling cannot be performed, then the lot can only be considered homogeneous to the extent that the spatial distribution is random. In this sense, if random sampling cannot be performed, the homogeneity of the lot depends on the spatial distribution.

For some lots consisting of bulk material, inhomogeneity means that several segments must be sampled from.

Sections 0 and 0 provide further guidance regarding the inspection of inhomogeneous lots consisting of bulk materials or discrete items, respectively.

3.2.5 *Distribution of the characteristic*

The options for sampling plans depend on whether the test results are measurements (variables data) or have nominal outcomes (attributes data). In some cases, variables data can be classified as binary outcomes, but this should only be done after careful consideration of the sampling options available as the sample size for attributes inspection can be much larger than for variables data.

In the case of variables data, the assumed statistical distribution of the measurements in the lot should also be specified, i.e. whether the characteristic is normally distributed, a compositional proportion, or follows some other distribution. If it is not possible to make an assumption regarding the distribution of the data, results can be classified as attributes (as long as measurement uncertainty is negligible [refer Section 3.2.8]), or plans based on the fractional nonconformance (FNC) method can be used (as long as measurement uncertainty is non-negligible [refer Section 5.2.6]).

However, the characteristic does not have to follow the assumed distribution exactly (and, in any case, it is difficult to verify conformance to a distribution based on a small sample size). In practice, it is sufficient that the assumed distribution provides a satisfactory model for the behaviour of the characteristic in the lot. However, if the actual distribution in the lot differs markedly from the assumed distribution, then the producer's and consumer's risks may exceed the allowed levels specified in the design of the plan.

A typical 'default' assumption in variables plans is that the characteristic follows a normal assumption.

It is important to note that in the case of attributes plans, the binomial distribution is always available as the 'default' assumption, and that departures from this assumption regarding the type (shape) of the distribution will have very little impact on the producer's and consumer's risks.

Sections 0 and 0 provide further guidance regarding the inspection of inhomogeneous lots consisting of bulk materials or discrete items, respectively.

Prior knowledge of the distribution of a characteristic

In acceptance sampling, acceptance/rejection of a lot is decided on the basis of a sample (the set of individual items or increments taken from the lot). The relationship between the probability of acceptance (upon application of a given sampling plan) and the quality level of the lot is determined on the basis of prior knowledge regarding the distribution of the characteristic within the lot.

This means that prior knowledge is required *even in connection with the inspection of isolated lots.* In other words, the inspection of isolated lots does not mean that no prior information is available. On the contrary, prior information is always required. Sometimes the prior information takes the form of (tacit) assumptions based on experience and expert judgement. For example, a typical 'default' assumption in variables plans is that a characteristic follows a normal distribution.

If the actual distribution in the lot differs markedly from the assumed distribution, then the producer's and consumer's risks may exceed the allowed levels specified in the design of the plan. There are two ways in which the actual distribution can differ from the distribution which was assumed on the basis of prior knowledge:

- the type (shape) of the distribution. For example, the assumption is that the distribution is normal whereas, in fact, the distribution is lognormal; and
- the parameters of the distribution. For example, it is assumed that the lot standard deviation is the same as the (underlying) process standard deviation, whereas in fact it is twice as large.

It is important to note that in the case of attributes plans, the binomial distribution is always available as 'default' assumption, and that departures from this assumption regarding the type (shape) of the distribution will have very little impact on the producer's and consumer's risks.

3.2.6 *Lot standard deviation*

In the context of these guidelines, the population under consideration is the lot itself rather than the underlying process. For this reason, the role which the *process* standard deviation $\boldsymbol{\sigma}$ plays in the ISO 3951[viii] standards is now played by the lot standard deviation. The lot standard deviation can be represented by either its true value $\boldsymbol{\sigma}$ (sigma) or by an estimate (often denoted $\boldsymbol{s}$) of $\boldsymbol{\sigma}$.

The lot standard deviation is relevant only for variables plans, particularly for characteristics that are normally distributed or follow distributions, such as the lognormal distribution,[ix] that are related to the normal distribution.

For a given characteristic, the lot standard deviation is a measure of the random variation of the characteristic within the lot under inspection. Its estimate, however, may be affected by components of analytical or sampling uncertainty.

It is expected that for isolated lots the lot standard deviation will usually be calculated from the test results obtained during the inspection. Notwithstanding, there are cases where the lot standard deviation may be known, especially when the lot has been produced by a process with a known process standard deviation. This can be adopted as lot standard deviation. In such cases, the sample size of the sampling plan can be considerably reduced.

If the process standard deviation is known, it is important to consider whether it was obtained on the basis of a sufficiently large number of data to ensure it provides a reliable characterization of the variation within the process.

[viii] See note iii above.

[ix] For lognormally-distributed characteristics, the logarithms of the 'measurements' are normally distributed.

Note: In acceptance sampling, the lot standard deviation is always based on a simple random sample. However, in principle, other sampling procedures may be applicable, such as those described in Annex C.2 of the EURACHEM[x]/CITAC guide to measurement uncertainty arising from sampling.[4] This guide describes several procedures for the calculation of sampling uncertainty. It does not describe procedures for acceptance sampling.

3.2.7 *Measurement uncertainty*

In connection with lot inspections, it is important to determine whether the analytical components of measurement uncertainty – including the uncertainty which arises from subsampling from the laboratory sample (see Section 5.2.6) – can be considered negligible. This is typically done by considering the ratio of the analytical uncertainty and the lot standard deviation. If the analytical component of measurement uncertainty cannot be considered negligible, it should be taken into consideration in the acceptance criterion.

Adjustment for the analytical component of measurement uncertainty in acceptance sampling is discussed in more detail in Section 5.

The lot standard deviation already represents the variation of the characteristic of interest within the lot and any further uncertainty arising from the sampling procedure. For this reason, in determining whether an adjustment is necessary, only the analytical component of measurement uncertainty needs to be considered.

The term *measurement error* should not be used, as the term has been superseded by the focus on uncertainty across JCGM,[xi] ISO and EURACHEM[xii] standards and guides, as reflected in the *Guidelines on Measurement Uncertainty* (CXG 54-2004)[2] and as adopted in the present guidelines.

3.2.8 *Lot size*

Lot size is not normally an input required for the design of sampling plans intended to control both the consumer's and producer's risks in acceptance sampling. However, specification of the lot size is required for attributes plans applied to small lots and it is an input in the sampling plans described in the ISO 2859[xiii] and ISO 3951[xiv] standards (see Sections 0, 0 and Appendix II).

[x] A network of organizations in Europe having the objective of establishing a system for the international traceability of chemical measurements and the promotion of good quality practices.
[xi] The Joint Committee for Guides in Metrology (JCGM).
[xii] See note x above.
[xiii] See note ii above.
[xiv] See note iii above.

4. SAMPLING PLANS

4.1 Selection of sampling plans

The following table provides direction to the relevant sections within these guidelines:

Table 2. Direction to the relevant part for the selection of sampling plans

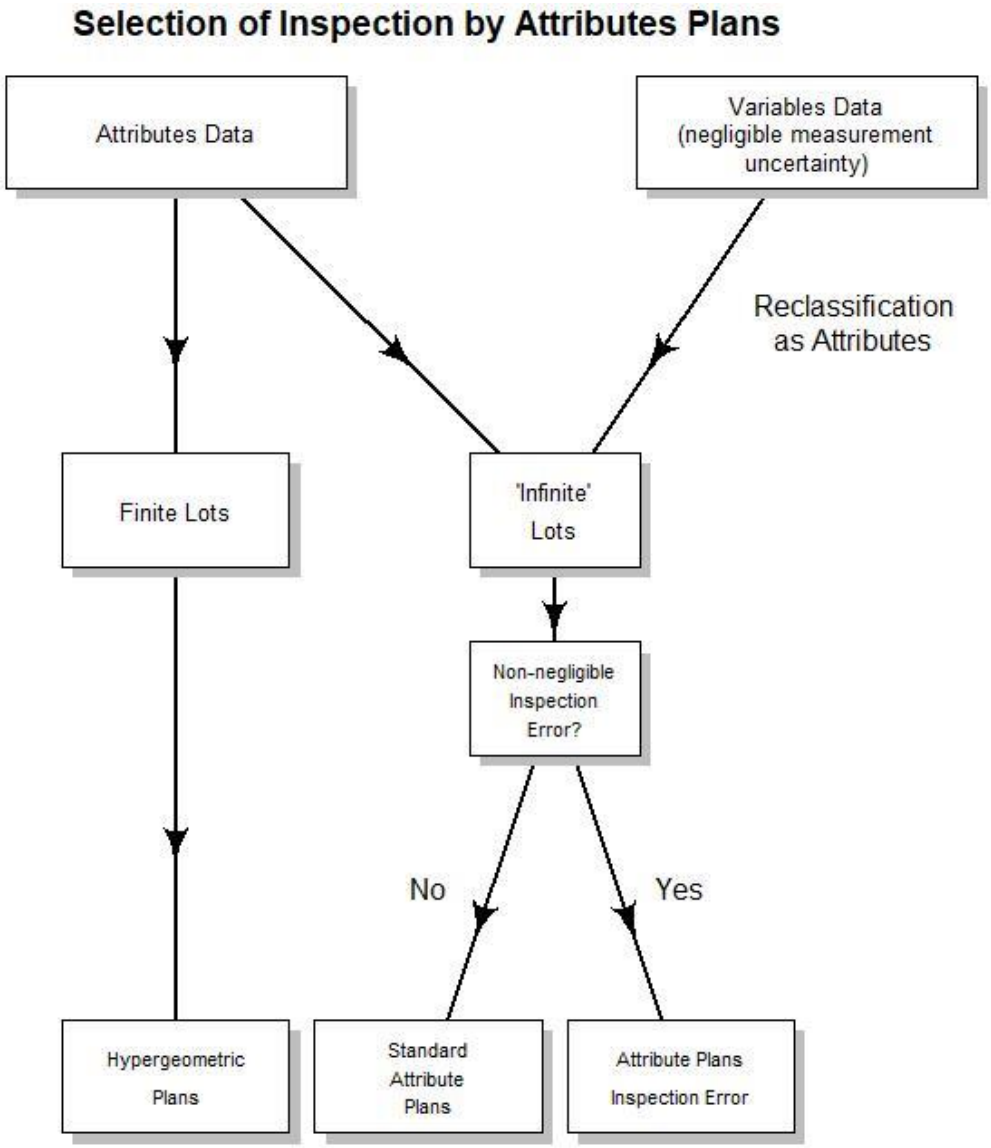
Homogeneous lots				
Data type	**Nature of provision**	**Distribution**	**Negligible** measurement uncertainty	**Non-negligible** measurement uncertainty
Attributes	Minimum or maximum	Not applicable	Inspection by attributes plans (Section 4.2) Appendix II Table 8.4.1	Known inspection errors (Section 5.1.1)
Variables	Minimum or maximum	Normal	Inspection by variables plans (Section 4.3) Appendix II Table 8.4.2	Repeatability error (no laboratory bias) (Section 5.2.6)
				General measurement uncertainty (Sections 5.2.5, 5.2.7, 5.2.8)
				Fractional nonconformance plans (Section 5.2.8)
	Minimum or maximum	Non-normal	Classification to attributes (Section 4.3.3)	Fractional nonconformance plans (Section 5.2.8)
Variables	Minimum or maximum	Compositional proportions	Plans for compositional proportions (Section 4.4.10)	Not included
	Average level	Not applicable	Plans for average level (Section 4.3.5)	Not included
Inhomogeneous lots (bulk materials)				
Attributes	Minimum or maximum	(blank)	Attributes plans (Section 4.4.6)	
Variables	Minimum or maximum	(blank)	Variables plans (Section 4.4.9)	
	Average level	Not applicable	Plans for average level (Section 4.4.8)	

4.2 Inspection by attributes plans

4.2.1 *Introduction*

These plans are usually referred to as attributes sampling plans. They are the simplest type of single sampling plan because the inspection results are classified into two possible outcomes – conforming or nonconforming. Because they are applicable to all sampling situations, they have become the benchmark that all other sampling plans can be compared against.

The following diagram shows the process for the selection of attributes sampling plans as it depends on the type of data and nature of the lot.



4.2.2 *Two-class attributes plans*

Two-class attributes plans are defined by two numbers: the sample size ***n***, the number of items to be taken from the lot under inspection and the acceptance number ***c***, the maximum number of nonconforming items allowed in the sample for acceptance of the lot. If the number of nonconforming items in the sample is less than or equal to ***c***, then the lot can be accepted. If the number of nonconforming items found is greater than ***c***, then the lot is rejected. In their most general form, the number of samples ***n*** and the acceptance number ***c*** for these plans are determined from specifications of the allowable consumer's and producer's risks. It should be noted that ***c*** need not be zero.

These plans can be used for either isolated lots or a continuing series of lots that consist of either discrete items or are bulk materials.

4.2.3 *ISO standards – attributes plans*

The ISO 2859[xv] series of standards provides sampling plans that are indexed by either CRQ or PRQ. The lot size is an input to the sampling plans in these standards as the sample size depends on the lot size.

The ISO 2859-2[5] plans are indexed by CRQ and are intended for the inspection of isolated lots consisting of discrete items. These plans are suitable for application in the field of food safety when it is not appropriate to explicitly control producer risks in the design of the plans.

Appendix II contains tables for inspection by attributes plans from ISO 2859-1.[6]

These plans are indexed by the PRQ.

[xv] See note ii.

4.2.4 *Plans for small lots (based on the hypergeometric distribution)*

If the sample size is large in relation to the lot size, some economy in the number of samples may be possible. As a rule, such economies are possible if the number of items, calculated assuming an infinite lot size, exceeds 10 percent of the lot size. For conceptually infinite lots, sampling plans based on the hypergeometric distribution are the same as the general two-class plans based on the binomial distribution.

4.2.5 *Zero-acceptance number plans*

Zero-acceptance number (ZAN) plans are a special case of two-class plans in which the acceptance numbers are set to ***c*** = 0. They are used in more critical situations such as for pathogens or foreign matter where only CR is considered directly and acceptance of lots demands that nonconforming items are not found in the inspection.

However, just because nonconforming items have not been found does not mean that they are not present in lots that have passed inspection. One disadvantage of ZAN plans is that they have poor discrimination between lots of good and poor quality, so they may not be generally applicable. The low sample numbers generally employed for microbiological applications enable high levels of consumer protection to be provided because of the offsets between the limits used in those plans and levels of contamination at which food might become unsafe (see Section 3.2.4).

ZAN plans for finite lots can also be designed based on the hypergeometric distribution.

4.2.6 *Three-class attribute plans*

In these plans, inspection results are classified into three classes, usually referred to as ‘good’, ‘marginal’ and ‘poor’ or ‘unacceptable’. This type of plan is frequently used in microbiological assessments. They have an advantage, relative to two-class plans, of providing better discrimination between good and poor quality; they have ‘steeper’ OC curves than two-class plans for the same number of samples.

Three-class plans are defined by four numbers (***n, c, m, M***) where:

- ***n*** is the number of samples to be taken;
- ***c*** is the maximum number of ‘marginal’ samples allowed for acceptance of the lot;
- ***m*** is the limit separating good quality from marginal quality samples;
- ***M*** is the limit above which samples are classified as ‘poor’; and
- samples with results lying between the numbers ***m*** and ***M*** are classified as marginal.

Lots are accepted provided that:

- none of the ***n*** samples is poor, having levels exceeding ***M***; and
- at most ***c*** of the samples are marginal, with levels between ***m*** and ***M***.

If ***m*** = ***M*** a three-class plan becomes a two-class plan.

Evaluation of these plans generally requires an assumption about the underlying distribution of the identified characteristic, such as the lognormal distribution for microbiological parameters. This might also apply to two-class plans, especially for microbiological plans.

Three-class plans for finite lots can be designed based on the hypergeometric distribution.

4.2.7 *Plans for variables data where an appropriate distribution is unknown*

If the underlying distribution of a measured characteristic within a lot is not known and we are not prepared to assume that the characteristic can be adequately described by the normal or some other distribution, then the only recourse available is to classify the results as conforming or nonconforming with respect to the specification limit and to use attributes plans. Note that this approach should be used only when measurement uncertainty is negligible.

4.2.8 *Attribute plans for multiple characteristics*

Attributes plans can be easily applied to multiple characteristics by classifying inspected items as nonconforming if any of the individual characteristics are nonconforming.

Obviously, it makes sense to apply a plan to multiple characteristics only if the individual characteristics are of similar ‘stringency’, i.e. if the same or similar plans would be used if the characteristics were inspected individually. These plans have the advantage, compared to the use of individual plans, of allowing better control of PR, of incorrectly rejecting lots of good quality.

4.3 Inspection by variables sampling plans

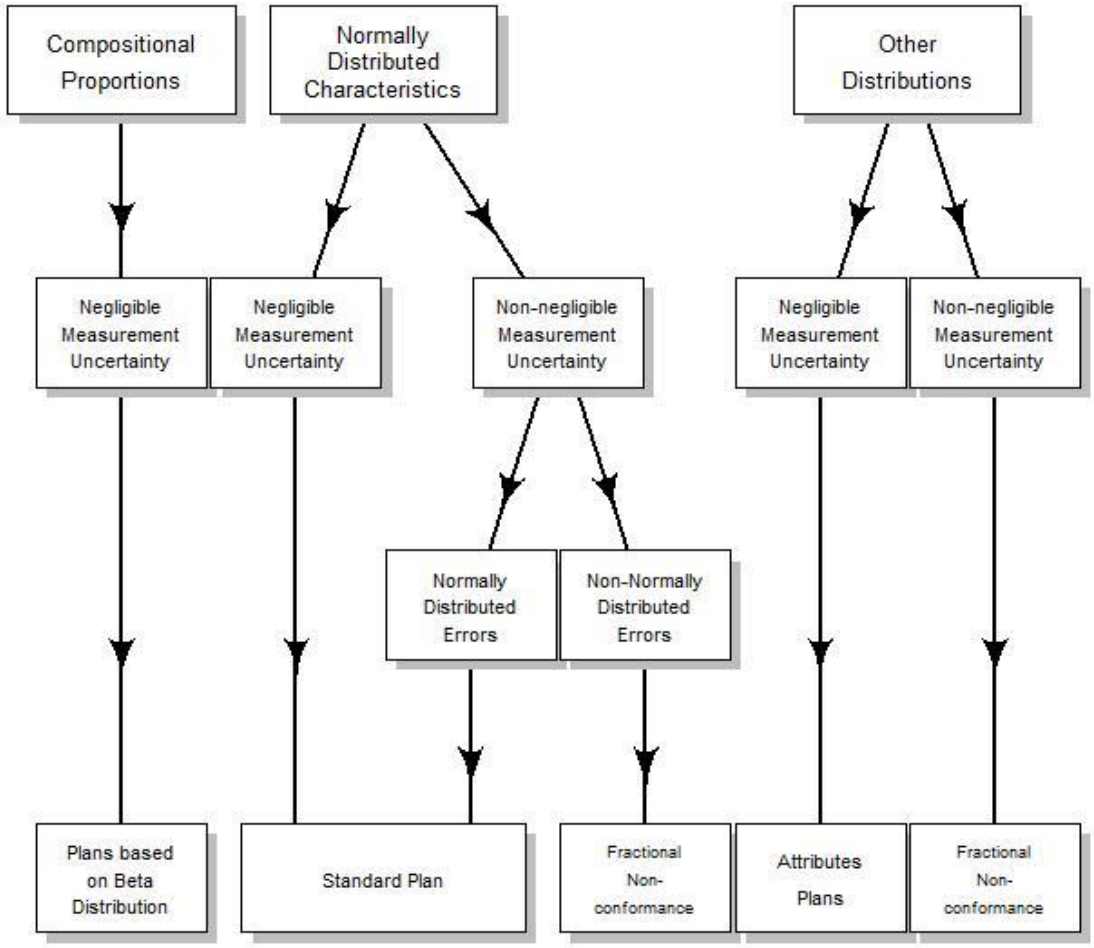
4.3.1 *Introduction*

If the underlying distribution of a measured characteristic is known, acceptance sampling can be performed directly on the measurements themselves. This often allows a considerable reduction in sample size.

For variables sampling plans, it is necessary to make an assumption regarding the distribution of the characteristic within the lot. While the normal (Gaussian) distribution is commonly adopted, for compositional proportions in bulk materials the beta distribution is more appropriate (though the normal distribution can serve as an approximation).

The following diagram shows the process for the selection of variables sampling plans:

Selection of Inspection by Variables Plans - Homogeneous Characteristics



4.3.2 *Advantages and disadvantages of variables sampling plans*

The advantages of variables sampling plans are:

- they offer the same protection with a smaller sample size than that required for attributes plans;
- there is feedback of data on the process which produced the units;
- there is more information available in waiver situations; and
- the extent of conformity of each unit is taken into account in the application of the plan.

The disadvantages are:

- the outcome is dependent on the appropriateness of the underlying distribution, that the assumed statistical distribution provides a satisfactory description for the behaviour of the characteristic within the lot;
- they are only applicable to one characteristic at a time;
- there may be a higher inspection cost per unit;
- a lot with no nonconforming units may be rejected by a variables plan, that occur when the average level lies too close to the specification limit, as measured in terms of the variation in the lot (lot standard deviation); and
- there is a possibility that no nonconforming units are found to show to the producer after rejection.

4.3.3 *Variables sampling plans*

Variables sampling plans are defined by two numbers: the sample size ***n***, the number of items to be taken from the lot under inspection, and the acceptability constant ***k***, the multiplier of the lot standard deviation ***S*** in the acceptance criterion.

A lot is accepted if $\bar{\boldsymbol{X}} + \boldsymbol{kS} \leq \boldsymbol{U}$ for an upper specification limit $\boldsymbol{U}$ or if $\bar{\boldsymbol{X}} - \boldsymbol{kS} \geq \boldsymbol{L}$ for a lower limit $\boldsymbol{L}$.

4.3.4 *ISO standards – variables plans*

The ISO 3951[xvi] standards provide sampling plans that are indexed by either CRQ or PRQ. The lot size is an input to the sampling plans in these standards as the sample size depends on the lot size.

The ISO plans indexed by CRQ are intended for the inspection of homogeneous isolated lots consisting of discrete items. These plans are more suited for provisions relating to food safety when it is not appropriate to explicitly control producer risks in the design of the plans.

Appendix II contains tables for inspection by variables plans from ISO 3951-1.[7] These plans are indexed by the PRQ.

The ISO 3951-6[8] standard also contains procedures that deal with non-negligible measurement uncertainty. This is discussed in more detail in Section 5.

4.3.5 *Plans for the average level in the lot*

In some cases, such as the net weight of packages, a limit applies to the average level, with the intention that the average level in the lot should not be less than the limit. In Codex, although an example of sampling plans for bulk materials, the plans for aflatoxins are also based on compliance of the average level. This is an example of the use of offsets (see Section 3.2.3).

It is usually assumed that the quality characteristic is normally distributed; the appropriateness of the distribution is less critical when compliance of the average level is being assessed. It is also usually assumed that there is a single specification limit, either a lower specification limit, ***L*** or an upper specification limit, ***U***.

When the lot standard deviation $\boldsymbol{\sigma}$ is known based on historical process data, the inspection plan for compliance of the average level to a minimum limit ***L*** is operated as follows:

1. take a random sample of size ***n*** and obtain the sample mean;
2. calculate $A = L + k \times \sigma$; and
3. if the sample mean $\bar{x} > A$ accept the lot; otherwise reject the lot.

The parameters of the plan are ***n*** and ***k***. Note that ***k*** does not denote the same quantity as in the usual variables plans. When the lot standard deviation $\boldsymbol{\sigma}$ is unknown, it is replaced with the sample standard deviation ***s***. The OC curve for this plan is less discriminatory than the plan when the standard deviation $\boldsymbol{\sigma}$ is known, and a greater sample size will be required to provide equivalent discrimination to that provided when the standard deviation is known.

[xvi] See note iii above.

4.4 Sampling of bulk materials

4.4.1 *Introduction*

Bulk materials are continuous, consisting for example of particles of different densities and sizes. It is impossible to consider a lot of a bulk material as a set of discrete items because there is no way of selecting the items in a way that is not biased when using simple random sampling.

Some general objectives of bulk sampling are:

- acceptance on a lot-to-lot basis;
- characterizing the material as to grade,[xvii] any need for further processing, and its destination;
- determination of weight or content for purposes of payment;
- determination of properties that must be known so that the end use will be appropriate; and
- experimentation and analysis to determine further sampling procedures and uses of the material.

Sampling units are created at the time of sampling by means of some kind of sampling device. The sampling units change depending on different factors such as how the device is employed, and the conditions that the device is used under.

In bulk sampling, a lot is seen as being composed of mutually exclusive segments.

Sometimes the segments are obvious, such as when the material comes in boxes or bags.

Other times the segments are not obvious, and so they have to be artificially created. One way of doing this is by superimposing imaginary grids over the material.

4.4.2 *Theory of sampling*

The theory of sampling provides a comprehensive approach to the design of sampling procedures, the aim of which is to obtain a sample for laboratory analysis whose composition is an unbiased estimate of the average level of a lot. However, this sample would not, by itself, be useful for assessing conformance of a lot to minimum or maximum specification limits as an additional allowance is required to compensate for variation in the lot to enable such assessments to be made.

4.4.3 *Terminology*

The special nature of sampling for bulk materials has led to the use of specific terminology, although this terminology varies between different fields, between authors, and also between different Codex committees. The *General Standard for Contaminants and Toxins in Food and Feed* (CXS 193-1995)[9] uses the following terminology.

Table 3. Bulk material terminology for sampling plans

Lot	An identifiable quantity of a food commodity delivered at one time and determined by the official to have common characteristics, such as origin, variety, type of packing, packer, consignor or markings.
Sublot	Designated part of a large lot in order to apply the sampling method on that designated part. Each sublot must be physically separate and identifiable.
Sampling plan	It is defined by a test procedure and an accept/reject limit. A test procedure consists of three steps: sample selection, sample preparation and quantification. The accept/reject limit is a tolerance usually equal to the Codex maximum level.
Incremental sample	A quantity of material taken from a single random place in the lot or sublot.
Aggregate sample	The combined total of all the incremental samples taken from the lot or sublot.
Laboratory sample	The smallest quantity of a food commodity comminuted in a mill or homogenized in an appropriate device. The laboratory sample may be a portion of, or the entire aggregate sample. If the aggregate sample is larger than the laboratory sample, the laboratory sample should be removed in a random manner from the aggregate sample in such a way to ensure the laboratory sample is still representative of the sublot sampled.

[xvii] Foods and other materials are often ranked according to their quality, with the different quality levels sometimes known as grades.

Test portion	A portion of the comminuted/homogenized laboratory sample. The entire laboratory sample should be comminuted in a mill or homogenized in an appropriate device. A portion of the comminuted/homogenized laboratory sample is randomly removed for analysis.

4.4.4 *Design of general sampling plans for bulk materials*

In the simplest case, such as the inspection of bulk materials of manufactured products, lots can often be considered homogeneous, allowing the standard attributes or variables plans to be used, with adjustment for analytical measurement uncertainty where appropriate.

On the other hand, some bulk materials, such as shipments of grains or other raw materials, cannot be considered homogeneous (see Section 0). Special techniques are required for this situation, but the statistical methods are complex and only an overview is provided in these guidelines.

Lot homogeneity is difficult to verify for bulk materials and generally requires large numbers of samples. Moreover, it is often difficult to perform random sampling from an entire lot of a bulk material. As a precaution, in cases where lot homogeneity can be neither assumed nor verified, lots should be treated as inhomogeneous.

The general approach to sampling inhomogeneous lots of bulk materials is that a lot is considered as a set of smaller segments (strata) each of which is more homogeneous than the entire lot. This allows the usual sampling procedures based on random sampling to be applied within each segment as inhomogeneity within each segment will have less effect.

The basic sampling and inspection procedure can be described as follows:

- segments, from which increments are to be taken, are chosen at random;
- several increments are chosen at random from each of the chosen segments;
- the increments from each segment can sometimes be combined to form a composite sample, which is thoroughly mixed;
- one or more subsamples are taken from each composite sample;
- these subsamples are tested; and
- acceptability of the lot is decided based on an acceptance criterion.

4.4.5 *Attributes plans for bulk materials*

The following points need to be considered in the design of attributes plans for bulk materials:

- inhomogeneity will be present and hence the standard attribute sampling plans for homogeneous lots will not be suitable as they do not provide adequate protection for consumers;
- inhomogeneity can be overcome either by allowing for the correlation within the batch in the design of the sampling plan or, alternatively, by splitting the lot into more homogeneous segments, and using stratified sampling techniques. Either way, a preliminary study is needed to estimate the correlation and the variation between segments; and
- the proposed plans should be validated using different statistical models for the behaviour of the level nonconforming within the lot, to ensure robustness against different levels of correlation.

4.4.6 *Variables plans for bulk materials*

Typically, the total observed variation within a lot of bulk materials consists of several components due to variation between and within segments, due to sample preparation (e.g. including subsampling), testing and other causes.

Sampling plans for bulk materials, especially cost-optimal sampling plans, can be designed most effectively with prior knowledge of the different components of variation that exist within lots; it is desirable that a preliminary investigation of the variation is carried out prior to the development of any plans.

A minimum of ten samples per segment is recommended to estimate the within lot variability, if the acceptance criterion involves averaging of multiple test results, laboratory samples should be tested at least in duplicate to allow estimation of the repeatability component of measurement uncertainty, unless an estimate is available from other sources such as a method validation study.

Example
The *General Standard for Contaminants and Toxins in Food and Feed* (CXS 193-1995)[9] shows the breakdown of the total variation for aflatoxins in tree nuts, with a focus on sampling, sample preparation and testing; the variation due to sampling includes both between and within segment variation. It should be noted that provisions for aflatoxins are expressed in terms of the average levels in a lot.

Table 1. Variances[a] associated with the aflatoxin test procedure for each treenut

Test procedure	Almonds	Hazelnuts	Pistachios	Shelled Brazil nuts
Sampling[b,c]	$S^2_s = (7\,730/ns)\ 5.759C^{1.561}$	$S^2_s = (10\,000/ns)\ 4.291C^{1.609}$	$S^2_s = 8\,000/ns)\ 7.913C^{1.475}$	$s_s^2 = (1\,850/ns)\ 4.8616C^{1.889}$
Sample Prep[d]	$S^2_{sp} = (100/nss)\ 0.170C^{1.646}$	$S^2_{sp} = (50/nss)\ 0.021C^{1.545}$	$S^2_{sp} = (25/nss)\ 2.334C^{1.522}$	$s_{ss}^2 = (50/nss)\ 0.0306C^{0.632}$
Analytical[e]	$S^2_a = (1/na)\ 0.0484C^{2.0}$	$S^2_a = (1/na)\ 0.0484C^{2.0}$	$S^2_a = (1/na)\ 0.0484C^{2.0}$	experimental $s_a^2 = (1/n)\ 0.0164C^{1.117}$ or FAPAS $s_a^2 = (1/n)\ 0.0484C^{2.0}$
Total variance	$S^2_s + S^2_{sp} + S^2_a$	$S^2_s + S^2_{sp} + S^2_a$	$S^2_s + S^2_{sp} + S^2_a$	$S^2_s + S^2_{sp} + S^2_a$

Notes:
a. Variance = S^2 (s, sp and 'a' denote sampling, sample preparation and analytical steps, respectively, of the aflatoxin test procedure).
b. ns = laboratory sample size in the number of shelled nuts, nss = test portion size in grams, na = number of aliquots quantified by HPLC, and C = aflatoxin concentration in µg/kg total aflatoxin.
c. Shelled nut count/kg for almonds, hazelnuts, pistachios and Brazil nuts is 773, 1 000, 1 600 and 185, resepctively.
d. Sample preparation for almonds, hazelnuts, and pistachios reflect Hobart, Robot Coupe, Marjaan Knatman and Turrax type mills, respectively. Laboratory samples were dry ground into a paste for each treenut except for Brazil nut that were prepared as a slurry Brazil nut/water 1/1 w/w.
e. Analytical variances reflect FAPAS recommendation for upper limit of analytical reproducibility uncertainty. A relative standard deviation of 22 percent, which is based upon FAPAS data, is considered, as an appropriate measure of the best agreement that can be obtained between laboratories. An analytical uncertainty of 22 percent is larger than the within laboratory uncertainty easured in the sampling studies for the four treenuts.

Since bulk materials are continuous, parts of each sample can be mixed to form a composite sample. This composite is then tested only once, rather than having to perform many tests on the individual samples. This is a physical way of creating a sample representing the average content per lot or segment. This averaging causes a reduction in the apparent variation, therefore adjustment of the acceptance criterion may be required for assessments against minimum or maximum limits.

Note however, that the use of composite sampling adds complexity to the design of a general sampling strategy due to the statistical complexity of modelling the mixing process; assuming that composites made up from many individual portions can be thoroughly mixed is possibly unrealistic.

4.4.7 *Variables plans for the average level*

Sampling plans for bulk materials are often used to assess compliance of the average level of a characteristic. In some cases, such as in the sampling plans for aflatoxins in the *General Standard for Contaminants and Toxins in Food and Feed* (CXS 193-1995),[9] these plans are used in conjunction with offsets (see Section 3.2.3) to provide consumer protection.

Other procedures for the inspection of the average level of a lot such as those in ISO 10725[10] are available that consider costs to derive plans that are economical to apply, although these plans might not be suitable in cases where a more precise determination of the average level is required.

Plans for the average level might also be applicable where the product is homogenized through blending or further processing.

4.4.8 *Variables plans for percentage nonconforming (minimum or maximum limits)*

The strategy is similar to the design of variables plans for the average level except that an additional allowance should be made for variation within the lot, obtainable from the statistical analysis described in Section 4.4.5. A simpler approach is to estimate within lot variation as the variation among the segments by taking one sample from each segment and testing those samples in duplicate to allow adjustment for measurement uncertainty, although this will not provide any information on other components of variation:

- the acceptance criterion has the same form as a conventional variables plan applied to homogeneous lots; and
- the number of samples ***n*** and the acceptability constant ***k*** can be found by trial and error, assessing the probabilities of acceptance against various alternative models for the behaviour of the characteristic in the lot. This should recognize that the formation of the segments might not reflect the disposition of nonconforming product within the lot.

4.4.9 *Variables plans for compositional proportions (measurement uncertainty negligible)*

Compositional characteristics are often quality measures for bulk materials. For example, the milkfat percentage with a minimum limit of 26 percent is a primary quality measure for whole milk powders.[xviii]

Compositional proportions, also referred to as mass fractions, are characterized by units of measure such as percent (of mass), mg/kg, µg/100 g and the like, which are, strictly speaking, 'dimensionless' numbers lying between 0 and 1.

Compositional proportions can be modelled using the beta distribution. Variables sampling plans based on the normal distribution can only be approximate for compositional proportions and can lead to a higher CR than desired.

Sampling plans for compositional proportions are defined by two parameters: ***m***, the number of samples to be taken from the lot, and ***k***, the acceptability constant defined in the same way as for the usual variables sampling plans. In order to design such plans, in addition to PRQ, CRQ, etc., an estimate of the 'precision parameter' for the beta distribution, denoted by $\boldsymbol{\theta}$, is required. This estimate can be obtained from the analysis of historical data.

When using these plans, the ***m*** samples are taken from the lot and can be tested individually or combined (blended, well mixed, etc.) to form a composite sample that needs to be tested only once.

The average level ***P*** is taken as either the average of the ***m*** results from the testing of the individual samples or the single result from the testing of the composite sample.

A feature of the beta distribution is that its standard deviation depends on the average level, enabling an assessment to be conducted using a single test of a composite sample taken from the lot. The standard deviation is calculated using the formula:

$$s = \sqrt{P(1-P)/\theta}$$

where $\boldsymbol{\theta}$ is the precision parameter for the beta distribution, estimated from historical data (see above).

The lot is accepted against an upper limit $\boldsymbol{U}$ provided $P + k \times s \leq U$ and similarly for a lower limit.

5. INSPECTION ERROR AND MEASUREMENT UNCERTAINTY

Inspection error relates to inspection by attributes, and measurement uncertainty relates to inspection by variables.

Non-negligible analytical measurement uncertainty and inspection error have the potential to affect the probabilities of acceptance of a sampling plan. Accordingly, non-negligible analytical measurement uncertainty or inspection error should be taken into account in sampling inspection.

It has been shown theoretically that analytical measurement uncertainty and inspection errors affect PR more than they affect CR, i.e. the increase in PR (rejecting a lot of acceptable quality) exceeds the increase in CR (accepting a lot of unacceptable quality). Accordingly, in the interests of fairness, it is important that appropriate allowances are made for non-negligible measurement and inspection errors.

Acceptance sampling plans can be designed to allow for non-negligible analytical measurement uncertainty and inspection error.

[xviii] *Standard for Milk Powders and Cream Powders* (CXS 207-1999).

5.1 Attributes plans

In the context of attributes plans, 'inspection error' refers to random errors of misclassifying conforming items as nonconforming and vice versa.

Inspection errors occur when testing an item for conformance and can be caused by human error, instrument error, or any other measurement-related errors.

There are two types of inspection errors:

- Type I errors (***e₁***) occur when conforming items are classified as nonconforming.
- Type II errors (***e₂***) are when nonconforming items are classified as conforming.

When inspection errors are present, they generally cause a greater increase in producer's risk than CR. For a single sampling plan, Type I errors (**e_1**) have a greater effect on the OC curve than Type II errors (**e_2**).

The true fraction nonconforming ***p*** and the observed fraction nonconforming ***p_e*** are related through the following equation:

$$p_e = e_1(1-p) + (1-e_2)p$$

The impact of inspection errors is particularly marked for ZAN plans.

5.1.1 *Known inspection errors*

If the misclassification errors are known, if precise estimates of the misclassification errors are available, for example from a method validation study, the estimates of the Type I and Type II errors can be used to design a sampling plan to control producer's and consumer's risks to specified levels. This will inevitably lead to increased sample sizes.

5.2 Variables plans

Measurement uncertainty provides information regarding the range of values that could reasonably be attributed to the measurand. As such, it constitutes an important measure of the quality or reliability of a test result.

For a more comprehensive discussion of measurement uncertainty, refer to the *Guidelines on Measurement Uncertainty* (CXG 54-2004).[2]

It should be noted that the concept of measurement uncertainty as usually understood (and as discussed in the *Guidelines on Measurement Uncertainty* [CXG 54-2004])[2] relates to a single determination performed on a single sample. This is appropriate for conformity assessment, but not for acceptance sampling (see Section 2.2). The same holds for the procedure illustrated in Figure 1 in the *Guidelines on Measurement Uncertainty* (CXG 54-2004).[2] In connection with acceptance sampling, it is important to take into account how the different measurement uncertainty components manifest themselves in the sampling and calculation procedures applied. This is discussed in Section 0, below.

The terms 'negligible' and 'non-negligible'[xix] are used to indicate whether or not allowances should be made for measurement uncertainty in acceptance sampling plans. In the ISO 3951[xx] series, measurement uncertainty is considered non-negligible if it is greater than 10 percent of the *process* standard deviation (SD). In connection with the inspection of isolated lots, the same criterion can be applied, but replacing the *process* SD with the *lot* SD (see Section 3.2.6). However, the only definitive way to assess whether an adjustment for measurement uncertainty is required is to examine the OC curve for the proposed sampling plan in the presence of measurement uncertainty (see Section 2.3.1).

5.2.1 *Measurement uncertainty*

In order to clarify the role of measurement uncertainty in acceptance sampling, it is necessary to draw a distinction between *analytical* measurement uncertainty and the *sampling component* of (the total) measurement uncertainty. We start by reproducing the following definition from Section 8 in the *Guidelines on Measurement Uncertainty* (CXG 54-2004):[2]

A laboratory sample is a sample as prepared (from the lot) for sending to the laboratory and intended for inspection or testing.

[xix] The term 'significant' is also used.
[xx] See note iii above.

Any sources which contribute to measurement uncertainty prior to the arrival of the laboratory sample in the laboratory can be considered components of sampling uncertainty:

- the sampling procedure and its implementation;
- the variation of the characteristic of interest within the lot;
- the person(s) performing the sampling;
- subsampling steps (leading to the laboratory sample); and
- contributions due to storage and transportation conditions (prior to the arrival of the laboratory sample in the laboratory).

Any sources which contribute to uncertainty within the laboratory can be considered components of analytical measurement uncertainty, for example:

- subsampling steps performed on the basis of the laboratory sample, such as taking a test sample, test portion, etc.;
- sample preparation;
- contributions due to storage conditions (in the laboratory);
- analytical steps; and
- laboratory procedures.

In determining measurement uncertainty, it is important to take account of all relevant contributions, including all sampling and analytical sources.

Role of measurement uncertainty in acceptance sampling

In acceptance sampling, the aim is to decide whether to accept or reject the lot under inspection via the application of an acceptance criterion. The application of the acceptance criterion often includes an estimate of the lot SD, which is a measure of the random variation of the characteristic within the lot under inspection. It is important to ensure the estimate of the lot SD is not affected by uncertainty sources. Accordingly, the role of measurement uncertainty in acceptance sampling can be described as follows:

> *Measurement uncertainty may affect the estimate of the lot SD. If this effect is non-negligible and thus impacts the consumer and producer risks, then the estimate of the lot SD must be corrected for the non-negligible measurement uncertainty.*

In theory, the estimate of the lot SD can be affected by both sampling and analytical components of measurement uncertainty. It should be noted, however, that while analytical uncertainty will always inflate the lot SD estimate, the effect of sampling components can be either to increase or decrease its value. For this reason, correcting the estimate of the lot SD for analytical uncertainty will always consist in 'subtracting' the uncertainty contribution and can thus be considered more readily achievable than a correction for sampling uncertainty components. The focus in this guidance document thus lies in correcting for non-negligible *analytical* uncertainty. Notwithstanding, it should be ensured sampling procedures are adequate. The use of statistically-based random sampling or validated sampling procedures is desirable. It should also be noted that any impact of analytical or sampling uncertainty on the lot SD estimate can be disregarded as long as the corresponding SD is less than 10 percent of the lot SD.

Procedures for correcting the lot SD for non-negligible analytical measurement uncertainty and sampling uncertainty are discussed in the following sections (see Section 5.2.6).

5.2.2 *General discussion of bias*

Measurement uncertainty consists, on the one hand, of components that reflect random effects (varying randomly with each test result) and, on the other hand, of components that reflect systematic effects (remaining constant across test results).

A systematic effect is commonly referred to as a bias.

In principle, if a bias is observed, it is corrected for; and it is the *uncertainty of the bias correction* which is taken into account in the measurement uncertainty.

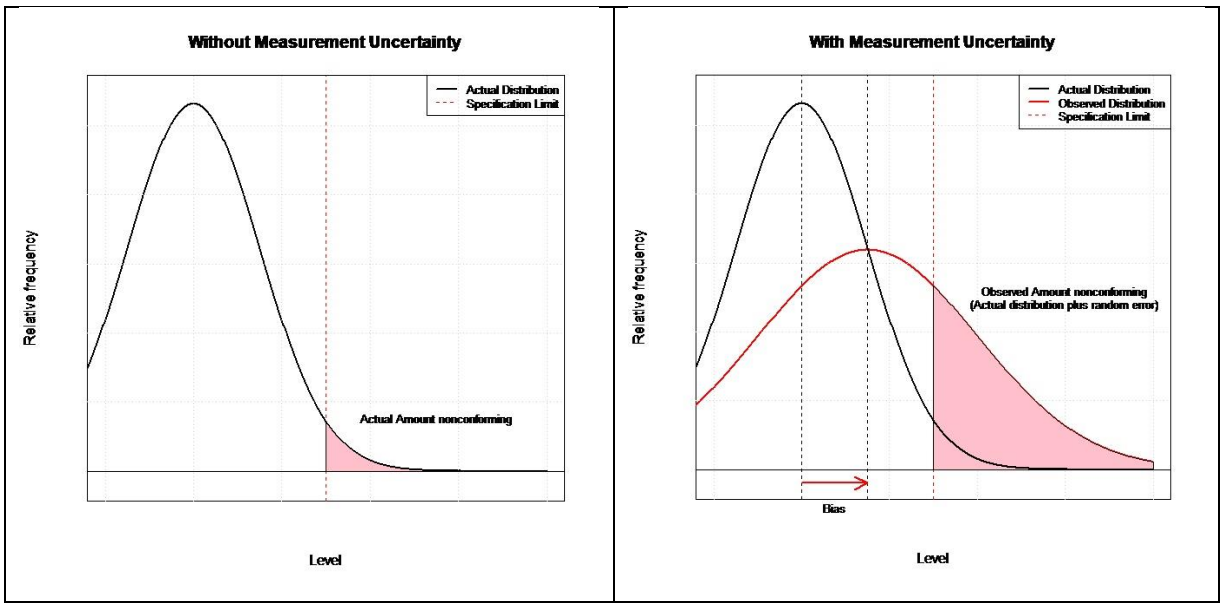
In practice, a bias may affect test results even after a bias correction is performed. This is the case, for example, if the bias correction is adequate for a given matrix, but not for another.

There may be various sources of bias. The analytical method itself may have a bias. In addition, the method bias may vary from one matrix to the next. In this sense, matrix effects (or a 'matrix bias') may be observed.

The method bias may vary from one laboratory to the next. In this sense, laboratory effects (or a 'laboratory bias') may be observed. Finally, there may also be a sampling bias, e.g. a given sampling procedure may consistently underestimate the lot mean or the lot SD.

It is often possible to obtain an estimate of the magnitude of a bias even in the absence of information regarding the 'true value'. For instance, the 'between-laboratory' component of reproducibility precision, calculated on the basis of data from a collaborative study, and typically expressed as a SD, characterizes the magnitude of the laboratory bias. Similarly, there are procedures for estimating laboratory bias on the basis of quality control data or proficiency test results which can be used to characterize the magnitude of the laboratory bias.

The following diagram shows the distribution and the percent nonconforming in a lot in the case where neither random effects nor bias (referred to as an 'error-free' plan), and the effect which random effects and bias can have on the observed distribution and the apparent percentage nonconforming in a lot. This diagram thus shows the effect that random effects and bias can have on the probability of acceptance of a lot, unless such effects are adequately accounted for.



5.2.3 *Top-down approach for determining measurement uncertainty: the ISO 5725-2[11] model*

In many cases, an estimate of analytical measurement uncertainty is supported by precision data from an inter-laboratory method validation study (collaborative study) calculated on the basis of the simple design from the ISO 5725-1[12] and ISO 5725-2[11] standards. This design allows two precision components to be calculated:

- one component reflecting *random effects* under near identical conditions within a given laboratory, referred to as the repeatability component; and
- one component reflecting *laboratory bias*, referred to as the between-laboratory component.

The underlying statistical model is not the most general model,[xxi] but many collaborative studies are conducted in accordance with ISO 5725-2[11] For this reason, the following sections will return to the two components of the ISO 5725-2[11] design.

Note:The 'between-laboratory' component in ISO 5725-2[11] characterizes the range of laboratory bias under repeatability conditions. ISO 5725-3 [13] includes other designs, which allow a separate estimation of repeatability precision, intermediate precision (factorial effects) and residual laboratory bias.

5.2.4 *The acceptance criterion*

The acceptance criterion in a variables plan often takes the form:

$$\bar{x} + k \cdot s \leq USL,$$

where $\bar{x}$ is the average value of the test results obtained from the inspection, s is their standard deviation and USL denotes the upper specification limit.

Ideally, the standard deviation s is a reliable measure of the variation of the characteristic of interest within the lot. However, in practice, s may include other components, such as, analytical measurement uncertainty.

[xxi] For common top-down approaches, see the *Guidelines on Measurement Uncertainty* (CXG 54-2004).

- The mean value $\bar{x}$ is calculated from several test results. When taking measurement uncertainty into account in the acceptance criterion, it is thus necessary to consider how averaging affects the different components of analytical measurement uncertainty.

As far as the two components from the ISO 5725-2[11] model discussed above:

- averaging across n test results will reduce the repeatability component by a factor of $sqrt(n)$; and
- however, averaging across n test results will not reduce the between-laboratory component.

In the absence of fundamental variability, the lot standard deviation from a *single* test result obtained from a *well-mixed* composite sample obtained from n increments is reduced by $sqrt(n)$.

5.2.5 *Laboratory bias in acceptance sampling*

In connection with acceptance sampling, the following should be noted:

- If information regarding laboratory bias is available in the form of a between-laboratory SD from an interlaboratory study conducted according to ISO 5725-2,[11] then measurements during lot inspection should be performed under repeatability conditions, with the bias, represented by the between-laboratory SD, taken into account in the sampling plan.
- Matrix effects (variation of bias across matrices within the scope of the method) can affect the test results differently in different laboratories (see the *Guidelines on Measurement Uncertainty* [CXG 54-2004]),[2] Sections 10, 12 and 15. This means that an estimate of the between-laboratory variation may be valid for a given matrix, but not for another. An estimate of the bias across different matrices can be obtained by means of an in-house experiment. If such an estimate is available, it should be taken into account in the sampling plan.

If an estimate of the between-laboratory SD is available, it is important to consider whether it constitutes a reliable characterization of the variation of laboratory bias, in the sense that the estimate was obtained on the basis of data from a sufficiently large number of laboratories (see the *Guidelines on Measurement Uncertainty* [CXG 54-2004][2] Sections 16, 17 and 18).

5.2.6 *Within-item variation*

For lots consisting of discrete items, one uncertainty source deserves special attention: **within-item variation**. Typically, one measurement value is obtained per item, and the lot SD is calculated on the basis of these item-specific values. Each measurement value is intended to represent the mean concentration of the given item. However, the lot SD calculated in this manner may be inflated by within-item variation. There are two cases to consider.

Case 1 – subsampling prior to the arrival of the sample in the laboratory

In this scenario, there is a subsampling step between item selection and the arrival of the laboratory sample in the laboratory, and this subsampling step causes non-negligible deviations between laboratory samples from one and the same item (if several laboratory samples were taken from the same item). Note that in this case, the lot SD will be inflated by a sampling (rather than an analytical) component of measurement uncertainty. Correcting for this type of overestimation of the lot SD presents practicability issues and is not typically contemplated. This case is mentioned here merely for the sake of completeness.

Case 2 – subsampling within the laboratory

In this scenario, subsampling inside the laboratory causes non-negligible deviations between test portions taken from the same laboratory sample (item). Conceptually, this component belongs to analytical rather than sampling measurement uncertainty. An estimate thereof can be obtained via a ‘duplicate’ experimental design, where two test portions per laboratory sample (item) are analysed. If a validation study is conducted on the basis of certified reference material, it may not be possible to obtain an estimate of this component. Moreover, depending on the context, this component may or may not be considered to belong to a given method’s precision. Accordingly, in some cases, an estimate for this component may not be available at all, or may only be available via studies conducted to determine sampling uncertainty rather than analytical uncertainty.

5.2.7 *Absence of laboratory bias*

In order to ensure unbiased estimates, the estimate of the lot SD must be corrected for any unwanted measurement uncertainty and subsampling components (as described under Case 2 in the previous section). In the absence of laboratory bias, it is possible achieve this via a relatively simple procedure.

If it can be assumed that:

- there is negligible bias;
- the characteristic follows a normal distribution in the lot under inspection; and
- repeatability effects follow a normal distribution;

then the following approach can be applied.

The standard deviation $\boldsymbol{s}$ is adjusted by 'subtracting' the standard deviation representing the repeatability component of measurement uncertainty $\boldsymbol{u}$:

$s_{adj}^2 = s^2 - u^2$. The adjusted SD is then used in the acceptance criterion:

$\bar{x} + ks_{adj} \leq USL$. If the measurement uncertainty is greater than $\boldsymbol{s}$, the adjusted standard deviation is set equal to zero.

If there is no subsampling variation, then the procedure described above is adequate.

If the lot SD is inflated by a subsampling component and $\boldsymbol{u}$ reflects this component, then the procedure described above is adequate.

If the lot SD is inflated by a subsampling component (as described under Case 2 in the previous section), and if $\boldsymbol{u}$ *does not* reflect this component, then another approach can be used to adjust the lot SD for both repeatability and the between-subsample variation. In particular, if every item is tested in duplicate, an adjustment for measurement uncertainty can be made for both subsampling variation and repeatability. In this case the observed standard deviation *s* calculated from all the data is adjusted by subtracting the quantity $½\boldsymbol{u}^2$ where $\boldsymbol{u}$ is the standard deviation of the differences between the results for each pair of duplicate samples:

$$s_{adj}^2 = s^2 - ½u^2.$$

5.2.8 *Presence of laboratory bias*

We consider the case that an estimate of between-laboratory variation is available, e.g. from a validation study previously conducted in accordance with ISO 5725.[xxii]

This estimate is considered a measure of laboratory bias and is taken into account in the sampling plan.

If the laboratory bias is relatively small, allowance can be made using the techniques described in Annex B of ISO 3951-6.[8] It is assumed that repeatability and laboratory-bias effects, as well as the characteristic, are normally distributed. While the acceptance criterion is of the same form as in the 'error-free' variables plans, in some circumstances it might not be possible to find a sampling plan (the number of samples $\boldsymbol{n}$ and the acceptability constant $\boldsymbol{k}$) that controls producer's and consumer's risks in the manner intended.

If the laboratory bias (i.e. the estimate of between-laboratory variation) is too large to apply the procedure from ISO 3951-6,[8] then an adjusted specification limit $\boldsymbol{USL_{adj}}$ should be calculated as $USL_{adj} = USL - q \cdot s_L$,

where $\boldsymbol{s_L}$ denotes the estimate of between-laboratory variation (expressed as a standard deviation) and $\boldsymbol{q}$ denotes the appropriate quantile. If an estimate of the variation of bias across matrices $\boldsymbol{s_{matrix}}$ is available, then the adjusted specification limit should be calculated as:

$$USL_{adj} = USL - q \cdot \sqrt{{s_L}^2 + {s_{matrix}}^2}.$$

5.2.9 *Fractional nonconformance*

If the characteristic does not follow a normal distribution (see Section 3.2.5), plans based on fractional nonconformance (FNC) can be used to allow for analytical measurement uncertainty.

The FNC for a sample can be thought of as the probability that the true value of the sample exceeds the specification limit, allowing for any measurement uncertainty present.

A sampling plan based on the FNC adjustment principle is defined by two numbers, $\boldsymbol{n}$, the number of samples to be taken and $\boldsymbol{Ac}$, the maximum acceptance limit for acceptance of the lot. These two numbers are determined in the same manner as for other types of plans, namely, by considering the allowable risks at PRQ and CRQ. Additional information on the ratio between measurement uncertainty and lot SD is also required for the design of these plans.

[xxii] ISO 5725. *Accuracy (trueness and precision) of measurement methods and results*.This ISO includes a series of standards (parts).

A lot is accepted provided the sum of the individual sample FNC values does not exceed the maximum acceptance limit.

$$\sum_{i=1}^{n} FNC_i \leq Ac$$

where $\boldsymbol{FNC_i}$ is the FNC value for the $\boldsymbol{i}^{\text{th}}$ sample ($i = 1 \ldots n$).

The use of FNC adjustment is preferred over approaches in which samples are classified as conforming or non-conforming against a specification limit or on a 'beyond reasonable doubt' basis taking measurement uncertainty in account. Such approaches are less economical in terms of sample numbers and might not be optimal in terms of controlling producer's and consumer's risks and need to be evaluated.

6. OTHER MATTERS RELATING TO SAMPLING

6.1 Physical sampling

The theory of sampling (see Section 4.4.2) relies on procedures that represent best practice for unbiased physical sampling from a lot. These sampling procedures should be observed with respect to each individual sample taken from a lot, and for any subsequent mixing and subsampling etc., noting that usually more than a single sample is required in acceptance sampling plans. Reference should be made to material-specific ISO or other standards for details of sampling procedures for different commodities. Adherence to specified sampling procedures might be a legislative or regulatory requirement for some commodities in some jurisdictions.

6.1.1 *Random sampling*

For lots consisting of discrete items, random sampling means that each item has an equal chance of being selected in the sample. The assumption of random sampling allows the operating characteristic to be calculated; deviating from random sampling might mean that the plan does not control the producer's or consumer's risks as might have been intended. In many cases systematic sampling, taking samples at regularly spaced intervals throughout a lot, will suffice as a substitute for true random sampling.

It is common for lots to be 'layered', individual items might be packed in cartons, there might be several (but the same number) of these smaller cartons packed into a larger carton, and several (but the same number) of the larger cartons packed on a pallet. Selecting a random sample of size ***n*** items would proceed as follows:

- select ***n*** pallets from the number of pallets in the lot (the same pallet can be selected more than once);
- select a random larger carton from the cartons on each side of the selected pallets;
- select a smaller carton from each of the larger cartons that have been selected; and
- finally, select an individual item from each of these smaller cartons – these constitute the sample which will be tested or examined.

For bulk materials taking a random sample is more difficult. Many lots of bulk materials can be considered as a collection of segments; stratified random sampling is used in which, in the simplest case, segments are selected at random from the total number of segments, then within each segment that has been chosen a random sample of increments is taken.

This is discussed in more detail in Section 4.4.

In principle, there is no need for random sampling for well-mixed fluids or bulk products; however random sampling might still be used as a precaution against inhomogeneity or for procedural reasons.

6.1.2 *Convenience sampling*

Convenience sampling is often referred to as pragmatic sampling. It involves taking samples, and sometimes only a single sample, from a part of a population that is convenient to sample and is often used due to low cost. It is a form of ad hoc sampling that is sometimes used in pilot testing.

There are usually more disadvantages than advantages with convenience sampling. There is a possibility of sampling error and lack of adequate representation of the population, and furthermore, use of convenience sampling might lead to disputes as it is neither a fair nor a valid procedure.

6.2 Inhomogeneous lots

While Section 3.2.4 discusses the *conditions under which* a lot can be considered homogeneous, this section addresses the question *how to handle* cases of inhomogeneous lots consisting of discrete items. For more information on sampling of inhomogeneous lots consisting of bulk materials, refer to Section 4.4.

Most sampling plans are based on the assumption that the lots are homogeneous. Use of these plans with inhomogeneous lots will usually increase producer's risks and consumer's risks, so that consumer protection may be compromised.

Lots may be inhomogeneous because inspection lots differ from manufacturing lots. Accordingly, one approach may be to split a given inhomogeneous inspection lot into sublots in line with production lots or other standardized manufacturing processes. Each of the sublots might then be sufficiently homogeneous to be inspected using standard attributes or variables sampling plans, inspecting each sublot with the same plan that would have been used for the entire lot, had it been homogeneous. However, lots should not be split into sublots based on results obtained from earlier testing.

APPENDIX I

GUIDE TO THE SELECTION AND DESIGN OF SAMPLING PLANS

1. Introduction

The concepts and criteria for sampling plans described in these guidelines are applicable to provisions in Codex standards. This Appendix provides a guide to the design of those sampling plans.

It has been structured in a way that allows users to follow the process for the design of a sampling plan from first principles to quickly identify options for sampling plans that are relevant to a particular situation in which sampling is to be undertaken.

Links are provided that allow users to quickly access further information about particular sampling options in the main document.

1.1 Starting point

The following examples are provided to assist in the design of sampling plans and should not be understood as a recommendation.

Example: Options for attributes sampling plans

In the following, the producer's risk (PR) is 5 percent and the consumer's risk (CR) is 10 percent. These values are commonly used.

Attribute sampling plans with producer's risk quality (PRQ), the quality level at which the lot of 6.5 percent may apply to commodity defects such as blemishes and other visual defects on fresh fruit.

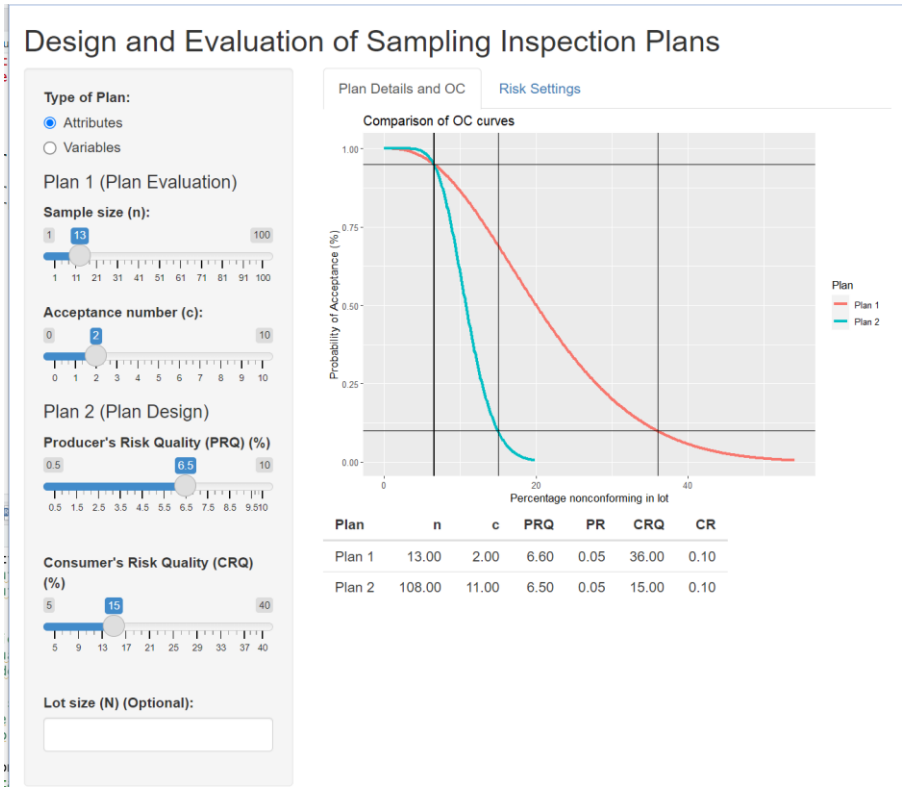
The PRQ of 6.5 percent means that lots containing 6.5 percent of nonconforming items will be accepted 95 percent of the time whereas, for example, a consumer's risk quality (CRQ) of 20 percent means that lots containing 20 percent of nonconforming items will be rejected 90 percent of the time.

The following table shows options for sampling plans for different levels of CR quality.

Table 4: Sampling plan options for PRQ = 6.5 percent

CRQ	PRQ	n	c
20%	6.5%	51	6
25%	6.5%	30	4
30%	6.5%	21	3
36%	6.5%	13	2

The operating characteristics for two of these plans is shown below; this shows the probability of accepting a lot with those plans at any quality level. The choice of sampling plan will depend on the probability of acceptance across the entire range of quality levels.



Example: Options for variables sampling plans

The provison for a compositional characteristic for a commodity specifies that the percentage content should not exceed a maximum limit. In this example, it is assumed that the measurement uncertainty is negligible and that the lot SD is known.

The following table shows options for variables sampling plans with a PRQ of 3.5 percent and for different levels of CR quality.

Table 5: Sampling plan options for PRQ = 3.5 percent

CRQ	PRQ	n	k
10%	3.5%	31	1.52
15%	3.5%	16	1.39
20%	3.5%	10	1.29
25%	3.5%	7	1.19
30%	3.5%	6	1.14
35%	3.5%	5	1.08

The operating characteristics for two of these plans is shown below; this shows the probability of accepting a lot with those plans at any quality level. The choice of sampling plan will depend on the probability of acceptance across the entire range of quality levels. It will also depend on whether the lot SD is known or unknown.

Design and Evaluation of Sampling Inspection Plans

Type of Plan:

- ○ Attributes
- ◉ Variables

Standard Deviation Type

- ◉ Known
- ○ Unknown

Plan 1 (Plan Evaluation)

Sample size (n):

1 [28] 100

1 11 21 31 41 51 61 71 81 91 100

k-constant (k):

1 [1.38] 3

1 1.2 1.4 1.6 1.8 2 2.2 2.4 2.6 2.8 3

Plan 2 (Plan Design)

Producer's Risk Quality (PRQ) (%)

0.5 [3.5] 10

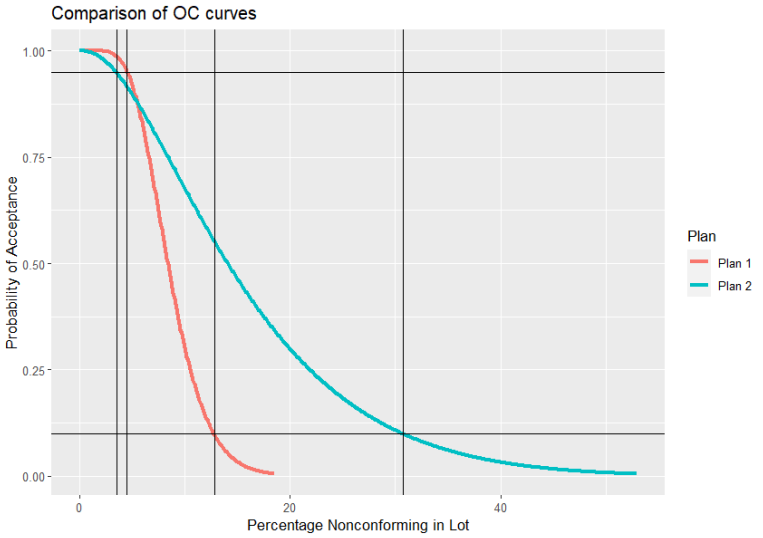
0.5 1.5 2.5 3.5 4.5 5.5 6.5 7.5 8.5 9.510

Consumer's Risk Quality (CRQ) (%)

5 [35] 40

5 9 13 17 21 25 29 33 37 40

Plan Details and OC | Risk Settings



Plan	n	k	PRQ	PR	CRQ	CR
Plan 1	28.00	1.38	4.50	0.05	12.80	0.10
Plan 2	5.00	1.08	3.50	0.05	35.00	0.10

1.2 Selection of options for sampling plans

A. Determine sampling plan options

Step 1. Type of data

Are the test results expressed as pass/fail outcomes (or equivalent) or are they measurements?

Pass/Fail (or equivalent) outcomes (attributes)	Go to step 2
Measurements (variables)	Go to step 3

Help on attributes data

Help on variables data

Step 2. Attributes data

Is the inspection error negligible or non-negligible?

Negligible	4.2.2 Two-class attributes ***plans***	PR & CR	
	CXG 50 4.2.3	CR only	ISO 2859-2
	CXG 50 Appendix II	PR only	ISO 2859-1
Non-negligible			
	CXG 50 5.1.1	Known inspection errors	

Help on design of attributes plans

Step 3. Variables data

Does the provision relate to compliance of the distribution or to the average level of the characteristic?

Step 3.a. Plans to assess compliance of the distribution

Is the characteristic normally distributed, a compositional characteristic or does it follow some other distribution?

Normally distributed	Go to step 4
Compositional proportion	Go to step 6
Some other distribution	Go to step 7

Help on design of variables plans

Step 3.b. Plans for the average level

Plans for the average level	Go to step 8

Help on provision

Help on average level

Step 4. Variables plans, normally distributed characteristics

Is measurement uncertainty negligible or non-negligible?

Negligible	CXG 50 4.3.3	PR & CR	
	CXG 50 4.3.4	CR only	ISO 3951-6
	CXG 50 Appendix 2	PR only	ISO 3951-1
Non-negligible	Go to step 5		

Step 5. Variables plans, normally distributed characteristics, non-negligible measurement uncertainty

Is the measurement uncertainty normally distributed or does it follow some other distribution?

Normally distributed	CXG 50 5.2.7	PR & CR	
	CXG 50 5.2.5	CR only	ISO 3951-6
Some other distribution	CXG 50 5.2.8	PR & CR	

Step 6. Compositional proportions

Is measurement uncertainty negligible or non-negligible?

Negligible	CXG 50 4.4.10	PR & CR
Non-negligible	Go to step 5	

Step 7. Characteristic is neither normally distributed nor a compositional proportion

Is the measurement uncertainty negligible or non-negligible?

Negligible	CXG 50 4.2.7	PR & CR
Non-negligible	CXG 50 5.2.8	PR & CR

Step 8. Provision is expressed in terms of the average level in a lot

Is the measurement uncertainty negligible or non-negligible?

Negligible	CXG 50 4.4.8	PR & CR
Non-negligible		
[no information provided]		

B. Specify stringency for the sampling plan (plans to assess compliance to minimum or maximum levels)

Consumer's risk quality level (CRQ)

What percentage nonconforming (quality level?) would you allow in lots that you would want to reject most of the time?	6.5%

Consumer's risk (CR)

What consumer's risk are you prepared to allow, i.e. how often would you want to accept lots containing 6.5 percent nonconforming?	10%

If the characteristic is a 'serious' food safety (or other) concern:

• it might not be appropriate to control producer's risks explicitly;

• use ISO plans (or alternatives) that control only the consumer's risk.

If the characteristic is not a 'serious' food safety or other concern, it is appropriate to also control the producer's risk.

Producer's risk quality level (PRQ)

What percentage nonconforming (quality level?) would need to be present in lots that you would want to accept most of the time?	5%

Producer's risk (PR)

What producer's risk are you prepared to allow, how often would you want to reject lots containing 5 percent nonconforming?	5%

C. Evaluate plan to determine plan parameters and calculate operating characteristic

Determine the number of samples and the acceptance number (attributes plans) or the acceptability constant (variables plans)

Supporting material

Context	Term	Explanation
Nature of the provision	Provision	A provision is a requirement for a commodity that must be met in order that the commodity conforms to the standard.

Nature of the provision	Overall distribution	Specification limits may be expressed as a minimum or a maximum limit (or both) applied to either the overall distribution of the characteristic in the lot, e.g. the percentage nonconforming quality level, or to the average level.
Nature of the provision	Average level	In some cases, such as the net weight of packages, a limit is set on the average level, with the intention that the average level in the batch should not be less than the limit. In Codex, although an example of sampling plans for bulk materials, the plans for aflatoxins are also based on compliance of the average level, to ensure that there is a small chance that the average level in a lot exceeds the maximum limit. It is usually assumed that the quality characteristic is normally distributed; the appropriateness of the distribution is less critical when compliance of the average level is being assessed. It is also usually assumed that there is a single specification limit, either a lower specification limit, L or an upper specification limit, U.
Types of data	Attributes	Data for which the test results have nominal outcomes or are measured on a scale, particularly binary outcomes such as pass or fail, and measurements classified as binary outcomes.
Types of data	Variables	Inspection by variables means that the outcomes of the measurements on each sample is a number, usually a decimal number. This is in contrast to attributes data where pass/fail outcomes are obtained or on a scale (sometimes described numerically, e.g. 1–5).
Type of sampling plan	Attributes plan	Inspection by attributes consists of examining an item, or characteristics of an item, and classifying the item as 'conforming' or 'nonconforming'. The action to be taken is decided by counting the number of nonconforming items or the number of nonconformities found in a random sample. An inspection by attributes sampling plan specifies the number of samples (n) and the maximum number of nonconforming items, referred to as the acceptance constant (c), for the lot to be accepted. The values of n and c are worked out from the specified levels of allowable risk.
Type of sampling plan	Variables plan	Inspection by variables plans use means and standard deviations (SD) calculated from the measurements (variables data) to make a decision about the acceptance of a lot. These plans are specified by the number of samples required to be taken (n) and an acceptability constant (k).
Measurement uncertainty		Parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand (i.e. the quantity intended to be measured). Measurement can consist of random and systematic components.
Lot standard deviation		A parameter, usually expressed as a SD, describing the variation of a characteristic within a lot.
Negligible measurement uncertainty		The situation where the measurement uncertainty (MU) is small in relation to the lot SD and does not need to be taken into account in the design of a sampling plan. Typically, MU is considered negligible if the SD representing the MU is less than 10 percent of the lot SD.
Non-negligible measurement uncertainty		Refers to cases where the MU is NOT negligible.
Standard deviation		SD is a measure of the amount of variation or dispersion in a

		set of values.
Known (true) standard deviation		Conceptually, the SD that would be found, for example, if every item in a lot was measured. In practice, SDs can be considered known if calculated using a reasonably large number of test results, typically 100–200. For a SD representing the longer-term variation of a process to be considered known, the process must be stable (consistent) over time.
Estimated (sample) standard deviation		A SD calculated from a smaller amount of data than required for the SD to be considered known.
Normal distribution		A statistical distribution commonly used in many branches of statistics to describe the variation of a measurement method under certain conditions or of a characteristic within a lot. A normal distribution is described by its mean (i.e. average level) and SD and follows a characteristic 'bell-shaped' curve.
Compositional proportion		A characteristic whose concentration within a lot can be expressed as a 'mass fraction', a number taking values between zero and one. Strictly speaking compositional proportions are dimensionless, and do not have proper units of measure, although it is common to express then using units such as percentages, parts-per-million (ppm) etc.
Producer's risk	PR	In general terms, PR is the risk that a lot of good quality will be rejected. More specifically, in the design of acceptance sampling plans, producer's risk is the probability of rejecting a lot that has a quality level equal to the producer's risk quality (PRQ) level.
Producer's risk quality level	PRQ	The quality level (percentage nonconforming in the lot) at which the probability of rejecting the lot is equal to the specified producer's risk (PR).
Consumer's risk	CR	Consumer's risk (CR) is the risk that a lot of poor quality will be accepted. More specifically, in the design of acceptance sampling plans, consumer's risk is the probability of accepting a lot that has a quality level equal to the consumer's risk quality (CRQ) level.
Consumer's risk quality level	CRQ	The quality level (percentage nonconforming in the lot) at which the probability of accepting the lot is equal to the specified CR.

APPENDIX II

ISO INSPECTION PLANS INDEXED BY PRODUCER'S RISK

1. ISO Inspection plans indexed by producer's risk – Introduction/Background

As noted in Sections 4.2.3 and 4.3.4, the sampling plans included in the ISO 2859[i] and ISO 3951[ii] standards differ from plans discussed elsewhere in these guidelines in that they have been designed to explicitly control either the producer's risk (PR) or the consumer's risk (CR), but not both, and use a lot size relationship to determine the required sample size.

1.1 Lot size versus sample size

Statistically, the lot size does not have an important role in determining protection to consumers and producers, whereas changes in the sample size does affect the protection afforded by any plan.

However, despite this, a lot size versus sample size relationship has been built into the design of the sampling plans appearing in the ISO standards. This relationship is arbitrary, although it has the general effect of reducing the risks of making incorrect decisions for larger lots, where the costs incurred from incorrect decisions will be greater. This relationship means that the ISO standards are applicable only to lots that consist of discrete items.

As a consequence of employing the sample size versus lot size relationship, ISO has designated that sampling plans indexed by producer's risk quality (PRQ), explicitly controlling the producer's risk, are intended for the inspection of a continuing series of lots and plans indexed by consumer's risk quality (CRQ), explicitly controlling consumer's risk, as being suitable for the inspection of isolated lots. However, this distinction is no longer relevant if both types of risk are considered in the design of plans.

1.2 Sampling schemes

The ISO standards indexed by PRQ employ sampling schemes, sets of sampling plans with different levels of inspection to ensure quality is effectively controlled. Sampling schemes employ switching rules for changing between inspection levels based on recent quality history. Typically, and in ISO standards, switching occurs between normal, tightened, and reduced inspection plans within each sampling scheme:

- normal inspection is used when the process is considered to be operating at, or slightly better than, the PRQ;
- tightened inspection uses stricter decision rules than those used in normal inspection. The main objective of using tightened inspection is to exert pressure on the producer when the quality is poorer than the PRQ by introducing a higher rate of rejection; and
- reduced inspection permits smaller sample sizes than those used in normal inspection. When the level of the submitted quality is sufficiently good, reduced inspection offers sampling economy.

Sampling schemes provide more comprehensive assurance than the use of individual sampling plans. However, switching rules are considered too complex to apply in international trade, and from a consumer's point of view in general, although it is possible to design a sampling plan that controls the producer's and consumer's risks to the same levels as an overall sampling scheme.

1.3 Table 6: Inspection by attributes plans in accordance with ISO 2859-1[6]

		Inspection level					
Lot size	**AQL**	**reduced**		**normal**		**tightened**	
(number of packages, each containing 1 or more units)		***n***	***c***	***n***	***c***	***n***	***c***
2–8	0.65%	8	0	8	0	8	0
	2.50%	2	0	5	0	8	0
	6.50%	2	0	2	0	3	0
9–15	0.65%	8	0	15	0	15	0
	2.50%	2	0	5	0	8	0
	6.50%	2	0	2	0	3	0

[i] ISO 2859: *Sampling procedures for inspection by attributes*. This ISO includes a series of standards (parts).
[ii] ISO 3951: *Sampling procedures for inspection by variables*. This ISO includes a series of standards (parts).

		Inspection level					
Lot size	**AQL**	**reduced**		**normal**		**tightened**	
16–25	0.65%	8	0	20	0	25	0
	2.50%	2	0	5	0	8	0
	6.50%	5	1	8	1	13	1
26–50	0.65%	8	0	20	0	32	0
	2.50%	2	0	5	0	8	0
	6.50%	5	1	8	1	13	1
51–90	0.65%	8	0	20	0	32	0
	2.50%	13	1	20	1	32	1
	6.50%	5	1	13	2	13	1
91–150	0.65%	8	0	20	0	32	0
	2.50%	13	1	20	1	32	1
	6.50%	8	2	20	3	20	2
151–280	0.65%	8	0	20	0	32	0
	2.50%	13	1	32	2	32	1
	6.50%	13	3	32	5	32	3
281–500	0.65%	50	1	80	1	125	1
	2.50%	20	2	50	3	50	2
	6.50%	20	5	50	7	50	5
501–1 200	0.65%	50	1	80	1	125	1
	2.50%	32	3	80	5	80	3
	6.50%	32	6	80	10	80	8
1 201–3 200	0.65%	50	1	125	2	125	1
	2.50%	50	5	125	7	125	5
	6.50%	50	8	125	14	125	12
3 201–10 000	0.65%	80	2	200	3	200	2
	2.50%	80	6	200	10	200	8
	6.50%	80	10	200	21	200	18
10 001–35 000	0.65%	125	3	315	5	315	3
	2.50%	125	8	315	14	315	12
	6.50%	80	10	200	21	200	18
35 001–150 000	0.65%	200	5	500	7	500	5
	2.50%	200	10	500	21	500	18
	6.50%	80	10	200	21	200	18
150 001–500 000	0.65%	315	6	800	10	800	8
	2.50%	200	10	500	21	500	18
	6.50%	80	10	200	21	200	18
500 001 and over	0.65%	500	8	1250	14	1250	12
	2.50%	200	10	500	21	500	18
	6.50%	80	10	200	21	200	18

If sample size ***n*** equals, or exceeds lot size, carry out 100 percent inspection.

1.4 Table 7: Inspection by variables plans from ISO 3951-1[7] (lot SD unknown)

		Inspection level					
Lot size	**AQL**	**reduced**		**normal**		**tightened**	
(number of packages, each containing 1 or more units)		***n***	***k***	***n***	***k***	***n***	***k***
2–8	0.65%	6	1.476	8	1.889	8	2.079
	2.50%	4	0.850	4	1.242	6	1.476
	6.50%	4	0.586	4	0.735	3	0.950
9–15	0.65%	6	1.476	11	1.889	15	2.079
	2.50%	4	0.850	4	1.242	6	1.476
	6.50%	4	0.586	4	0.735	3	0.950
16–25	0.65%	6	1.476	11	1.889	15	2.079
	2.50%	4	0.850	4	1.242	6	1.476
	6.50%	4	0.586	6	0.939	6	1.061
26–50	0.65%	6	1.476	11	1.889	15	2.079
	2.50%	4	0.850	9	1.323	6	1.476
	6.50%	4	0.586	6	0.887	9	1.218
51–90	0.65%	6	1.476	11	1.889	15	2.079
	2.50%	6	1.061	13	1.475	13	1.569
	6.50%	5	0.550	9	0.869	9	1.190
91–150	0.65%	6	1.476	11	1.889	15	2.079
	2.50%	9	1.218	13	1.426	18	1.682
	6.50%	7	0.507	14	0.935	14	1.147
151–280	0.65%	11	1.642	22	1.972	15	2.079
	2.50%	9	1.190	20	1.411	18	1.659
	6.50%	9	0.628	21	0.945	21	1.227
281–500	0.65%	17	1.769	30	2.079	28	2.153
	2.50%	14	1.147	30	1.471	27	1.636
	6.50%	14	0.601	33	1.036	32	1.225
501–1 200	0.65%	23	1.893	31	2.061	38	2.263
	2.50%	21	1.227	46	1.482	41	1.702
	6.50%	21	0.830	52	1.120	50	1.245
1 201–3 200	0.65%	24	1.862	48	2.043	40	2.237
	2.50%	32	1.225	69	1.552	63	1.702
	6.50%	33	0.954	79	1.195	78	1.281
3 201–10 000	0.65%	37	1.853	71	2.101	61	2.230
	2.50%	48	1.394	105	1.619	99	1.720
	6.50%	52	1.120	124	1.239	122	1.325
10 001–35 000	0.65%	54	1.904	108	2.104	89	2.279

		Inspection level					
Lot size	**AQL**	**reduced**		**normal**		**tightened**	
	2.50%	71	1.489	159	1.683	150	1.752
	6.50%	52	1.120	124	1.239	122	1.325
	0.65%	84	1.914	159	2.166	137	2.285
35 001–150 000	2.50%	105	1.619	247	1.716	233	1.785
	6.50%	52	1.120	124	1.239	122	1.325
	0.65%	117	2.037	239	2.220	214	2.300
150 001–500 000	2.50%	105	1.619	247	1.716	233	1.785
	6.50%	52	1.120	124	1.239	122	1.325
	0.65%	169	2.117	348	2.268	323	2.324
500 001 and over	2.50%	105	1.619	247	1.716	233	1.785
	6.50%	52	1.120	124	1.239	122	1.325

If sample size ***n*** equals, or exceeds lot size, carry out 100 percent inspection.

NOTES

[1] FAO and WHO. 2009. *Guidelines on Analytical Terminology*. Codex Alimentarius Guideline, No. CXG 72-2009. Codex Alimentarius Commission. Rome.

[2] FAO and WHO. 2004. *Guidelines on Measurement Uncertainty*. Codex Alimentarius Guideline, No. CXG 54-2004. Codex Alimentarius Commission. Rome.

[3] FAO and WHO. 2013. *Principles for the Use of Sampling and Testing in International Food Trade*. Codex Alimentarius Guideline, No. CXG 83-2013. Codex Alimentarius Commission. Rome.

[4] EURACHEM & CITAC. 2000. *Guide quantifying uncertainty in analytical measurement (Second Edition).* EURACHEM Secretariat. BAM. Berlin. www.eurachem.org

[5] ISO. 2020.*ISO 2859-2: Sampling procedures for inspection by attributes – Part 2: Sampling plans indexed by limiting quality (LQ) for isolated lot inspection*. Geneva. ISO.

[6] ISO.1999. *ISO 2859-1: Sampling procedures for inspection by attributes – Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection*. Geneva. ISO.

[7] ISO. 2022. *ISO 3951-1: Sampling procedures for inspection by variables – Part 1: Specification for single sampling plans indexed by acceptance quality limit (AQL) for lot-by-lot inspection for a single quality characteristic and a single AQL*. Geneva. ISO.

[8] ISO. Forthcoming. *ISO/DIS 3951-6: Sampling procedures for inspection by variables – Part 6: Specification for single sampling plans for isolated lot inspection indexed by limiting quality (LQ)*. Geneva. ISO.

[9] FAO and WHO. 1995. *General Standard for Contaminants and Toxins in Food and Feed*. Codex Alimentarius Standard, No. CXS 193-1995. Codex Alimentarius Commission. Rome.

[10] ISO. 2017. *ISO/IEC 17025: General requirements for the competence of testing and calibration laboratories*. Geneva. ISO.

[11] ISO. 2019. *ISO 5725-2: Accuracy (trueness and precision) of measurement methods and results – Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method*. Geneva. ISO.

[12] ISO. 1994. *ISO 5725-1: Accuracy (trueness and precision) of measurement methods and results – Part 1: General principles and definitions*. Geneva. ISO.

[13] ISO. 1994. *ISO 5725-3: Accuracy (trueness and precision) of measurement methods and results – Part 3: Intermediate measures of the precision of a standard measurement method*. Geneva. ISO.