CODEX ALIMENTARIUS COMMISSION E

Food and Agriculture Organization of the United Nations

World Health Organization

Viale delle Terme di Caracalla, 00153 Rome, Italy - Tel: (+39) 06 57051 - E-mail: codex@fao.org - www.codexalimentarius.org

Agenda Item 7 **CX/MAS 20/41/9**
March 2020

JOINT FAO/WHO FOOD STANDARDS PROGRAMME

CODEX COMMITTEE ON METHODS OF ANALYSIS AND SAMPLING

41st Session
Budapest, Hungary, 11 - 15 May 2020

REVISION OF THE GENERAL GUIDELINES ON SAMPLING (CXG 50 – 2004)

(Prepared by the EWG led by New Zealand and the United States of America)

Codex members and Observers wishing to submit comments at Step 3 on this document should do so as instructed in CL 2020/27/OCS-MAS available on the Codex webpage/Circular Letters:
http://www.fao.org/fao-who-codexalimentarius/resources/circular-letters/en/

Introduction

1. The 39th Session of the Committee on Methods of Analysis and Sampling (CCMAS39) agreed to start new work on the revision of the *General Guidelines on Sampling* (CXG 50-2004). This is set out in REP18/MAS Para 71, Appendices V (project document) and VI (prioritization areas of work). This new work was approved by CAC41 (REP18/CAC, Appendix VI).

2. CCMAS40 supported the continuation of work on the revision of CXG 50 in accordance with the prioritization of work as agreed by CCMAS39.[1]

3. CCMAS40 tasked an EWG chaired by New Zealand and co-chaired by the USA, to continue with the revision of CXG 50 and the further development of the supplementary document (e-book with sampling plan apps) taking into account written comments submitted (CL 2019/17-MAS) and comments and recommendations made during the session.

4. Throughout this paper, we will refer to the working draft revision document as the 'revised Guidelines'.

The process followed by the EWG

5. New Zealand issued an invitation to the EWG on 26th August 2019 to comment on the latest version of the revised Guidelines. A list of EWG participants is appended (Appendix V). New Zealand, as chair of the EWG reviewed these comments along with those from the CCMAS40 discussion and prepared a revised version of the Guidelines.

6. New Zealand provided these revised Guidelines to the USA (co-chair). Experts from both New Zealand and the USA corresponded in writing and together with the leads, discussed the revised Guidelines. We had hoped to provide a version to the EWG that presented a consensus between NZ and the USA. However, we eventually agreed that there was not sufficient time to discuss the US proposals and reach a consensus and be able to meet the project deadlines for CCMAS41.

7. Accordingly, a jointly-signed email was sent to the EWG on 24th December 2019 and the following documents were posted on the EWG forum:

- latest revised version of the Guidelines, taking account of some of the USA comments;
- responses (to date) to the comments and questions from the CL 2019/17- MAS; and
- an update to the e-book for sampling plan apps, issued as a WORD document.

In the jointly-signed email we advised:

- that following discussion with the Codex Secretariat, references to the e-book and sampling plan apps were removed from the latest version of the revised Guidelines. We proposed to provide the e-book

[1] Full discussion and decisions are in REP19/MAS, paras 67 - 80

and the link to the sampling plan apps outside of the revised Guidelines. Possibilities include publishing it as an electronic information document; and

- that due to timing, this latest version of the revised Guidelines did not take into account many of the comments from the co-chair (USA); they would post their comment on the EWG platform. This would provide an opportunity for consideration of the USA comments and proposals by the wider EWG.

8. The EWG was requested to provide comments (consistent with the terms of reference and with the scope as described in the project document (Appendix V of REP18/MAS)) by 7 February 2020. Six country EWG members provided comprehensive comments.

9. Following this process, New Zealand produced a summary of the comments including a written response to each (issued on the forum site). Based on the comments, we either updated the revised Guidelines or identified the comment as one to be included in the next version of the Guidelines (which is intended to be considered at CCMAS41).

10. The USA co-chair advised that the USA would be providing more technical information (for on-going discussion) to New Zealand but not in time for the EWG report.

Summary of the discussions

11. This summary of the discussions is set out in 2 parts; one on the discussions between New Zealand and the USA as chair and co-chair respectively, and one on the EWG consultation.

Discussions between New Zealand and the USA:

12. The USA provided:

- commentary on technical areas in the revised Guidelines, including that where sampling plans have been prescribed within Codex, attributes sampling plans have generally been used, the need to allow for measurement uncertainty and the inclusion of the material on reinspection; and
- A proposed top-level outline for the revised Guidelines.

13. New Zealand identified and responded to the technical areas identified by the USA:

- the focus on attributes plans is too restrictive and might exclude other valid, more economical methods for sampling inspection;
- the relevance of the FAO/WHO 2016 Guide[2];
- that allowance for measurement error is required in cases where measurement error is significant; and
- that control of Producer's Risk should be taken into account explicitly in the design of sampling plans.

14. New Zealand did not respond to the proposed top-level outline for the revised Guidelines. We considered that this proposal represented enough of a change from the approved project document that it would need wider discussion among the EWG and CCMAS itself.

15. New Zealand and the USA plan to have further engagement on the technical areas prior to the CCMAS41 meeting and an update will be presented to the CCMAS41.

EWG consultation:

16. The EWG consultation provided comments from 6 participants for administrative and technical updates to the revised Guidelines. As noted, the USA also provided comment that included a proposal for a top-level outline for a revised Guidelines.

17. In response to these comments, areas of improvement included the removal of redundant text and definitions deemed unnecessary, the inclusion of material for some key areas such as physical sampling and random sampling, and the removal of content considered unimportant or outside the scope of the project document. Some changes to the revised Guidelines have been made as a result of the EWG consultation.

18. In response to this, identified technical areas in the summary of comments and identified key areas in revised Guidelines will be posted on the forum for further consideration.

19. Discussion at CCMAS41 should focus on these technical areas in the revised Guidelines, supported by the summary of comments.

[2] Statistical Aspects of Microbiological Criteria Related to Foods. A Risk Managers Guide.
Food and Agriculture Organization of the United Nations, World Health Organization, Rome, 2016

20. New Zealand, as Chair of the EWG, recommends discussion should also consider the proposals for the revised Guidelines the USA provided to the EWG.

Conclusions and recommendations

21. The revised Guidelines (presented in Appendix I) are intended to represent the work as outlined in the project document and the prioritization list. They describe the design and evaluation of sampling plans for the international trade of food.

22. The revised Guidelines are intended primarily for use by Codex commodity committees responsible for developing sampling plans for provisions in Codex standards and by governments responsible for import or export inspection of foods. However the Guidelines are applicable quite generally and could be used by any party engaged in the trade or sale of foods. For instance they could be used by any two parties at any stage of the supply chain, with appropriate consideration of the fairness of the transaction; or they could be used by a single party, for instance by a processor using a sampling plan for end-product verification.

Recommendations

23. The Committee is invited to consider the proposed draft revised Guidelines (Appendix I) and the supplementary e-book (content and location) (Appendix II).

24. Based on the questions raised on the current revised Guidelines, it is recommended that CCMAS also consider the following points:

a) The key technical areas identified by New Zealand in response to the USA commentary (Appendix IV):
   - the focus on attributes plans;
   - the relevance of the FAO/WHO 2016 Guide[3];
   - allowance for measurement error; and
   - control of Producer's Risk.

b) The key technical areas the USA provided (Appendix III):
   - that the majority of sampling plans used in Codex are based on attributes sampling;
   - that allowance for measurement uncertainty is not required;
   - that material on reinspection should not be included in the guidelines; and

c) The USA proposal for a top-level outline of a revised Guidelines (Appendix III):
   - content.

[3] Statistical Aspects of Microbiological Criteria Related to Foods. A Risk Managers Guide. Food and Agriculture Organization of the United Nations, World Health Organization, Rome, 2016

Appendix I

Proposed draft revision of the General Guidelines on Sampling (CXG 50-2004)
(for comments through CL 2020/27-MAS)

1 Purpose of the Guidelines

The General Guidelines on Sampling (CXG 50 - 2004) will be referred to as 'the Guidelines' in this document.

These Guidelines describe the design and evaluation of sampling plan for the international trade of food commodities.

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

In Codex, sampling plans, in conjunction with methods of analysis, are intended as a means of verifying that foods comply with provisions such as composition, limits for contaminants and pesticide residues and microbiological criteria that are part of Codex standards.

Sampling[4] therefore has 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 a role in avoiding or removing difficulties which may be created by diverging legal, administrative and technical approaches to sampling and by diverging interpretation of results of analysis in relation to lots or consignments of foods, in the light of the relevant provision(s) of the applicable Codex standard.[5]

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

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. The main aim of sampling is to ensure that the consumer receives product of acceptable quality. In addition, sampling plans should be designed to provide a high rate of acceptance of compliant product.

The 'outcome to be achieved' therefore defines allowable risks for the consumer and the producer.

Commodity committees should define sampling plans for provisions in Codex standards.

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.

The design of these plans should be based on the principles described in these Guidelines.

These Guidelines provide recommendations on the design of sampling plans that achieve these aims.

1.1 Basis of the Guidelines

Throughout these Guidelines the two parties involved in a transaction in the food chain are referred to as the *'producer'* and the *'consumer'*. The terms 'producer' and 'consumer' are conventional and may apply to a range of different operators in the food chain, such as a grower, manufacturer, supplier, exporting country, processor, on-seller, and customer or importing country.

Sampling involves the selection of small quantities from a lot, and drawing conclusions about the lot from results of inspection or testing of the samples. Sampling reduces costs and avoids loss of food through destructive tests, but it inevitably creates risks for both consumers and producers.

These Guidelines describe the inputs needed for the design of sampling plans, in order to accurately define the sampling situation, the acceptable levels of and the acceptance conditions of a lot. This is covered in section 3.

These Guidelines also provide information on the design and evaluation of sampling plans in sections 4 and 6 respectively. They will also help both consumers and producers design sampling plans that are appropriate for inspections of lots or consignments of food for acceptance according to the inputs as

[4] Sometimes referred to as "sampling inspection" to distinguish from usual interpretation of sampling meaning the process of physically taking samples from a lot.
[5] "Principles for the Establishment or Selection of Codex Sampling Procedures", Codex Procedural Manual.

described. They are intended to assist interpretation of results of analysis of foods when making a decision on the disposition (acceptance or rejection) of food when some characteristic of the food is subject to such a provision.

The Guidelines are based on the principles expressed in the *Principles for the Use of Sampling and Testing in International Food Trade* (CXG 83-2013).

1.2 Application of the Guidelines

The Guidelines are intended primarily for use by Codex commodity committees responsible for developing sampling plans for provisions in Codex standards and by governments responsible for import or export inspection of foods. However the Guidelines are applicable quite generally and could be used by any party engaged in the trade or sale of foods. For instance they could be used by any two parties at any stage of the supply chain, with appropriate consideration of the fairness of the transaction; or they could be used by a single party, for instance by a processor using a sampling plan for end-product verification.

It should be noted that Codex sampling plans might not directly address the producer's perspective. Producers should be aware that the purpose of these plans is to specify the outcome to be achieved in terms of allowable risks. When a sampling plan or an outcome is clearly set out, producers are able to devise appropriate control procedures to achieve them.

The Guidelines also provide information on the evaluation of sampling plans obtained from other sources, to allow issues of fairness to be investigated.

1.3 Codex commodity committees

In some situations, such as when measurement error is significant, it might not be possible to specify a standard plan, suitable for general use. To overcome this commodity committees can specify criteria in terms of allowable risks that sampling plans are expected to achieve so that users can develop plans specific to their situation.

However commodity committees should not just specify outcomes to be achieved without actually deriving sampling plans and assessing their fitness for purpose and the implications on producers in terms of fairness.

2 Concepts of Sampling

2.1 Approach to sampling

In the context of sampling, risk[6] refers to the probability of making an incorrect decision about a lot of product, of either incorrectly accepting a lot of poor quality or of rejecting a lot of good quality product.

There are three possible approaches to sampling:

a. 100% inspection, involving inspection of all (i.e.100%) of the product;

b. sampling based on the principles of probability; and

c. *ad hoc* inspection, that is, a sampling plan without a statistical basis.

The risks and costs associated with each of these three options can be considered:

For **approach (a)**, it is clear that 100% sampling 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. Also, the presence of measurement error means that it is still not possible to provide a 100% guarantee, even if all items in the lot are inspected.

Approach (b) has the disadvantage of higher risks as compared to approach (a), since some product will not be inspected. However by using the probability approach the risks can be calculated and a sampling plan chosen that ensures these risks are controlled to desired levels. It also has the advantage of practicability and lower costs.

Approach (b), the probability approach, is described in detail later.

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

[6] The use of 'risk' here relates to probabilities of incorrect acceptance or rejection of product. This is different to the way that risk is usually understood in Codex, i.e. probability of an adverse health event; where the term refers to 'public health risk'.

lead to illness or unwarranted rejection that in turn, could lead to the imposition of fines or penalties, trade sanctions or loss of access to markets. The risks associated with such plans should be evaluated where possible. Decisions on acceptance or rejection should not be made solely on the basis of these plans such a plan except by mutual agreement of the consumer and producer based on an understanding of those risks.

The approach to sampling should be based on control of the levels of assurance provided and the costs to the parties involved in the transaction.

2.2 The probability approach

2.2.1 Control of risks

It is not possible to provide 100% assurance that all product in a lot complies with a specification when sampling is used. There are two types of risks that can occur:

- the risk that product of unsatisfactory quality will be accepted (consumer's risk); and
- the risk that acceptable quality product will be rejected (producer's risk).

Sampling plans should be designed to control the risks to desired levels, i.e. they should take account of the principle of fitness for purpose[7]. Such control provides assurance, over the longer term, across many lots (i.e. in terms of probability).

The risks are expressed in terms of a level non-conforming and the associated chance of acceptance or rejection at that level. Commonly, the producer's and consumer's risks are specified in terms of the Producer's Risk Quality (PRQ) and the Consumer's Risk Quality (CRQ) respectively, for example:

- producer's risk - 5% chance of rejection at a PRQ of 1% non-conforming (or equivalently, 95% chance of acceptance at 1% non-conforming); and
- consumer's risk - 10% chance of acceptance at a CRQ of 5% non-conforming.

Once the PRQ and CRQ, along with their associated probabilities of rejection and acceptance respectively are specified, a sampling plan, allowing no more than these levels of risk, can be developed. In some cases, where measurement error is significant, additional information may be required.

The PRQ and CRQ along with their associated probabilities of acceptance are two fundamental inputs in the design of sampling plans.

Information note

Probability and what it means in sampling

Variation is present everywhere; raw materials vary in their composition, manufacturing process vary and, as a consequence, the products manufactured by those processes will also vary.

Therefore, when we take a set of samples from a lot of product, we do not expect those samples to be of the same composition. Further, the presence of measurement error 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 to always be the same; there will be some variation of those results.

Variation causes uncertainty when we attempt to make decisions about the compliance of a lot to a specification limit; at any level non-conforming some lots might be accepted and some might be rejected.

However, if we describe the variation of the product and of the measurement process statistically, we can predict the expected outcome in any given situation, at any level non-conforming for any given sampling plan. In acceptance sampling this expected outcome can be expressed as the average rate of acceptance (or success rate) over a long-series of inspections of lots having the same level non-conforming. This average rate is more commonly known as the probability of acceptance and can lie between zero (lots with that level non-conforming are never accepted) and one (lots are always accepted).

In acceptance sampling the probability of acceptance for a particular plan depends on the level non-conforming in a lot, the decision criterion for that sampling plan and possibly, in the case of significant measurement error, on the bias and variation inherent in the measurement process. In

[7] See *Principles for the Use of Sampling and Testing in International Food Trade* (CXG 83-2013), Principle 6.

practice, the level non-conforming in a lot is not known beforehand but it is possible to calculate the probability of acceptance for any assumed level non-conforming in a lot.

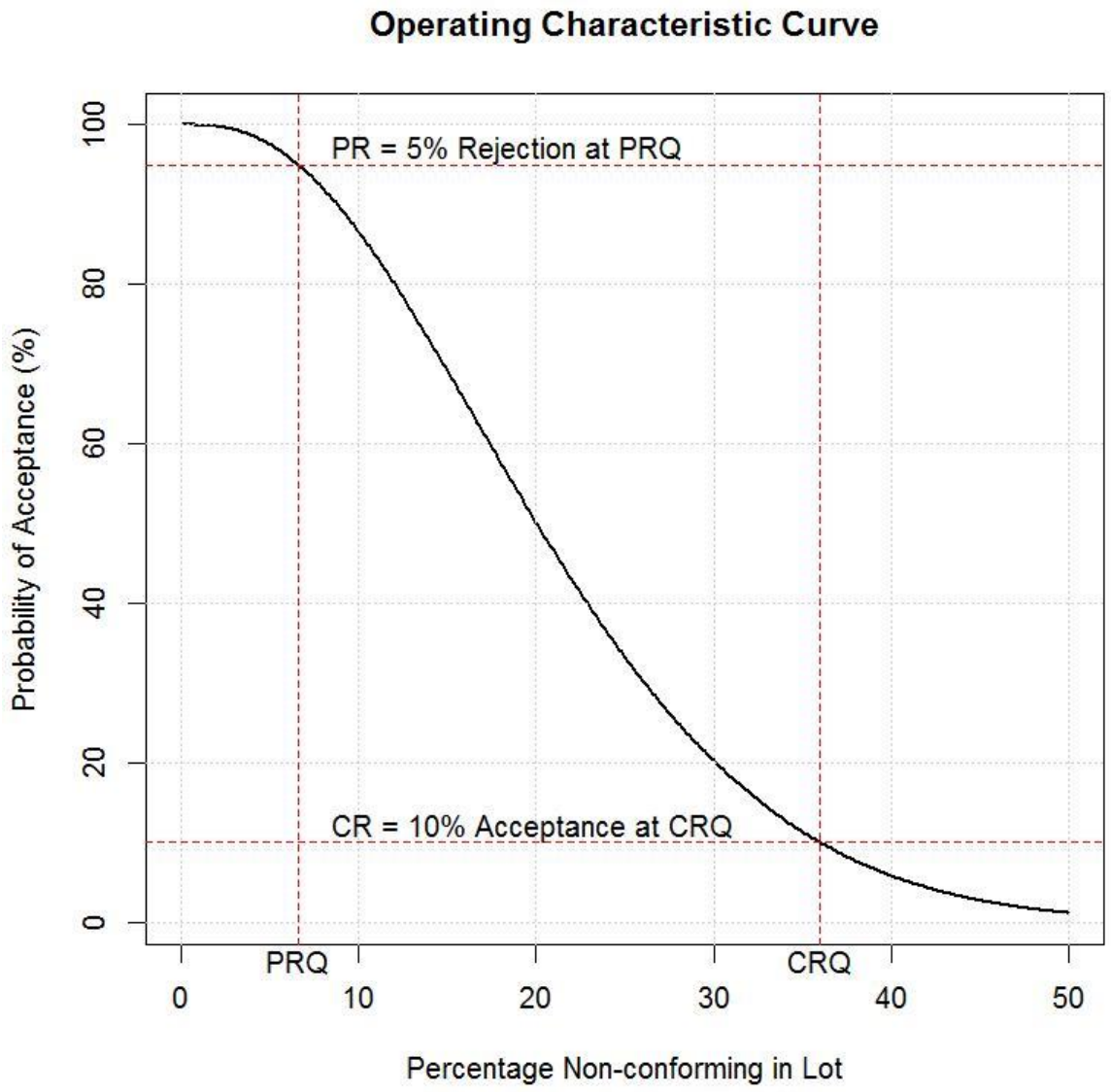
The relationship between the probabilities of acceptance and the assumed levels non-conforming for a sampling plan is known as the Operating Characteristic, usually shown graphically by an Operating Characteristic curve.

Operating Characteristic (OC) curve:

This is defined in ISO 3534 as:

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

The following diagram is an example of an Operating Characteristic curve.



The diagram shows points on the Operating Characteristic that are fundamental to the design of sampling plans:

- the Producer's Risk (PR) is the probability of rejection of a lot considered to be of good quality, having a low level non-conforming;
- the Producer's Risk Quality (PRQ) is the level non-conforming in a lot corresponding to the Producer's Risk;
- the Consumer's Risk (CR) is the probability of acceptance of a lot considered to be of unsatisfactory quality, having an unacceptable level non-conforming; and

- the Consumer's Risk Point (CRQ) is the level non-conforming in a lot corresponding to the Consumer's Risk.

Specification of these four parameters, defining two points on the Operating Characteristic, allows a sampling plan, the number of samples and the acceptance parameter, to be found.

Note that the choice of these points is arbitrary, but is conventional to define points on the curve corresponding to Producer's and Consumer's risks. The choice of Producer's and Consumer's risks, of 5% rejection and 10% acceptance respectively, is also arbitrary, but those values have also tended to become default values, so that designers of plans need to be concerned only with selecting appropriate levels for the Producer's and Consumer's Risk Quality.

Although specification of these parameters enables the design of sampling plans, it is important to consider the probability of acceptance across the entire range of levels non-conforming, to ensure in particular that there is suitable discrimination between lots of good quality and those of poor quality.

These issues are discussed more fully in the following sections.

2.3 Acceptance sampling

Acceptance sampling is the process in which samples are taken from a lot and decisions are made concerning the disposition of that lot, whether the lot is accepted or rejected, based on the results from the testing or examination of those samples.

An acceptance sampling plan specifies the number of samples to be taken and how they are to be taken, the procedure used to test or examine those samples, and the acceptance criterion, based on the results from the testing of those samples, used to decide whether a lot should be accepted.

In general acceptance sampling is used in order to:

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

Most sampling plans are based on the assumption that the lots to which they are applied are homogeneous, so that lots are either accepted or rejected in entirety. A discussion of the effects of inhomogeneity and ways of overcoming it are discussed in section 7.2.3.

2.3.1 Random sampling

Acceptance sampling plans are usually based on the assumption that samples are selected randomly from a lot. For lots consisting of discrete items, this 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 (say) 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' pellets from the number of pallets in the lot;
- 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; 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.

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.

3 Inputs to the Design of Sampling Plans

3.1.1 Introduction

There are two types of inputs to the design of sampling plans; **administrative inputs** such as descriptions of the food or Codex provision the sampling plan applies to and **statistical inputs**, which are those required for the design the sampling plans.

3.1.2 Administrative inputs to the design of sampling plans

The context in which the food will be sampled should be clearly set out, since the sampling plan is specific to a particular context.

For the purpose of the Guidelines, the context for the sampling plan should include consideration of the following points:

Inputs	Description
Provision in a Codex Standard [Refer definitions]	A requirement for a commodity that must be met in order that the commodity conforms to the standard.
An identified food or group of foods	The sampling plan should relate to an identified food or group of foods.
Use of food	Whether the food is intended for direct consumption or used as an ingredient, it's content in the final food and the nature of any further processing steps.
Identified characteristic(s)	A characteristic is the attribute in the commodity to which the provision relates.
Basis of the criteria	The criteria may be a minimum or a maximum limit.
Codex Procedural Manual	Information relating to the scope or field of application and the type of sampling (e.g. bulk or unit).

3.1.3 Generic statistical inputs to the design of sampling plans

The inputs to the design of a sampling plan to assess compliance of products with the specification should state:

- **whether the specification applies to every item in a lot*, or to the average in a lot, or the proportion non-conforming (inferences to be made to lots or processes);**
- **the appropriate acceptable quality levels to be used (levels of risk to be accepted); and**
- **the acceptance conditions of a lot controlled, in relation to the qualitative/quantitative characteristic determined on a sample (decision rules).**

**Requiring compliance of every item in a lot is not the same as requiring that every item in the sample must comply. There are some critical parameters for which 100% inspection is required but no Codex commodity standards require this level of inspection.*

Inputs	Description
Form of the food	The food may be in the form of discrete items or a bulk material.
PRQ, CRQ and associated risks	The PRQ, CRQ and associated risks.
Composition of the lot*	Whether the lot consists of discrete items or is a bulk material. Refer to the example below.
Identified characteristic(s)	Whether the test results are measurements or binary outcomes e.g. pass or fail, including measurements classified as binary outcomes. NB: This is not exactly the nature of the characteristic but more how those results are used as inputs into the criterion of the sampling plan. The nature of the characteristic also includes the statistical distribution for variables data, specifically whether it is normally distributed, a compositional proportion or something else. The nature of the characteristic, its raw state (attributes or variables) and its distribution, if variables, will determine the form of the plan.

Inputs	Description
Lot size	This is not normally an input required for the design of sampling plans intended to control the consumer's and producer's risks in acceptance sampling. However specification of the lot size is needed for attribute plans applied to small lots.
Measurement error	Whether measurement error is significant and if so values for the repeatability, reproducibility and possibly bias describing the measurement error.
Type of sampling plan	The type of sampling plan to be used, chosen from the options available depending on the above factors. This defines the form of the acceptance criterion; the design process will determine the sample size (n) and the 'acceptance parameter' (c or k) in the acceptance criterion.

PRQ, CRQ and associated risks

These define the stringency of the sampling plan, the rate of acceptance of lots in terms of the level non-conforming present in those lots. Usually it is necessary to specify only the PRQ and CRQ as inputs to the design of a sampling plan as the associated risks to producers and consumers are typically assumed to be 5% and 10% respectively. On this basis the PRQ would be defined as the level non-conforming in lots which would be rejected 5% of the time and similarly the CRQ as the level at which the lots would be accepted 10% of the time. However different risk settings can be used if required.

The setting of the PRQ and CRQ needs to also take account of issues such as:

- whether more stringent plans should be used for foods for direct consumption (for relevant parameters); and
- relativity among parameters, for example not having plans for composition that are more stringent than those used for food safety.

In the interests of fairness, stringency should be in keeping with the perceived risks associated with failure and relativity among different parameters. The following example [adapted from the International Commission on Microbiological Specifications in Foods (ICMSF)] shows an approach that could be used to set allowable levels of consumer's risks across different parameters. Each parameter would be ranked according to the rating scale below and then the levels of allowable risk and associated levels non-conforming would be assigned. The process could be extended to also include producer's risk.

Acceptance sampling plans are designed based on the principles of probability to control the risks of making incorrect decisions about a lot, of accepting product of unacceptable quality and of rejecting product of good quality, to desired levels.

Example: Stringency

Risk rating	Severe	Serious	Moderate	Indicator	Utility
Level non-conforming	1%	5%	8%	10%	20%
Consumer's risk (allowable probability of acceptance)	1%	1%	5%	5%	5%

Example: Composition of the lot

There needs to be consideration whether a lot should be regarded as comprised of discrete items or as a bulk material for the purposes of the particular sampling plan. This classification depends on the characteristic being considered, for example for blemishes on the skins of oranges the lot would be considered to consist of discrete units i.e. the oranges.

For kiwifruit, the minimum maturity requirements mean they must have reached an appropriate degree of maturity, in accordance with characteristics of the variety, to allow for development of satisfactory organoleptic characteristics.

The fruit must have attained a degree of maturity of at least 6.2° Brix or an average dry matter content of 15%.

In real situations:

- the maturity specifications are a 'minimum requirement' but are treated as a grade standard in the same way as, for instance, rot in apples
- the test method is not specified but both Brix and dry matter would be by refractometer.

The lot could be considered to be a bulk material in which case it might be possible that a composite sample could be taken and used to perform the inspection.

On the other hand the lot could be considered as consisting of discrete items and we might wish to assess where 90% of the kiwifruit have attained the required maturity status. In this case an attributes plan involving classification of measurements to attributes data would be appropriate.

Therefore careful consideration has to be made of the purpose of the sampling in terms of the assurance to be provided.

The first case might be more suited to kiwifruit intended for processing applications such as jam making, whereas the latter where the kiwifruit are destined for direct consumption by consumers.

4 Design of Sampling Plans

The Codex Procedural Manual and the *Principles for the Use of Sampling and Testing in International Food Trade* (CXG 83 - 2013) state that 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'*.

The Procedural Manual recommends that when commodity committees have included provisions on sampling in a Codex commodity standard, these should be referred to the Committee of Methods of Analysis and Sampling for endorsement along with certain information relating to the sampling plan.

4.1 Design of sampling plans for various situations

This section discusses the design of sampling plans for various situations, depending on:

- the inputs for the design of a sampling plan (section 3); and
- the nature of the lot under inspection, the nature of the measurements made, the presence of significant measurement error and other considerations.

This section provides information on the types of sampling plans that are appropriate for different situations. Sampling plans for each of these situations can be derived from apps or the statistical literature, including standards in some cases. The remainder of this 'Design of Sampling Plans' section discusses the design of sampling plans for various situations, depending on the nature of the lot under inspection, the nature of the measurements made, the presence of significant measurement error and other considerations.

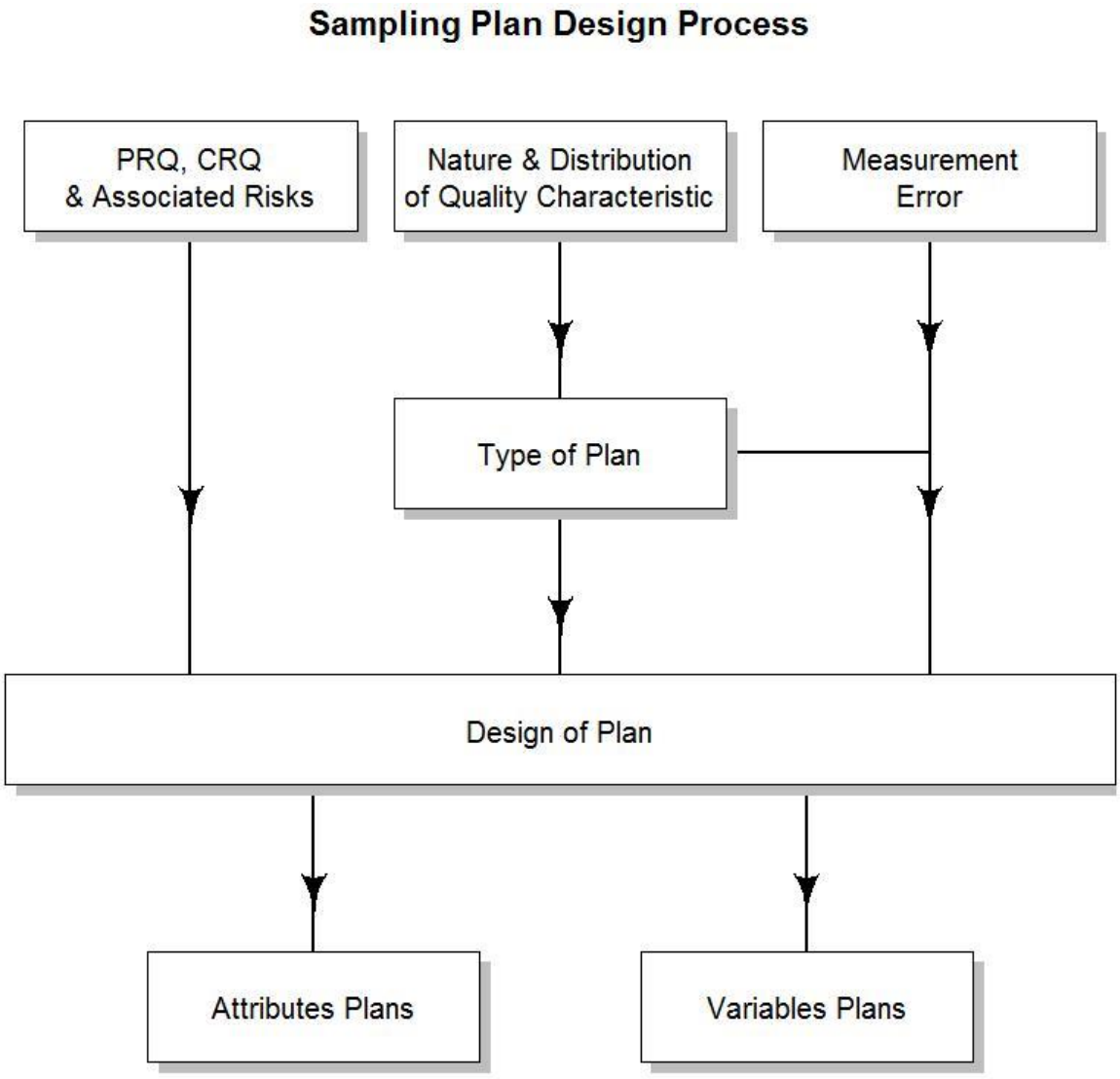
4.1.1 Roadmap

The following Roadmap provides references within these Guidelines:

		Lot Type			
		Discrete Lots	**CXG 50 reference**	**Bulk Materials**	**CXG 50 reference**
Measurement Error	**Plan Type**				
Negligible Measurement Error	Attributes	ISO 2859	9		
		General Design	4.2.3		
	Variables	ISO 3951	4.3.3	Beta distributed parameters	4.4.1 - 4.4.3
		General Design	4.2.3	Compositional proportions	4.4.2 – 4.4.3
	Attributes				

Significant Measurement Error	Variables	ISO3951-1 Annex O	4.3.3	ISO 3951-1 Annex O	4.4.2
		ISO 3951-6	4.3.3	ISO 3951-6	4.4.2
		Repeatability adjustment	4.3.3	Repeatability adjustment	
		Fractional Non-Conformance (FNC)	4.2.3	FNC	4.4.2

4.1.2 Flow diagram



4.2 Sampling plans for inspection by attributes

4.2.1 Introduction

These plans are usually referred to as Attribute Sampling Plans. They are the simplest type of single sampling plan because the inspection results are classified into only two classes of outcomes, e.g. conforming or non-conforming. Because they are applicable to all sampling situations, they have become the benchmark that all other sampling plans can be compared against. Attribute Sampling Plans are also applied to categorical results, where outcomes are measured on a scale.

4.2.2 Form of Attributes Sampling Plans

There are several different types of attributes plans. The simplest cases, where the inspection results are classified into two classes, are referred to as 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 non-conforming items that can be found in the sample. If the number of non-conforming items in the sample is less than or equal to 'c' then the lot can be accepted. If the number of non-conforming items found is greater than 'c' then the lot is rejected.

Plans based on the hypergeometric distribution

If the sample size for an attributes plan is large in relation to the lot size some economy in the number of samples may be possible. As a general rule, such economies are possible if the number of samples is greater than 10% of the lot size. For conceptually infinite lots sampling plans based on the hypergeometric distribution are the same as those derived using the binomial distribution.

Zero-acceptance number (ZAN) plans

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 where only consumer's risk is considered directly. These plans are often used in metrological applications or situations such as pathogens or for foreign matter where acceptance of lots demands that non-conforming items are not found in the inspection.

However it should be noted that just because non-conforming 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 good and poor quality product, so 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 large offsets between the limits used in those plans and levels of contamination at which food becomes unsafe.

Three-class Attribute Plans

In these plans inspection results are classified into three classes, usually referred to as 'good', 'marginal' and 'poor' (or similar descriptors). 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 i.e. they have 'steeper' Operating Characteristic 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, those for which inspection results between m and M, allowed for acceptance of the lot; and
- the numbers m and M define the range within which results are classified as marginal. If m=M the three class plan becomes a two-class plan.

Lots are accepted provided:

- none of the n samples has a level exceeding M; and
- at most c of the samples have levels between m and M.

Evaluation of these plans generally requires an assumption to be made about the underlying distribution of the identified characteristic, such as the log-normal distribution for the microbiological parameters used in the ICMSF plans.

Attribute Plans for Multiple Characteristics

Attribute plans can be easily applied to multiple characteristics by classifying inspected items as non-conforming if any of the individual characteristics are non-conforming. 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 producer's risk, of failing product of good quality.

4.2.3 Design of Attributes Sampling Plans

Refer to the inputs section for baseline information.

There are four situations:

Attribute data in the absence of significant measurement error

This requires specification of only the PRQ, CRQ and their associated risks.

Attribute data in the presence of significant measurement error

In this case measurement errors relate to the probabilities of misclassification, of misclassifying conforming items as non-conforming, and vice versa. In general, larger sample sizes are required to overcome the effects of misclassification. The effect of inspection errors is more serious for zero acceptance number sampling plans.

Measurements (variables data) classified as attributes in the absence of significant measurement error

Use standard attributes plans.

Measurements (variables data) classified as attributes in the presence of significant measurement error

In this context, measurement error refers to repeatability, reproducibility and bias. Plans based on the Fractional Non-conformance (FNC) principle can be used. FNC can be loosely interpreted as the probability that the true value of the sample tested lies outside the limit.

Allowance for random measurement errors is made using reproducibility standard deviations; bias adjustments made where appropriate.

NB: It is more economical, in terms of sampling and testing, to use plans based on FNC rather than classifying results as pass or fail with respect to the specification limit and using an attributes sampling plan.

4.3 Sampling plans for inspection by variables

4.3.1 Introduction

These plans are usually referred to as **Variables Sampling Plans**. If the underlying distribution of individual measurements is known, acceptance sampling can be performed directly on the measurements themselves. This often allows a considerable saving in sample size but we need to know the probability distribution of the underlying measurements. The Gaussian or normal distribution is commonly adopted as the distribution of the measurements. For compositional proportions in bulk material, the beta distribution is more appropriate but the normal distribution can serve as an approximation.

The advantages of variable sampling plans are:

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

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 (assuming a distribution to be able to make a decision on the disposition of the lot, not to estimate the quality or the parameters underlying that distribution)
- variables sampling plans are only applicable to one characteristic at a time;
- there may be a higher inspection cost per unit;
- there may be higher clerical cost per unit due to the calculations involved;
- a lot with no non-conforming units may be rejected by a variables plan; and
- there is a possibility that no non-conforming units are found to show to the producer after rejection.

4.3.2 Form of Variables Sampling Plans

In variables plans, the mean ($\bar{X}$), is compared with the acceptance limit in a similar way to the attributes plans but, in order to allow for the variability in the lot, the sample standard deviation 'S' is computed.

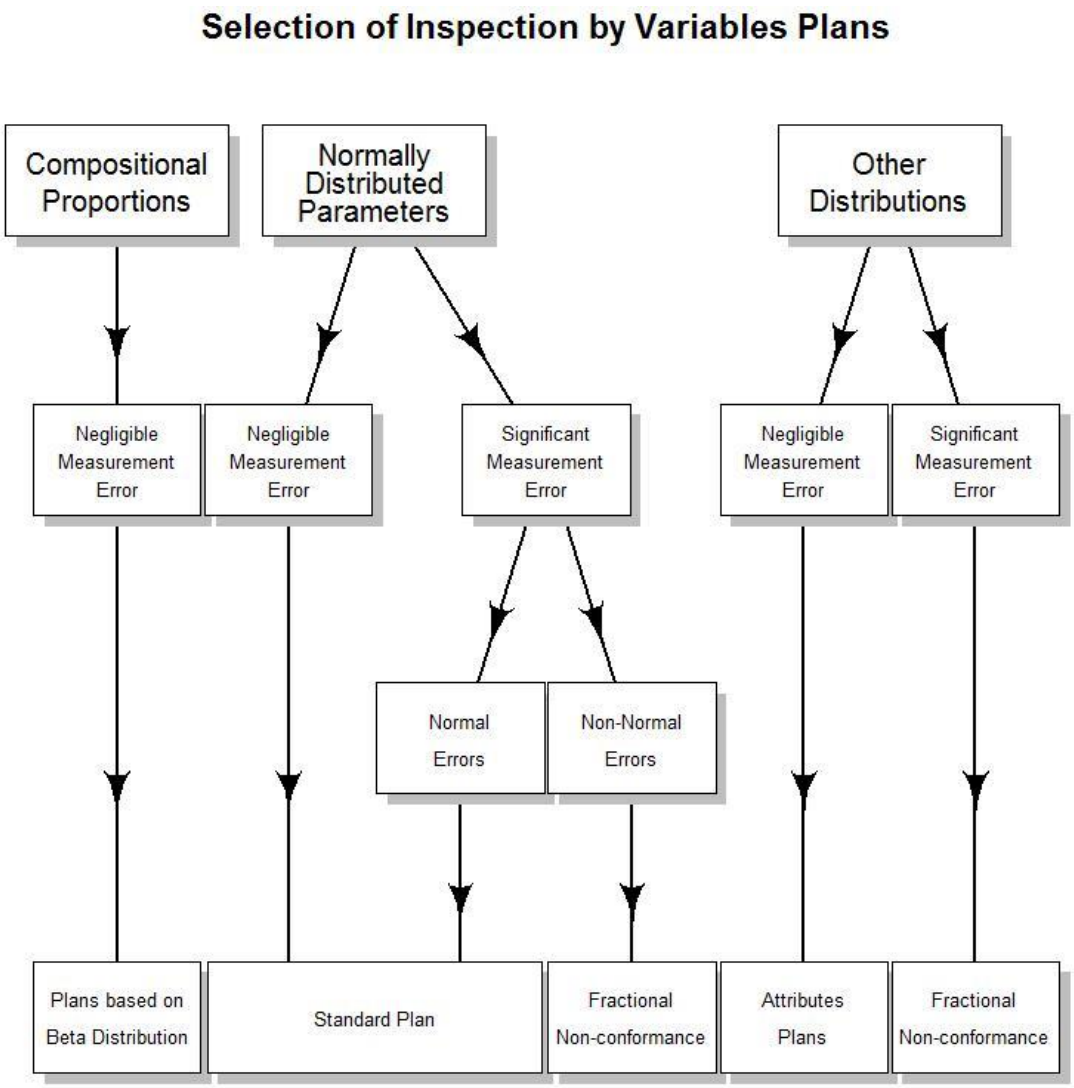
Variables sampling plans are defined by two numbers, the sample size 'n', the number of items to be taken from the lot under inspection and the acceptability constant 'k', the multiplier of the standard deviation in the acceptance criterion.

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

4.3.3 Design of Variables Sampling Plans based on the normal distribution

Refer to the inputs section for baseline information.

Refer to the following flow chart for the design of Variables Sampling plans:



There are three situations based on the normal distribution as follows. There is discussion on other plans, for example plans based on beta distribution in other parts of this document.

Negligible measurement error

This requires specification of only the PRQ, CRQ and their associated risks. It is assumed that the characteristic follows a normal distribution in the lot under inspection – if this is not the case, a simple but possibly inefficient option is to classify the measurements as attributes with respect to the specification limit.

Usually the standard deviation representing variation in the lot is calculated using the results from testing of the lot under inspection. These plans are referred to as 's-plans' where 's' refers to the estimated value of a standard deviation, rather than the true value [σ – 'sigma'] that this estimate represents. However in some cases, such as when there has been a long history of stable performance of a manufacturing process, the true value of standard deviation (σ) representing variation in the process can be considered known. This leads to so-called 'sigma-plans' (σ-plans) that have the advantage, relative to s-plans, that fewer samples are needed to control risks to the same levels as that required for the corresponding s-plans.

The σ-method is also applicable to the other situations for inspection by variables discussed below.

Significant Repeatability Measurement Error

This situation is also dealt with in ISO 3951-1 Annex O, assuming that measurement error also follows a normal distribution a plan allowing for repeatability error has the same acceptability constant (k value) but requires a larger sample size than the 'error' free plan based on the same control of risks.

Sampling plans such as those in ISO 3951 Annex O plan that allow for measurement error depend on the 'error variance ratio', the square of the ratio of the repeatability standard deviation to the process standard deviation.

Measurement error is often considered significant when this ratio exceeds 1% (i.e. when the measurement error standard deviation is more than 10% of the process standard deviation). However this does not take non-repeatability measurement error into account – this is discussed in detail in ISO 3951-6 [currently under development].

It is assumed that the characteristic follows a normal distribution in the lot under inspection. If not, plans based on Fraction Non-conformance can be used.

Significant General Measurement Error

In this context, measurement error refers to repeatability, reproducibility and bias. This is dealt with in ISO 3951-6. It is assumed that repeatability and reproducibility, as well as the identified characteristic, are normally distributed.

4.4 Sampling of bulk materials

4.4.1 Introduction

Bulk materials are continuous, consisting for example of particles of different density and sizes etc. An example is milk powder. It is impossible to view bulk materials present in a lot as a set of distinct objects because there is no way of selecting the items one by one in a way that is not biased when using simple random sampling. This is where a different methodology is introduced, which brings with it sampling bias and non-representativeness.

Some general objectives of bulk sampling are:

- acceptance on a lot-to-lot basis;
- characterize the material as to grade, any need for further processing, and its destination;
- control during processing;
- 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. These factors include things such as how the device is employed, and the conditions that the device is used under.

In bulk sampling, the lots of bulk material are 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. Other means of real or synthetic division can also occur.

Bulk materials being continuous means parts of samples can be mixed together to form a composite. This composite then gets tested only once, rather than having to do many tests on the individual parts. This is a physical way of creating a composite sample representing the average content of lot.

Since bulk materials are continuous, parts of samples can be mixed together to form a composite. This composite then gets tested only once, rather than having to do many tests on the individual parts. This is a physical way of creating a composite sample representing the average content of lot.

Composite sampling is commonly used for bulk products. Variables sampling inspection plans based on the beta distribution fit well for this context.

4.4.2 Design of general sampling plans for bulk materials

Refer to the inputs section for baseline information.

In the simplest case, such as the inspection of bulk materials of manufactured products, the standard attributes or variables plans can be used. However for some bulk materials, for example bulk shipments of grains, the variation within a lot cannot be satisfactorily described by a single standard deviation. Special techniques are required for this type of situation. However the statistical techniques are complex and it is only possible to provide an overview in these guidelines – see section 4.4.4.

Inspection by Attributes Plans

As noted in section 4.1.1, attributes plans are considered the benchmark as they can be used in any situation, including for the inspection of lots of bulk materials. The process described in section 4.1 can be used to design sampling plans for attributes plans for the inspection of bulk materials.

When measurement error is significant, plans are based on the FNC, as described in section 4.1.3, are suitable in this context as it is more economical to use plans based on FNC rather than classifying results as pass or fail with respect to the specification limit and using an attributes sampling plan.

Inspection by Variables Plans based on the normal distribution:

Variables plans such as those described above and in ISO 3951 can be used for bulk materials. The methods described in ISO 3951-1 Annex O and ISO 3951-6 can also be applied to the inspection of bulk materials in situations where measurement error is significant.

NB: The plans in ISO 3951 are designed to control either the producer's risk, or the consumer's risk, so they might not be immediately suitable.

ISO 3951 assumes that the characteristics under inspection are normally distributed. This might not apply for bulk materials in general; in particular variables sampling plans based on the normal distribution can only be approximate for compositional proportions and their use can involve higher consumer's risks than desired.

4.4.3 Sampling plans for compositional proportions (where measurement error is negligible)

Compositional characteristics are often quality measures for bulk materials. For example, the percentage fat with a minimum limit of 26% is a primary quality measure for milk powders. Compositional proportions, also referred to as mass fractions, are characterized by units of measure such as percentages [by mass], mg/kg, µg/100g and the like, which are, strictly speaking, 'dimensionless' numbers lying between 0 and 1.

Compositional fractions in a lot of manufactured product 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 higher consumer's risks than desired.

In these plans 'm' increments/subsamples are taken from a lot under inspection, with the value of m determined in the design of the plans. These samples may be tested individually or combined to form a composite which is tested once.

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 addition to the PRQ, CRQ etc. to design these plans we also need an estimate of the 'precision parameter' for the beta distribution, denoted by θ, which can be obtained from the analysis of historical data.

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

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

A feature of the beta distribution is that the standard deviation depends on the mean 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 θ is the precision parameter for the beta distribution, estimated from historical data (see below).

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

4.4.4 Special sampling plans

[Under development]

5 Outputs of the Design of Sampling Plans

The designed sampling plan should include key information

A commodity committee might choose an alternate approach, specifying a minimum of the PRQ, CRQ and associated risks but possibly also including details of assumed values measurement error and other parameters on which the design of the plan has been based.

5.1.1 Key information

The designed sampling plan should include key information:

- details of the proposed sampling plan including the purpose (for inspection of the provision in the commodity), the number of samples to be taken and details of the acceptance criterion;
- reference to the source of the sampling plan;
- the Operating Characteristic;
- justification for the choice of PRQ, CRQ especially where these might not align with usual expectations;
- the inputs used in the design of the plan including values of parameters such as measurement errors or presumed known standard deviations representing lot variation; and
- considerations about the impact on producers, producers risks, what plans producers might have to use etc. relating to fitness for purpose.

Of course the outputs from the design process differ from those needed in a submission to CCMAS or other parties for that sampling plan to be considered for approval. In this case details of the inputs and possibly some justification for the choice of values for those inputs would also be required along with details of the sampling plan itself. There should also be some consideration of the impact of the proposed plan on producers, both in terms of rejection of product and the plans producers themselves might have to use.

6 Evaluation of Sampling Plans

A sampling plan that is designed based on specified inputs needs review to ensure that it is fit for purpose and meets the Codex requirements of fairness.

Evaluation should consider both fitness for purpose as well as the expected components of a sampling plan intended to provide for fair and valid sampling.

It is possible that the initial plan might not meet these requirements so that the design of a suitable sampling plan becomes an iterative process; in particular it may be necessary to re-define the allowable risks to ensure that the plan is fit for purpose and practical to apply.

6.1 Fitness for Purpose

6.1.1 Introduction

In the general sense, 'fitness for purpose' entails cost, practicality and fairness.

6.1.2 Practicality

The *Principles for the Use of Sampling and Testing in International Food Trade* (CXG 83 – 2013) states that sampling and testing procedures selected should be fit for their intended purposes:

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

The aims of sampling are:

- to see that the consumer receives product having the required quality; and
- to control producer's risk for the non-acceptance of good quality product.

These determine whether a sampling plan is fit for purpose. However, other issues such as the cost and practicality associated with the use of sampling plans should also be considered.

The designer of sampling plans should consider the total cost when making decisions about the practicality of sampling plans, taking account of the costs of incorrectly accepting or incorrectly rejecting lots as well as the cost of testing.

Sampling plans can be designed to control overall cost, specifically the costs associated with acceptance of non-conforming lots and the rejection of compliant lots, but costs associated with sampling and testing, which are usually smaller, and other costs can also be taken into account.

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

- 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; and
- the use of 'indifference' plans that are designed around the level of defects at which there is 50% acceptance, rather than based on PRQ, CRQ. This leads to plans having more manageable sample sizes.

6.1.3 Fairness

As described in CXG 83, fairness must necessarily involve consideration of both consumer's and producer's risks, to avoid situations such as the following:

Sampling plans having inappropriate stringency, not commensurate with the situation e.g. plans for assessment of composition that are more stringent than those for food safety.

High producer's or consumer's risks that may arise due to plans not designed from specifications of those risks.

Plans not based on statistically valid principles, e.g. failure to allow for either sampling or measurement errors or inappropriate allowances made for these errors.

Use of single sampling plans, including those chosen from sampling schemes, might be unfair, even though producer's and consumer's risks have been specified in their design, for example:

- use of the same sampling plan by the producer in a situation of deteriorating quality could result in increased consumer's risk (even assuming that only product that passed the producer's assessments was received by the consumer); and
- there is always a chance that product of good quality may fail a consumer's inspection.

NB: these points (in bold) would be or should be taken into account in the design process. However evaluation might also apply to a plan taken from the literature.

However fairness 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 the consumer. For example, in the interests of fairness, one should ensure that producers are not exposed to unreasonable costs in terms of increased costs of sampling and testing, loss of yields, or excessive rejection of their products in order to achieve compliance.

6.2 Endorsement of sampling plans

6.2.1 Introduction

This Guideline is intended to provide guidance to commodity committees on the development of sampling plans. Subsequent CCMAS endorsement of sampling plans is based on the information provided by commodity committees along with expertise to judge the validity and appropriateness of the proposed plans.

Codex commodity committees should provide CCMAS with key information for consideration of endorsement of sampling plans.

This key information is set out in section 5.

CCMAS will consider the proposal in keeping with the principles outlined in CXG 50 and should be in a position to endorse the sampling plan presented, whether the plan is sourced from CXG 50, other Codex standards, standard development organizations (preferably ISO), scientific journals or publications, so long as the sampling plan that will meet the requirements of the commodity committee to demonstrate 'fair and valid sampling procedures are used when food is being tested for compliance with a particular Codex commodity standard'.

These key parameters include the inputs to design of the sampling plan and the evaluation of the plan.

In some cases a rationale might have to be provided to support decisions made (e.g. decisions on stringency that might not align with usual expectations) or values of parameters describing measurement errors.

6.3 Document and communicate sampling plans

Sampling plans or outcomes leading to sampling plans, such as the PRQ and CRQ and the associated risks, should be included in the commodity standards, either in an individual standard or a general standard such as CXS 234.

7 Implementation of Sampling Plans

This section deals with aspects concerning the practical application of sampling plans.

CXG 83: sampling involves a three stage process [physical] sampling procedures, testing of those samples and the assessment of lot conformance. The implementation of the sampling plan should include the physical sampling procedures, re-inspection and dealing with sampling plan problems.

7.1 *Physical sampling procedures*

[Under development]

This may include:

- Sampling procedures were not part of the agreed list of actions. However, we will include a section on the principles of sampling procedures based on CXG 50 as well as reference to international standards on sampling procedures such as ISO 707 | IDF 50 Milk and milk products – Guidance on sampling.
- NMKL Procedure 12 also deals with physical sampling procedures.

7.1.1 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 nearby and convenient. It is a non-probability sampling and sometimes used in pilot testing.

It is a quick method of sampling that is readily available, often being at a low cost.

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 be subject to challenge being neither a fair nor a valid procedure.

7.2 Dealing with disputes over sampling

Often the first indication of possible issues with sampling and testing occurs when there is a dispute over the assessment of product. Disputes between parties in trade may occur for many reasons including differences in the testing between the laboratories concerned; the existence, appropriateness and statistical validity of the sampling plan used to assess the product; the allowances made for general measurement error and within-lot variation product variation; differences in physical sampling procedures; differences in composition of the samples tested due to product inhomogeneity or changes occurring during storage and/or transport of the product.

Only after all other possible causes, other than sampling, have been ruled out can be established that the problem lies with the absence of, inappropriately set or incorrect application of sampling plans.

Some of these issues are discussed below.

7.2.1 Absence of or inadequately defined plans

Clear definition of sampling plans is needed for transparency, especially to provide clarity around acceptable risks so producers can design appropriate sampling plans to ensure that expectations of consumers are met.

Disputed lots provide motivation for the development of sampling plans which, in accordance with CXG 83, should have been developed before trade commenced. The same process should be followed in the design of the sampling plan; within-lot product variation and measurement error would be standard inputs in to the design of a sampling plan.

7.2.2 Re-inspection

Sampling inspection plans usually assume that a random sample is taken from the lot. When random sampling of pre-packaged commodities from large containers is difficult, physical sampling may be done poorly. Hence it is natural for the producers or consumers to occasionally suspect or dispute the sampling done.

When the original inspection results are suspect, the provision for lot re-inspection can be incorporated. Re-inspection is done when a lot was rejected on the first inspection, but the lot is resubmitted for acceptance inspection so that a new sample can be taken to make a decision. This process can be repeated but the design of the sampling plan depending on the number of re-inspections allowed.

In addition, the use of sampling plans having relatively small sample sizes can result in high risks of making incorrect decisions, so re-inspection plans should be used in the interests of fairness.

Re-inspection schemes are particularly useful for zero acceptance number sampling plans. It is well known that the zero acceptance number plans generally involve a higher risk to the producer. Hence use of a re-inspection scheme allows the producer to opt for re-inspection of the lot when there is good process history to believe that the quality of the lot is indeed good but a lot has been rejected due to poor sampling or problems with measurement. Variables sampling plans with large ‘k’ values such as k= 2 can also be harsh on producers. These plans also involve small sample sizes. Re-inspection can also be used to reduce the producer’s risks.

7.2.3 Inhomogeneous lots

Inhomogeneous lots might occur because inspection lots differ from manufacturing lots or other reasons.

The term homogeneous is used to describe lots that are similar in nature, having similar quality throughout. The level non-conforming in a homogeneous lot varies at random around the same average level non-conforming throughout the lot. Acceptance sampling plans are generally intended to be applied to lots that are homogeneous.

Acceptance inspection and compliance testing often necessitate levels of protection for both the consumer and the producer that require large sample sizes relative to lot size. A given sample size can, however, be made to apply to several lots jointly if the lots can be shown to be homogeneous. This reduces the economic impact of a necessarily large sample size. If the lots are not homogeneous, then this is unable to occur.

Lot heterogeneity usually increases producer’s and consumer’s risks, so that consumer protection may be compromised when an inspection lot is not homogeneous.

For a non-homogeneous lot of unknown status, one approach may be to split that lot into sub-lots in line with production lots or other standardized manufacturing processes. Each of the sub-lots might then be sufficiently homogeneous to be inspected using standard attributes or variables sampling plans.

8 Other Sampling Plans

This part provides guidance on sampling plans not already mentioned. Some sampling plans are not included such as 'compliance of the average level' as this parameter is not included in any Codex provision.

Sampling plans not included in these Guidelines could be taken.

8.1 *Microbiological sampling plans*

[For review as to whether this stays in the Guidelines]

Microbiological plans are referred to in GL 50 (section 3.2) and there is reference to GL 50 as a source for the development and selection of sampling plans in the 'Principles and Guidelines for the Establishment and Application of Microbiological Criteria related to Foods (CXG 21-1997)*'.*

Further comments on this:

- *two and three class attributes plans are commonly used in microbiological applications, often involving classification of micro counts as conforming or non-conforming with respect to a limit in two class plans, or to acceptable, marginal and unacceptable in three class plans;*
- *two class plans are simply the usual attributes plans discussed in section 4.1; and*
- *another complication is that quite often M is not the level at which food becomes unsafe for consumption, there is generally a safety margin between M and the unsafe level.*

If we're going to proceed we possibly need a tool to evaluate 3-class plans.

8.2 *Compliance of the average level*

[For review as to whether this stays in the Guidelines]

There is increased use of acceptance sampling plans in connection with compliance testing to Government standards and validation tests etc.

There may be only one example of this, for weights by the average quantity system.

Which, as far as we are aware, is not included in any Codex provisions.

8.3 Double sampling plans and equivalent (single) plans

8.3.1 Introduction

Double sampling plans are often referred to as two stage plans. Product is initially inspected using the first stage plan; if the lot is of very good quality or very poor quality (relative to the current PRQ and CRQ respectively), the lot is either accepted immediately or rejected immediately.

If the lot is neither accepted nor rejected at the first stage, then a further inspection is carried out using the second stage plan. The second stage plan need not be the same as the first stage plan, depending on the design of the plan.

While they are more complex to administer, two stage plans have the advantage, relative to single stage plans, that overall fewer samples need to be taken and tested that for the corresponding single stage plan having the same stringency (same PRQ and CRQ).

For a specified double sampling plan it is possible to design an equivalent single sampling plan, by matching the OC curve of the double plan with that of the single sampling plan. Note that for attributes plans it might not be possible to match the OC curves exactly.

8.3.2 Form of the sampling plan

The input will be a specified double sampling plan, based either on either inspection by attributes or inspection by variables.

8.4 *Conformity testing*

[For review as to whether this stays in the Guidelines]

Conformity testing, also known as evaluation of conformity or compliance testing, is used to assure that an entity meets a specific requirement and/or regulatory standard. In this context the entity refers to the sample actually tested.

The objective of conformity assessment differs from that of acceptance sampling - acceptance sampling uses a limited number of samples to determine whether to accept or reject a lot of some product

whereas, in contrast, in conformity testing the inference is limited to the 'entity' i.e. the sample tested. In other words, conformity testing is a procedure for making a decision about the particular sample whereas the proportion nonconforming in a lot is the main quality measure of interest in acceptance sampling.

Examples of conformity testing include the test of the concentration of some trace elements in the blood of employees for their health evaluation, the analysis of an athlete's urine to detect abuse of xenobiotic anabolic steroids, testosterone and doping etc.

The specification for the quantifiable characteristic, such as the maximum allowable concentration of a drug or trace element in blood for normal people, is called as a limiting value (LV) in the conformity testing protocols. The LV could be understood as either a minimum value (Lower Limit or Lower Specification Limit) or a maximum value (Upper Limit or Upper Specification Limit), or both. The interval containing all permissible values of the characteristic is called the region of permissible values. A conformity testing protocol provides assurance of conformity by checking whether the measurement of interest falls within the region of permissible values or not. Conformity can be declared if and only if the whole uncertainty interval is located within the region of permissible values.

Measurement and sampling uncertainties, including metrological traceability, become crucial for the declaration of conformity, especially when the measured value is close to the set limiting value.

Measurement uncertainty is usually reported as an uncertainty interval, given in the form of a confidence interval. The common practice for conformity testing is to compare the measurement uncertainty interval around the measurement result with the region of permissible values.

The main disadvantage of the conformity testing procedure is that in many cases, inconclusive results will be obtained even though a sample is conforming but due to measurement errors, the uncertainty interval includes the limiting value.

ISO 10576-1 International Standard recommends performing the conformity test as a two-stage procedure. The rules for asserting conformity or nonconformity are:

- *assurance of conformity: The uncertainty interval is inside the region of permissible values;*
- *Assurance of non-conformity: The uncertainty interval is included in the region of non-permissible values; and*
- *inconclusive result: The uncertainty interval includes LV.*

The ISO 10576 Standard does not encourage reduction of measurement errors by design and hence poorer measurement systems will produce more inconclusive results. Hence producers may be forced guard-band in order to reduce the incidence of inconclusive results from measurements.

Alternatively, a two-stage conformity testing procedure based on FNC has been found to reduce the probability of incorrect declarations of conformity or of inconclusive results for nonconforming entities (Type II error) when the number of test samples is greater than one and this superiority becomes more significant when the sample size increases.

9 General Information

9.1 Explanation of ISO sampling plans

9.1.1 Introduction

The two standards ISO 2859 Sampling procedures for inspection by attributes and ISO 3951 Sampling procedures for inspection by variables are the two principal ISO standards dealing with sampling inspection. These standards are based on the following principles and assumptions:

- they are applicable to lots consisting of discrete items;
- the sample size is determined according to the lot size;
- the standards describe sampling schemes, i.e. sets for sampling plans for normal, tightened and reduced inspection, with switching rules based on recent quality history to swap between those plans;
- the sampling schemes are designed to specifically control either the producer's risk, or the consumer's risk, but not both; and
- it is assumed that measurement error is negligible in the construction of most of these plans although ISO 3951 does contain some information relating to adjustment for repeatability error.

9.1.2 Lot size vs sample size

Statistically, the lot size itself does not have an important role in determining protection to consumer and producer whereas changes in sample size does effect on the protection afforded by a plan.

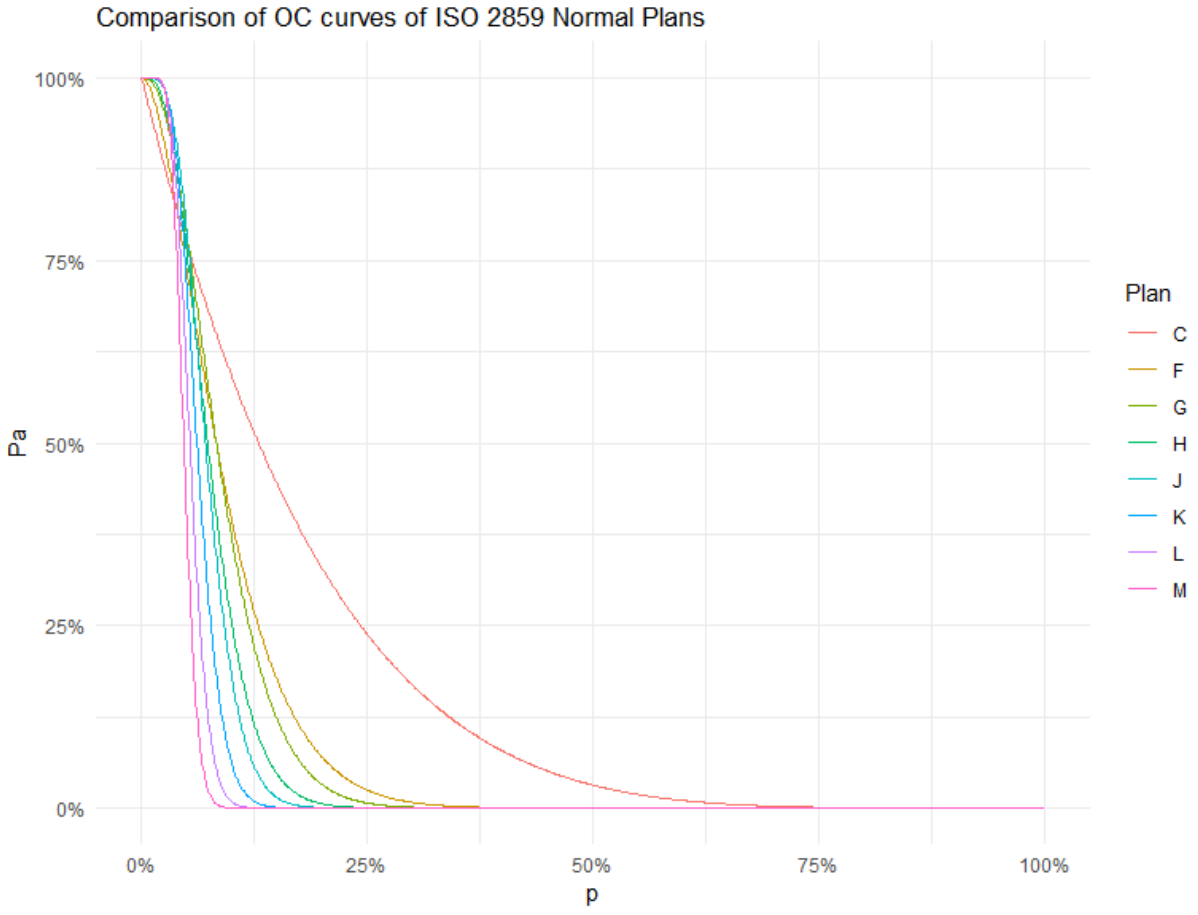
However despite this, a lot size versus sample size relationship has been built in to the design of the sampling plans appearing in the ISO standards. This relationship is arbitrary, and has been changed over time, although it has the general effect of reducing the risks of making incorrect decisions for larger lots, where the costs incurred from incorrect decisions will be greater.

To achieve this, the designers of the ISO plans have chosen not to explicitly control either the consumer's risk in the plans based on PRQ or the producer's risk for plans based on the CRQ.

Sampling schemes indexed by PRQ do not fix the consumer's risk at a constant level such as 5%. The consumer's risk will decrease only for large lot sizes. The table given below shows ISO 2859 normal single sampling plans for a PRQ of 2.5% (Level II General Inspection):

Lot size range	Sample Code	(n , c)	Producer's Risk		Consumer's Risk	
			Level non-conforming	Probability of Rejection	Level non-conforming	Probability of Acceptance
16-25	C	(5,0)	2.5%	0.119	36.9%	0.10
91-150	F	(20,1)	2.5%	0.088	18.1%	0.10
151-280	G	(32,2)	2.5%	0.045	15.8%	0.10
281-500	H	(50,3)	2.5%	0.036	12.9%	0.10
501-1200	J	(80,5)	2.5%	0.015	11.3%	0.10
1201-3200	K	(125, 7)	2.5%	0.014	9.2%	0.10
3201-10000	L	(200, 10)	2.5%	0.013	7.6%	0.10
10001-35000	M	(315, 14)	2.5%	0.014	6.3%	0.10

The following graph shows the OC curves of the above selection of plans. The producer's and consumer's risks differ significantly for these plans and the selection is solely guided by the lot size.



As a consequence of employing a sample size versus lot size relationship, ISO has designated that sampling plans indexed by PRQ, explicitly controlling the producer's risk, are intended for the inspection of a continuing series of lots and plans indexed by 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.

9.1.3 Sampling schemes

ISO standards employ sampling schemes, sets of sampling plans with different levels of inspection to ensure quality is effectively controlled. Sampling schemes also contain 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 plan is the plan used when the process is considered to be operating at, or slightly better than, the Acceptance Quality Level (AQL).

Tightened inspection plan is a plan for using stricter acceptance criteria 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 AQL by introducing a higher rate of rejection.

Reduced inspection plan is a plan which 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.

However switching rules are considered too complex to apply in international trade, particularly from a consumer's point of view.

It is possible to design an equivalent [single] sampling plan that controls the producer's and consumer's risks to the same levels as an overall sampling scheme.

9.2 *Guidance for producers*

[For review as to whether this stays in the Guidelines]

The provision of 'fair and valid sampling' means consumers, including commodity committees should ensure that plans they develop not only provide for satisfactory control of producer's risks but also that the plans that producers use to ensure that consumer expectations are met, do not place unnecessary or unfair burden on producers by way of sampling and testing, and impacts due to lot rejection or on product yields etc.

Acceptance sampling is a tool used to manage the public health risk management tool for food. Key points are:

- *use of acceptance sampling should be considered in relation to other public health risk management tools that may be used[8];*
- *the producer may need to apply more stringent sampling plans than the consumer in order to avoid an unacceptable risk of rejection upon inspection by a consumer; and*
- *food growers or manufacturers may apply acceptance sampling.*

[8] Principles for the Use of Sampling and Testing in International Food Trade (CXG 83-2013), Guidelines for Food Import Control Systems (CXG 47-2003) and its Appendix, Principles and Guidelines for Imported Food Inspection Based on Risk

10 Appendix: Definitions

Acceptance Sampling Plan

An **Acceptance Sampling Plan** is one intended for determining the acceptance or the rejection of a lot. The plan specifies:

- the number of samples to be taken and how those samples are to be taken from a lot;
- how those samples will be tested; and
- the criterion, based on the test results obtained, used to determine whether the lot is accepted or rejected.

ISO 3534: 4.3.3

Acceptance sampling plan - plan which states the sample size(s) (1.2.26) to be used and the associated criteria for lot (1.2.4) acceptance.

Consumer and Producer

The terms ***'producer***' and '***consumer***' are conventional and may apply to a range of different operators in the food chain, such as a grower, manufacturer, the manufacturer's own quality control system, supplier, exporting country, processor, on-seller, or importing country.

Consumer's Risk (CR)

Consumer's Risk is the probability of wrongly accepting a lot that is not of acceptable quality. It is a point on the OC curve corresponding to a predetermined and usually low probability of acceptance.

Consumer's Risk Quality (CRQ)

Consumer's Risk Quality is the quality level of a lot or process which, in an acceptance sampling plan, corresponds to a specified consumer's risk.

Inspection by Attributes

The **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 non-conforming 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.

Inspection by Variables

The **Inspection by Variables** starts with selecting a sample of a number of items and measuring dimensions or characteristics so that information is available not only on whether a dimension, for example, is within certain limits but on the actual value of the dimension. The decision whether or not to accept a lot is made on the basis of calculations of the average and the variability of the measurements.

An inspection by variables sampling plan specifies the **number of samples (n)** and an **acceptability constant (k)**. A lot is accepted against an upper specification limit if the acceptance criterion 'average result + k * the standard deviation of results' does not exceed the upper limit, and similarly for a lower limit. In other words the acceptance criterion is based on the average value $\bar{x}$ and the standard deviation of the results from the testing.

The values of n and k are worked out from the specified levels of allowable risk.

Producer's Risk (PR)

Producer's Risk is the probability of wrongly rejecting a lot that is of acceptable quality. It is a point on the OC curve corresponding to a predetermined and usually high probability of acceptance.

Producer's Risk Quality (PRQ)

Producer's Risk Quality is the quality level of a lot or process which, in an acceptance sampling plan, corresponds to a specified producer's risk.

Provision, Characteristic, Standard

[Under development]

*A **provision** is a requirement for a commodity that must be met in order that the commodity conforms to the standard.*

*A characteristic is the **attribute** in the commodity to which the provision relates.*

*A **standard** is a set of provisions relating to a commodity, all of which must be met in order that the commodity conforms to the standard.*

Example

Fat in Whole Milk Powder (WMP) must exceed 26%.

Identified food or group of foods e.g. Milk powders and Cream Powders (Standard for Milk Powders and Cream Powders (CXS 207-1999)).

The attribute is the 'characteristic' in the commodity to which the provision relates e.g. fat.

Provision is the requirement that must be met e.g. must exceed 26%.

Sampling

Sampling is the procedure used to draw one or more items from a population (or lot) and intended to serve as a basis for a decision about the population (or lot).

Note: Sampling should not be confused with *physical sampling* e.g.

- IDF50/ISO 707 *Milk and milk products -- Guidance on sampling;* and
- ISO 7002 *Agricultural food products—Layout for a standard method of sampling from a lot.*

Information note

Representative samples

The word representative has no meaning, statisticians do not use this word …why not use sampling procedures that are dictated by statistical theory, with the advantages of less cost, and with meaningful, calculable tolerances?

- W.E. Deming

Sampling error

[Under development]

Sampling Plan

A **sampling plan** is one according to which one or more samples are taken from a lot in order to obtain information about or possibly reach a decision about that lot.

Information note

Confidence

The term 'confidence' is often used in conjunction with sampling plans. However while it is a statistical term, in reality it has nothing to do with acceptance sampling. It is easier to express risks in terms of probabilities of acceptance or rejection.

Confidence can be associated with consumer's risk, for instance 95% confidence [that the lot is of satisfactory quality] means there is only 5% chance of acceptance.

However confidence does not work well with producer's risk

Sampling scheme

A **sampling scheme** defines what data will be obtained and how. Precision and systematic sampling error are two principles that guide the choice of sampling scheme.

Appendix II

The proposed draft supplementary e-book
(for comments through CL 2020/27-MAS)

Codex Sampling

[This document provides an update of the e-book that sits alongside the GL50 guidelines. It contains additional information on sampling and apps for the design and evaluation of sampling plans.

For technical and other reasons it has not been possible to re-issue the e-book electronically at this time. The previous version of the e-book can be viewed here: GL50 e-book v1.]

Chapter 1 Introduction

Acceptance sampling is the methodology giving the procedures by which decisions to accept or not accept (a lot or a series of lots usually) are based. Acceptance sampling depends on the results of the inspection of samples.

Acceptance sampling is preferred when: testing is destructive, the cost and time for 100% inspection are high, or there are limitations of the work force.

There are some disadvantages of acceptance sampling. These include the risk of accepting bad lots, or rejecting good lots. Acceptance sampling does not provide any direct form of quality improvement, it simply accepts or rejects lots.

The Codex Procedural Manual and the Principles for the Use of Sampling and Testing in International Food Trade' (CXG 83-2013) (GL 83) state that 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'.*

Fairness can only be established by consideration of both the consumer's and the producer's risks. This revised CXG 50 contains sections covering:

- Concepts of sampling;
- Guidance to design a sampling plan for foods;
- Sampling plan tools (containing links to apps of sampling plans tools, rather than the larger document full of tables, plots and formulas);
- Other identified technical information e.g. measurement error, sampling of bulk materials, sampling of non-homogeneous lots (refer to REP18/MAS Appendix VI: Prioritization); and
- Links to other sources of scientifically valid sampling plans.

The sampling plan tool (refer 2.3.1) allows **for control of *both* consumer's and producer's risks** as part of the design. This tool will also produce an Operating Characteristic (OC) curve. The OC curve is an important component of sampling plan design as it is used to gauge the protection to the consumers and producers. The Codex Procedural Manual says *'a commodity committee should, whenever possible, provide information to CCMAS for each sampling plan relating to the scope or field of application, the type of sampling (e.g. bulk or unit), sample sizes, decision rules, details of plans (e.g. Operating Characteristic curves), inferences to be made to lots or processes, levels of risk to be accepted and pertinent supportive data'.*

Commodity committees can use the sampling plan tool, and the resulting OC curve, to understand the important components of sampling plan design including the levels of consumer's and producer's risks.

Codex commodity committees are responsible for developing Codex provisions – and need to be aware of how sampling plans will perform in regard to Codex provisions. The sampling plan tools can be used to demonstrate the OC curve that comes from selection of a combination of Producer's Risk Quality (PRQ) and Consumer's Risk Quality (CRQ), the number of samples ***n***, the acceptance number ***c*** or the acceptability constant ***k***, and the resulting consumer's and producer's risks.

Chapter 2 Concepts of Sampling

2.1 The purpose of sampling

The main aim of sampling inspection is to ensure that the customer receives product of the required quality and to ensure that products are safe, while remembering that financial resources are limited and the cost of the product must also reflect any costs associated with sampling and testing.

In addition, the Codex Procedural Manual and the Principles for the Use of Sampling and Testing in International Food Trade' (CXG 83-2013) (GL 83) state that 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'.*

The choice of sampling plan depends on the level of protection against poor quality products to be provided to the consumer, whilst also ensuring suitable fairness to producers, in recognition of fair practices in food trade and the nature of measurements associated with the testing for the provision.

2.1.1 What are the ways that sampling inspection can be carried out?

There are three possible ways that sampling inspection can be carried out:

a. 100% inspection;

b. sampling design based on probability, application on statistics; and

c. ad hoc inspection, that is, a sampling plan without a statistical basis.

For **Approach (a)**, it is clear that 100% sampling is not feasible due to the prohibitive cost of testing and in addition, there might not be any product left to sell. Also, the presence of measurement error means that it is still not possible to provide a 100% guarantee, even if all items in the lot are inspected.

Approach (b) has the disadvantage of higher risks as compared to approach (a), some product might not be inspected. However by using the probability approach, the risks can be calculated and a sampling plan can be chosen to ensure these risks are controlled to the desired levels. It also has the advantage of practicability and lower costs. Another important point is to be realistic about the level we wish to control risks, if it is to be achievable.

Approach (c) is often 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 or rejection of foods. The probabilities associated with such plans should be evaluated where possible. Decisions on acceptance or rejection should not be made solely on the basis of these plans.

Approach (b) - the probability approach

Clearly an intended 100% guarantee cannot be provided when sampling methods are used, as not all the product will be inspected. This means that there are two types of risks that can occur:

- the risk that product of unsatisfactory quality will be accepted (Consumer's Risk); and
- the risk that good quality product will be rejected (Producer's Risk).

However, if we specify how we want to control these risks, we can design a sampling plan that ensures these risks are not exceeded.

In practice, the producer's and consumer's risks are specified in terms of the Producer's Risk Quality (PRQ) and the Consumer's Risk Quality (CRQ) respectively. Once these are specified, along with their associated probabilities of rejection and acceptance respectively, a sampling plan, allowing no more than these levels of risk can be developed.

2.1.2 Key definitions

The **Operating Characteristic (OC) curve** is a curve showing, for a given sampling plan, the probability of acceptance of a lot as a function of its actual quality, typically expressed by the fraction non-conforming.

Producer's Risk (PR) Producer's Risk is the probability of wrongly rejecting a lot that is of acceptable quality. It is a point on the OC curve corresponding to a predetermined and usually high probability of acceptance.

Consumers' Risk (CR) is the probability of wrongly accepting a lot that is not of acceptable quality. It is a point on the OC curve corresponding to a predetermined and usually low probability of acceptance.

The **Producer's Risk Quality (PRQ)** previously called the Acceptance Quality Level or the Acceptable Quality Level, (**AQL**) is the quality level of a lot or process which, in an acceptance sampling plan, corresponds to a specified producer's risk.

The **Consumer's Risk Quality (CRQ)** is quality level of a lot or process which, in an acceptance sampling plan, corresponds to a specified consumer's risk. The CRQ was previously called the Limiting Quality or Limiting Quality Level, denoted by LQ or LQL.

A **Sampling Plan** is a plan according to which one or more samples are taken from a lot in order to obtain information about or possibly reach a decision about that lot.

An **Acceptance Sampling Plan** is a Sampling Plan intended for determining the acceptance or the rejection of a lot.

The **Inspection by Attributes** method consists of examining an item, or characteristics of an item, and classifying the item as 'conforming' or 'non-conforming'. The action to be taken is decided by counting the number of non-conforming items found in a random sample.

An inspection by attributes sampling plan specifies the **number of samples**, ***n***, and the maximum number of non-conforming items, referred to as the **acceptance constant**, ***c***, for the lot to be accepted.

The **Inspection by Variables** method starts with selecting a sample of a number of items and measuring dimensions or characteristics so that information is available not only on whether a dimension, for example, is within certain limits, but on the actual value of the dimension. The decision whether to accept a lot is made on the basis of calculations of the average and the variability of the measurements.

An inspection by variables sampling plan specifies the **number of samples**, ***n***, and an **acceptability constant**, *k*. A lot is accepted against an upper specification limit if the acceptance criterion "average result + *k* x their standard deviation" does not exceed the upper limit, and similarly for a lower limit.

2.2 Different Sampling Plan Design Approaches

Commodity committees need to understand that there are different approaches to the design of suitable sampling plans. When sampling plans are presented to CCMAS, the basis for the design of these sampling plans needs to be clear. The key parameters behind, and required for the approval of a sampling plan include the producer's and consumer's risks. Approval might also involve consideration of practicality, fitness for purpose and potential unfairness to one of the parties.

There is no one-size-fits-all sampling plan design that applies. What is important is that the approach used is science-based, with sound statistical backing. In practice, sampling plans may be based on industry practice. However, the choice of plans should still be made with knowledge of the associated risks, bearing in mind that the main purpose of sampling is to ensure that the customer receives product of satisfactory quality.

2.3 Endorsement by CCMAS of sampling plans from different sources

The Codex Procedural Manual 'General instructions for the Selection of Methods of Sampling' says that sampling methods described in CXG 50 or elaborated by international organizations are preferred, and provides as guidance, different types of sampling plans and procedures.

The Codex Procedural Manual also says *'a commodity committee should, whenever possible, provide information to CCMAS for each sampling plan relating to the scope or field of application, the type of sampling (e.g. bulk or unit), sample sizes, decision rules, details of plans (e.g. Operating Characteristic curves), inferences to be made to lots or processes, levels of risk to be accepted and pertinent supportive data'.*

CCMAS endorsement of sampling plans is based on the information provided, and expertise to judge the validity of the plan. Commodity committee design of a sampling plan is also based on criteria for the design, as well as expertise to apply the criteria to a suitable sampling plan to demonstrate *'fair and valid sampling procedures are used when food is being tested for compliance with a particular Codex commodity standard'*.

To aid the design of a sampling plan by the commodity committee, and to help with the provision of the basis for sampling plans, the OC curve can be used. The sampling plan tools we have developed

provide an opportunity for commodity committees to compare different sampling plan criteria, based on the requirements of commodity standards.

CCMAS will be in a position to endorse sampling plans whether the plan is sourced from CXG 50, ISO or another source, so long as the plans meet the requirements of the commodity committee and it can be shown the plan satisfies the principles adopted by Codex, that *'fair and valid sampling procedures are used when food is being tested for compliance with a particular Codex commodity standard'*.

2.3.1 Apps to demonstrate acceptance sampling

There are 14 apps available in the newly developed R package called nzcodex. This R package can be launched in RStudio environment.

The nzcodex R package containing all the apps and sampling inspection tool documentation can be downloaded at the following link.

nzcodex

The package will launch shiny applications (apps) or tools. Some apps are intended for demonstrating risk assessment principles while other apps are to design sampling inspection plans on statistical risk assessment principles.

App1 is about design and evaluation of sampling plans. This app can be used to examine the OC curves before creating and using a sampling plan as the different curves can be compared. The app can be used to investigate either attributes sampling plans or variables plans. In the attributes sampling plan, there is the option to change the sample size and the acceptance number for plan 1 (the purposive plan). For plan 2 (the designed plan), the PRQ, CRQ, producer's risk, and consumer's risk are all to be entered. Once the parameters are chosen, the two OC curves can be compared. Variables sampling plans are similar except there is a *k*-constant instead of an acceptance number. There is also an additional parameter, which is whether the standard deviation is known or unknown. The two OC curves can again be compared for the variables sampling plan.

App2 demonstrates the effect of lot size. This app allows you to see the impact that lot size and sample size have on the OC curves. There are two curves, belonging to finite and infinite lots. The OC curve for the infinite lot doesn't change, but the finite lot OC curve changes depending on the plan parameters. Each of the sample size, acceptance number, lot size and producer's and consumer's risks can be altered to see what effect the changes have on the OC curves.

App3 is created to demonstrate variables plan for averages. There are different parameters that can be selected. These include whether the standard deviation is known, whether the specification limit is upper or lower, and what this particular specification limit is. Also, if the standard deviation is known, then the value is entered. The sample size and *k*-constant are also selected, along with the producer's and consumer's risks. The OC curves will be different depending on whether the standard deviation was known or not, and these curves can be compared. This app might not be included in the final document.

App4 is about repeat testing. It is designed for users to be able to explore the effect of repeat testing. Only items which are classified as non-conforming are able to be re-tested. There are two OC curves, one for repeat testing and the other for single testing. When the plan parameters: sample size, acceptance number, maximum number of tests and the chances of misclassification (percentage conforming as non-conforming, and percentage non-conforming as conforming) are altered, these OC curves change and can be compared. The producer's and consumer's risks can also be selected. It can be seen what effect repeat testing has on the PRQ and CRQ levels, in order to determine its impact.

App5 explores the effect of lot heterogeneity. The sample size, acceptance number and correlation parameter are all chosen. In addition, the producer's and consumer's risks can be decided upon. When these details are changed, the graph of the OC curves will change. There are separate curves for homogeneous and heterogeneous cases which can be compared. In order to compensate for lack of homogeneity, the sample size can be increased to reduce the consumer's risk in general.

App6 is about resampling. This app allows you to select either an attributes or a variables plan. For the attributes plan the sample size, the number of reinspections to be carried out, and what the acceptance number can be changed. For variables plans the sample size, number of reinspections, and the *k*-constant can be altered, along with whether the standard deviation is known or not. Once these are decided upon, the OC curve changes. The OC curve plot shows the difference between a single inspection, and what the resampling scheme would look like, given by the two different curves on the

plot. The producer's risk and consumer's risk are also detailed. Therefore, this app can be used to explore what impact a resampling inspection scheme has.

App7 relates to the effect of inspection errors on risks. The sample size and acceptance number are chosen for the attributes plan, along with the producer's and consumer's risks. Then the chance of misclassifying conforming as non-conforming, and non-conforming as conforming is also selected. The OC curve is given dependent on what is selected for these variables. There are two OC curves, one for with inspection error, and the other for no inspection error. Both of these curves can be compared as the parameter values change to see what effect the inspection error has on risks.

App8 facilitates the implementation of the FNC plans. FNC stands for fractional non-conformance. Quality characteristics observed in industrial processes are not always free from measurement errors and so FNC plots were developed. FNC refers to the probability of an error-prone observation breaching the specification limits. They assess the probability of conformance when measurements are error-prone. This app lets the user input a dataset to work with. The user can select a fractional acceptance number and view the associated FNC values for given repeatability standard deviation, and error variance ratio. The app computes the sum of individual FNC probabilities and compares with the specified acceptance number so that the lot acceptance or rejection decision can be made.

App9 is about conformity testing, i.e. whether the true value of the sample tested complies with a limit. This app looks at the probability of declaring conformity for both the FNC and ISO methods. A sample size, the significance levels for the limiting value (LV) and FNC and the variance ratio each need to be selected. Plots are then displayed which show the probability of conformity and non-conformity for ISO two-stage and FNC testing procedures to compare. It is assumed that both these conformity testing procedures have equal samples that are tested.

App10 is about sampling plans for compositional proportions. This app allows the user to change the PRQ and CRQ levels, along with the *U* or *L* (upper or lower specification limit) and theta value (the 'precision parameter' describing the variation for the beta distribution). Changing these inputs allows users to see what will happen to the OC Curves (which is a way of describing the behavior of a sampling plan). OC curves for plans based on both the beta and normal distributions are shown, and can be compared.

App11 is about compressed limit sampling inspection plans. This app is a tool for matching compressed limit sampling inspections plans. The underlying distribution can be chosen from normal, gamma or Weibull. The sample size and acceptance number are also chosen for the reference single sampling plan. The matched compressed limit plan will then be given based on the different values. The compression constant is given in order to know what factor the sample size and acceptance number was compressed by.

App12 demonstrates optimum guard-banding analytics for a user dataset that can be uploaded. The upper or lower specification limit can be specified. The model to be used is selected out of normal, beta and minimum cost. The data is then processed to show the FNC plot, acceptance control chart plan, and the optimum guard-banding levels.

This app might not appear in the final document.

App13 is a tool that matches single and two-stage microbiological sampling plans based on compressed limits. The underlying distribution is chosen from Poisson, Poisson-lognormal, Poisson-gamma or Lognormal distributions. The single plan has a sample size, acceptance number and microbiological limit to be selected. This is then developed further with the two stage plan which has a matching method of two different approaches. These two stage plans include sample sizes, acceptance numbers, rejection numbers, and a compressed limit. Plots are then produced, which change depending the parameters selected.

App15 enables the user to design a variables sampling inspection plan that is adjusted for the repeatability SD of measurement errors. This app particularly shows that the acceptability constant *k* constant must be smaller depending on the size of the repeatability SD.

App16 compares the FNC based inspection plans with the variables plans adjusted for repeatability type measurement error. FNC inspection plans are particularly useful when the normal distribution does not hold for the underlying quality characteristic.

App17 is about FNC plots. This apps lets the user input a data set to work with to obtain FNC charts which can be useful to monitor a short-run production process. The false alarm rate and *k*-constant are

chosen, along with either the upper or lower specification limit. Then either the beta or normal model needs to be specified with the value for the process standard deviation, and the PRQ level. Once these are all included, the data is plotted and the control chart can be viewed. This app relates more to process control and might not be included in the final version.

Supplementary technical notes and examples are also given in the apps.

Chapter 3 Designing sampling plans

Sampling inspection plans are usually designed to protect the interests of the producer and the consumer. This is accomplished by specifying quality levels and the associated risks of acceptance and rejection. The most popular indices for the design of a plan are the PRQ, CRQ and the associated producer's and consumer's risks.

3.1 Broader issues

A guideline on the need to consider the broader issues of sampling inspection is given in Bicking (1967):

1. Clarify the purpose of sampling.
    - What population will the sample be taken from?
    - What information is needed about the population?
    - What is the criteria that lot acceptance will be based on?
2. Specify the population and investigate the history of a lot.
    - Does the process that produced the lot come from a state of control?
    - Does the lot size agree with the expectations of the producer and the consumer?
    - Are the methods of handling and storage being considered properly when the lot size is determined?
3. Look at the measurement error.
    - Separate the measurement error and sampling error.
    - Compare these two sources of error.
4. Think about what the within-lots and between-lots variances due to different processes are.
5. List the instructions for sampling, ensuring to protect against the following issues:
    - any lack of clarity in the purpose of sampling;
    - any lack of instructions that aren't specific enough;
    - not providing methods that check sampling error, reliability and bias; and
    - methods that would be unsuitable when handling the sample.
6. Control the sampling operation.
    - Make sure the samplers are well trained.
    - Do check samples to control the operation of the plan.
7. Ensure that the sampling instructions are reviewed and alter any changes that are required for the process.

The sampling inspection plan then needs to be agreed upon and passed onto whoever is responsible to carry out the plan. In order for this acceptance sampling plan to be effective, it involves more than just selecting and applying rules. Inspection should also include: good data, quick information and incentives for the producer to provide quality at satisfactory levels.

3.2 Administration of sampling plans

An acceptance sampling plan is an important aspect to the overall approach of maximising the quality at a minimum cost. Acceptance sampling plans need to change in order to consider current results, and any history of inspections that have been performed. This process is known as acceptance control, because it involves selection, application, and modification of the acceptance sampling procedures in order to adapt to a changing inspection environment. Inspection results allow you to accept or reject

individual lots as they are found, but they are also beneficial for any future production planning for the producer. This is because it can be decided whether the process needs any alterations in order to eliminate any issues.

Before a sampling plan is used, the quality levels need to be determined. The consumer seeks to try minimize the total cost of purchase, inspection, assembly and the eventual service. It is not reasonable for the consumer to expect quality levels that are higher than the previous quality levels in the industry. The producer needs to select an acceptable level that is for all intended customers at the prices they are willing to pay, rather than setting individual quality levels for each customer. Both the consumer and the producer need to understand the quality levels. Then the producer is responsible to carry out inspection that is sufficient to assure conformance.

Process data should be analyzed over a long enough period of time to assess the overall level of performance. This is what is used to set the Producer's Risk Quality (PRQ). There are some instances where the PRQ will differ from the state-of-the-art process average. Some of these instances include (i) handling a class of nonconformities instead of a single quality characteristic, or (ii) there is an urgency of demand for a product. In the end, the quality levels are often decided by economic considerations.

The basic principle of administration of sampling inspection is that there is a need for simplicity and practicality. Methods and procedures need to be safe, sure, swift and simple. In order to be successfully used in industry, acceptance sampling and all forms of administration need to be as uncomplicated as possible.

The text by Schilling and Neubauer (2008) may be consulted for more details on the administration of acceptance sampling.

3.3 Design of sampling plans and risk management

By the phrase 'designing a plan` it is meant fixing the parameters of the sampling plan such as the sample size *n* and acceptance number *c* so that the plan is fit for purpose. In designing a sampling plan one has to accomplish a number of different purposes. Hamaker (1960) has listed the following purposes as the important ones:

1. To strike a proper balance between the consumer's requirements, the producer's capabilities, and the inspector's capacity.
2. To separate bad lots from good.
3. Simplicity of procedures and administration.
4. Economy in number of observations.
5. To reduce the risk of wrong decisions with increasing lot size.
6. To use accumulated sample data as a valuable source of information.
7. To exert pressure on the producer or supplier when the quality of the lots received is unreliable or not up to standard.
8. To reduce sampling when the quality is reliable and satisfactory.

Hamaker (1960) also cautioned that the above aims are partly conflicting and not all of them can be simultaneously realized.

The first four purposes are particularly critical. The designed sampling plan must explicitly quantify the producer's and consumer's risks. Some of the published sampling inspection procedures such as those in ISO 3951 place more emphasis on reducing the producer's risk with increasing lot size. This is to encourage large scale production and lot formation. For international trade and particularly for food products, consumer's risk control is particularly important in addition to simplicity of operation and transparency and fairness in reducing the risks to both producers and consumers.

Chapter 4 Routine attributes and variables sampling plans

4.1 What information is needed to design the sampling plan?

We have developed a toolbox of apps for the design of attributes and variables sampling plans for routine inspection. This tool will assist commodity committees with the design of a sampling plan to ensure fair practices in food trade. The tool can be enhanced for example, to allow for measurement error.

The guidance for the selection or design of suitable sampling plans is based in statistical theory. Use of these tools allows for the statistics to sit in the background.

These tools help guide the design of appropriate sampling plans by showing the operating characteristic (OC) curve to demonstrate the performance of the plan. The tool also allows the plan to be designed from the Producer's Risk Quality (PRQ) and Consumer's Risk Quality (CRQ). The OC curve shows the probability of accepting a lot versus the fraction non-conforming in that lot for a given sample size and acceptance number.

The tools for the design of plans can be used by specifying both the PRQ and CRQ, from which it will work out the number of samples ***n*** and the acceptance number ***c*** for attribute plans, or the ***n*** and the acceptability constant ***k*** for variables plans. This means the Consumer's Risk Quality must be specified as part of the plan design.

The tools also provide the option to move away from the usual approach in which it is assumed that the PRQ and CRQ are associated with probabilities of acceptance of 95% and 10% respectively. In general when measurement-error is negligible you need to specify any two points on the operating characteristic, i.e. two quality levels and the associated probability of acceptance or rejection at those levels, in order to determine ***n*** and ***c*** (or ***k***). However 95% acceptance is usually associated with good quality and 10% acceptance with poor quality so it seems easier to specify levels representing what is good quality that should be accepted most of the time and what is poor quality that should be rejected most of the time.

The input parameters in this tool allow the probabilities of acceptance, or levels out of specification corresponding to specified levels of acceptance, to be calculated.

4.2 Single Sampling Plans for Attributes

The simplest single sampling plan is one done by attributes. This is because the inspection results are classified into only two classes of outcomes. Because it is applicable to all sampling situations, it has become the benchmark that all other sampling plans can be compared against. It can be used in different ways in inspection. This includes: counting the number of non-conforming items found in a sample (Poisson distribution), evaluating the proportions non-conforming from large lots (binomial distribution), or from individual lots (hypergeometric distribution).

It is relatively simple to implement attributes sampling plans. A random sample of size ***n*** is taken from a lot of size ***N***, which can be very large or infinite. The number of non-conforming items found is compared to the acceptance number ***c***. If the number of non-conforming items is less than or equal to ***c***, then the lot can be accepted. However if the number of non-conforming items found is greater than ***c***, then the lot is rejected.

There are many different charts and tables that can be used to determine a single-sampling attributes plan. Chapter 5 in Schilling and Neubauer (2008) contains explanations of many of these and demonstrates how they can be implemented.

4.3 Single Sampling Plan for Variables

If the underlying distribution of individual measurements is known, acceptance sampling can be performed directly on the measurements themselves. This often allows a considerable saving in sample size but we need to know the probability distribution of the underlying measurements. The Gaussian or normal distribution is commonly adopted as the distribution of the measurements. For compositional proportions in bulk material, beta distribution is more appropriate but the normal distribution can often serve as an approximation.

In variables plans, the mean $\bar{x}$ is compared with the acceptance limit in a similar way to the number of nonconforming units, ***d***, being compared to an acceptance number, ***c***, in the attributes plans. In order to adjust for the variability in the lot, the sample standard deviation ***S*** is computed. The quantity $\bar{x} \pm kS$ is then compared with the lower, *L* or upper, *U* specification limits. The lot acceptance criterion is $\bar{x} +$

$kS \leq U$ or $\bar{x} - kS \geq L$. This method of operation of variables plan is known as the *k*-method to control the fraction nonconforming *p*. The variables plan can be selected for a given PRQ level, producer's risk, CRQ level, and consumer's risk.

Schilling and Neubauer (2008) explains some of the advantages and disadvantages of variables sampling plans.

The advantages of variable sampling plans are:

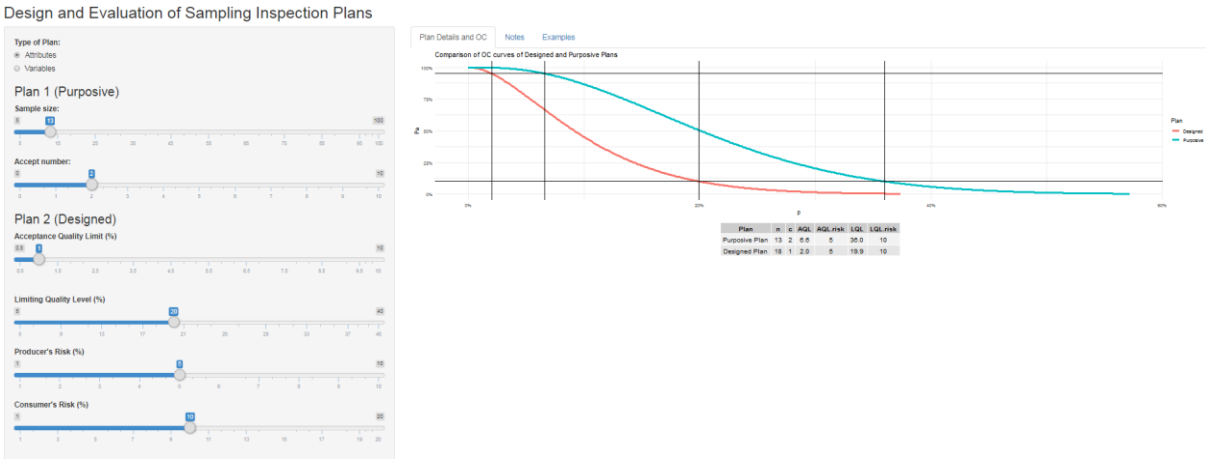
1. it offers the same protection with a smaller sample size than what is required for attributes;
2. there is feedback of data on the process which produced the units;
3. there is more data available in waiver situations;
4. there is the extent of conformity of each unit given weight in application of the plan; and
5. there is an increased likelihood that any errors in measurement will be detected.

The disadvantages are:

1. the results are dependent on the underlying distribution of measurements assumption being correct;
2. variables sampling plans are only applicable to one characteristic at a time;
3. there is a higher inspection cost per unit;
4. there is a higher clerical cost per unit;
5. there is a possibility of no non-conforming units being found to show to the producer after rejection; and
6. unfortunately, a lot with no non-conforming units may be rejected by a variables plan.

For more details on variables sampling plans, consult Chapter 10 of Schilling and Neubauer (2008).

4.3.1 App implementing attributes and variables sampling plans



This app can be accessed via the link: **App1**

4.3.2 Examples

4.3.2.1 Attribute Inspection Plans

Assume that a sampling plan with sample size *n=5*, and accept number *c=0* is used to inspect a lot of button type mushrooms. The button style be defined as whole mushrooms, with attached stems not exceeding 5mm in length, measured from the bottom of the veil.

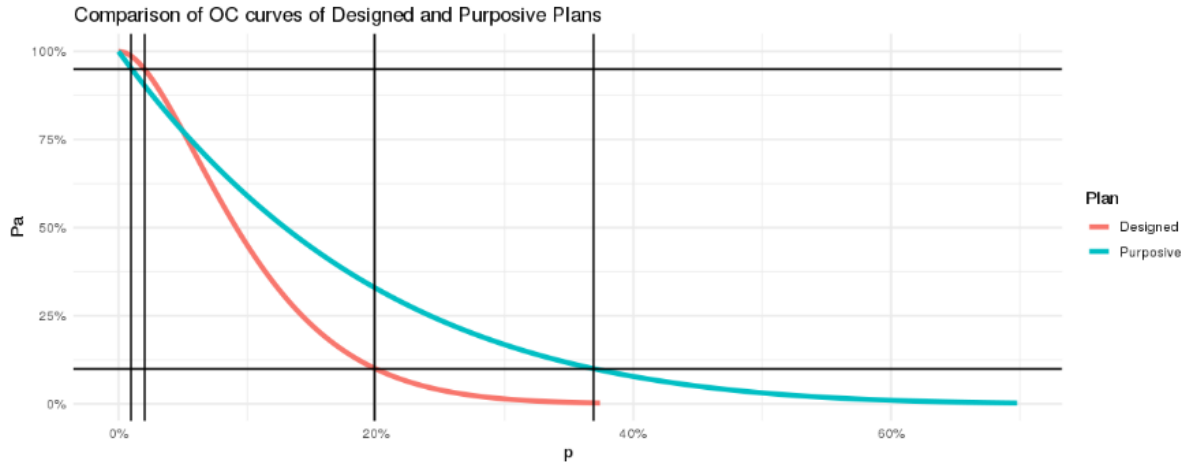
A number of quality measures are inspected after opening randomly sampled five cans from the lot. These measures often include:

1. flavour (normal flavour and odour but free from other flavours or odours foreign to the product);
2. texture and character (based on mushroom units with detached caps or stems etc.);

3. contamination free (Lead (Pb) below 1 mg/kg etc.); and

other quality characteristics listed in the appropriate Codex Standard.

For this *(n=5, c=0)* plan, the consumer's risk is about 40%. Using the app, a better balance between producer's and consumer's risk can be achieved. The designed plan *(n=18, c=1)* plan is able to discriminate well between a good and poor quality lot. Many other sampling inspection plans can be obtained so that consumer's risk is lower than the risk under *(n=5, c=0)* plan.



Plan	n	c	AQL	AQL.risk	LQL	LQL.risk
Purposive Plan	5	0	1	5	36.9	10
Designed Plan	18	1	2	5	19.9	10

4.3.2.2 Variables Inspection Plan

Solubility is an important quality characteristic for instant coffee. An upper specification limit *U=30* seconds is set for the time to dissolve instant coffee readily in boiling water with moderate stirring.

Assume that currently a variables sampling plan with *n=10* and *k=1.5* is employed in order to reduce the consumer's risk as well as improve the discrimination, the sample size can be increased to *n=13* and *k* is adjusted to *k=1.638* using the app by trial and error.

Suppose however we assume a lot size *n* of 8000, and we want a sampling plan with Producer's Risk Quality PRQ = 2.5% and that the standard deviation is unknown, i.e. it is not known from historical data but will be estimated from the results obtained from the sampling and testing. Assume also, for the purposes of this example, that measurement error is negligible.

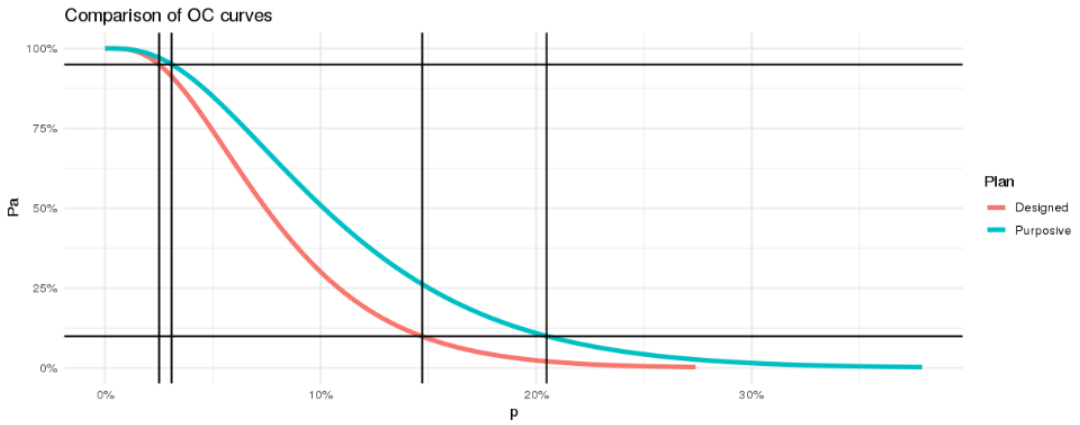
The variables inspection plan currently employed for the quality assurance of solubility was obtained from Table 14 of CXG 50 corresponding to a lot size of 8000, and Acceptance Quality Limit, AQL = 2.5%. This plan calls for taking n = 15 samples and fixes the acceptability constant k = 1.30. This variables plan is operated as follows:

1. obtain the solubility time in seconds for each of the 15 samples taken; and
2. the lot would be accepted against an upper specification limit (USL) as long as

$$\bar{x} + kS \leq USL$$

where $\bar{x}$ is the average of the test results and *S* is their standard deviation.

The performance of this plan for various nonconforming levels of true solubility can be assessed using the OC curve displayed in the app. Using the tool, we can specifically evaluate the consumer's risk. Note that in this example the PRQ and the sample size (*n*) have been specified so that the Consumer's Risk Quality (CRQ), indicating the level of Consumer's Risk, is intrinsically determined. The following screenshot shows the OC curve of the current purposive plan along with a new designed plan with *n=22* and *k=1.463*.



Plan	n	k	AQL	AQL.risk	LQL	LQL.risk
Purposive Plan	15	1.300	3.1	5	20.5	10
Designed Plan	22	1.463	2.5	5	14.7	10

The designed plan has a lower CRQ of about 15% instead of the CRQ of about 20% for the current plan. The new plan is more discriminatory at other solubility quality levels, having lower probabilities of acceptance at higher levels non-conforming.

Chapter 5 Some Issues of Routine Inspection

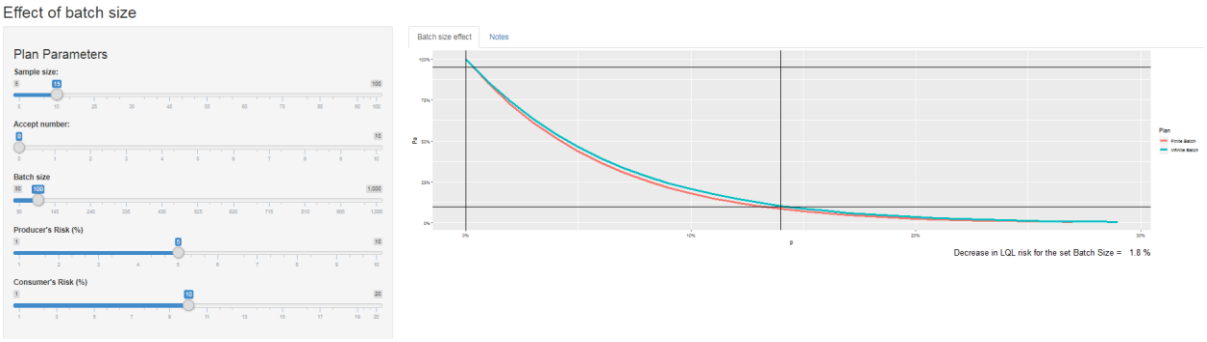
In this section, some of the commonly encountered issues such as the relationship between sample size and lot size are discussed. Resampling and retesting (not the same thing) are also discussed. Resampling is used to reduce the producer's risk when random sampling of the lot is difficult whereas retesting is a way of overcoming inaccuracy of test results due to measurement uncertainty. If measurement errors are expected to dominate, sampling inspection plans can be adjusted for measurement errors. This adjustment can be done fairly to protect both the producers and consumers. This topic is covered in detail in a later section (Section 9).

5.1 Lot size vs sample size

The OC curve is a fundamental tool for assessing the consumer's and producer's risks in acceptance sampling. The effect of lot size on the OC curve is minimal when only a small proportion of the lot is sampled for testing. This means that the risks won't change dramatically with lot size unless the sampling fraction is large. The absolute sample size is rather important and it largely determines the protection afforded by a plan.

Schilling and Neubauer (2008) may be consulted for more discussion on why the lot size itself doesn't have an important role in determining protection to consumer and producer.

5.1.1 App to demonstrate the effect of lot size



This app can be accessed via the link: **App2**

5.2 Explanation of ISO, CXG 50 sampling plans

ISO Standards employ sampling schemes, switching between normal, tightened and reduced inspection to effectively control quality, but these are generally not workable in international trade.

Normal inspection plans are used when the process is considered to be operating at, or slightly better than, the PRQ.

Tightened inspection plans use stricter acceptance criteria 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.

Reduced inspection plans permit smaller sample sizes than those used in normal inspection. When the level of the submitted quality is sufficiently good, reduced inspection offers sampling economy.

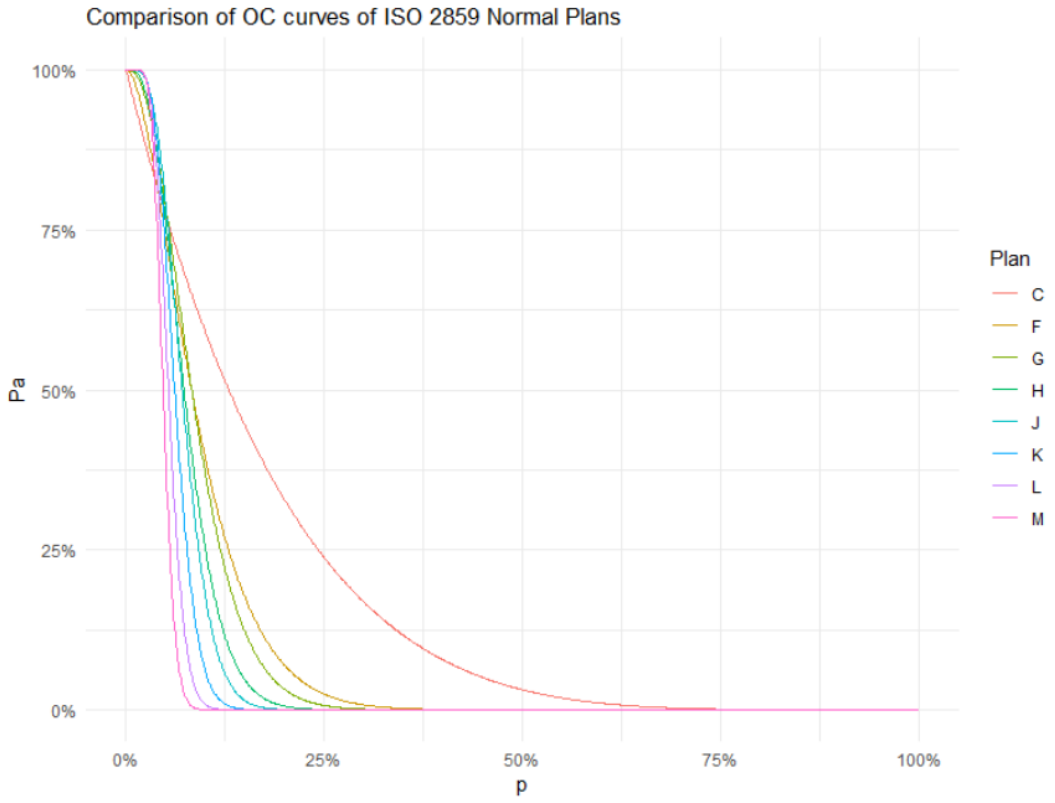
Switching rules govern the switching between normal, tightened and reduced inspection based on recent inspection history.

Even though it is desired that a product which has a history of good quality is one that will need less inspection compared to a product which has no history, or a history of poor quality, importers do not always have good access to the quality history data of the exporters and hence it is not possible to accurately determine whether the exporter's process average fraction nonconforming compares with the set PRQ. Sampling schemes are administratively complex and more suitable to sampling inspection of regular localized supply.

The OC curves of normal, tightened and reduced plans offer different levels of protection to the producer and consumer. The full effect of a sampling scheme is realized only when all the rules for switching to and from normal-tightened-reduced plans are fully implemented. In other words, the overall or steady state OC curve of a sampling scheme can only be realized for a very long series of lots of consistent quality. Hence sampling schemes are suitable only when producers and consumers enter a long term supply arrangement for hundreds lots. Sampling schemes indexed by PRQ do not fix the producer's risk at a fixed level such as 5%. Producer's risk will become smaller for only for large lot sizes. The table given below shows ISO 2859 normal single sampling plans for an AQL of 2.5 (Level II General Inspection):

Lot size range	**Sample Code**	**(*n*, *c*)**
16-25	C	(5,0)
91-150	F	(20,1)
151-280	G	(32,2)
281-500	H	(50,3)
501-1200	J	(80,5)
1201-3200	K	(125, 7)
3201-10000	L	(200, 10)
10001-35000	M	(315, 14)

The following graph shows the OC curves of the above selection of plans. The producer's and consumer's risks differ significantly for these plans and the selection is solely guided by the lot size.



5.2.1 Equivalent sampling plans

It is possible to design attribute or variables sampling plans equivalent to ISO, CXG 50 sampling schemes. The steady state or composite OC curve of the sampling scheme must be consulted to obtain the PRQ and CRQ values for the set producer's and consumer's risk. The routine single attribute or variables plans can then be obtained. For an example of this approach to obtaining equivalent plans by matching the OC curve at two-points of the OC curve, consult Chapter 11 in Schilling and Neubauer (2008).

5.3 Provision for Reinspection or resampling

Sampling inspection plans usually assume that a random sample is taken from the lot. When random sampling of prepackaged commodities from large containers is difficult, physical sampling may be done poorly. Hence it is natural for the producers or consumers to occasionally suspect or dispute the sampling done.

When the original inspection results are suspect, the provision for lot reinspection or resampling can be incorporated. Reinspection or resampling is done when a lot was rejected on the first inspection, but it is resubmitted for acceptance inspection so that a second new sample can be taken to make a decision on the lot.

The resampling scheme, discussed in Govindaraju and Ganesalingam (1997), is implemented as follows:

The operating procedure for this is:

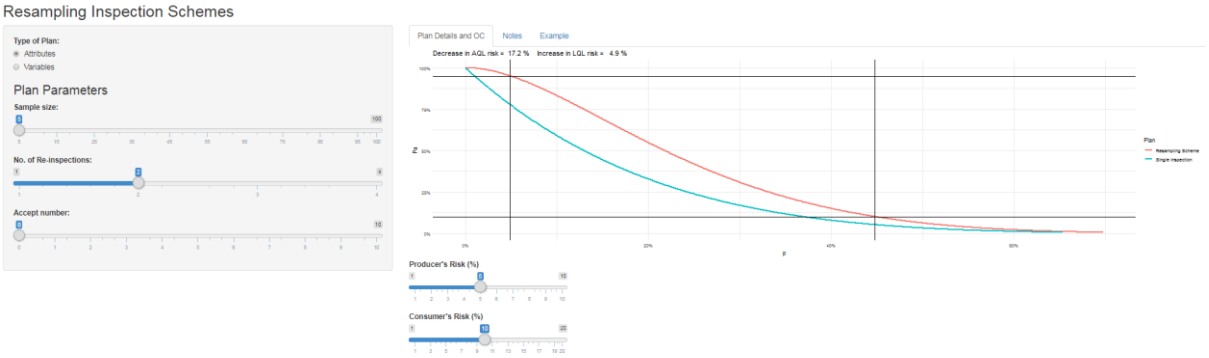
1. Do an original inspection (for example: a single sampling plan with sample size, ***n*** and acceptance number, ***c***.)
2. Given that this original inspection was not accepted, apply the same plan ***m*** times and reject the lot if it is not accepted on ***(m-1)****st* re-submission.

Resampling schemes are particularly useful for zero acceptance number sampling plans. It is well known that the zero acceptance number plans generally involve higher risks to the producer. Hence resampling schemes allow the producer to opt for reinspection of the lot when there is good process history to believe that the quality of the lot is indeed good but rejected due to poor sampling.

Variables sampling plans with large ***k*** values such as ***k=2*** can also be harsh on the producers. These plans also involve a small sample size. Resampling can also help here to reduce the producer's risks.

5.3.1 App to implement resampling plans

The following app deals with resampling for both attribute and variables plans. The user can adjust the number of resampling or reinspections allowed. It should be noted that the app controls the overall risk and hence the plan design still controls the consumer's risk at the set level.



This app can be accessed via the link: **App6**

5.3.2 Further Details of Resampling tool

At the true fraction non-conforming ***p***, the OC function of the resampling scheme that allows ***m*** resampling or reinspection becomes:

$$P_A(p) = 1 - \left(1 - P_a(p)\right)^m$$

where $P_a(p)$ is the OC function of the original (single) inspection plan. Both attribute and variables plans can be implemented under the resampling scheme.

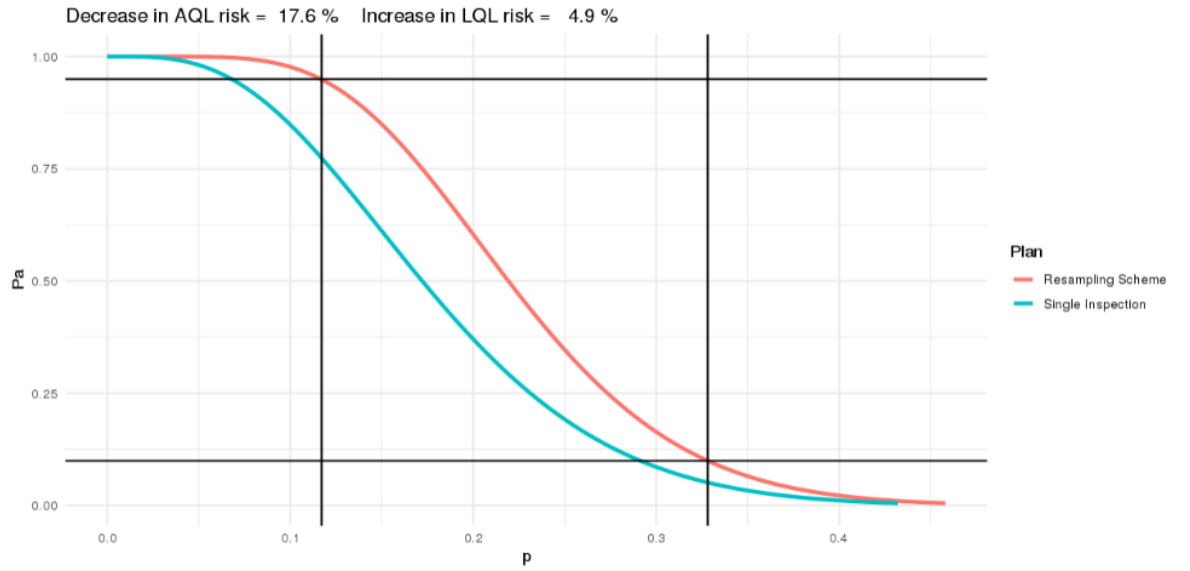
The main advantage of the resampling scheme is the greater reduction of the producer's risk, particularly for zero acceptance number single sampling plans or variables plan with large acceptability constant ***k***.

This app allows users to design resampling schemes by trial and error examination of the OC curve to control the overall Producer's risk and Consumer's risk at the desired levels.

5.3.3 Example

Quick frozen fillets are slices of fish of irregular size and shape which are removed from the carcass of the same species of fish suitable for human consumption. A sample unit is the primary container or for individually quick frozen products is at least a 1 kg portion of the sample unit. Sampling and testing for odor and flavor characteristic can be difficult. Occasionally the result may be disputed by the producer because sensory and physical of fillets are harder to assess and re-examination may be required in the event of a dispute.

Assume that sampling inspection of lots is done using a single sampling attributes plan *n=21*, *c=3* corresponding to PRQ=6.5 (in percent). In the event of reinspection, how much additional risk to the consumer occurs and how much producer's risk can be reduced because of resampling can be examined using this tool (see the figure below).

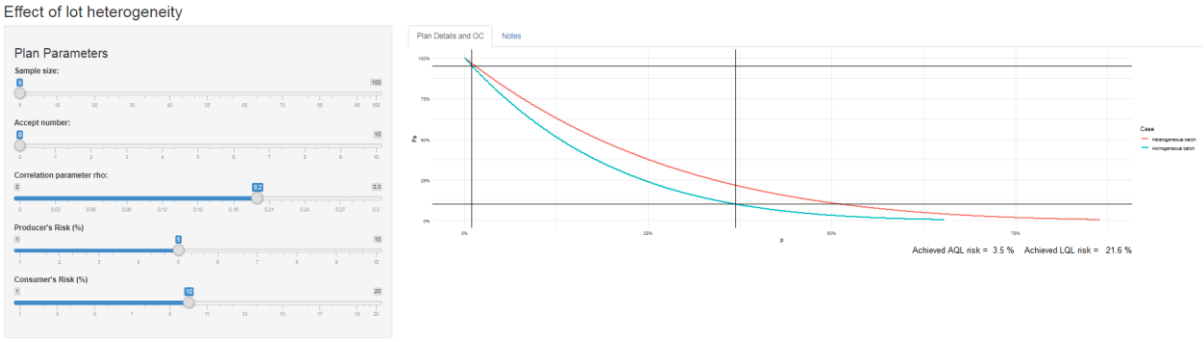


5.4 Inhomogeneous lots

Homogeneous lots is a way of describing that the lots are similar in nature. Non-homogeneous lots are therefore not similar.

Acceptance inspection and compliance testing often necessitate levels of protection for both the consumer and the producer that require large sample sizes relative to lot size. A given sample size can, however, be made to apply to several lots jointly if the lots can be shown to be homogeneous. This reduces the economic impact of a necessarily large sample size. If the lots are not homogeneous, then this is unable to occur.

The effect of lot heterogeneity on producer's and consumer's risks is demonstrated using the following app.



This app can be accessed via the link: **App5**

This app demonstrates the effect of lot heterogeneity on producer's and consumer's risks. If a lot is homogeneous, the fraction non-conforming is a constant say *p*. When a random sample of size *n* is taken from a lot, the number of non-conforming units *d* follows the binomial distribution which the OC curve is drawn for the homogeneous case.

When the lot is heterogeneous, *p* is assumed to follow the beta distribution with parameters α and β. For the beta-binomial model, the R package VGAM function pbetabinom() can be used to compute the OC curve. The parameters α and β are reparametrized in terms of *p*, and ρ, the correlation parameter. When ρ, the regular binomial case is obtained.

This app will show that the risks can be higher in general, and the consumer's protection may be compromised when the inspection lot fails to be homogeneous.

Chapter 6 Compliance of the average level

Instead of controlling the proportion non-conforming *p*, variables plans can also be used to control the mean (i.e. the average) level.

Single Specification Limit Plans: It is assumed that the quality characteristic *X* is normally distributed. The single specification limit plans also assume that either a lower specification limit, *L* or an upper specification limit, *U* is specified.

When the lot standard deviation σ is known based on historical process data, the inspection plan is operated as follows:

1. Take a random sample of size *n* and obtain the sample mean $\bar{x}$.
2. Let $A = L + k\sigma$. If x>A, accept the lot; otherwise reject the lot.

In the case of an upper specification limit U, the acceptability constant A is set as $A = U - k\sigma$, and the acceptance criterion is reversed as $\bar{x} < A$. The parameters of the plan are n and k (or A).

When the lot standard deviation σ is unknown, it is replaced with the sample standard deviation *S*. The Operating Characteristic (OC) performance for known and unknown σ plans will differ. The plan based on *S* will call for a greater sample size or else the OC curve of the unknown σ plan will be less discriminatory. See Schilling and Neubauer (2008) for more information.

The following app demonstrates how the OC curves are drawn for sampling plans for controlling the average level. Note that the probability of acceptance is plotted against the true mean *instead of the fraction nonconforming* p*.

6.0.1 Example

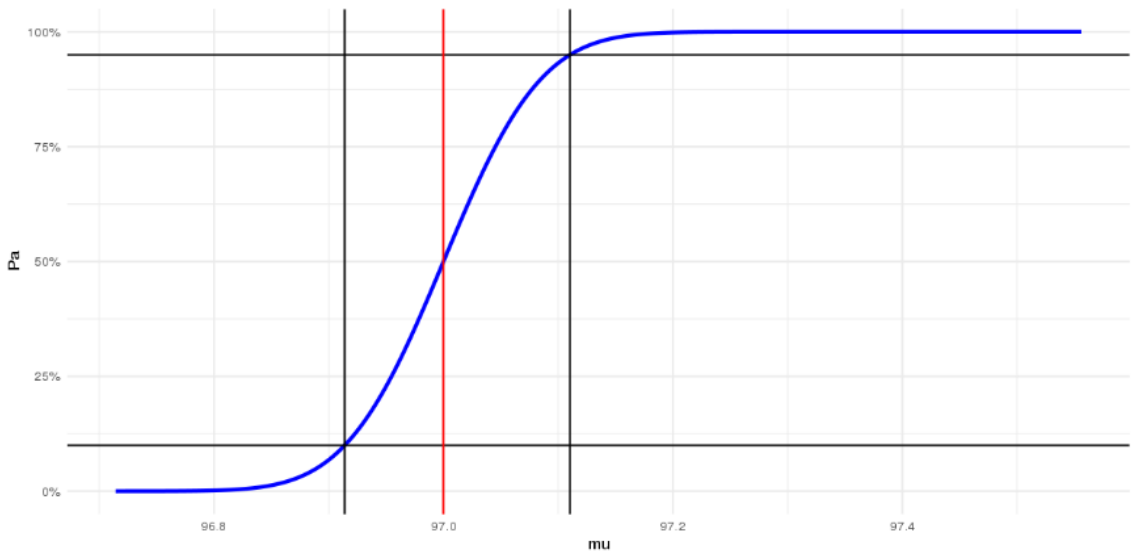
The Codex Standard for food grade salt prescribes that the average content of NaCl shall not be less than 97% on a dry matter basis, exclusive of additives.

One of the lot acceptance criterion is that the average NaCl in a sample of size *n* should be at the minimum level specified, 97%, or more.

Assume that a sample of size 80 is taken for a lot of size 8,000. The sample mean NaCl must be 97% or more for acceptance of the lot. The OC performance of this acceptance criterion is evaluated setting *k=0* in the app.

Even though normal distribution is often a poor fit for compositional proportions, the acceptance criterion is just based on the mean NaCl level and hence the OC curve of the variables plan based on normal distribution can be employed to assess the risks.

The following is a screenshot of the OC curve produced by the tool for *n=80*, *L=80*, *k=0* for the true SD σ=0.6.



6.1 Average Quantity System

The International Recommendation OIML R 87 Edition 2016 (E) *Quantity of product in prepackages* is based on the following three main principles:

1. If Q_{nom} is the nominal prepackage quantity, q_i is the actual quantity of the i^{th} prepackage, then the error for the i^{th} prepackage $e_i = Q_{nom} - q_i$. In a random sample of size *n* drawn from the lot whose prepackage quantity is normally distributed with mean *and standard deviation σ, it is ensured that the lot is rejected when* $\mathrm{e}_{\mathrm{avg}} < \mathrm{c}$ *where* c is a constant found satisfying:

$$Pr(e_{avg} < c|\mu = Q_{nom}) = 0.005$$

In other words, the *c* constant is a parameter for the **test of average requirement** which mainly protects the interest of the producer. The producer's risk of rejecting the lot, whose true mean is at the nominal value, is controlled.

For consumer's protection, the probability of rejection is at least 0.9 for unacceptable lots with $\mu < Q_{nom} - 0.74\sigma$. The consumer's risk for unacceptable lots will be no greater than 10%.

The sample size *n* must satisfy the inequality

$$\frac{n(N-1)}{N-n} \geq \frac{t_{0.9,n-1} - t_{0.005,n-1}}{0.74}$$

in order to meet the producer's and consumer's risks pertaining to the test for average requirement.

2. **T1 error control**. Firstly, a parameter *T* is defined in such a way that T is a parameter that ensures that the percentage of prepackages with $q_i < Q_{nom} - T$ is no greater than 2.5 %.

The T1 error is when the individual package error is less than *-T* but equal to or greater than *-2T*.

For the sample of *n* prepackages, d_{T1}, the number of prepackages failing T1 criterion is limited to c_{T1} or less.

In other words, the attribute plan $(n,\ c_{T1})$ is applied to control the proportion of prepackages not conforming to the T1 error criterion.

3. **T2 error** control. Individual prepackages with errors less than -2T are called T2 error prepackages, which are extremely short compared to the nominal Q_{nom}. The lot is rejected in the event of a T2 error. In other words, a zero acceptance number attributes plan is employed to control the proportion of prepackages not conforming to the T2 error criterion.

The International Recommendation OIML R 87 Edition 2016 (E) aims to control both the average quantity as well as the proportion of packages which may be too *short* (having weight or content less than the nominal level).

Sampling plans for small finite size lots are given which are based on the hypergeometric distribution. Large lots are modelled using the normal distribution but adjusted for the small sample size using the Student's *t* distribution.

OIML R 87 Edition 2016 does not evaluate the performance of the sampling inspection scheme using OC curves. The statistical basis for the proposal assumes normality and heavily relies on the Student's t distribution for finding various constants. The limits for T1 and T2 errors serve as a protection limit but their efficacy is not fully evaluated. As a result, this quantity assurance scheme may not achieve the same producer's and consumer's risks for various lot sizes. (It is necessary to resort to simulation studies in order to evaluate the producer's and consumer's risks).

Chapter 7 Bulk materials

7.1 Introduction

Bulk materials are continuous, and are made of particles of different density and sizes etc. An example is milk powder. It is impossible to view bulk materials present in a lot as a set of distinct objects because there is no way of selecting the items one by one in a way that is not biased when using simple random sampling. This is where a different methodology is introduced, which brings with it sampling bias and non-representativeness.

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. These factors include things such as how the device is employed, and the conditions that the device is used under.

In bulk sampling, the lots of bulk material are 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. Other means of real or synthetic division can also occur.

The following general objectives of bulk sampling were described by Bicking (1978).

1. Characterize the material in place as to location, amount and value.
2. Characterize the material as to grade, any need for further processing, and its destination.
3. Control during processing.
4. Acceptance on a lot-to-lot basis.
5. Determination of weight for purposes of payment.
6. Determination of properties that must be known so that the end use will be appropriate.
7. Experimentation and analysis to determine further sampling procedures and uses of the material.

Schilling and Neubauer (2008) may be consulted for further references on bulk sampling inspection plans.

Bulk materials being continuous, parts of a samples can be mixed together to form a composite. This composite then gets tested only once, rather than having to do many tests on the individual parts. This is a physical way of being able to average the samples that are composites.

7.2 Sampling Inspection Plans for Compositional Proportions

Compositional characteristics are often quality measures for bulk materials. For example, the percentage protein is a primary quality measure for milk products, and a minimum protein limit such as 34% is set for milk powders. Compositional fractions in a lot of manufactured product can be modelled using the beta distribution. Variables sampling plans based on the normal distribution can only be approximate for compositional proportion and plans based on the normal distribution can involve higher consumer's risks than desired.

Composite sampling is also commonly used for bulk products. Sampling inspection plans based on the beta distribution fits well for this context. Variables sampling inspection plans based on the beta distribution to control the fraction nonconforming product can be designed.

Let the total bulk material amount sampled be *G* (such as 100g, 200 ml). *G* can be expressed as a multiple of the standard or primary unit mass g. Let *m=G/g* (which need not be an integer). The quantity *m* is similar to the sample size *n* fixed for discrete or non-bulk items. Let the random variable $\hat{\mu}$ be the mean compositional fraction for amount *G*. Note that $\hat{\mu}$ can be a single measurement based on a well-mixed composite, and need not be the arithmetic mean of *m* measurements of individual test samples. The distribution of $\hat{\mu}$ can be fitted using the beta distribution instead of crudely approximating with normal.

A variables plan based on the beta distribution is implemented as follows:

1. obtain *m* primary samples or increments of bulk material and form a composite of amount *G;*
2. test the sample(s) and estimate the compositional fraction $\hat{\mu}$ as the average level;
3. estimate the standard deviation $\hat{\sigma} = \sqrt{\hat{\mu}(1-\hat{\mu})/\theta}$ where θ is the known precision parameters found from past data; and
4. accept the lot if $\hat{\mu} - k\hat{\sigma} > L$ where *L* is the lower specification limit. For upper specification limit *U*, the acceptance criterion becomes $\hat{\mu} + k\hat{\sigma} < U$.

If past history is unavailable, a conservative (or a smaller) value of θ can be used.

7.2.1 Example

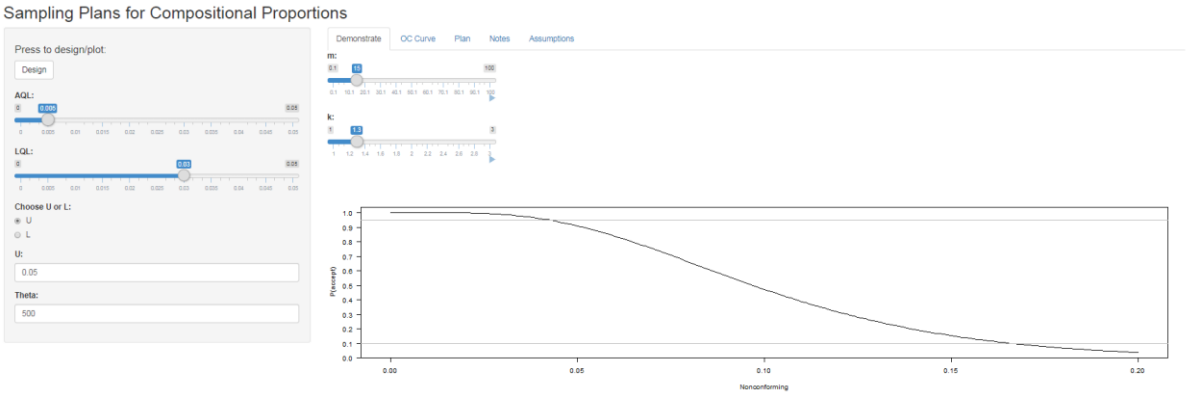
Milk powder production process generally involves only a very small amount of variability. A conservative value of θ=10000 can be employed to implement the sampling plan based on the beta distribution.

1. Using *m=24* subsamples, a final composite is formed.
2. The estimated protein composition of 33.2% is obtained after lab test.
3. The SD is estimated as $=\hat{\sigma} = \sqrt{\hat{\mu}(1-\hat{\mu})/\theta} = \sqrt{0.332(1-0.332)/10000}$=0.00471. For L=32.4% and *k=1.3*, $\hat{\mu} - k\hat{\sigma} = 0.332 - 1.3 * 0.0015 = 32.6\%$ which is greater than the lower limit L=32.4%. The lot is therefore accepted.

The plan parameters *G* (or *m*) and *k* can be determined for given two-points of the OC curve. Let p_1 be the Producer's Risk Quality (PRQ) for the proportion nonconforming compositional fraction, and p_2 be the Consumer's Risk Quality (CRQ). Let α and β be the producer's and consumer's risks respectively corresponding to p_1 and p_2. This two point design imposes the conditions $P_a(p_1) = 1-\alpha$ and $P_a(p_2) = \beta$. The amount *G* or *m* controls the variability in the estimates $\hat{\mu}$ and $\hat{\sigma}$ while *k* mainly influences the achieved producer's and consumer's risks.
For implementation of the beta plan, an online tool is available. The user will need to input the usual PRQ, CRQ, producer's and consumer's risk as well as the precision θ to design the beta plan. That is, the number of primary increments to be taken and the *k* constant to be used will be computed by the tool. This tool will also show the OC curve of a given plan with *m* and *k* so that the risks can be evaluated graphically.

7.2.2 App for the design of beta sampling plan

The main limitation of the beta plan tool is that it does not incorporate the measurement error in the determination of the compositional fraction using the composite sample.



This app can be accessed via the link: **App10**

7.3 Sampling plans for Compositional Means

If the lot is homogeneous, sampling plans to control the average level discussed in Section 7 can be used to control the average compositional levels.

It is however necessary to ensure that the normal distributional assumption is satisfactorily met using historical data.

When the lot is heterogeneous, two stage sampling inspection plans are recommended in the literature. For details, see Schilling and Neubauer (2008).

The ISO Standard ISO 11648-1:2003 deals with many of the non-manufactured bulk materials, including particulate matter etc. For food export inspection, these procedures are of limited use. This is because lack of homogeneity in food quality characteristics can be detrimental for consumer protection.

Chapter 8 Measurement and Inspection errors

8.1 Measurement Errors for Numerical Data

The term measurement errors relates to numerical measurements on the quality characteristic of interest. The following definitions relating to measurement errors are based on ISO-5725.1 (1994).

Trueness is the closeness of agreement between the mean of a large number of test results and the accepted reference value. Trueness is normally expressed in terms of bias.

Bias is the difference between the expectation of the test result and the accepted reference value.

Precision is the closeness of agreement between test results. Precision is necessary because tests that are performed in what appear to be identical circumstances, often don't give identical results. This is because of random errors that are present in all measurements, which cannot all be controlled.

Repeatability is the minimum variation in results. By the term **repeatability conditions**, it is meant that results are obtained using the same method, in the same lab, by the same person, and with the same equipment within a short time frame.

Repeatability standard deviation is the standard deviation of the results from the test that are gained under the conditions listed previously. The term **repeatability limit** refers to the expected value equal or less to the difference of two test results with 95% confidence (probability), given that the repeatability conditions were met.

Reproducibility is the maximum variability in results. The **reproducibility conditions** are where test results are obtained using the same method, but with different labs, different people and different equipment.

The **reproducibility standard deviation** is a measure of dispersion of the distribution of results obtained under the reproducibility conditions. Likewise, the **reproducibility limit** is also the same as the repeatability limit, except it is based on results that are obtained by the reproducibility conditions.

In reproducibility, the conditions do contribute to the random variability of the test results, however, in repeatability the conditions don't contribute to the systematic variability of test results. Hence, repeatability and reproducibility are the two extremes of precision.

Accuracy combines both trueness and precision, and is known as the total displacement of a result from a reference value due to random and systematic effects.

An **outlier** is a value which is inconsistent with other members from a set.

Error is the difference between the measured value and the true value of what is being measured. Errors can be either random or systematic. Random errors are uncorrelated, but they affect the results of the repeated measurements. Some examples are: whether they are repeatable, whether they are reproducible, and whether they are stable. Systematic errors are different, in that they affect all measurements taken in the same way and can be identified when the random errors are small. Some examples are: accuracy, bias, and drift.

In order for measurements to be made in the same way, there needs to be a measurement method that has been standardized (to eliminate as many differences as possible). This requires a document that has full details on how the measurements will be carried out.

Accuracy measures can be determined by a series of test results reported by different labs. An accuracy experiment can be considered a practical test of the adequacy of a standard measurement method. The results found in an accuracy experiment will show how effective the standardization of the method was.

The metrological objective is to produce reliable test results which can later help to make good decisions. On the other hand, acceptance sampling inspection aims to make good decisions on the lot given that there are measurement error related issues.

Sampling inspection plans can be designed and adjusted when measurement and classification errors of random kind are present. This adjustment can be done for both attribute and variables type sampling plans. The term **inspection error** is used to mean the random errors of classifying a conforming items as nonconforming and vice versa. For example, certain sensory tests are subjective in nature, and even a trained analyst is expected to cause inspection errors. In the next section, a procedure for adjusting the single sampling attribute plans for inspection errors is briefly described.

8.2 Measurement Error Adjustment

Hahn (1982) presented simple methods of removing the measurement errors from the observed numerical data. Even though the examples given Hahn (1982) related to net weight assurance for containers, they can be extended to a general situation. The mathematical theory on the effect of repeatability error and bias on the OC curve of the *k*-method plan is discussed in Owen and Chou (1983).

Let Y_i, i = 1, 2,…, n be *n* apparent or observed measurements. Let $\bar{x}$ and S_y be the sample mean and standard deviations respectively.

Let X_i, i = 1, 2,…, n be the true but unknown levels corresponding to these measurements. Under the simple additive error model $Y = X + Z$ where Z are the measurement errors, the variances are decomposed as

$$\sigma_Y^2 = \sigma_X^2 + \sigma_Z^2$$

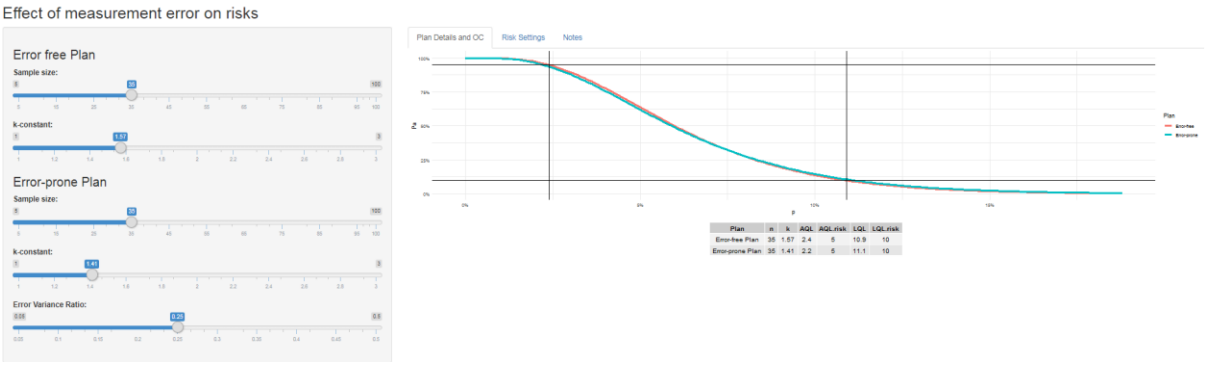
Let the ratio σ_Z^2 / σ_Y^2 be called as the error variance ratio. It is assumed that we have good knowledge of this ratio based on past measurement system studies.

The OC curve of the *k*-method variables plan can be adjusted for given error variance ratio. The acceptability constant will be smaller when adjusted for the repeatability SD in general.

In order to adjust for the bias, the actual measurements can be converted to bias adjusted measurements and then the variables plan can be applied.

8.2.1 Tool for adjusting variables plans

This tool requires the user to specify the error variance ratio, so that the OC curves of the variables plan with and without measurement errors can be compared and adjusted for the measurement errors.



The default settings shows how the *k* constant becomes smaller in the presence of measurement errors. This app can be accessed via the link: **App15**

8.3 Inspection errors (Attribute Plans)

Schilling and Neubauer (2008) details some of the reasons for inspection inaccuracy as follows:

1. Willful errors which include: criminal acts, and falsification to make it more convenient for the inspector.
2. Intermediate errors due to: bias, rounding off etc. Failure to call a defect when it is close to the specification limit falls into this category also.
3. Involuntary errors due to: blunders, fatigue, or other human imperfections

Inspection errors are caused when testing a unit of inspection for its conformance. The sources of inspection errors include human error, instrument error, or any other measurement related errors. Type I errors are when a true conforming unit is placed as apparently nonconforming. The type II errors are when a true nonconforming unit is placed as apparently conforming.

The impact of these two types of inspection errors on the OC curve has been studied by many. When inspection errors are present, they generally increase the producer's risk when compared to the consumer's risk. The impact of inspection error is particularly greater for zero acceptance number plans.

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

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

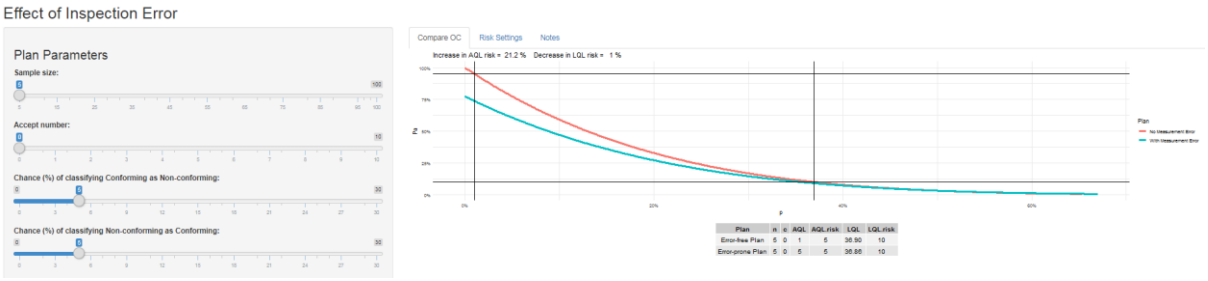
e_1 is the probability of classifying a conforming item as nonconforming and

e_2 is the probability of classifying a nonconforming item as conforming.

It is established in the literature that e_1 is more important than e_2. In other words, the OC curve of the single sampling plan is largely a function of e_1 and e_2 does not have much effect the performance of the plan.

8.3.1 Tool for attribute plan inspection error adjustment

This tool allows a comparison of the OC curves of single sampling attribute plans with and without inspection errors. For example, conformity testing procedures (see Section 1.23) are based on 95% confidence intervals so that, e_1, the Type I error probability of misclassifying a conforming item as nonconforming, is fixed as 0.05 by design. As a result, the attribute classification using a conformity test procedure will hugely increase the producer's risk when the measurements are closer to the specification.



This app can be accessed via the link: **App7**

Unlike variables plans, adjustment for the inspection errors cannot be done for attribute inspection plans. The remedy lies in repeated testing so that the overall Type I misclassification errors becomes small. This is discussed in the next section.

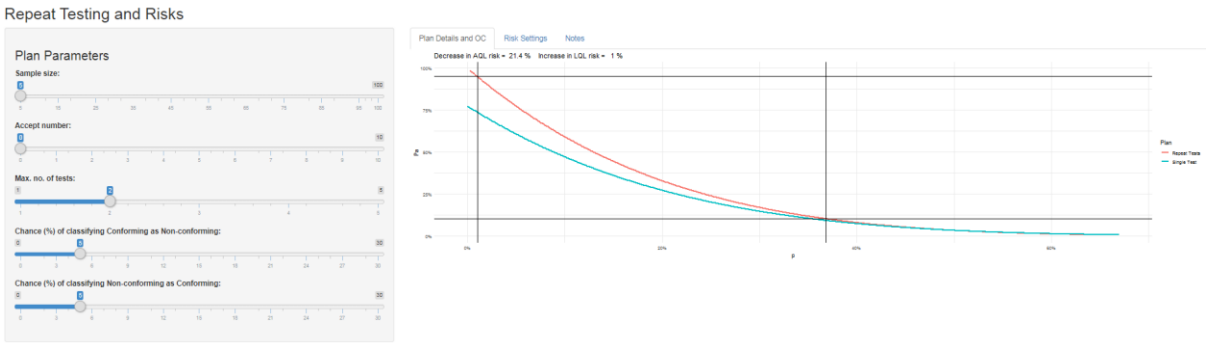
8.4 Repeated Testing

One of the approaches to mitigate the impact of inspection and measurement errors is retesting. If an item is found to be nonconforming, it can be tested again. This is because production of nonconforming units is expected to be in a smaller proportion and only occasionally will retesting be used. Even though conforming units can be re-tested, this strategy is often not beneficial due to economic and other reasons. It is more important to try control the Type I inspection error (of classifying conforming items as nonconforming) because the lot quality is generally good rather than bad. Therefore, it makes more sense to re-test the items that are apparently nonconforming as compared to the items that are apparently conforming. This re-test of an item can be done up to a maximum of ***m*** times. This means that each sampled item will have a maximum of ***m*** chances for conformance. There needs to be the assumption that testing will not degrade the quality of the item. If a sample is of non-discrete type physical material such as powder, then it is assumed that ***m*** homogeneous sub-samples can be made for every unit of the sample.

If classification errors are large, retesting of nonconforming items is necessary to reduce the adverse impact of inspection errors on the producer's risk. The presence of inspection error affects the consumer's risk but it can be compensated for by adjusting the sample size slightly. However, the use of repeat tests is particularly essential to avoid rejection when the quality is in the parts per million range.

8.4.1 App to evaluate repeated testing

A single sampling attribute plan can be evaluated using the tool shown below. The OC curves of the single sampling plan with and without inspection errors are shown. The plan parameters can be adjusted so that the risks are maintained at the desired PRQ and CRQ.



This app can be accessed via the link: **App4**

8.5 Fractional Nonconformance Inspection

The term fractional nonconformance or **FNC** refers to the probability of an error-prone observation breaching the specification limits.

An observed measurement *Y* is classified with certainty as conforming or not for given specification limits only when there are no measurement errors. Analytical testing of fat content etc. involves

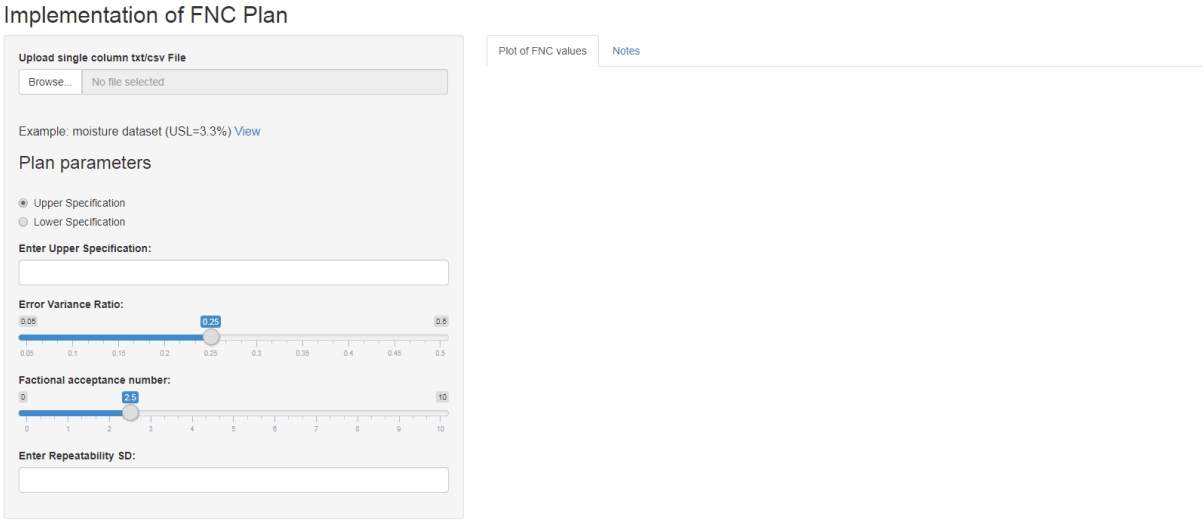
considerable measurement uncertainty, often up to half of the observed variation. The distribution of the measurement errors (*Z*) can be fairly well ascertained using past calibration studies. Measurement error uncertainty results only in an estimated probability of conformance of a unit. The probability of nonconformance of an individual unit based on the error-prone measurement is defined as the *fractional* nonconforming unit. The following figure illustrates the concept of fractional nonconformance. Given the measurement error distribution, the probability of breaching the upper specification limit, $\hat{p}_{iu}$ is the FNC value.

Even though the observed measurement is below the Upper Specification Limit, it still has a small chance of nonconforming because of the measurement errors.

As an example, consider five numerical measurements of a weight characteristic (100.5, 100.7, 100.2, 100.6, 100.4). If the measurement error distribution is known to be normally distributed with mean zero and standard deviation 0.25, i.e. N(0, 0.25), the probabilities of these five measurements falling below the lower specification limit of *L=100* are (0.023, 0.003, 0.212, 0.008, 0.055). The sum of all the FNC values, $\sum \hat{p}_{iu}$ is given by 0.3. This sum can be compared with a fractional acceptance number such as 0.5. This approach is similar to comparing the number of nonconforming units *d* with the acceptance number *c* in the attribute plan. The plan can also be implemented using the mean FNC which can be compared with the maximum allowable fraction nonconforming.

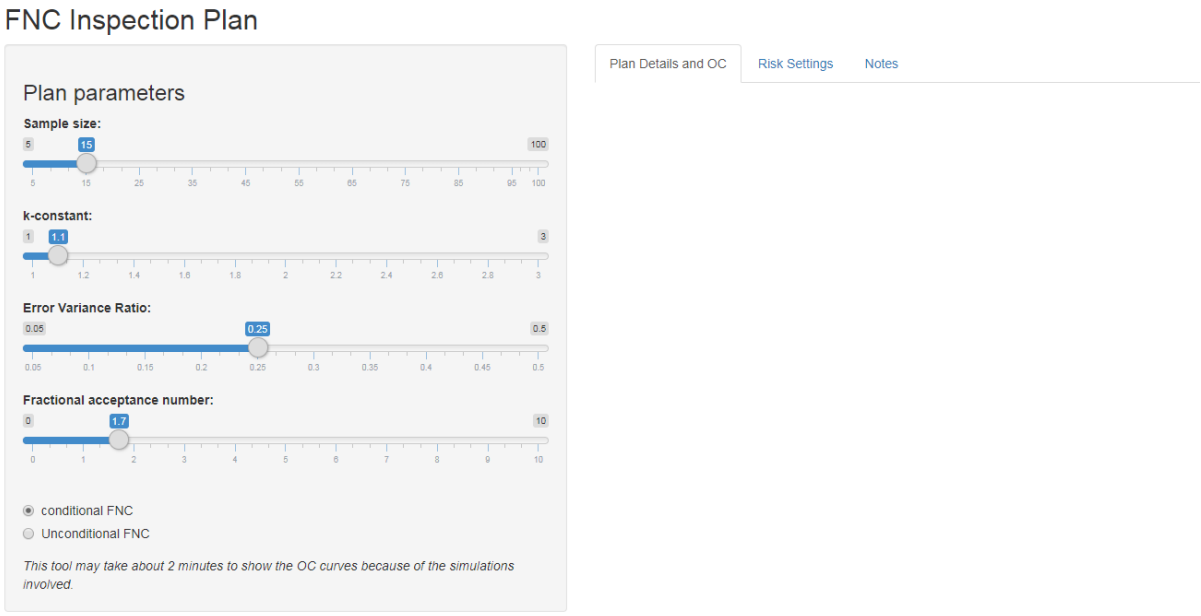
A conditional version of the FNC can also be defined. The probability distribution of the measurement error *Z* conditional on the given observed measurement value *y* is used to obtain the conditional FNC, $\hat{p}_{ic}$ values. The additional knowledge that an apparent measurement has been made and its distance from the sample mean contains additional information on its nonconformance.

An app to implement FNC based plans is shown below. This tool requires the user to input the data and select an acceptance number. The repeatability standard deviation and the error variance ratio must be specified. The tool will then show the FNC probabilities and the decision of acceptance or rejection of the lot.



This app can be accessed via the link **App8**

A further shiny tool to evaluate the OC curves of the FNC based plans and measurement error adjusted variables plans is shown below:



This app can be accessed via the link: **App16**

The conditional FNC based adjustment for measurement error is more powerful because it provides better discrimination between good and poor quality lots.

The main advantage of the FNC inspection plan is that the plan can be used even when the underlying quality characteristic is not normally distributed. On the other hand, variables plan require the underlying distributional assumptions to be met. If normality does not hold, the OC curve of the variables plan based the normal assumption is not fully trustworthy.

In Section 1.25, a further tool to compute and chart the FNC values for user data is available.

8.6 Conformity Testing

Conformity testing, also known as evaluation of conformity or compliance testing, is used to assure that an entity meets a specific requirement and/or regulatory standard. In this context the entity refers to the sample actually tested.

The objective of conformity assessment differs from that of acceptance sampling - acceptance sampling uses a limited number of samples to determine whether to accept or reject a lot of some product whereas, in contrast, in conformity testing the inference is limited to the 'entity' i.e. the sample tested. In other words, conformity testing is a procedure for making a decision about the particular sample whereas the proportion nonconforming in a lot is the main quality measure of interest in acceptance sampling.

Examples of conformity testing include the test of the concentration of some trace elements in the blood of employees for their health evaluation, the analysis of an athlete's urine to detect abuse of xenobiotic anabolic steroids, testosterone and doping etc.

The specification for the quantifiable characteristic, such as the maximum allowable concentration of a drug or trace element in blood for normal people, is called as a limiting value (LV) in the conformity testing protocols. The LV could be understood as either a minimum value (Lower Limit or Lower Specification Limit) or a maximum value (Upper Limit or Upper Specification Limit), or both. The interval containing all permissible values of the characteristic is called the region of permissible values. A conformity testing protocol provides assurance of conformity by checking whether the measurement of interest falls within the region of permissible values or not. Conformity can be declared if and only if the whole uncertainty interval is located within the region of permissible values.

Measurement and sampling uncertainties, including metrological traceability, become crucial for the declaration of conformity, especially when the measured value is close to the set limiting value. Measurement uncertainty is usually reported as an uncertainty interval, given in the form of a confidence interval. The common practice for conformity testing is to compare the measurement uncertainty interval around the measurement result with the region of permissible values.

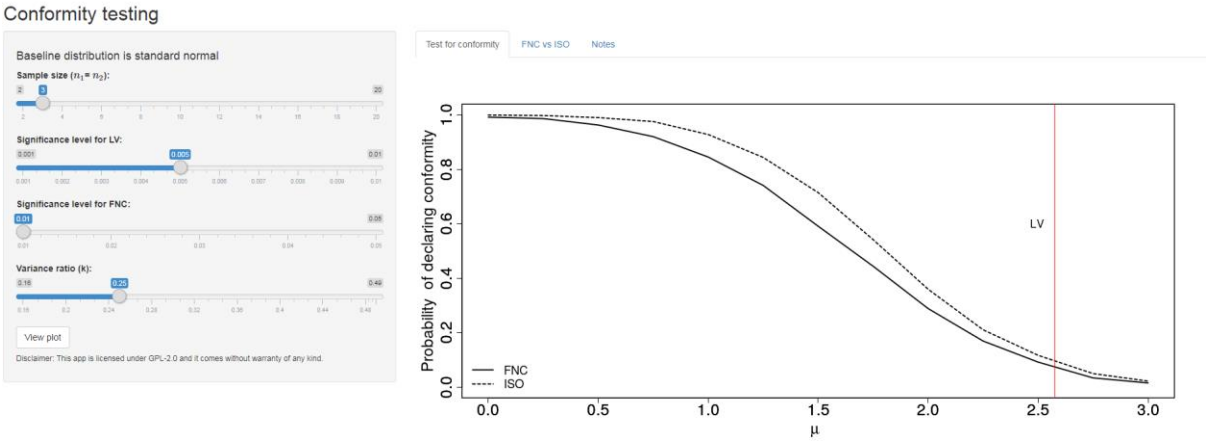
ISO 10576-1 International Standard recommends performing the conformity test as a two-stage procedure, which was initially proposed by Holst, Thyregod, and Wilrich (2001). The rules for asserting conformity or nonconformity are:

- assurance of conformity: the uncertainty interval is inside the region of permissible values;
- assurance of non-conformity: the uncertainty interval is included in the region of non-permissible values; and
- inconclusive result: the uncertainty interval includes LV.

The main disadvantage of the conformity testing procedure is that in many cases, inconclusive results will be obtained even though a sample is conforming but due to measurement errors, the uncertainty interval includes the limiting value.

The ISO 10576-1 Standard does not encourage reduction of measurement errors by design and hence poorer measurement systems will produce a higher proportion of inconclusive results. Hence producers are forced to guard-band in order to reduce the inconclusive result for a measurement. In other words, the conformity testing procedures, not being acceptance sampling procedures, do not aim to make a decision on the lot but only concerned with the risk that the measured sample is conforming or not.

A FNC based two-stage conformity testing procedure is found to reduce the probability of incorrect declaration of conformity or inconclusive result for nonconforming entities (Type II error) when the number of test samples is greater than one and this superiority becomes more significant when the sample size increases. A new tool that looks at the probability of declaring conformity for both the FNC and ISO methods of assessing conformance in the presence of measurement errors is presented below.



This app can be accessed via the link: **App9**

A sample size, the significance level for the limiting value (LV), the significance level for FNC and the variance ratio each need to be selected. Plots are then displayed which show the probability of conformity and non-conformity for ISO two-stage and FNC testing procedures to compare. It is assumed that both these conformity testing procedures have equal samples that are tested.

Chapter 9 Special purpose Sampling Inspection Tools

The following is a quick summary of certain special purpose sampling inspection plans and the tools for implementing them. These tools may not have universal applications.

9.1 Microbiological sampling plans

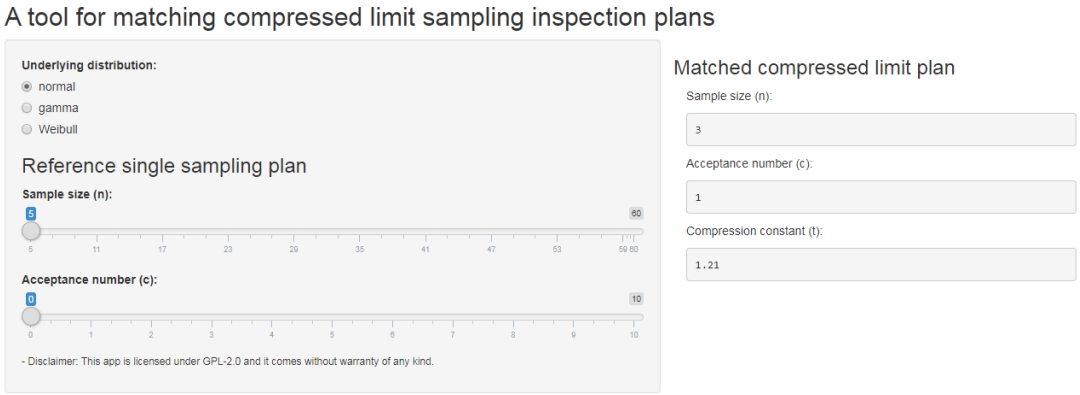
Sampling inspection for food safety forms a special class of acceptance sampling plans. The International Commission on Microbiological Specifications for Foods (ICMSF, the Commission) has formulated a number of sampling inspection plans for food safety.

Their online tools are available at the ICMSF sampling plan tools website.

The ICMSF plans evaluate the producer's and consumer's risks using OC curves. It is essential to consult these tools to assess the discriminatory performance of the chosen plan.

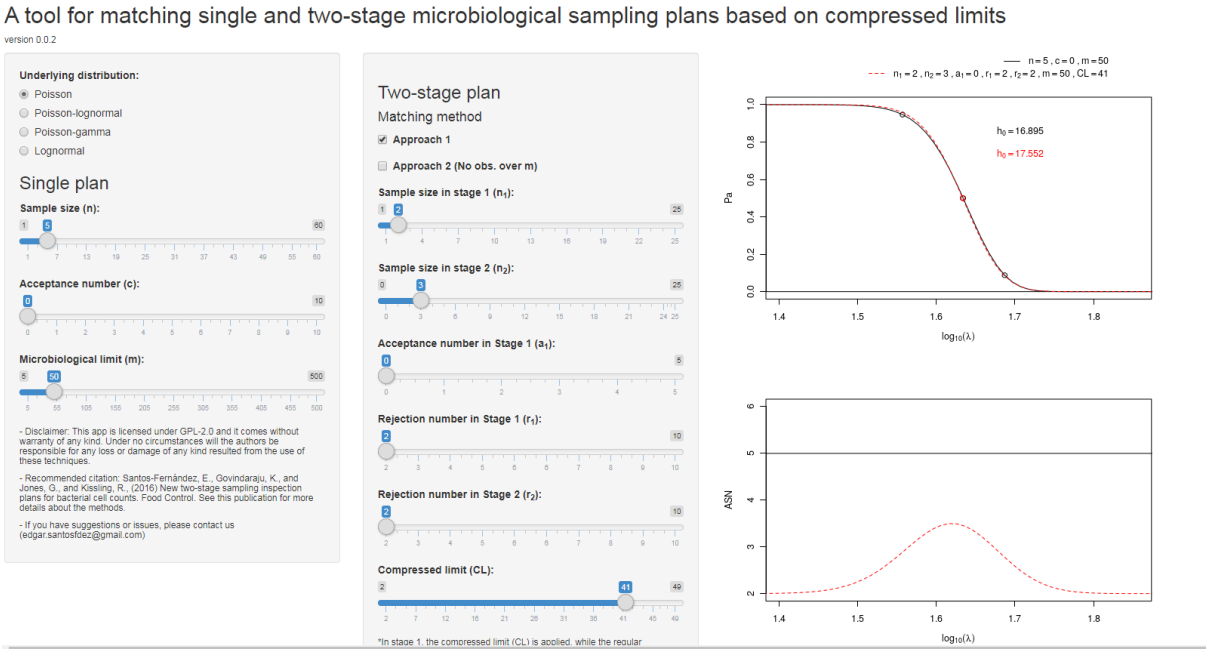
Other more newly developed online apps for food safety inspection are given below.

Instead of the actual microbiological limit or specification limit, a compressed limit can be employed. This compressed limit is usually set as the good manufacturing practices (GMP) limit. It is possible to match the plan based on the original limit with the plan based on the compressed limit. The use of the compressed limit plan is generally economical in terms of test costs. The following app designs such compressed limit plans. Empirical knowledge of the underlying distribution is necessary to implement such compressed plans. The following app will be useful to design such compressed limit variables plans.



This app can be accessed via the link: **App11**

The next app is used to design a double sampling plan employing a compressed limit in the first stage of inspection to reduce testing costs. These plans were designed for sanitary characteristics where the bacterial count generally fits a Poisson or a mixed-Poisson distribution. Both plans showed improved consumer protection with substantially smaller average sample size.



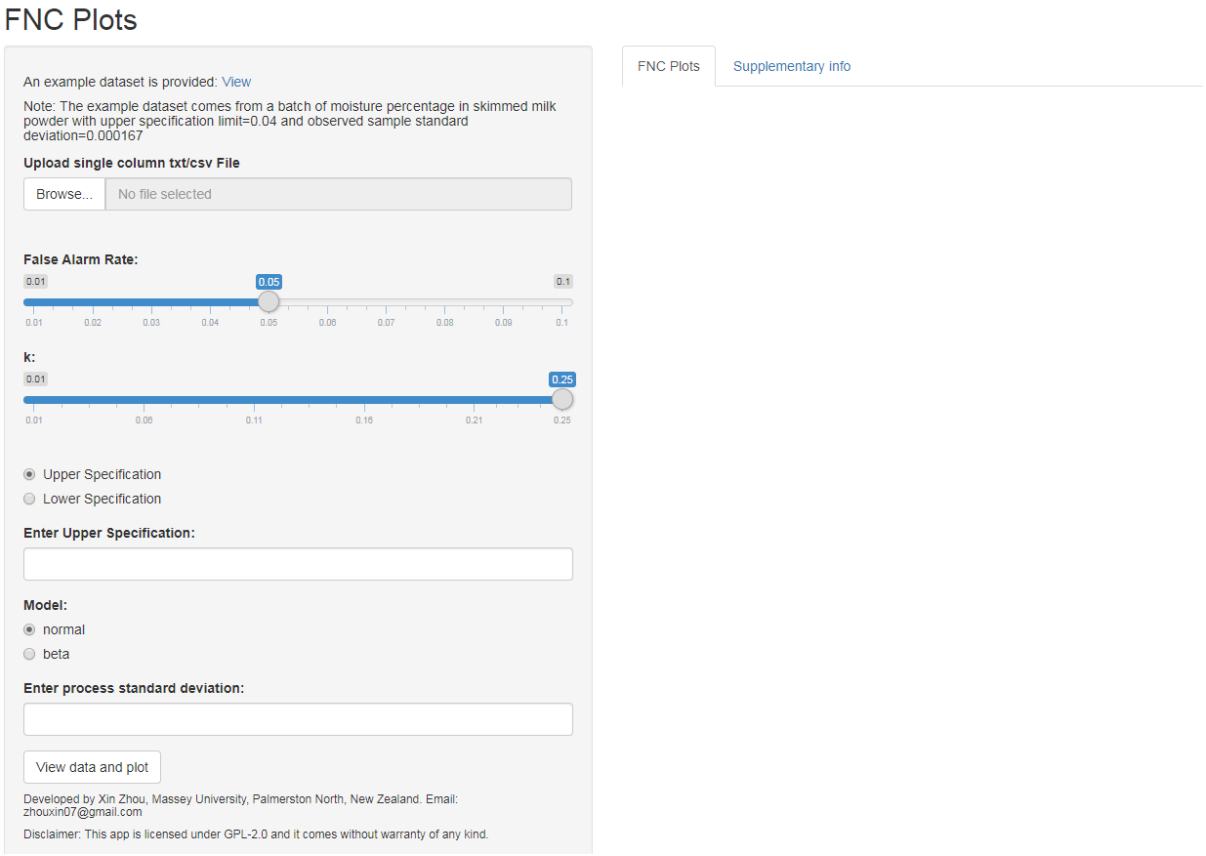
This app can be accessed via the link: **App13**

9.2 Fractional Nonconformance Charting

In Section 1.22, the fractional nonconformance or **FNC** inspection plans were presented. Recall that FNC refers to the probability of an error-prone observation breaching the specification limits. The FNC statistic quantifies the probability of conformance when measurements are error-prone. The following tool allows the user to compute the FNC values for their data and also chart them.

This tool is particularly useful for producers to see the variability in FNC during production periods and exercise control over the fractional nonconformance.

A tool for FNC assessment is shown below:

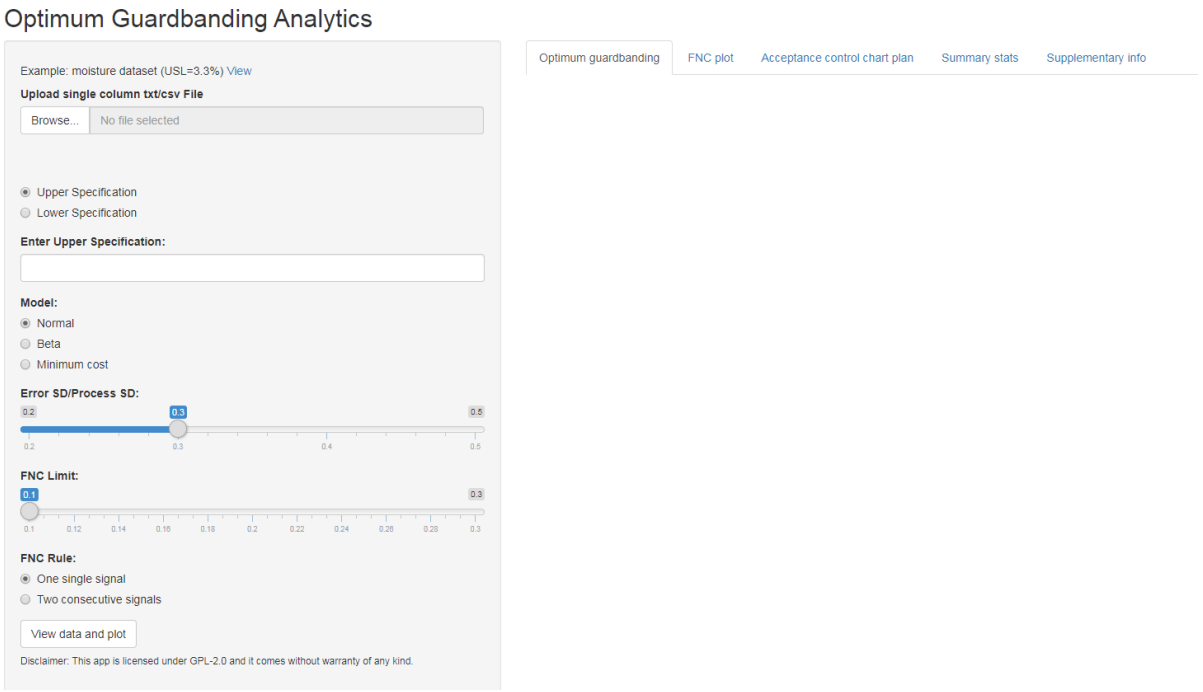


This app can be accessed via the link: **App17**

This app lets you input a data set to work with. The false alarm rate and *k*-constant are chosen, along with either the upper or lower specification limit. Then either the beta or normal model needs to be specified with the value for the process standard deviation, and what the PRQ is. Once these are all included, the inspection data is plotted and can be viewed for assessment of quality.

9.3 Acceptance Control Chart Inspection Plan

When measurement errors are dominant, guard-banding at the producer's end can become unfair for the producer. FNC refers to the probability of an error-prone observation breaching the specification limits. The design of sampling plans and conformity testing using the FNC statistic is given in the following app. An acceptance control chart plan is designed and demonstrated in this tool.



For more details about the app, supplementary notes are given in the app itself. The app can be accessed via: **App12**

Chapter 10 Summary

Ad hoc or convenience sampling involves a sample being taken from a part of a population that is nearby and convenient. It is non-probability sampling, and is sometimes used in pilot testing. There is no basis for it, other than the samples being available for testing. Even though convenience sampling can be cost effective, it is not possible to quantify the producer's and consumer's risks for such plans. The potential sampling error and lack of representation of population render them very inaccurate due to the bias this ad hoc sampling inspection carries.

Sampling inspection plans for routine inspections are often single sampling plans. Both attribute and variables sampling plans intended for routine inspection assume that physical sampling is done correctly and no errors are present in testing or measuring the variables of interest.

The evaluation of the risks to the producer and consumer for such routine sampling plans can be done using their Operating Characteristic (OC) curves. Both the producer and consumer must fully aware of the risk or chances that good quality lots are rejected as well as poor quality lots being wrongly accepted as good. It is necessary to control these risks with the appropriate choice of the sample size and set an acceptance criterion accordingly.

It is also important to recognize that the routine plans may fail in the presence of excessive measurement or inspection errors. Routine sampling plans can be adjusted for measurement uncertainty and then the risks can be evaluated. The OC curve again serves as the appropriate tool for making this risk assessment.

With the advancement of software technology, it is easy to evaluate the underlying risks quantitatively using online web tools. A number of such tools and examples are presented in this document.

Appendix III

The comments provided by the USA to the EWG including the draft top-level outline for revised Guidelines

(for information and general consideration – see CL 2020/27-MAS)

US Comments to GXG50 EWG February 4, 2020

The US thanks New Zealand for all of their work and effort in drafting of the circulated document. The material in CXG50 and in describing sampling is complicated and it is a difficult task to try and make such material accessible and useful to a nonexperts. The US fully supports continued work on the updating of CXG50. We have reviewed the draft and have a number of comments and suggestions for consideration by the EWG. Out comments are separated in three formats. First, directly below, are bulleted comments about the overall approach and some specific sections and recommendations. Second, is a proposed outline that we feel captures many of the topics that should be captured in the update of CXG50, some but not all that are present in the current draft. The outline attempts to follow the format of the current CXG50, but expands in areas that may be lacking. Finally, we have also added specific comments to the draft pdf version. In some cases, the inserted comments and the bulleted list may overlap.

- The revision was expected to correct certain parts of CXG50 for ease of understanding and use; however, the proposed draft is an entirely different document that misses much of the practical guidance and useful information found in CXG50.
- The advanced methods introduced are confusing and likely unnecessary for most readers. Users need to understand the sampling plans selected from the Guidelines.
- A simple Excel spreadsheet can replace many of the existing tables in CXG50.
- A simple table, such as Table 2 in CAC/GL 33 (below), would illustrate zero-acceptance number sampling plan performance versus the number of samples required, and would cover many user's needs.

Incidence of non-compliant residues in the lot %	Minimum number of samples required to detect a non-compliant residue with a probability of:		
	90%	*95%*	*99%*
90	*1*	*-*	*2*
80	*-*	*2*	*3*
70	*2*	*3*	*4*
60	*3*	*4*	*5*
50	*4*	*5*	*7*
40	*5*	*6*	*9*
35	*6*	*7*	*11*
30	*7*	*9*	*13*
25	*9*	*11*	*17*
20	*11*	*14*	*21*
15	*15*	*19*	*29*
10	*22*	*29*	*44*
5	*45*	*59*	*90*
1	*231*	*299*	*459*
0.5	*460*	*598*	*919*
0.1	*2302*	*2995*	*4603*

- A discussion of the practical limitations of sampling is needed. A marginally practical zero-acceptance number plan using 29 samples can only reject (with 95% confidence) lots that have 100 non-compliant units per 1,000. This level of non-conformity is far above food safety and suitability objectives.

- The revision should discuss how food safety and quality objectives are achieved through HACCP and other production control systems and how routine sampling should only be used when no better tool is available to assure objectives are met.
- The draft recommends using 95% confidence for accepting good lots, but only 90% confidence for rejecting bad lots, but doesn't clarify why these confidences are chosen or what the ramifications are of that choice. Why these may be adjustable, previous experience with Codex Committees indicates that the "reference" or initial set parameters will always be used, without understanding or regard to the applicability.
- The revision should consider removing ISO terminology that has confused Committees in the past (e.g. AQL/LQ and PRQ/CRQ) and use plain language. Codex Committee on Fish and Fishery Product members used AQL = 6.5 in standards believing this provided consumer protection, but it only protects the producer. Additionally, sampling plans in CCCF have simply referenced an AQL = 6.5.
- The proposed draft seems to suggest re-inspecting rejected lots. This is a statistically invalid approach because it offers greater opportunity to accept a lot that may be out of compliance. The Guidelines on lotting and re-inspection found in FAO/WHO 2016 should be used.
- The proposed draft seems to suggest changing the definition of an unacceptable lot during selection of a sampling plan (e.g. Section 3.1.3). However, a lot that should be rejected is determined before selecting a sampling plan. It is based on risk assessment and food safety/quality objectives, and sometimes food security.
- The guidance for bulk material was confusing to the US. Bulk material only affects how random samples are drawn. When the purpose of sampling is similar, sampling plans should not change for product received in packaged or bulk form.
- The discussion of composite samples would benefit from further elaboration. Composites samples are used in different ways for different purposes, unrelated to bulk or packaged material.
- The proposed draft overemphasizes 'variables sampling plans', which have limited applicability in Codex and confuse Codex committees.
- The revision should discuss the need for 'distribution-free' sampling plans (i.e. attribute plans) for typical isolated trade lot sampling. Trade lots usually contain product produced on different days and from different raw material sources, resulting in unknown and irregular distributions of hazard and defects.
- The scope of the Guidelines should not include 'statistical process control'. Also, end-product testing by manufacturers is ineffective and should be discouraged. Codex sampling guidelines are predominantly used for border inspection and other receiving oriented situations.
- It is unusual to include measurement error in sampling plans. Limits in standards are based on the reference methods prescribed, and measurement error should not be accounted for twice. In addition, measurement error is generally insignificant compared to the variation between sample units, and health related provisions contain safety factors that are orders of magnitude greater than the measurement error.
- The revision should avoid reference to ISO standards because they are not designed for, or appropriate for, isolated food lot acceptance testing, and they regularly change.
- The Guidelines should avoid reference to ICMSF example microbiological cases based on health risk. They are not generally applicable for several reasons, such as the presumption of a 10-fold difference between "m" and "M", and the presumption of continuous lot-by-lot inspection from the same source.
- Both CXG50 and the proposed draft cover homogeneity confusingly. Homogeneity does not affect attribute sampling plans, and for variables sampling plans, heterogeneity is modelled by the assumed normal (or other) distribution.
- The proposed draft is a bit redundant overall and repeats content from other standards. For example, the proposed draft discusses the principle of fairness 21 times, however the principle

of fairness is already covered in Codex Principles for the use of Sampling and Testing in International Food Trade (CXG 83-2103).

- The many “risk” terms found in the draft (below) cause confusion between health risk and statistical probability, and are unnecessary.

  > *Perceived risks, risk setting, control of risks, level of risk to be accepted, sampling creates risk, allowable risks, higher risks, unacceptable risks, risk rating, risk of sampling options, CRQ/PRQ associated risks, specified risks, level of risk to be accepted, risk of incorrect decision, some options are high risk, risk of sampling options.*

- We may misinterpret the intent (see bullet above), but as written we do not agree with the concept that sampling plans define ‘acceptable consumer risk’. Acceptable risk levels for food safety and suitability generally range from one incident per thousand servings to one incident per million servings. These levels cannot be detected with practical sampling.
- Several new draft sections contain subject areas with low applicability and usefulness that could be replaced with more practical material from CXG50.
- The revision should include guidance on different sampling objectives and how these relate to sampling frequency and stringency.
- The revision should include guidance on selecting an appropriate type of sampling plan based on the type of provision and the sampling objective.
- The revision should discuss the relevant sample unit amount, and the advantages of many smaller sample units over fewer larger sample units (see FAO/WHO 2016).
- The revision should discuss random sampling methods and discourage combining lots prior to sampling.
- The Guidelines should not cover double, multiple, sequential, and other sampling schemes. These are often misapplied and are purposely excluded from CXG50.
- Per discussions in CCMAS, the revision should avoid examples, and we do not support the examples in the proposed draft.
- The proposed draft sometimes changes the intended audience, inappropriately addressing only commodity committees, and not governments and industry.
- The requirements for approval of sampling plans by CCMAS should be in the *Procedural Manual* and not duplicated or changed through revision of CXG50.
- Microbiological sampling is covered by CCFH standards and FAO/WHO 2016, and these documents should be referenced for that purpose.

Reference: FAO/WHO 2016. *Statistical Aspects of Microbiological Criteria Related to Foods*.

Link: http://www.fao.org/3/a-i3996e.pdf

US Proposed CXG50 Revision Top-level Outline

(for information and general comment – see CL 2020/27-MAS)

Preamble

- Brief (few paragraphs at most) focus on aspects of what document is and why it is needed
- Should be understandable to audience with limited statistical training and relatively jargon free

Key terms and definitions (Brief definitions)

- This could be an appendix
- Reference CXG 72 when appropriate

1. Introduction

- Prerequisite Codex documents (Procedural Manual, CXG 83)
- Target audience: Codex committees, Governments, Industry.
- Scope
  - Food hazards, suitability for consumption, and quality
  - Focus on border inspection and other receiver-oriented situations
    - Sampling raw materials, or finished products, at receiving
    - Does not cover statistical process control (SPC)
    - Discourage end-product testing, other than required by buyer/importer.
  - Focus is on 'Attribute Sampling Plans'
    - Does not cover multiple, sequential, or switching sampling schemes

      [Variables plans are inappropriate for Codex standards, however could be carefully handled if CCMAS believes necessary.]

2. Basic concepts

- Reasons for sampling
  - Acceptance sampling
    - Acceptability of lots with unknown control history
    - Routine sampling (lot-by-lot)
    - Use when no better tools are available
    - Practical limitations cannot assure consumer protection
  - Verification sampling
    - Verification of supplier system performance (not lot acceptance)
    - Intermittent sampling (between-lot)
    - Rigorous plan - results used for follow-up
- Provision criteria
  - Food safety/quality objective
    - Health hazards (chemical, physical, biological)
      - Typical acceptable risk: 1:1,000 to 1:1,000,000
    - Suitability for human consumption (nutrition, filth, decomposition)
      - Acceptable risk: similar to health hazard
    - Quality defects (e.g. composition, color, texture, size, uniformity, net weight)
      - Acceptable risk: 1:100 to "Six Sigma" (3.4 defects per million)
  - Decision unit amount (weight, volume, count)
    - The sample unit size that the provision applies to
    - Determines the variance measured [discuss heterogeneity of food hazards/defects]
    - Relates to provision objective
      - For acute hazards, the sample unit amount corresponds to a single serving amount
      - For quality defects it is based on end-user acceptance and generally corresponds to a retail package size
  - Decision unit limits (for conformity)
    - Maximum/minimum (continuous variable, count, proportion)

        - Presence/absence
    - Lot acceptance
        - Maximum/minimum proportion of nonconforming units in lot, per provision objective
        - [Maximum/minimum average level in lot (variables sampling plans)]
- Sampling probability and plan performance
    - Samples are used to determine lot compliance with the provision
    - Subset used to estimate the actual lot parameters
    - OC curve plots probability of accepting/rejecting lot vs the level of non-compliant units in the lot.
        - [“Consumer risk-point”]
        - [“Producer risk-point”]
- Making decisions about lots (see Part 2 FAO/WHO 2016).
    - What is a lot?
        - How defined
        - Why not redefined after sampling
        - Why invalid to re-sample
    - Between-lot testing and lot-by-lot testing

3. Two-class attribute sampling plans

- Why predominantly used
    - Universally applicable
    - Reliable performance (distribution-free)
- Binomial distribution (presence/absence, over/under)
- Zero-acceptance number plans
    - Discus why appropriate for most situations
    - Design
        - Proportion (%) non-compliant units in lot, above which the lot is unacceptable, per provision
        - Select sample size to detect unacceptable lots with high probability (confidence) (e.g. 95% for defects, 99%+ for hazards) [“consumer risk-point”].
        - When applicable, also check the level of non-compliant units that will be accepted with high probability under the plan and determine if this level is achievable by the supplier using GMPs or may result in too many rejections of acceptable lots.
        - Discuss sample size limitations and compromised performance
    - General formula and hypergeometric formula
    - Table (CXG-33, Table 2)
    - App link
- Two-class plans using acceptance number
    - Used when suppliers cannot meet zero-acceptance number plans (increase discrimination, steepen OC curve)
    - Require more samples.
    - How designed
    - Formula
    - [example table/OC curve, C=0,1,2]
    - App link

4. Special situation sampling plans

- Three-class attribute sampling plans
    - Design (determining “marginal” level)
    - App link
- [Plans that assume known distribution and known standard deviation.]
    - Why inappropriate for Codex sampling plans
        - Require unchanging stable standard deviation
            - Trade lots generally from mixed production lines, mixed dates, mixed raw material sources have unstable standard deviations
            - Increased standard deviation (often bad lots) increases probability of accepting bad lots

- Decreased standard deviation (often good lots) decreases probability of accepting good lots
- Not appropriate for detecting economic fraud (mixed lots)
- Lot rejections are to be expected when all sample units are compliant.
- Attribute plans not affected by distribution

- Variables sampling plans
  - Used within production facility and like situations
    - Continuous series from same production line and raw material
    - Standard deviation monitored regularly and stable or estimated from the lot samples

5. Random sampling

- Sampling packaged goods
  - Importance for plan validity
  - Random numbers
- Sampling bulk material
  - Decision/sample unit = “increment”
  - Same principle as discrete units
  - Methods used

6. Composite samples

- Use with two-class and three-class sampling plans
  - Used to save analytical cost when lots usually acceptable
  - Methods and requirements
  - Importance of composite homogeneity
- [Use with variables sampling plans]
  - Used to save analytical cost, when known standard deviation
  - Importance of composite homogeneity

7. Sample handling

- Drawing analytical sub-samples and importance of homogeneity
- Environment and time
- Traceability

8. References

- FAO/WHO Statistical Aspects of Microbiological Criteria Related to Foods
- FAO/WHO Public Health Risks of Histamine and other Biogenic Amines from Fish and Fishery Products
- Recommended Methods of Sampling for the Determination of Pesticide Residues for Compliance with MRLS CXG 33-1999

[Only cite FAO/WHO publications. ISO and ICMSF publications are inapplicable and source of past problems with CXG-50.]

[Less applicable material can be consolidated/replaced with needed material to keep document manageable, such as, endorsement, measurement error, disputes, fitness, documentation, communication, Poisson, beta, conformity testing, convenience sampling, ISO problems, boxes]

Appendix IV

The New Zealand response to four key areas identified by USA (for information)

Key areas

We have identified four key but inter-related areas on which to comment:

- the focus on attributes plans;
- reference to the FAO/WHO 2016[9] document, referred to as "FAO/WHO 2016" in the text below;
- allowance for measurement error; and
- control of Producer's Risk.

Focus on Attribute Plans

NZ does not agree that the principal focus of the guidelines should be on attributes plans. We feel this is too restrictive. Our intention, based on the current GL50 and the TOR, is that the guidelines should present a range of possible sampling plan options along with supporting guidance about when it is appropriate to use each type of plan.

While we agree that attributes plans are robust to distributional assumptions about the behavior of quality characteristics, there are problems with attributes plans. They:

- are possibly inefficient, needing large numbers of samples to operate when, in many cases, variables plans requiring fewer samples can be used legitimately. It is not necessary that the parameter follows the assumed distribution exactly, merely that the assumed distribution provides a reasonable approximation describing the behavior of the parameter within each lot; and
- do not allow for measurement error adjustment, in cases where measurement error is significant (see below).

It is worth noting that although attributes plans do not necessarily rely on distributional assumptions in their creation, it is often necessary to make assumptions about underlying distributions in their evaluation, such as the log-normal distribution in some of the plans in FAO/WHO 2016.

Reference to FAO/WHO 2016 Document

NZ considers that the process described for the design of sampling plans does not follow conventional statistical practice and should not be used as a model for the design of sampling plans for determining compliance with commodity standards. For the revised GL 50, NZ has adopted the standard statistical approach that underpins the GL83 guidelines in which sampling plans are designed from specifications of producer's and consumer's risks. This approach is used in the apps included in the e-book and R package in order that both consumer's and producer's risks can be controlled to appropriate levels, in keeping with the principle adopted by Codex that fair and valid procedures should be used for the assessment of foods.

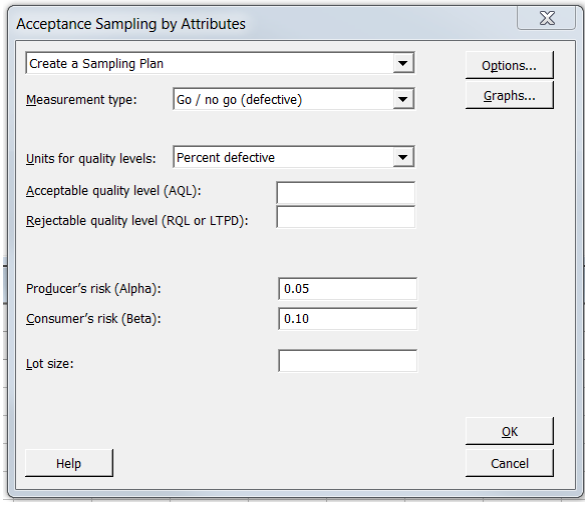
Several comments have been made around the apparent inappropriateness of the ISO plans. We agree that some aspects of the ISO plans render those plans as possibly unsuitable for use by Codex but nevertheless, much of the basis behind the ISO plans is still statistically valid.

ISO plans have employed a sample size versus lot size relationship, apparently at the request of the user base and submissions made in the consultation process. However this relationship is not statistically based and setting sample sizes based on the lot size enables either the producer's risk or the consumer's risk (but not both) to be controlled explicitly.

However NZ has abandoned the ISO approach in favor of the standard statistical approach in which both producer's and consumer's risks are controlled and there is no relationship between sample size and lot size, except where sample size is a large fraction of the lot size.

This standard statistical approach is implemented in Minitab (a well-known statistical package) – the following shows the dialog box relating to the design of attributes plans.

[9] Statistical Aspects of Microbiological Criteria Related to Foods. A Risk Managers Guide.
Food and Agriculture Organization of the United Nations, World Health Organization, Rome, 2016



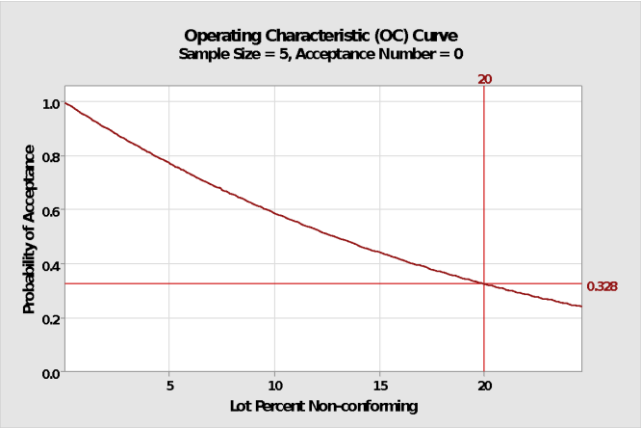
The standard approach also appears, for example, in the find.plan routine of the R package AcceptanceSampling, the author of which, Andreas Kiermeier, was part of the FAO/WHO review team.

FAO/WHO 2016 proposes a trial and error approach to the design of sampling plans, varying the number of samples n and the acceptance number c using a sampling tool until the Operating Characteristic is considered 'satisfactory', i.e. having acceptable risk, acceptable in the view of the user, of accepting or rejecting product at a certain level non-conforming or average level. On the other hand NZ is recommending the standard statistical approach where sampling plans are designed from pre-specified levels of acceptable producer's and consumer's risks – if we know these risks why not use these to design plans directly rather than using trial and error.

Selecting the number of samples n based on perceived risk is not scientific and setting c=0 is somewhat cosmetic - FAO/WHO 2016 makes the point several times that this does not assure consumers that accepted lots are free of non-conforming product. This methodology provides only limited control of consumer's risk as essentially the power of attributes plans to detect non-conforming product depends on the sample size.

The mathematical (and statistical) approach is that we have two unknowns, the number of samples n and the acceptance number c, and by choosing two points on the operating characteristic, traditionally a point representing producer's risk and a point representing consumer's risk, we obtain two simultaneous equations in (n, c) that can then be solved to determine the sampling plan.

Finally in this section we note that attributes plans employing low numbers of samples, such as the (n=5, c=0) plans in FAO/WHO 2016 are, in principle, poor in their ability to discriminate between good and poor quality lots. The following shows the Operating Characteristic for this plan showing, as an example, a 33% chance approximately of accepting lots in which 20% of the product is non-conforming. This would obviously be unsatisfactory for a food safety application indicating that a mitigation is provided from the large offset between the limits m and/or M and the levels at which food becomes unsafe to consume. This mitigation is not mentioned in FAO/WHO 2016 – the document possibly provides a misleading impression about how sampling plans are designed.



Measurement Error

It is not correct to assume that methods are chosen by Codex on the basis of their fitness for purpose, so that measurement error adjustment is not required. Our view is that the guidelines on measurement uncertainty do not provide the guidance needed (and in terms of assessment relate to conformity assessment rather than sampling inspection). The agreed prioritization list provides a mandate to include this – along with bulk materials measurement error adjustment was seen as one of the major omissions in the current GL50.

Codex does not consider fitness for purpose of an analytical (or microbiological) method as part of method endorsement. Fitness for purpose of a test method is not only a property of the analytical method itself but depends also on the purpose, how test results are used, in particular in acceptance criteria associated with sampling plans. However no such evaluation is carried out but, in any case, the significance of measurement error in sampling depends on the 'error variance ratio', the ratio between the measurement error and true product standard deviations (see section 8.2 of the e-book).

Control of Producer's Risk (PR)

As above, NZ has based the revision of GL50 on the standard statistical approach in which both producer's risks and consumer's risks can be controlled to appropriate levels.

The suggestion that PR need not be considered as there is a high probability of acceptance of good quality product is not necessarily valid – it also depends on the level non-conforming at which the probability of acceptance is suitably high. Indeed the Acceptance Quality Level (AQL), now called the Producer's Risk Point (PRP) is often considered as the quality level that is acceptable as a process average.

It should also be remembered that not all parameters relate to food safety, so in the interests of fairness (a principle adopted by Codex), the design of sampling plans used for commodity defects and compositional standards of identity and other parameters should take producer's risks into account. Fairness and consideration of producer's risk might also apply to food safety parameters as in general consideration of producer's risk as well as consumer's risk is necessary so that sampling plans achieve suitable discrimination between good and poor quality product, which is not provided by c=0 plans.

However decisions on acceptable levels of risk are matters to be decided by the Codex Commodity Committees; the purpose of the GL50 Guidelines is only to provide a toolbox to assist commodity committees and others with the design of plans appropriate to each provision in Codex standards. With this in mind NZ considers that the GL50 guidelines should be written in the most general way, based on sound statistical principles, considering both producer's and consumer's risk.

Appendix V

List of participants

Chair
Susan Morris
Ministry for Primary Industries – NZ

Roger Kissling
Fonterra - NZ

Co-chair
Gregory Noonan
US FDA

Argentina
Punto Focal Codex
Agroindustry Secretariat
Argentina

Australia
Richard Coghlan
National Measurement Institute - Australia

Karina Budd
Department of Agriculture and Water Resources

Brazil
Ligia Lindner Schreiner
ANVISA

Ana Claudia Marquim Firmo de Araújo
Carolina Araújo Viera

Canada
Thea Rawn

China
Stephen CHUNG Wai-cheung*
Centre for Food Safety

Egypt
Mariam Barsoum Onsy
Egyptian Organization for Standardization & Quality

France
Jean-Luc Deborde

Germany
Katrin Franks
BVL

Dorina Bodi
German Federal Institute for Risk Assessment

Vásárhelyi Adrienn*
National Food Chain Safety Office

Hanow Kamiran Aron

Attila Hucker
MTKI Kft.

Attila Nagy DVM
National Food Chain Safety Office (NÉBIH)

Kasza Viktor
National Food Chain Safety Office
Zsuzsa Farkas, PhD
National Food Chain Safety Office (NÉBIH)

India
National Codex Contact Point of India
Anoop A Krishnan

Rajesh R Nair
National Dairy Development Board

Iran
Samaneh Eghtedari*
ISIRI

Japan
Codex Japan

Hidetaka Kobayashi*
Ministry of Agriculture, Forestry and Fisheries

Takahiro Watanabe

Takanori Ukena

Mexico
Tania Daniela Fosado Soriano
Secretaría de Economía

The Netherland
Sanneli Kingma
MVO- the Dutch Oils and Fats Industry
Netherlands

Nigeria
Ibitayo Femi James
Federal Ministry of Agriculture and Rural Development

Norway
Norwegian Codex Contact Point
Norwegian Food Safety Authority

Stig Valdersnes
Institute of Marine Research

Perú
Juan Carlos Huiza Trujillo
Digesa(direccion general de salud ambiental)

Gloria Castillo Vargas
Instituto Nacional de Calidad - INACAL

Thailand
Rungrassamee Mahakhaphong
ACFS

Codex Thailand
Chanchai Jaengsawang

Republic of Korea
Yoo Min
Ministry of Food and Drug Safety(MFDS)

United Kingdom
Bhvana Parmar

Uruguay
Pedro Friedrich
Laboratorio Tecnologico del Uruguay

USA
Clarke Beaudry
US Food and Drug Administration

John Bowers
US Food and Drug Administration

Timothy Norden
United States Department of Agriculture

ICGMA
Nancy Wilkins

NMKL
Nina Skall Nielsen