E

CODEX ALIMENTARIUS COMMISSION

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 6.1 **CX/MAS 25/44/8**

April 2025

JOINT FAO/WHO FOOD STANDARDS PROGRAMME

CODEX COMMITTEE ON METHODS OF ANALYSIS AND SAMPLING

44th Session

Virtual

5 – 8 May and 14 May 2025

INFORMATION DOCUMENT: *GENERAL GUIDELINES ON SAMPLING* (CXG 50–2004) – E-BOOK WITH SAMPLING PLANS APPLICATIONS

(Prepared by the EWG led by New Zealand and co-chaired by Germany)

Codex Members and Observers wishing to submit comments on the recommendations in this document should do so as instructed in CL 2025/18-MAS available on the Codex webpage/Circular Letters: https://www.fao.org/fao-who-codexalimentarius/resources/circular-letters/en/

Introduction

1. The 42nd session of the Codex Committee on Methods of Analysis and Sampling (CCMAS42, 2023) agreed to:

   a. forward the revised *General guidelines on sampling* (CXG 50-2004) to CAC46 (2023) for adoption at Step 8; and

   b. re-establish the electronic working group (EWG), chaired by New Zealand and co-chaired by Germany, to continue working on the information document namely the e-book with the sampling plans applications for consideration by CCMAS43 (2024).[1]

2. CCMAS43 was updated on the progress of this work and following discussions[2], agreed to:

   a. continue developing the information document; and

   b. establish an EWG, chaired by New Zealand and co-chaired by Germany, to continue developing the information document taking into account the discussion at CCMAS43 and all written comments submitted to the meeting for comments and consideration by CCMAS44 (2025).

3. It was further noted that information documents do not go through the Step Procedure, as opposed to Codex standards or guidelines, and as such, information documents are not adopted by the Codex Alimentarius Commission but remain available for internal use by the Committee or for public consultation on the Codex webpage following agreement by the Committee. It was also noted that information documents could be considered as living documents subject to revisions when necessary. Therefore, further updates to information documents could be made in future if required. Information documents are published on the CCMAS webpage.[3]

4. This paper focuses on the information document as per the terms of reference for the EWG indicated in paragraph 2. The other term of reference related to the review of sampling plans is addressed in CX/MAS 25/44/9.

5. This working document contains the following:

   a. **A summary of the process since CCMAS43 including the EWG process and recommendations to CCMAS44**

   b. **Appendix I: Information document for the *General guidelines on sampling* (CXG 50-2004)**. The information document represents the work set out by CCMAS42 and CCMAS43. In Part 1 it includes practical examples of sampling plans and information to support design of sampling plans for isolated

[1] REP23/MAS, paragraph 81 (i-ii)
[2] REP24/MAS, paragraphs 33, 35-37
[3] REP23/MAS paragraph 71

lots and sampling plan apps and in Part 2 it includes more detail on statistical information, measurement uncertainty and Bayesian approaches.

c. Appendix II: Summary of the EWG consultation on the information document for the *General guidelines on sampling* (CXG 50-2004)

d. **Appendix III: EWG participants**

EWG registration and consultation

6. The EWG registration on the information document was sent out in September 2024 using the online Codex forum. The EWG Chair and co-chair worked closely and provided an update to the EWG later in the year. There were 18 member countries registered. The list of participants is contained in Appendix III.

7. The EWG Chair and co-chair worked closely to review and update the information document based on the discussion and comments provided to CCMAS43. A summary of the changes has been set out in Appendix II.

8. The latest version of the sampling plan app "App1" has been included for consideration by the EWG. It was clarified that this app, for homogeneous lots, simplified the process for evaluating and designing sampling plans. A PowerPoint presentation had been included for consideration by the EWG that showed the process for designing sampling plans using App1. It was suggested that more resources, such as video clips, may be provided at a later date.

9. The consultation with the EWG was sent out in December 2024 using the online Codex forum. This included the CXG 50-2004 information document, a link to App1 and a presentation on using CXG 50-2004 and App1 for the development of sampling plans.

10. The EWG was requested to review and provide comments on:

- the content of the document; and
- the proposal to finalise the information document, App1 and related presentation/tutorial at CCMAS44.

11. Questions were answered through the forum or discussed through virtual meetings.

12. The EWG consultation closed in January 2025. Detailed submissions were received from Australia, Canada, and Japan and the EWG Chair's and co-chair's responses are contained in Appendix II. In summary, the EWG Chair and co-chair agreed with most of the technical comments received and made changes accordingly. A response has been provided to a concern regarding the inclusion of Bayesian plans. There was consensus within the EWG that the current content is acceptable, and to recommend CCMAS44 to agree on the publication of the information document.

13. A need for sampling plans employing low sample size was expressed in CCMAS42. Bayesian plans could allow a considerable reduction in sample size and were thus discussed during the workshop held in CCMAS43. Many delegations expressed a strong interest in Bayesian plans following this workshop and it is proposed that Bayesian plans be discussed during the CCMAS44 plenary. It should be noted that a technical report on the application of Bayesian methods to acceptance sampling will soon be published by the International Organization for Standardization Technical Committee 69 Sub-committee 5 Working group 10 (ISO TC 69 SC 5 WG 10).

Conclusion

14. Following consultation with the EWG, the information document has been updated. The intent is to conclude the work at CCMAS44 and publish a technical document that provides background information and apps for developing sampling plans for international trade of food commodities.

Recommendation

15. CCMAS44 is invited to:

i. consider the information document and agree to its publication on the CCMAS webpage;

ii. note that as other Apps are developed, they will be forwarded for CCMAS' consideration for inclusion to the list of Apps in the information document; and

iii. note that other supporting resources e.g. webinars will be made available on the CCMAS webpage.

Appendix I

Information document for the *General guidelines on sampling* (CXG 50-2004)

(For comments through CL 2025/18-MAS)

Contents

1 Introduction 5

2 Design of sampling plans 7

2.1 Principles behind the design of sampling plans 7

2.1.1 Producers and consumers 7

2.1.2 Producer's and consumer's risks 7

2.2 Design of sampling plans 8

2.2.1 Overview of the design process 8

2.2.2 Process for the design of sampling plans 9

2.2.3 Use of the same sampling plan by the producer and the consumer 11

2.2.4 CCMAS endorsement of sampling plans 12

2.3 Apps for the design and evaluation of sampling plans 13

2.3.1 Description of apps 13

3 Case studies (examples for specific scenarios) 13

3.1 Examples on the Use of Attributes plans 14

3.1.1 Example: Attributes plan with $\boldsymbol{c > 0}$ 14

3.1.2 Example: Attributes plan with c=0 16

3.1.3 Example: Attributes plans based on AQL of 6.5% 19

3.2 Examples for Variables plans 21

3.2.1 Example: Variables plan with negligible MU 21

3.2.2 Example: Variables plan with non-negligible MU with no laboratory bias 25

3.2.3 Example: Variables plan with non-negligible MU with laboratory bias 28

3.2.4 Example: Fractional Nonconformance plans 29

3.3 Lots consisting of bulk materials 30

3.3.1 Example: Aflatoxin sampling plans due to Whitaker (2006) et al. 30

3.3.2 Example: Plans based on the beta distribution 31

3.4 Other sampling plans 34

3.4.1 Example: ISO sampling plans – 6.5% AQL 34

3.4.2 *Ad hoc* plans 35

4 Background to Acceptance Sampling Plans 41

4.1 Attributes plans 41

4.2 Variables plans 41

4.2.1 Basis for calculations in App1 42

4.3 Calculating measurement uncertainty from precision estimates 43

4.4 Combined attributes-variables plans 43

4.5 Multi-stage plans 44

4.6 Lots consisting of bulk materials 45

4.6.1 Example: Variables Plans for homogeneous lots—negligible MU 45
4.6.2 Example: Variables plan with non-negligible MU with no laboratory bias 46
4.6.3 Example: Variables plan with non-negligible MU with laboratory bias 46
4.6.4 ISO 10725 46
4.6.5 Aflatoxin sampling plans 46
4.6.6 General plans for assessment against minimum or maximum levels 53
4.7 Plans for microbiological assessment 53
4.7.1 3-class attribute plans 53
5 Statistical appendix 55
5.1 Background for the main (attributes & variables) sampling plans 55
5.1.1 Calculating acceptance probabilities – attributes plans 55
5.1.2 Derivation of formulas for variables plans 56
5.1.3 Within-item variability 58
5.2 Understanding ISO plans 59
5.2.1 Attributes plans constructed in terms of unity value 59
5.2.2 Variables plans constructed in terms of the producer's risk 61
5.3 Acceptance sampling versus conformity assessment 61
5.3.1 Definitions 62
5.3.2 Positions in ISO standards 62
5.3.3 Positions in the literature 62
5.3.4 Discussion 63
5.4 The role of measurement uncertainty in acceptance sampling 63
5.4.1 Specifying the measurand 63
5.4.2 Sampling uncertainty versus acceptance sampling (terminological clarification) 64
5.4.3 Effects of analytical and sampling uncertainty in acceptance sampling 65
5.5 Bayesian plans 68
5.5.1 Prior distributions 68
5.5.2 Conformance probability approach 69
5.5.3 Utility approach 71
5.5.4 Bayesian plans: glossary of terms 83
5.6 References 84

1 Introduction

The purpose of this document is to provide further information on the sampling plans referred to in CXG 50, including background and examples for each of the main types of sampling plan, as well as additional information on other sampling plans including Bayesian plans. A link to the app1 for the design and evaluation of these sampling plans is included. Links for new apps will be included when they become available.

The document consists of two parts:

Part 1 consists of Section 2 and Section 3 and contains general information relating to the design of sampling plans, including examples:

Section 2.1 deals with principles behind the "classical" approach to sampling plans based on specification of producer's and consumer's risks, to allow for any level nonconforming in a lot.

Section 2.2 contains information about the design process, including suggestions on the use of pre-defined sampling plans, such as ISO plans, as well as specifications of allowable risks, as a starting point.

Section 2.3 describes the different apps that were provided with the original package and provides a link to an on-line app for the design of attribute and variables sampling plans with and without measurement uncertainty.

Section 3 contains case studies showing the main types of sampling plans mentioned in CXG 50, including some where measurement uncertainty is non-negligible.

Sections 3.1 and 3.2 deal with various types of attributes and variables plans, including an explanation of the basis underlying the plans in the ISO2859 and ISO3951 standards.

Section 3.3 discusses sampling for bulk materials with a particular focus on the plans for mycotoxins described in the *General standard for contaminants and toxins in food and feed* (CXS 193-1995).

Section 3.4 covers other sampling plans. Examples include attribute sampling plans with an AQL of 6.5 taken from ISO, and ad hoc plans using small numbers of samples formed into composite samples for testing.

Part 2 consists of Section 4 and Section 5 and contains more background on sampling plans including a statistical appendix:

Section 4 includes statistical derivation of attributes and variables plans when measurement uncertainty is negligible and references the sampling plans in Part 1.

Section 4.3 discusses measurement uncertainty and its role in acceptance sampling.

Section 4.6.5 covers the basis for sampling plans for Mycotoxins derived by Whitaker et al. that are special cases of plans for bulk materials and outlines the statistical complexity for the future development of sampling plans for bulk materials.

Section 4.7 contains information about other sampling plans, including 3-class attribute plans, used for microbiological assessments.

Section 5.5.2 discusses Bayesian plans based on a risk-based approach. More specifically, the plans are based on the concepts of specific consumer risk and conformance probability from JCGM 106. An overview of Bayesian risks is also provided. This approach was developed in ISO TC 69 SC 5 WG 10 and is described in a technical report as well as in a separate publication.

Section 5.5.3 discusses Bayesian plans based on a utility-based approach. Standard plans are provided for the practitioner. This approach was developed in ISO TC 69 SC 5 WG 10 and is described in a technical report as well as in a separate publication.

Note:

Some Excel formulas in the text (and in the Excel file provided) use the English style, with decimal points and comma separators.

PART ONE

General information relating to the design of sampling plans

Examples and case studies

2 Design of sampling plans

2.1 Principles behind the design of sampling plans

2.1.1 Producers and consumers

Depending on the nature of the transaction, a producer could include either:

- A manufacturer, supplier or seller of a food product or ingredients, or
- A regulatory agency providing assurance for exported product to an importing country agency.

and a consumer could include:

- A customer purchasing the food product or ingredient for the manufacture of other food products, or
- An importing country regulatory agency seeking to provide assurance to the individual consumers living in that country, or
- An exporting country regulatory agency providing official assurance to an importing country agency acting on behalf of the importing country, or
- An individual purchasing a food product, although individuals would not normally be able to carry out lot-based inspections of foods, or
- A manufacturer purchasing ingredients for the production of a food product.

2.1.2 Producer's and consumer's risks

Acceptance sampling plans always carry intrinsic risks that a lot of poor quality will be incorrectly accepted or that a lot of good quality will be incorrectly rejected. These two risks are generally referred to as the consumer's risk and the producer's risk, respectively.

However, by following statistical principles sampling plans can be designed to control these risks to allowable levels. This is achieved by specifying a particular producer's risk quality level, the PRQ, and a particular consumer's risk quality level, the CRQ, along with a corresponding producer's risk (PR), the probability of rejecting a lot with quality level equal to the PRQ, and a consumer's risk (CR), the probability of accepting a lot with quality level is equal to the CRQ, respectively. Once these four parameters, the (PRQ, CRQ, PR and CR), are specified the sampling plan is uniquely determined and the probability of acceptance and therefore the producer's and consumer's risks at any quality level can be calculated.

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

The *Principles for the Use of Sampling and Testing in International Food Trade* (CXG 83-2013) recommends that ideally, producers and consumers should agree on a sampling plan prior to its use. However, direct collaboration/negotiation between producers and consumers on the sampling plan to be used or the way in which it will be used might not always be possible.

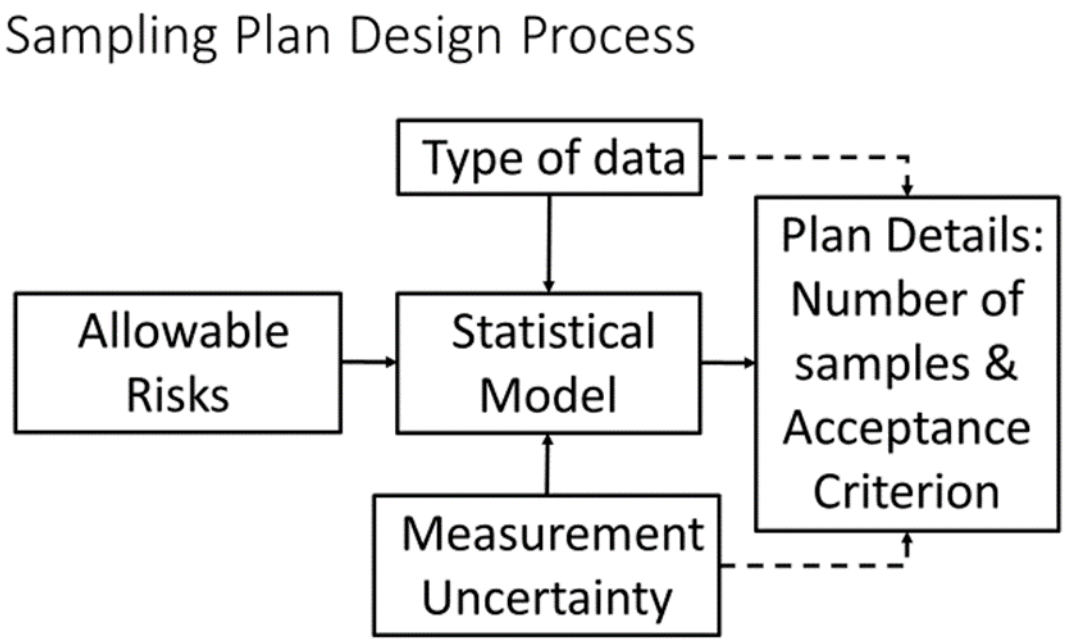
This is the traditional approach to the design of sampling plans. In this approach, risks are calculated on the basis of a known quality level, e.g. "if a lot contained $X\%$ of nonconforming items, then the probability of acceptance would be P."

However, it is also possible to introduce a Bayesian framework which allows other definitions of risk, involving the use of a prior distribution based e.g. on past inspections of a characteristic in a product. Bayesian methods can potentially allow the user to achieve a considerable reduction in sample size.

2.2 Design of sampling plans

2.2.1 Overview of the design process

Figure 1 Sampling Plan Design Process



This diagram shows a high-level view of the design process, showing the fundamental inputs to the design as reflected in the workflow in the *General Guidelines on Sampling* (CXG 50-2004).

Specification of the allowable risks is a key input, the producer's risk and the consumer's risk might both be specified or only one of those risks might be specified. In the ISO plans only the consumer's risk would be specified for the inspection of isolated lots of incoming goods whereas plans based on only the producer's risk might be used in the context of a long-term supply contact between a manufacturer and a customer.

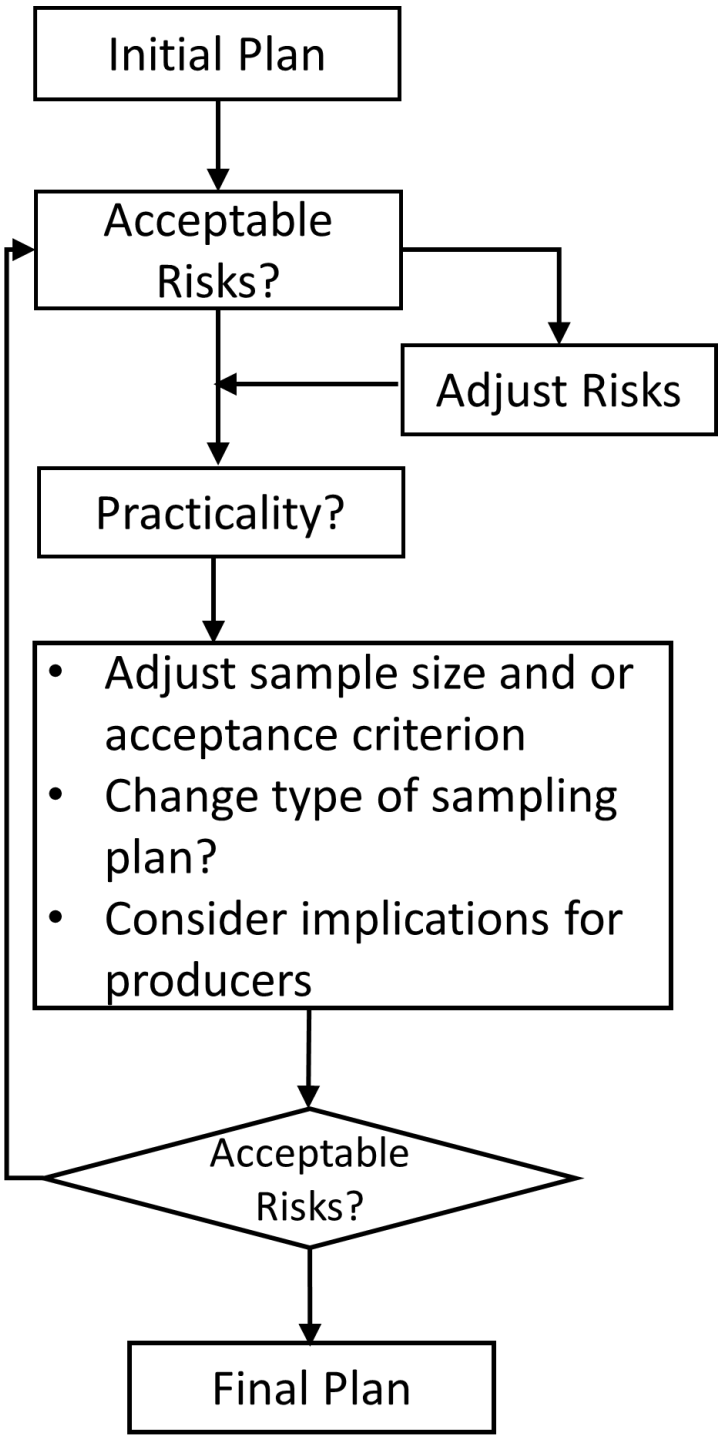
The type of data and non-negligible measurement uncertainty will determine the statistical model that is used to work out the details of the sampling plan. For example, if one has attributes data the statistical model is based on the binomial distribution, whereas if the one has variables data the model could be based on the normal distribution or in the case of a compositional proportion the beta distribution, or possibly some other distribution (not dealt with in CXG 50-2004).

The type of data and the presence of non-negligible measurement uncertainty also determine the form of the acceptance criterion; in the simplest case where measurement uncertainty is negligible an attributes plans is specified by the number of samples n and the acceptance number c, but a variables plan is specified by the number of samples n and the acceptability constant k. The form of the acceptance criterion could be more complex for variables plans where measurement uncertainty is non-negligible.

2.2.2 Process for the design of sampling plans

This diagram shows a process that might be used to design a sampling plan.

Figure 2 Process for Design of Sampling Plans



Step 1 Select an Initial Plan as a starting point:

The design of sampling plans requires specification of the allowable consumer's and producer's risks following which sampling plans are designed using statistical methodology.

However, it is often difficult for designers of plans to decide on appropriate levels of allowable risks [that should possibly be decided jointly by both the producer and the consumer], so that the following process is suggested as a way of proceeding.

The starting point could be a plan from either ISO 2859-1 or ISO 3951-1, depending on whether one has attributes or variables data. In CXG 50-2004 Appendix 1 plans with selected PRQ and CRQ levels were used as starting points.

The following plans based on these ISO standards could be used as starting points for the design of plans. In these plans the producer's risk (PR) is 5% and the consumer's risk (CR) is 10%.

Table: Attributes plans from ISO 2859-1 for PRQ = 6.5%

PRQ	c	n	CRQ
6.5%	0	2	68.4%
6.5%	0	3	53.6%
6.5%	1	5	58.4%
6.5%	1	8	40.6%
6.5%	2	13	36.0%
6.5%	3	20	30.4%
6.5%	5	32	27.1%
6.5%	7	50	22.4%

Table: Attributes plans from ISO 2859-1 for PRQ = 1.5%

PRQ	c	n	CRQ
1.5%	0	8	25.0%
1.5%	0	13	16.2%
1.5%	1	20	18.1%
1.5%	1	32	11.6%
1.5%	2	50	10.3%

Table: Variables plans from ISO 3951-1 for PRQ = 2.5% (σ-method)

PRQ	k	n	CRQ
2.5%	1.115	3	35.4%
2.5%	1.240	6	23.7%
2.5%	1.419	8	16.7%
2.5%	1.366	8	18.1%
2.5%	1.370	12	15.9%
2.5%	1.439	16	13.2%
2.5%	1.456	21	12.0%
2.5%	1.533	29	9.76%
2.5%	1.606	42	7.95%

Step 2 Examine the OC curve:

A plan taken directly from a standard might not necessarily be suitable for a particular application, it might be too stringent or not stringent enough. Users need to consider whether the acceptance probability, the proportion of lots that will be accepted in the longer term by the plan, is acceptable at various levels nonconforming that might occur, for example:

- Is the probability of acceptance of lots containing 10% (or 5% or 20%) of nonconforming product acceptable?

Step 3 Adjust the Risks to the desired levels.

Step 4 Consider the practicality of the proposed sampling plan:

A key consideration is the number of samples that will need to be taken and tested for each lot that is inspected, and the expected number of lots that will be inspected in any year.

In general sample numbers can be economized by:

- Increasing the CRQ or decreasing the PRQ, or both
- Increasing the producer's and/or the consumer's risks PR and CR (both might be increased)
- Use of indifference quality plans for commodity characteristics (refer CXG 50-2004 3.2.2)
- Requiring a lower stringency at the individual lot level in favour of assurance in the longer term.

Specific ways to economize sample numbers include:

- Not assessing compliance of the lot on an individual basis but treating the product in the lot as a bulk material and inspecting the lot as a whole rather than focusing on compliance of the individual items
- Use of variables plans instead of attributes plans where applicable.
- Use of known lot standard deviations, if they are known.
- Use of plans based on the beta distribution for compositional characteristics.
- Use of offsets (including offsets to allow for non-negligible laboratory bias) (refer CXG 50-2004 3.2.3)
- Bayesian plans may be another way sample numbers might be reduced.

Step 5 Examine the OC curve to check that the risks are acceptable.

Step 6 Adopt the sampling plan or return to step 3 and repeat the process.

Measurement uncertainty should also be allowed for if it is non-negligible.

The resulting plan should be reviewed to ensure it will meet users' expectations and, where appropriate to ensure that it is fair to producers – choice of a suitable sampling plan should focus on the control of risks and the total cost, especially the costs of incorrect acceptance and incorrect rejection of lots, rather than just the cost of sampling and testing.

However, if multiple characteristics are inspected when assessing compliance to a standard, there is a risk that the producer's risk of inappropriate rejection will increase with the number of characteristics inspected. This risk can be mitigated by reducing the producer's risk on each of the individual sampling plans, so that the overall producer's risk is not excessive. This measure should be applied only among 'similar' characteristics, such as compositional characteristics.

In the interests of fairness, designers of plans should also take account of the measures that the producer may have to take to ensure compliance, remembering that in Codex, CXG 50-2004 will be applied mainly to compositional characteristics and commodity defects (refer CXG 50-2004 3.2.2 Fairness).

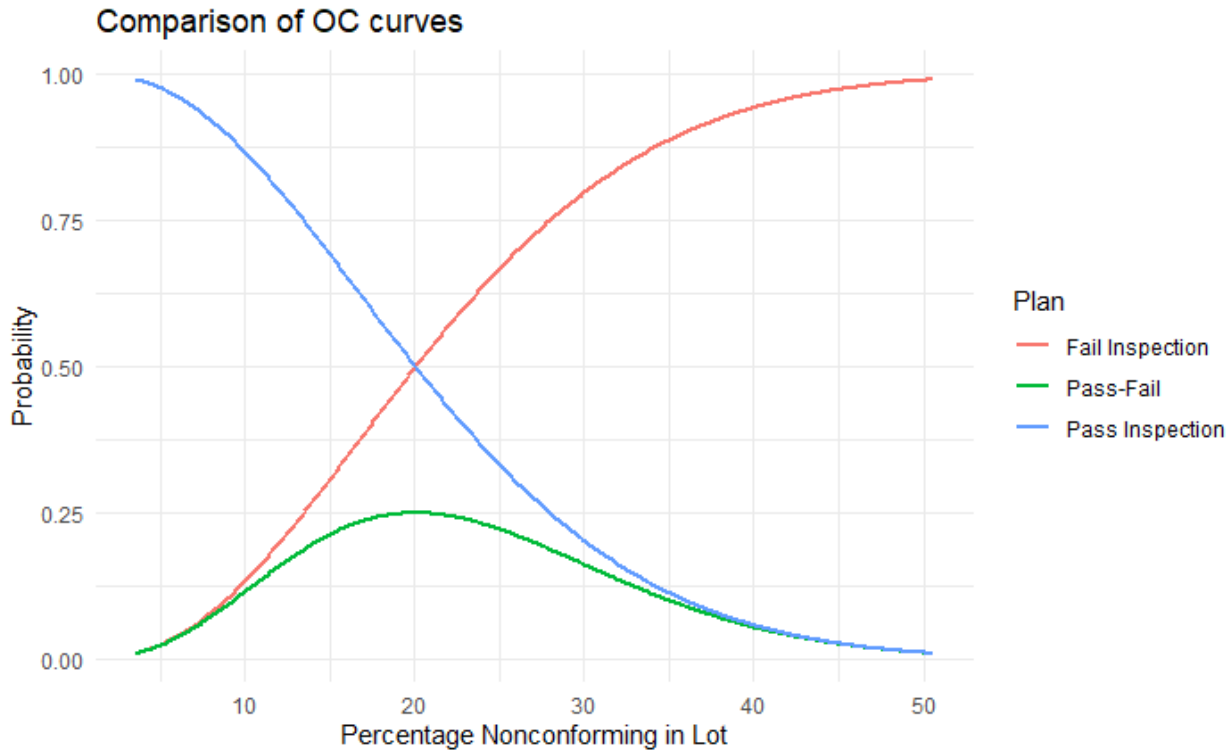
2.2.3 Use of the same sampling plan by the producer and the consumer

It is not always appropriate for a producer to use the same sampling plan as that used by a consumer, particularly if the consumer is also going to inspect the lot.

The following plot shows the probability of acceptance of a lot for a given plan [in **blue**] for the inspection and the probability of rejection using the same plan [in **red**]. The **green** line shows the probability that a lot will be accepted in the producer's inspection but then be rejected in the consumer's inspection. The maximum probability of a lot being accepted on an initial inspection and then being rejected on a subsequent inspection using the same plan is 25% that occurs when the probability of acceptance by the plan is 50%.

This shows that producers are potentially disadvantaged in adversarial sampling situations, where both the producer and the consumer both test the product; that should be considered in the design of plans to ensure fairness.

Figure 3 OC curves In Adversarial Sampling



The OC curves on the plot have been calculated as follows.

- The probability of acceptance (the blue line) is the usual calculation, described in section 3.1 for attributes plans and section 3.2 for variables plans for cases when the measurement uncertainty is negligible. If measurement uncertainty is non-negligible it is necessary to use the app.
- The probability of rejection (red line) is easily calculated since if a lot is either accepted or rejected:

$$Probability\ of\ rejection\ =\ 1\ -\ Probability\ of\ acceptance$$

- The probability of acceptance of a lot and subsequent rejection upon reinspection (the green line) is the product of the two probabilities:

$$Probability\ (Pass - Fail)\ =\ Probability\ of\ acceptance\ \times\ Probability\ of\ rejection$$

Obvious mitigations include:

- the producer using a plan that reduces the risk of rejection if the lot is inspected by a consumer,
- for the producer to operate at a quality level that ensures a lower rate of rejection if the lot is inspected by the consumer,
- consumers might rely on the producer's inspections rather than inspect the lots themselves.

2.2.4 CCMAS endorsement of sampling plans

Commodity Committees might propose sampling plans for provisions, or they might propose outcomes for sampling plans in terms of the maximum allowable producer's and consumer's risks for the inspection of a provision. This means that there is often more than one option for the sampling plan that could be used. The latter approach, specifying the outcome, is needed when measurement uncertainty is non-negligible as the plan will depend on the lot standard deviation, that will vary among producers.

The OIML International Recommendation R087[4] for the average quantity system in prepackaged is an example where maximum allowable producer's and consumer's risks are specified:

Producer's Risk

The probability of rejecting a lot whose true mean weight is equal to or exceeds the label quantity e.g. weight should be at most 0.5%.

[4] OIML Recommendation (OIML R 87): Quantity of product in prepackages, International Organisation for Legal Metrology, Paris (2016)

Consumer's Risk

The probability of accepting a lot whose true mean weight is less than the label quantity e.g. weight by more than a specified amount (not provided here) should be at most 10%.

These risk specifications are used in two ways to design plans for the inspection of quantities by weights, in a variables-plan to check compliance of the average weight and in an attributes plan to check that there is not an excessive proportion of deficient packages, weighing less than the label weight by more than a certain amount, in the lot.

2.3 Apps for the design and evaluation of sampling plans

This section contains a brief description of each of the three apps provided with the CXG 50-2004 package, along with links from where they can be run. References to the relevant sections in CXG 50-2004 are provided and further information can be found in this document.

The apps all follow the same general format and have been designed to:

- *Plan Evaluation* of a specified sampling plan to calculate the probability of acceptance in terms of the percentage nonconforming[5] in a lot and to show the Operating Characteristic (OC) curve, and to calculate the producer's and consumer's quality levels PRQ and CRQ corresponding to specified producer's and consumer's risks (with default values of 5% and 10% respectively)
- *Plan Design*, the design of a sampling plan, working out the number of samples and the acceptance criterion, the acceptance number 'c' or the acceptability constant 'k', from specifications of the producer's and consumer's risk quality levels and their associated probabilities of rejection and acceptance, respectively.

2.3.1 Description of apps

App1 relates to the design and evaluation of attributes sampling plans and variables plans for normally distributed characteristics, including situations where measurement uncertainty is also normally distributed. The app can be used to examine and compare the producer's and consumer's risks and the OC curves for different sampling plan options.

In attributes plans, the app can evaluate a sampling plan specified by a sample size 'n' and an acceptance number 'c' or design a plan based on the specified values of the PRQ, CRQ, producer's risk, and consumer's risk. The OC curves and the producer's and consumer's risks are shown for both plans.

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 app also takes account of measurement uncertainty with the values of the components of measurement uncertainty and the lot standard deviation where these values have been specified.

This app has been deployed on the shinyapps.io server:

https://codex-testing.shinyapps.io/codex-testing-SamplingPlan/

Links for following apps will be included when they become available.

App2 relates to Fractional Nonconformance Plans and allows users to evaluate a specified plan, specified by the number of samples and the maximum allowable value of the individual sample FNC values for acceptance of the lot, or to design a sampling plan from specifications of consumer's and producer's risks.

See section 3.2.4 for further details.

App3 relates to plans for the assessment of lots for compositional characteristics. These plans are based on the beta distribution so that it is possible to apply the plan based on a single test of a composite sample formed from a specified number of increments. The app can evaluated a specified plan or design a plan based on specified risks. Refer to section 3.3.3 for details.

3 Case studies (examples for specific scenarios)

These examples follow the step-by-step design process described in CXG 50-2004 Appendix 1 and illustrate the use of the apps.

[5] The proportion of nonconforming items expressed as a percentage. Also called the proportion nonconforming in the ISO 2859 and 3951 standards.

3.1 Examples on the Use of Attributes plans

3.1.1 Example: Attributes plan with $c > 0$

In general, if both the consumer's and producer's risks are specified in the design of the plan, as might be appropriate for non-food safety characteristics such as commodity defects, it is unlikely that the acceptance numbers, the c values, will be zero. It should be noted that rather large sample sizes (and large acceptance numbers) might be needed for plans where the operating ratio (CRQ/PRQ) is small.

Example: Browning in Milk powder

- A customer found higher than usual levels of browning (discoloration) in a lot of WMP. The customer advised that they could accept the powder provided no more than 20% of the powder was nonconforming, as it would still be usable.
- The manufacturer wanted to control the risk of rejecting product that might still be usable, so that the product should be accepted [most of the time] if there was 10% nonconforming in the lot.

Key steps in the step-by-step design process

1. Attributes or variables data?

Excessive browning is an example of attributes data, samples are classified as either PASS or FAIL when compared against a reference powder.

2. Inspection error negligible or non-negligible?

It is assumed that inspection error is negligible in this example.

3. Set consumer's risk quality level (CRQ):

The customer has advised that the powder is still usable even if the level nonconforming was 20%, so that, for the purposes of the inspection, the consumer's risk quality level could be set at 20%.

4. Set producer's risk quality level (PRQ):

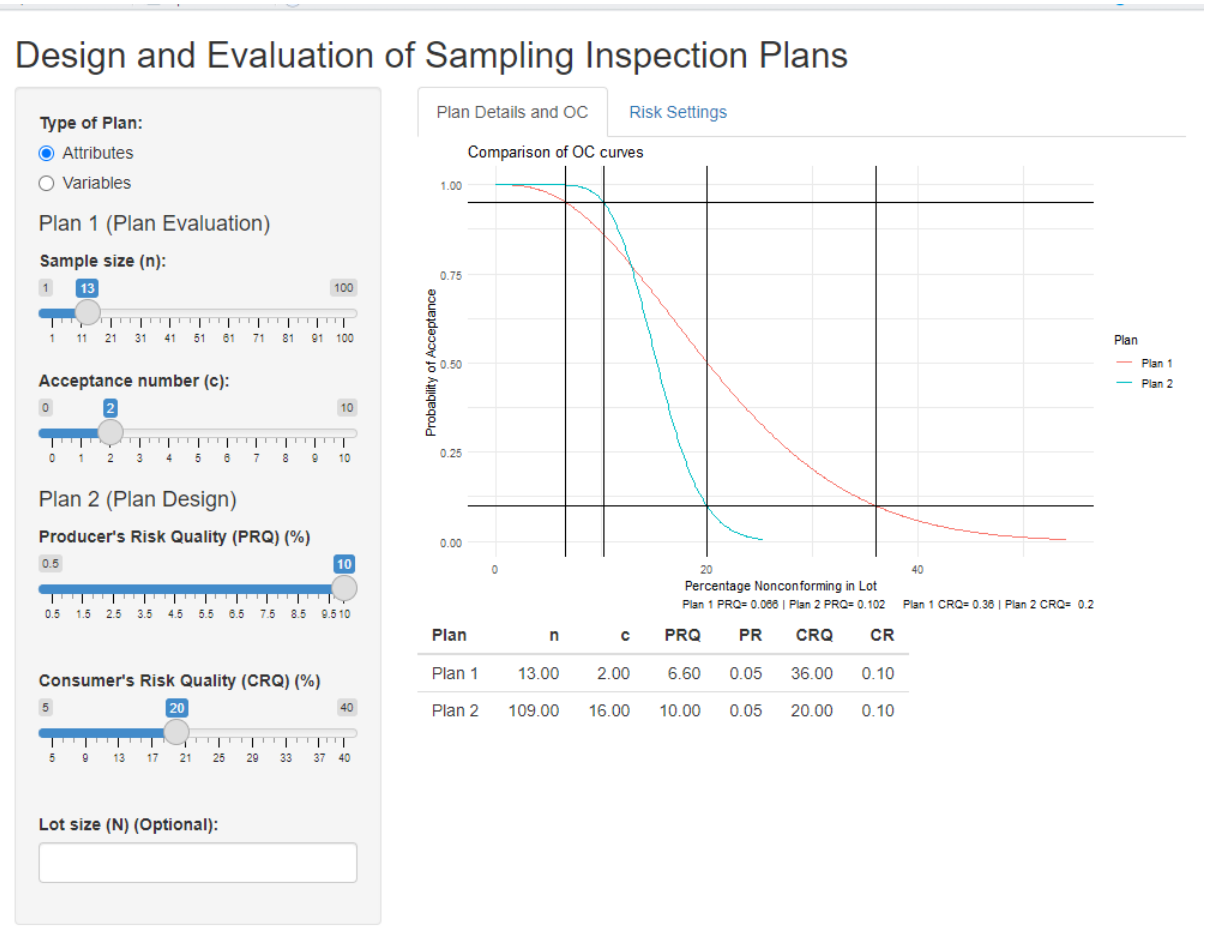
The PRQ can be set anywhere below the CRQ, noting that the smaller the operating ratio $\frac{CRQ}{PRQ}$, the larger the number of samples required to be taken.

Various values of the PRQ can be tried to assess the required sample size. Some possible options are:

PRQ	CRQ	n	c
5%	20%	38	4
10%	20%	109	16
15%	20%	500	88

For the purposes of this example, a producer's risk quality level of 10% was used; a PRQ of 5% is too stringent considering that the powder is still usable if up to 20% of the lot was nonconforming, and 500 samples is too many to take and evaluate.

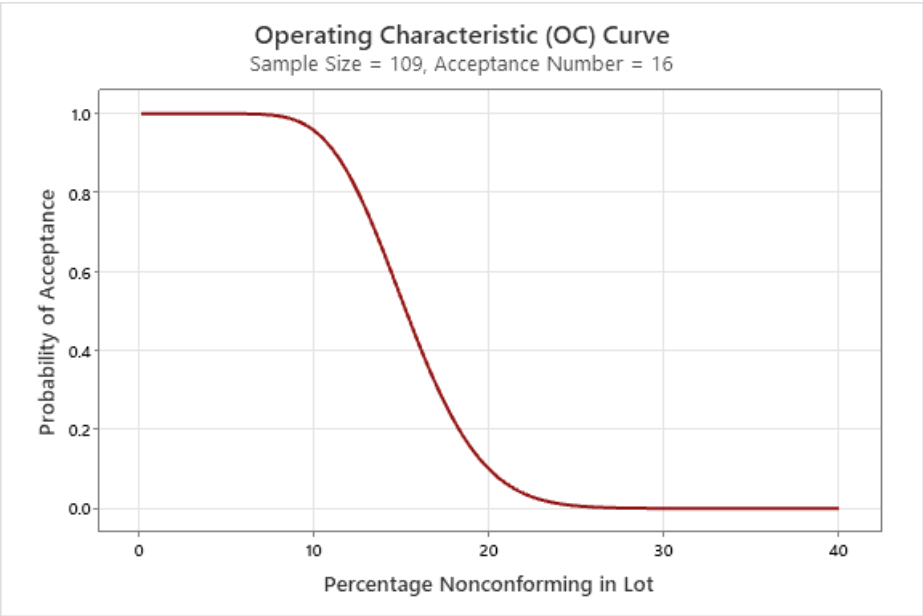
Figure 4 Design of plan for dispute resolution



Plan	n	c	PRQ	PR	CRQ	CR
Plan 1	13.00	2.00	6.60	0.05	36.00	0.10
Plan 2	109.00	16.00	10.00	0.05	20.00	0.10

The required sampling plan, required to control the consumer's and producer's risk to the specified levels is (n = 109, c=16) i.e. n=109 samples are taken, and the lot is accepted provided no more than 16 of those samples is nonconforming.

Figure 5 OC curve of selected plan - dispute resolution



The Operating Characteristic shows a 95% chance of accepting the lot when the level nonconforming is 10% (i.e. the PRQ), a 50% chance approximately of accepting the lot when the level nonconforming is 15% and a 10% chance of accepting the lot when the level nonconforming is 20% (the CRQ).

3.1.2 Example: Attributes plan with c=0

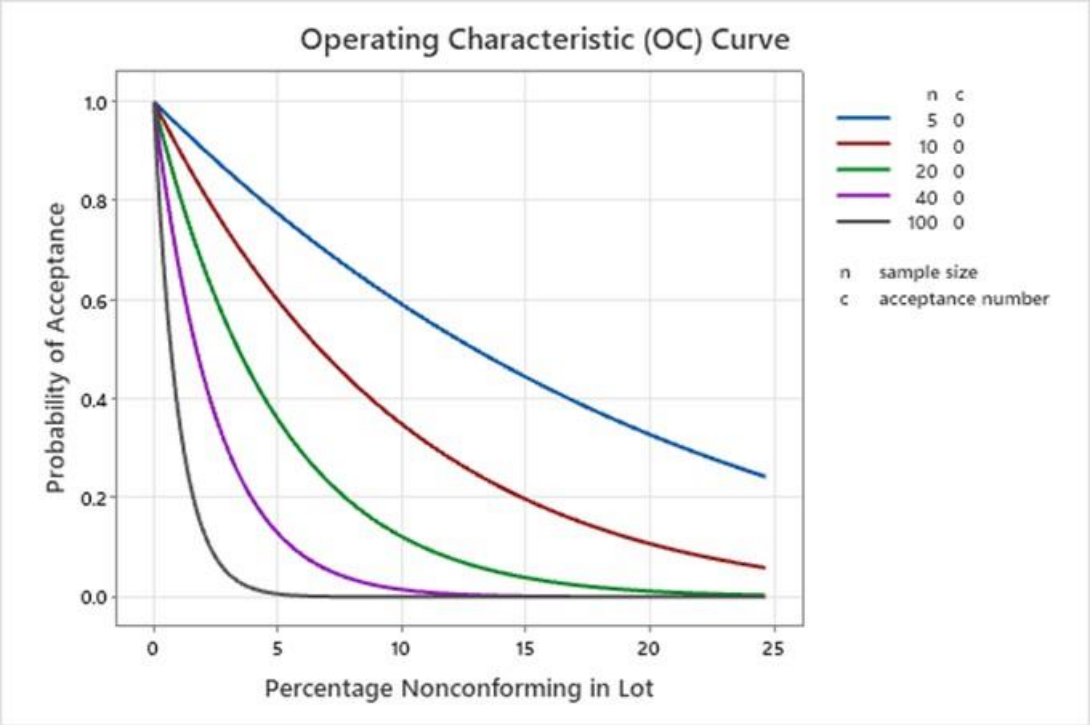
Refer to CXG 50-2004 4.2.5 Zero-acceptance number plans.

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

ZAN plans are commonly used, apparently based on the philosophy of zero defects and the perception that if c>0 then lots containing nonconforming product are being accepted.

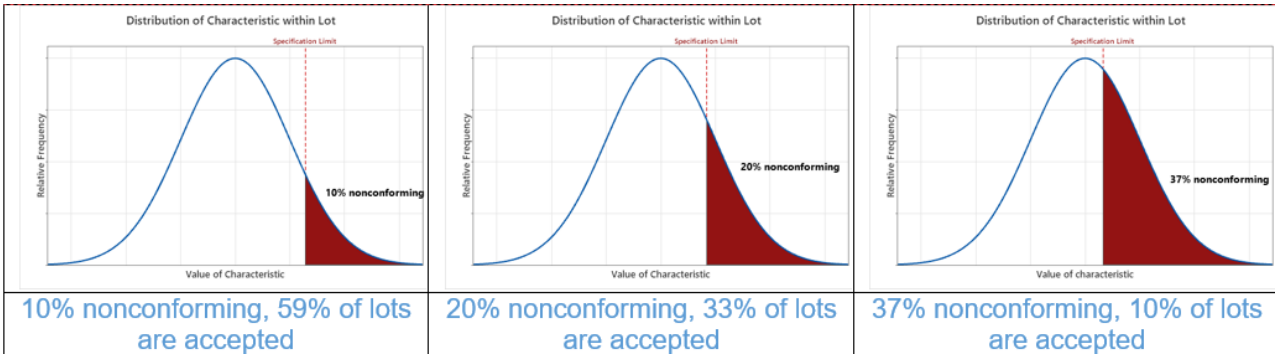
The following shows the Operating Characteristics for several ZAN plans:

Figure 6 OC curves for ZAN plans



However, ZAN plans cannot guarantee that lots that have passed inspection do not contain nonconforming items; no matter which plan is used there will always be a risk of accepting a lot containing some level of nonconforming product. The following shows the chance of accepting a lot for various levels nonconforming, using the (n=5, c=0) sampling plan.

Figure 7 Risks with n=5, c=0 plans



The number of samples n can be calculated directly using the formula:

$$CR = (1 - CRQ)^n \text{ or } n = \log(CR) / \log(1 - CRQ)$$

Typical outcomes that are often expressed in terms of the quality in the lot are as follows:

- If we select 60 items, taken at random from a 'lot', and find none of those items nonconforming, then we can claim with 95% confidence that no more than 5% of ALL the items in the lot are nonconforming.

- If we select 150 items, taken at random from a 'lot', and find none of those items nonconforming, then we can claim with 95% confidence that no more than 2% of ALL the items in the lot are nonconforming.
- If we select 300 items, taken at random from a 'lot', and find none of those items nonconforming, then we can claim with 95% confidence that no more than 1% of ALL the items in the lot are nonconforming.

If one or more nonconforming items has been found, it is still possible to make a statement about the quality level within the lot.

The Excel file pexact.xlsx included in the package can be used to calculate the 95% confidence intervals for the level nonconforming in a lot, or the total number of defects in a lot, for any number of nonconforming items or defects found in the sample, noting that an individual item may have more than a single defect:

- The sheet Binomial calculates 95% confidence intervals for the level of individual items conforming in the lot overall.
- The sheet Poisson calculates 95% confidence intervals for the number of defects in the lot overall. These limits can be converted to rates by dividing by the number of items examined.

Excel formulas for the calculation of the lower and upper 95% confidence limits for the two cases are given in Section 4.1.1.

Examples:

Binomial Case

If n=60 items were examined and c=2 of those 60 items were found nonconforming, then the estimated percentage of nonconforming items in the lot is 2/60 = 3.33% and, with 95% confidence, the level nonconforming in the lot lies between 0.41% and 11.53%.

Poisson Case

If 60 items were examined and 5 defects were found, with possibly more than one defect found on a single item, then with 95% confidence the number of defects in the lot lies between 1.62 and 11.67. Equivalently, these numbers could be expressed as defect rates of 1.62*100/60 = 2.7 to 11.67*100/60=19.45 defects per 100 items.

Example: Inspection for Foreign Matter

It is suspected that a lot is contaminated with foreign matter but that the contamination is not believed to be a food safety concern. However, it is known that the intended customer will not accept product in which any foreign matter has been found, so that a zero-acceptance number (ZAN) plan should be used.

Since the contamination is not a food-safety issue, it was decided to design a plan based on a consumer's risk (CR) of 5%, at a consumer's risk quality level (CRQ) of 3%.

The number of samples *n* can be calculated directly using the formula given above:

$CR = (1 - CRQ)^n$ or $n = \log(CR) / \log(1 - CRQ)$

Using the second formula we have:

$$n = \frac{\log(CR)}{\log(1 - CRQ)} = \frac{\log(0.05)}{\log(0.97)} = \frac{-2.9957}{-0.0305} = 98$$

Therefore, the lot is accepted provided none of the 98 samples inspected contains any foreign matter contamination. In practice one might use n=100 for simplicity.

3.1.2.1 (n=1, c=0) sampling plans

These sampling plans, often used by classifying variables data as attributes, are commonly used for inspection of contaminants and more widely, with or without allowance for MU.

For contaminants these plans rely on an assumption of homogeneity and possibly also on the usually large offsets between the decision limits used in those plans and food safety levels, so that allowance for MU may not be necessary. However, there is a considerable risk of incorrectly accepting a noncompliant lot if that lot was not homogeneous.

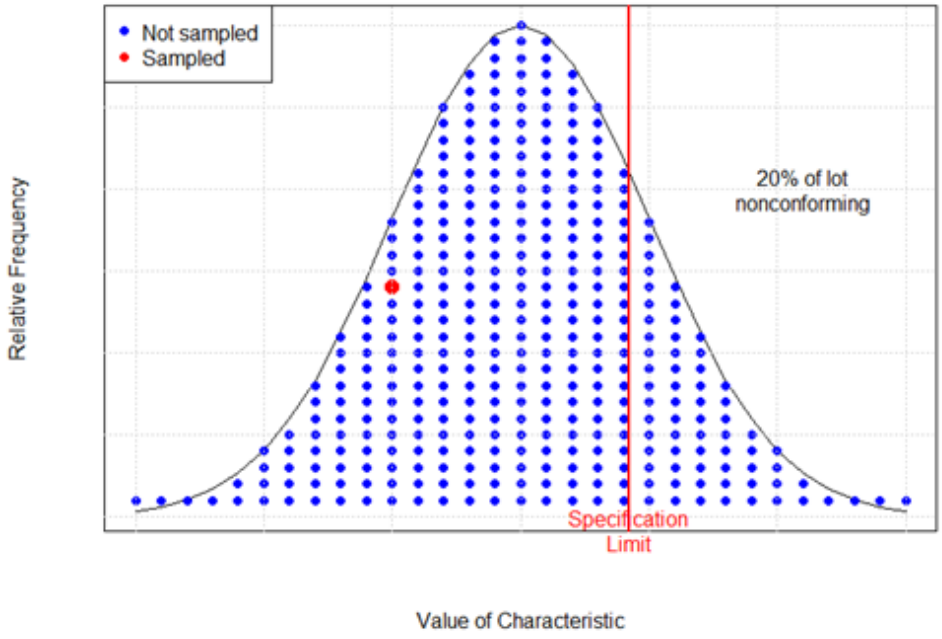
For other characteristics (n=1, c=0) plans are often used but might not provide the desired level of assurance to consumers. This might be due to:

- Attempting to minimize the cost of testing
- Ignoring of the principles of sampling

- Performing a spot check on compliance of the lot, that is potentially unfair to producers, without any intention of protecting consumer's risk; this practice often leads to complaints, there being a wide perception that if nonconforming items/samples are found on inspection then the entire lot is nonconforming

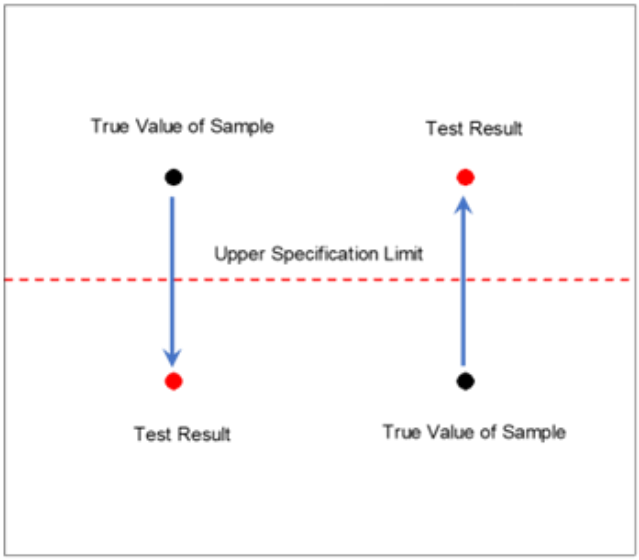
The fundamental problem with (n=1, c=0) plans is that decisions on acceptance or rejection are not necessarily related to the quality of the lots (Refer CXG 50-2004 Section 2.2). The following diagrams show the potential risks when using (n=1, c=0) sampling plans. The first plot shows the risk of making an incorrect decision due to sampling error, when 20% of the lot is nonconforming there is an 80% chance of not finding a nonconforming sample if only one sample is taken, assuming measurement uncertainty is negligible.

Figure 8 Risk due to sampling uncertainty for n=1 plans



The second plot shows the risk of making an incorrect decision due to analytical measurement uncertainty.

Figure 9 Risk due to measurement uncertainty in n=1 plans



The (n=1, c=0) plans have been extended to include allowance for MU [See CXG 54-2004 Figure 1 where MU uncertainty intervals are included to show the decision process.].

These plans, however, cannot simply be extended to lot inspection by including sampling components in the MU while allowing the PR and CR to be controlled to specified levels. Also, given that a single result is an estimate of the mean of a lot, use of this adjustment amounts to assessing compliance of the mean level of

the lot by comparing it against a maximum or minimum limit for the entire distribution. This comparison remains inappropriate regardless of whether sampling uncertainty is allowed for or not.

Sometimes guard-bands are applied, but their use may be unfair to producers.

If the characteristic is a compositional proportion, then provided measurement uncertainty is negligible, it is possible to design a sampling plan that can control both producer's and consumer's risks, but which requires only a single test of a composite sample to be performed.

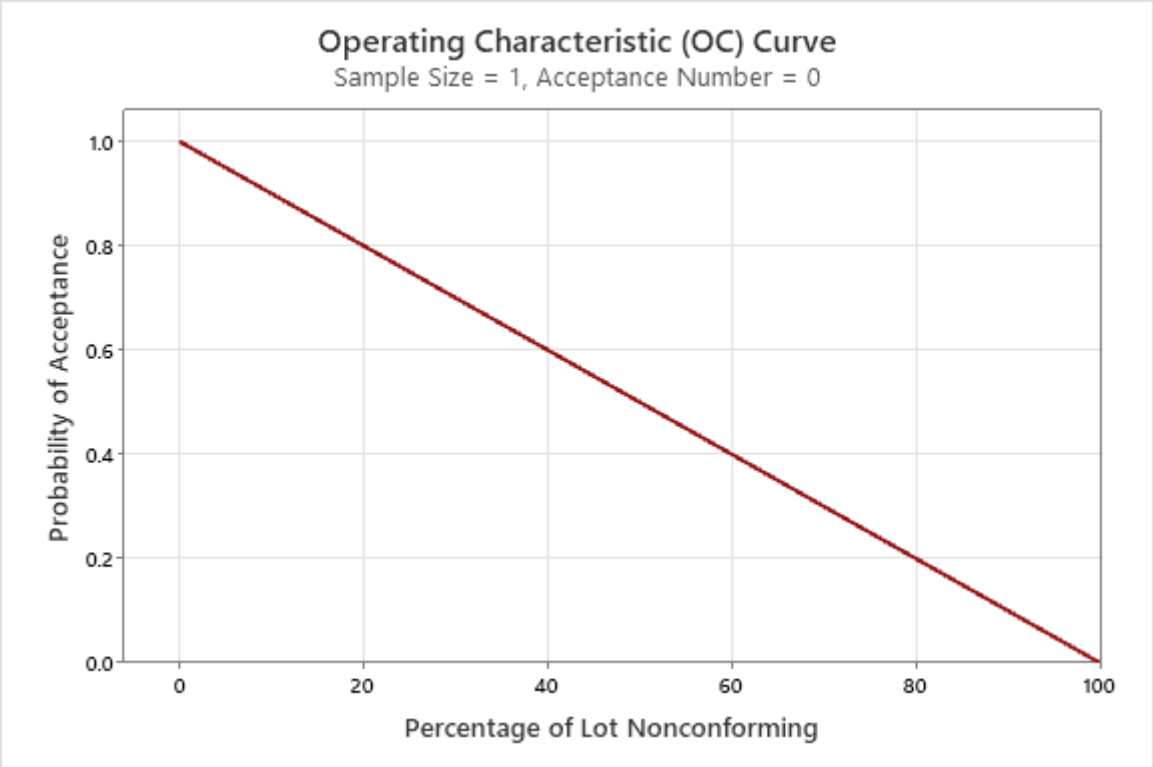
Refer to the example in section 3.3.2.

Section 5.5 discusses Bayesian sampling plans that allow plans controlling producer's and consumer's risks to be designed while requiring only small sample sizes.

Operating Characteristic Curve (n=1, c=0) sampling plan

The ideal Operating Characteristic Curve, i.e. where analytical measurement uncertainty is negligible, for this plan is shown below – basically this says that if a percentage P of the product in a lot is nonconforming then the probability of acceptance, $p.acc = 1 - P$.

Figure 10 OC curve for (n=1, c=0) plan, negligible MU



This and the points above show that there is a high risk of making an incorrect assessment of a lot using these plans, and indeed any plan based on a small number of samples. This suggests that confirmatory sampling and testing using on a more discriminatory sampling plan should be undertaken prior to raising a dispute.

3.1.3 Example: Attributes plans based on AQL of 6.5%

The CXS 3-1981 Standard for Canned Salmon has three provisions based on sampling that must be met for acceptance of a lot.

Section 8 defines defective sampling units (cans) in terms of foreign matter, odour and flavour, texture, discolouration exceeding 5% of the net contents and objectionable matter.

A lot is accepted provided:

1. the total number of defectives as classified according to Section 8 does not exceed the acceptance number (c) of an appropriate sampling plan with an AQL of 6.5%;

2. the total number of sample units not meeting the form of presentation as defined in Section 2.3 does not exceed the acceptance number (c) of an appropriate sampling plan with an AQL of 6.5%;

3. the average net weight and the average drained weight where appropriate of all sample units examined is not less than the declared weight or drained weight as appropriate, and provided there is no unreasonable shortage in any individual container.

In this standard, Section 7.1 contains the following information:

SAMPLING	
(i)	Sampling of lots for examination of the final product as prescribed in Section 3.3 shall be in accordance with an appropriate sampling plan with an AQL of 6.5%.
(ii)	Sampling of lots for examination of net weight shall be carried out in accordance with an appropriate sampling plan meeting the criteria established by the CAC.

The following options for sampling plans are from ISO 2859-1999, CXG 50-2004 Appendix II ISO INSPECTION PLANS INDEXED BY PRODUCER'S RISK from ISO 2859-1 Plans indexed by AQL.

Using the table, for a lot of size N=500 cans, a sample size of n=50 cans would be required at the normal inspection level, with the lot rejected if more than 7 of those cans were nonconforming, that is if they contained foreign matter, objectionable matter, atypical odour or flavour, abnormal texture, or discolouration exceeding 5% of the net contents.

Presumably, in this plan, each of the 50 cans is examined for each of the listed defects with a can classified as nonconforming if the criterion is breached for any one of them.

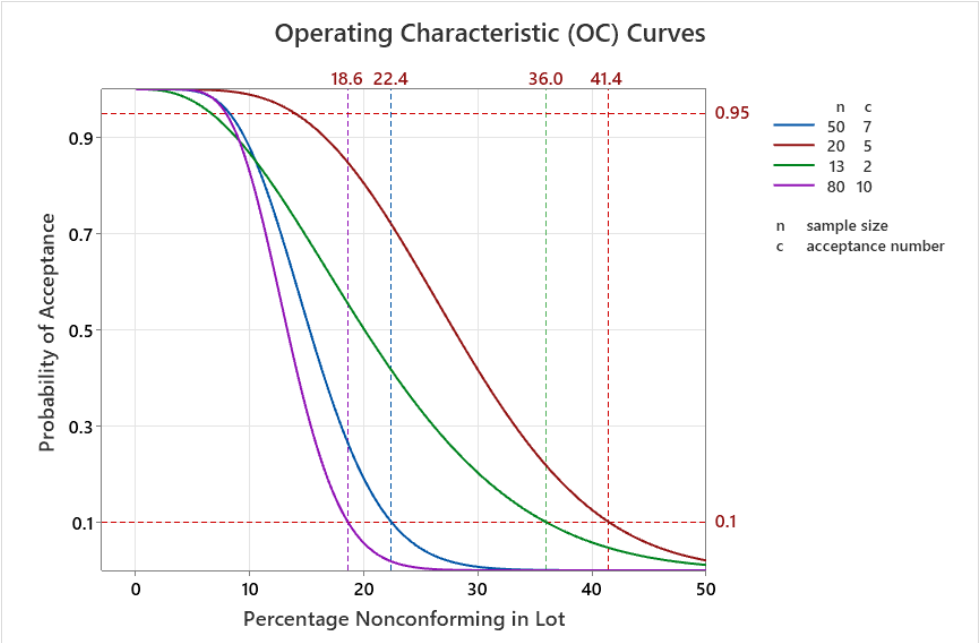
The consumer's risk quality level (CRQ) for this plan (n=50, c=7) is 22.4% meaning that there is a 10% chance of accepting a lot in which 22.4% of the cans are nonconforming. A decision should be made whether this is an acceptable level of risk before this plan is used.

This number of samples could be considered excessive, especially in terms of the overall lot size, although, as noted in CXG 50, the sampling fraction does not play a role in the design of sampling plans except possibly for quite small lots.

If it was decided to use the reduced inspection plan, (n=20, c=5) the consumer's risk quality level would increase to CRQ = 41.4%. If the sample size versus lot size relationship was disregarded and if the plan (n=13, c=2) was applied, then the CRQ would increase to 36%. On the other hand, if the plan (n = 80, c=10) was used, then the consumer's risk quantity level CRQ would decrease to 18.6%.

Operating characteristic curves for the four options are:

Figure 11 OC curves for ISO attribute plans

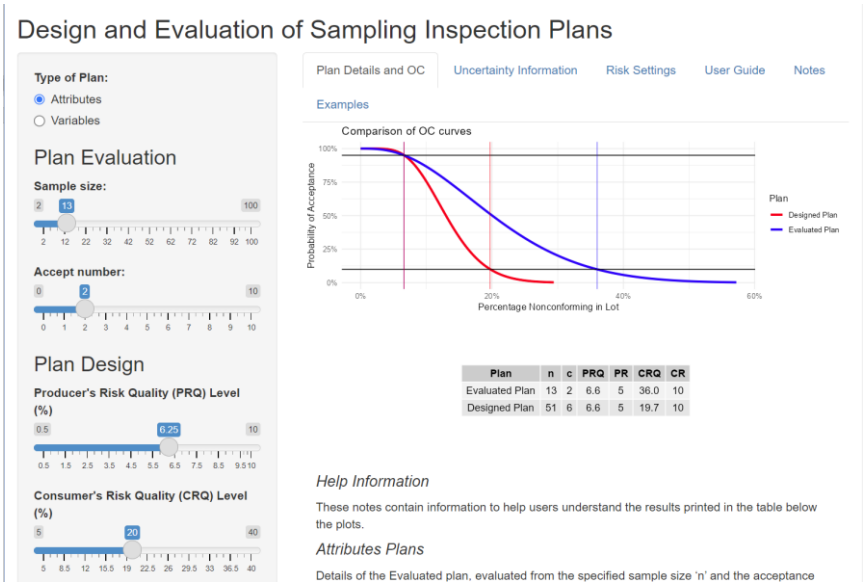


Extension

Suppose we wanted to modify this plan to provide a greater level of assurance to customers against accepting lots containing high levels of nonconforming product and that we were happy to keep the PRQ level at 6.5% but set the CRQ at 20%. This is easily accomplished using the app that shows the required plan is (n=50,

c=6). The output from the app is shown below. Note that unlike the ISO tables, the plan does not depend on the lot size.

Figure 12 OC curve for modified plan based on ISO plan



3.2 Examples for Variables plans

3.2.1 Example: Variables plan with negligible MU

Fat in Whole milk powder

The examples for variables plans are based on the provision for fat in whole milk powder from Codex Standard CXS 207-1999. The provision states that the milkfat content should exceed 26%.

From Codex Standard CXS 234 the analytical test method for fat in whole milk powder is the Röse-Gottlieb method (ISO 23318/IDF 249). The method contains the following precision data:

The repeatability limit is 0.2% (0.2 percentage points)
The reproducibility limit is 0.3% (0.3 percentage points)

Using these values, the repeatability standard deviation is calculated as:

$$\sigma_r = 0.2/(1.96 \cdot \sqrt{2}) = 0.2/2.77 = 0.072\ (percentage\ points)$$

With a similar calculation for the reproducibility standard deviation:

$$\sigma_R = 0.3/\left(1.96 \cdot \sqrt{2}\right) = 0.3/2.77 = 0.108\ (percentage\ points)$$

Using these values, the 'between-laboratory' standard deviation is calculated as:

$$\sigma_L = \sqrt{\sigma_R^2 - \sigma_r^2} = \sqrt{0.108^2 - 0.072^2} = 0.081\ (percentage\ points)$$

Refer to the Guidelines on Analytical Terminology CAC/GL 72-2009 for more information about the repeatability and reproducibility.

Key steps in the step-by-step design process

The design process involves finding the number of samples n, and the acceptability constant k, in order that the risks of incorrect decisions can be controlled the levels specified by the producer's risk (PR) and consumer's risk (CR) at their corresponding quality levels PRQ and CRQ respectively.

In this context the plan evaluation feature of the app might not be relevant but still might be useful for comparison of the designed plan with other options.

1. Attributes or variables?

Fat is a measured characteristic so this is an example of variables data.

2. Does the provision relate to the average value or to the entire distribution, i.e. to a maximum or minimum allowable level for the characteristic in the lot?

The provision specifies a minimum limit so relates to the entire distribution – 'most' of the product in the lot should comply.

3. Does the distribution of the characteristic follow a normal or some other distribution?

For the purposes of this example, we assume that lot consists of 1000 cans of milk powder and that the milkfat content of the cans in the lot is normally distributed. This is a reasonable assumption if the cans are produced using powder from a manufacturing process in a state of statistical control.

4. Is measurement uncertainty negligible or non-negligible?

If the standard deviation representing the variation of the characteristic in the lot was $\sigma = 0.3$, then the error-variance ratio is $(0.072/0.3)^2 = 0.058$, that, being less than 10%, means it would not be necessary to allow for measurement uncertainty in the design of this plan.

5. Specify the stringency required for the sampling plan.

Consumer's Risk Quality (CRQ):

What of percentage nonconforming cans would you be prepared to allow in lots that you would want to reject most of the time?

Consumer's Risk (CR):

How often would you want to accept such lots? (Usually 10%)

Producer's Risk Quality (PRQ)

What percentage of nonconforming cans would you be prepared to allow in lots that you would want to accept most of the time?

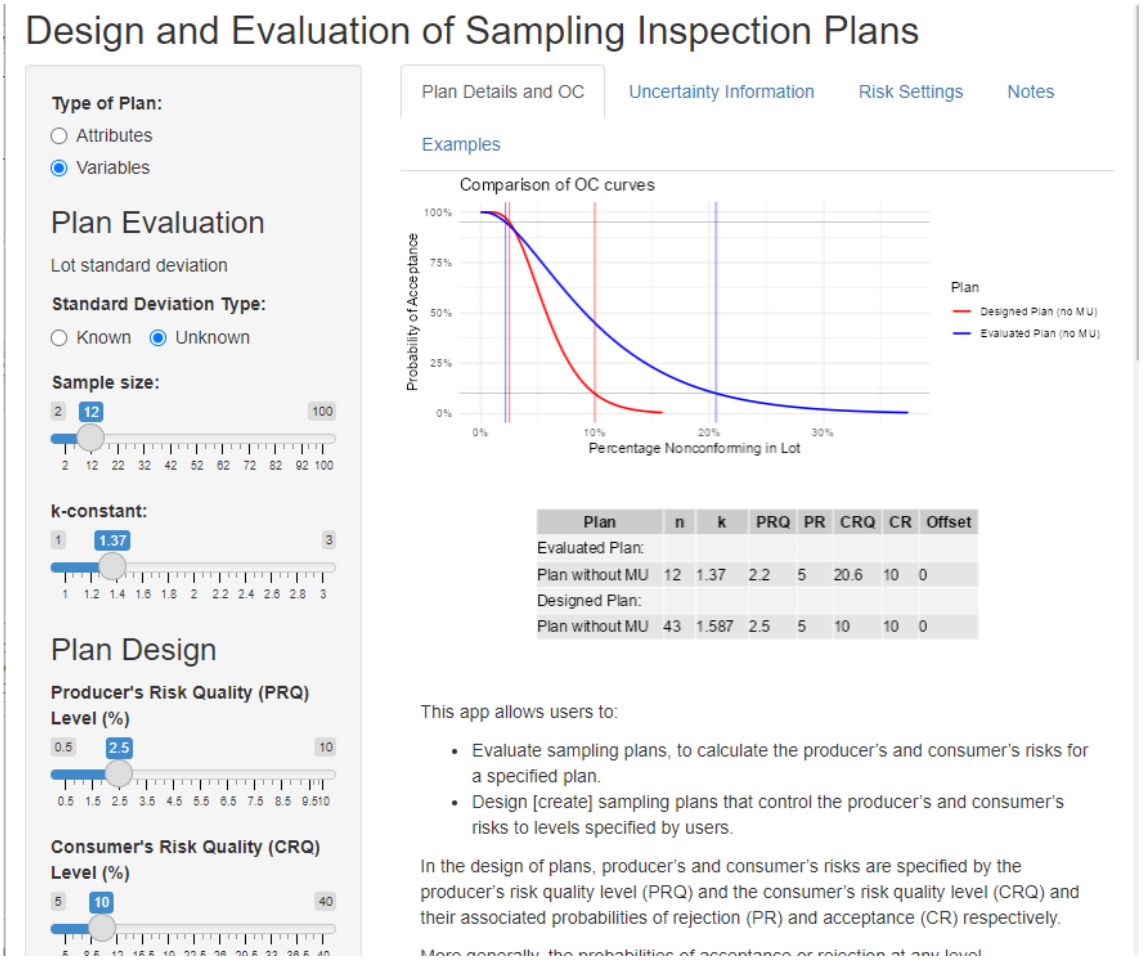
Consumer's Risk (CR):

How often would you want to reject such lots? (Usually 5%)

In this example the CRQ was chosen as 10% and the PRQ as 2.5%, with the CR and PR unchanged. This means that the plan will have:

- A 10% chance of accepting a lot in which 10% of the product is nonconforming.
- A 5% chance of rejecting a lot in which 2.5% of the product is nonconforming.

Figure 13 OC curve for variables plan - negligible measurement uncertainty



The plan required to control risks to the specified levels is (n=43, k=1.59), i.e. 43 samples need to be taken from the lot and tested. The lot is accepted provided the average and the standard deviation of the results meet the acceptance criterion:

$$\bar{x} - \mathrm{k} \times s \geq 26$$

where:

- $\bar{x}$ is the average of the 43 individual results and 's' their standard deviation,
- k is the acceptability constant, k=1.59 in this example.
- It is assumed that the measurements are expressed as percentages e.g. moisture of 5% on a weight/weight basis.

Note that the ISO plan for PRQ = 2.5% for a lot size N=1000 (Sample Code J, Inspection Level II) is (n=12, k=1.370) if the lot standard deviation is known and (n=46, k=1.482) if the lot standard deviation is unknown. Both plans have an actual PRQ of about 3.4% and a CRQ of about 11.3%.

Negligible measurement uncertainty

The validity of the assumption of negligible measurement uncertainty can be checked using the app, as follows.

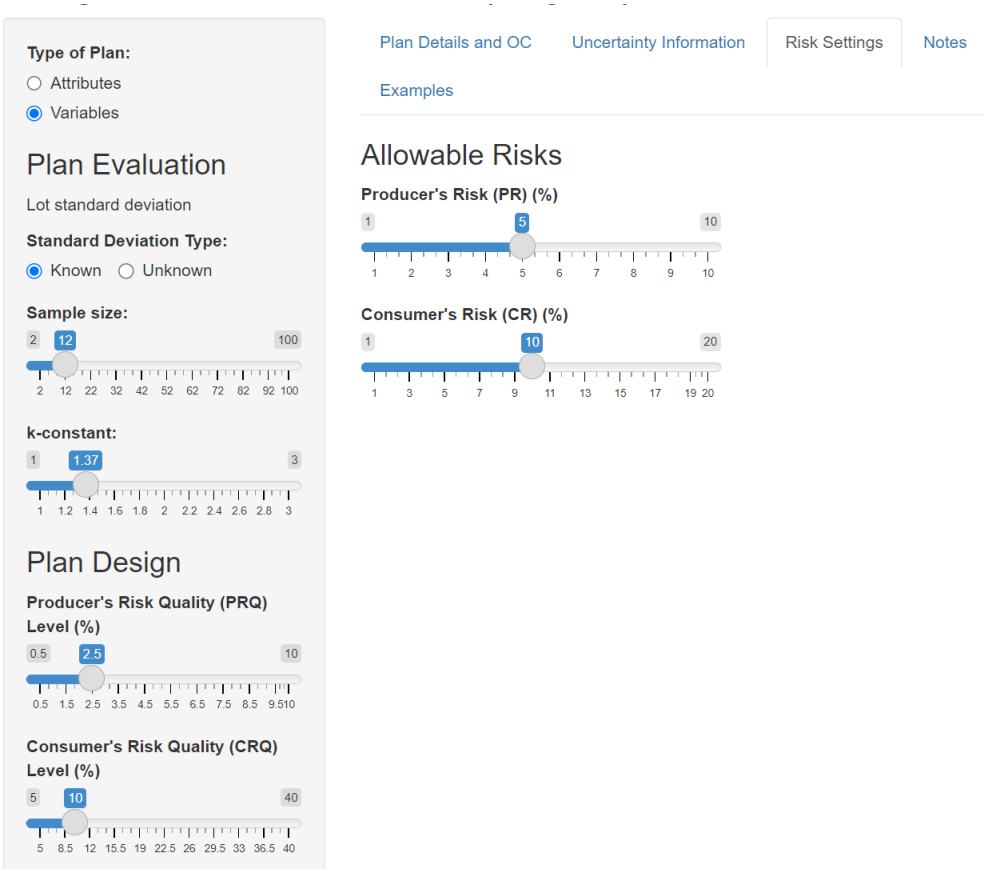
Step 1.

Open the app and select 'Variables' as the type of plan.

Step 2.

In the Plan Design section of the Plan Details and OC tab, set the Producer's Risk Quality level (PRQ) to 2.5% and the Consumer's Risk Quality level (CRQ) to 10%. There is no need to change the producer's and consumer's risks from their default settings, 5% and 10% respectively, at least for this example.

Figure 14 Negligible MU example - setting allowable risks



Step 3

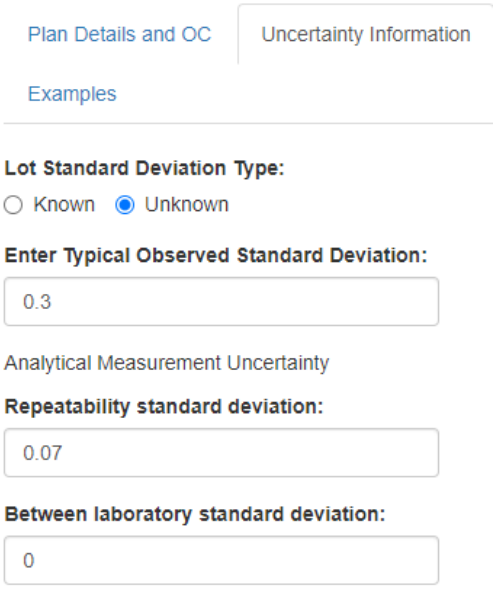
Go to the Uncertainty Information tab.

Select the “Lot Standard Deviation Type” to “Unknown”.

In the same window set the typical value of the observed standard deviation to 0.30 using the spin button ( ) or by entering the value directly and the repeatability standard deviation to $\sigma_r = 0.07$ similarly.

Note that the app has been designed so that if the lot standard deviation is known that value is used but if it is unknown, a typical value of the observed total standard deviation inclusive of repeatability error is entered.

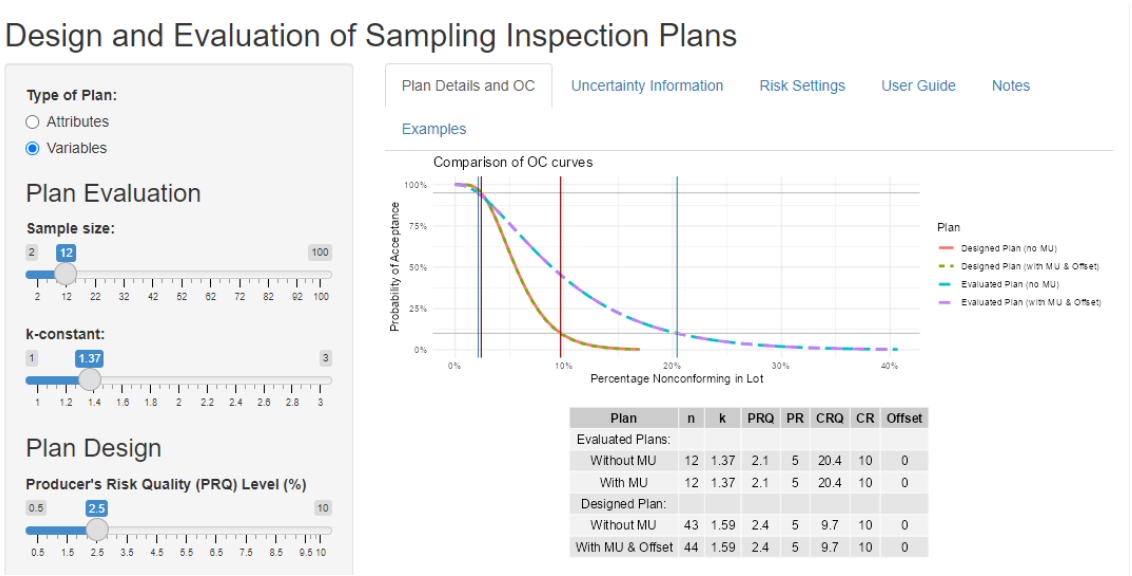
Figure 15 Variables Plan example - setting uncertainty information



Step 4

Return to the Plan Details and OC tab to view details of the plan and the OC curve. The app shows the OC curves and details of the plans including the risks,

Figure 16 Variables plan example - output



Plan	n	k	PRQ	PR	CRQ	CR	Offset
Evaluated Plans:							
Without MU	12	1.37	2.1	5	20.4	10	0
With MU	12	1.37	2.1	5	20.4	10	0
Designed Plan:							
Without MU	43	1.59	2.4	5	9.7	10	0
With MU & Offset	44	1.59	2.4	5	9.7	10	0

The last row of the table shows that the designed plan with the specified risks, allowing for the measurement uncertainty specified is (n=44, k=1. 59), the same as the plan where there is no measurement uncertainty shown in the line above. This differs slightly from the evaluated plan with the same n and k due to rounding of the k value within the app.

Notice also that although less than 10%, the error variance ratio of 0.058 means that one would expect the sample number to increase from 43 to $43 \times (1 + 0.058) = 45.5$ that, when rounded up gives n=46. However, for the unknown lot standard deviation case Hahn's adjustment (CXG 50 section 5.2.7) has been applied, meaning that although allowance must be made for the unknown standard deviation, it is not necessary to adjust the lot standard deviation for the measurement uncertainty. In any case, there is little difference between the plans in their sample numbers or risks and the 'error-free' plan can be used.

The lot is accepted provided the average and the standard deviation of the results meet the acceptance criterion:

$$\bar{x} - 1.59 \times s \geq 26$$

where:

- $\bar{x}$ is the average of the 43 individual results and 's' their standard deviation.
- It is assumed that the measurements are expressed as percentages e.g. a fat level of 26.5% on a weight/weight basis.

Although it is unnecessary in this example, the calculations assume that Hahn's adjustment has been applied to the observed total standard deviation, calculated from the inspection data. The adjusted standard deviation can be calculated using the formula:

$$s_{adj}^2 = s_{obs}^2 - \sigma_r^2$$

provided the right-hand side is greater than zero, otherwise the value of the adjusted standard deviation is taken as zero.

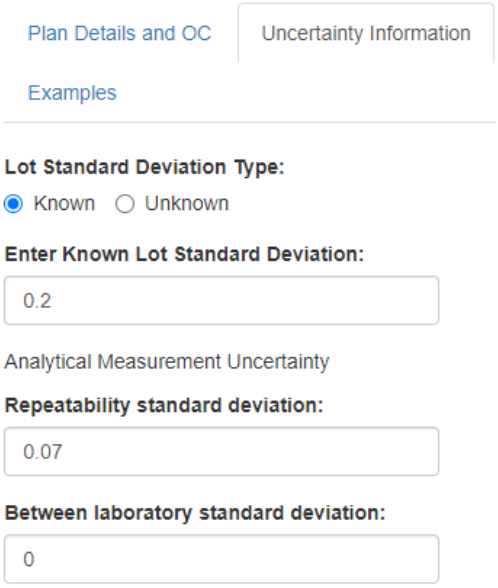
3.2.2 Example: Variables plan with non-negligible MU with no laboratory bias

It is assumed that the specified producer's and consumer's risks are set at the same levels and the between laboratory component of the analytical measurement uncertainty is negligible. However, in this example the lot standard deviation is assumed known.

If the lot standard deviation was $\sigma = 0.2$, the error-variance ratio is $(0.072/0.2)^2 = 0.13$ and, being greater than 10%, suggests that to allow for measurement uncertainty the number of samples should be increased to $19 \times (1 + 0.13) = 21.5 = 22$ after rounding.

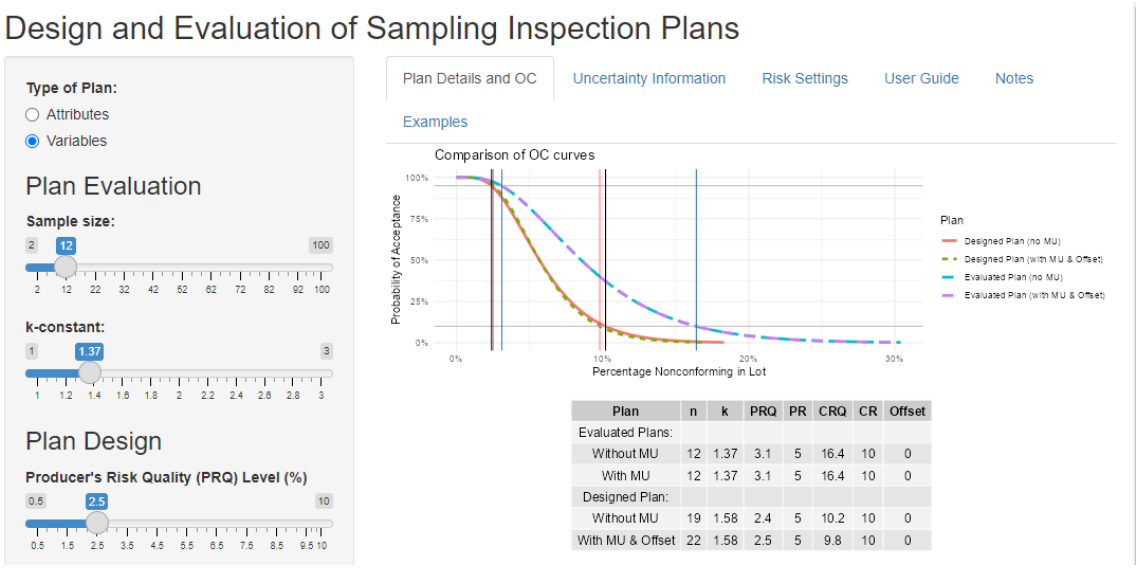
The app is used in the same way as in the example above, except that the lot standard deviation type is set to known (and, of course, a different value of the lot standard deviation is entered).

Figure 17 Variables plan example - non-negligible MU



The last row of the table below the OC curves confirm the calculation above, that there is a modest increase in sample size to n=22.

Figure 18 Variables Plan design - non-negligible MU



Alternatively, in the known standard deviation case, or if the error variance ratio γ is known, say from a measurement error study if the lot standard deviation is unknown, the acceptability constant k can be reduced to compensate for the increase in variability without the need to increase the sample size:

$$k^* = k/\sqrt{1+\gamma}$$

where:

k is the acceptability constant for the original plan,

k* is the acceptability constant for the modified plan.

Suppose the variables sampling plan (n=23, k=1.19) is being used to assess compliance of a particular characteristic having an upper limit of U = 10 and we have obtained test results as follows:

9.92, 9.85, 10, 9.62, 9.94, 10.02, 9.87, 9.8, 9.87, 9.95, 10.05, 10.03, 9.57, 9.83, 9.93, 9.93, 9.89, 9.79, 9.97, 9.96, 9.92, 9.83, 10.05

It is known from a previous measurement study that the error-variance ratio, the ratio of the repeatability variance to the lot standard deviation variance, is 0.25. Recall from CXG 50-2004 that the variance is the square of the standard deviation.

If the assessment of compliance proceeds in the usual way, the mean value of the results is m=9.90, the standard deviation s = 0.12, so that the acceptance criterion has a value of 9.90 + 1.19*0.12 = 10.04, and the lot should not be accepted.

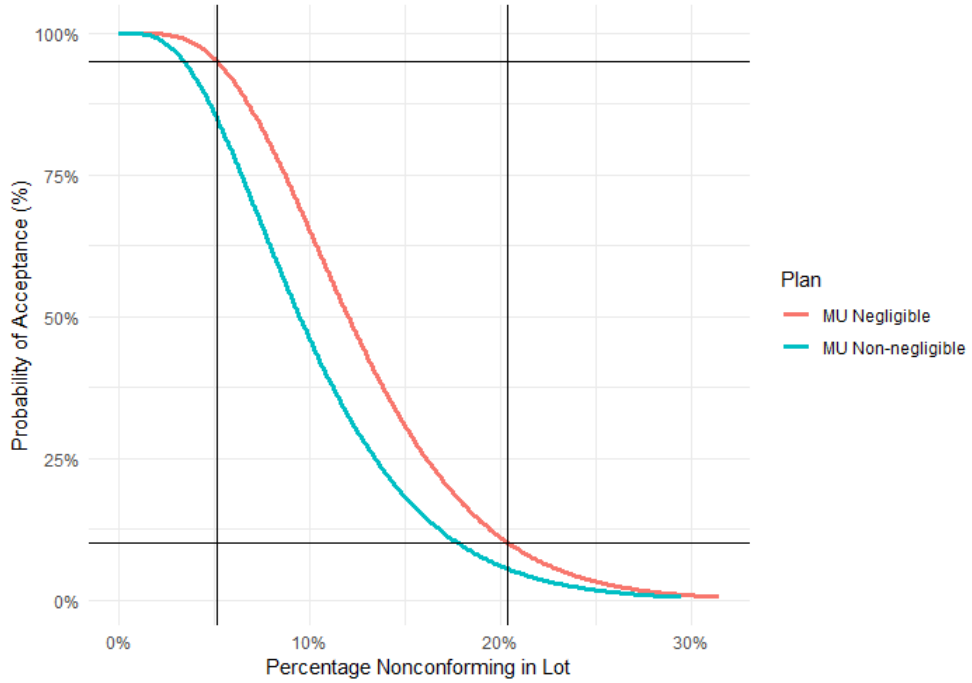
However, it is possible that measurement uncertainty has caused the lot to fail inspection. Hahn's adjustment can be applied to adjust the observed standard deviation for measurement uncertainty using the repeatability standard deviation is known from method validation. Supposing the repeatability standard deviation is $\sigma_e = 0.10$, the adjusted observed standard deviation s_{adj} is calculated by:

$$s_{adj}^2 = s_{obs}^2 - s_e^2 = 0.12^2 - 0.10^2 = 0.0044$$

so that the adjusted standard deviation $s_{adj} = 0.066$ and the updated value of the acceptance criterion is 9.90+1.19*0.066 = 9.98 and the lot can be accepted.

The OC curves below show that the probability of acceptance at any given percentage nonconforming in a lot will be less when repeatability-type measurement uncertainty is present.

Figure 19 OC curves with and without MU (no bias)



Plan	n	k	PRQ	PR	CRQ	CR
Error-free Plan	23	1.19	5.1	5	20.4	10
Error-prone Plan	23	1.19	3.4	5	17.8	10

Another way of overcoming non-negligible, repeatability measurement uncertainty is to increase the sample size; ISO3951-1: 2013 gives the formula:

$$n^* = n(1+\gamma)$$

where:

n is the sample size for the original plan in which measurement uncertainty is negligible,

n* is the sample size for the modified plan, and

γ is the error-variance ratio.

Alternatively, if the error variance ratio γ was known, the acceptability constant k can be reduced to compensate for the increased variability without the need to increase the sample size:

$$k^* = k/\sqrt{1+\gamma}$$

where:

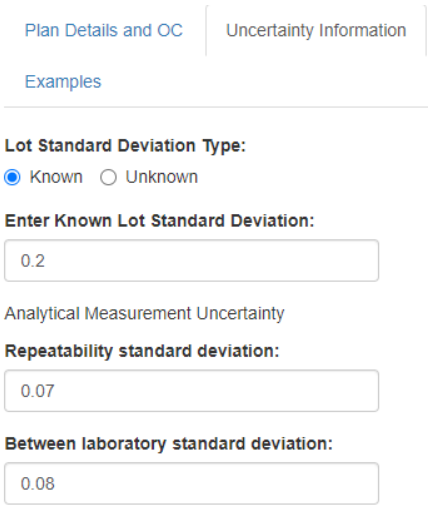
k is the acceptability constant for the original plan,

k* is the acceptability constant for the modified plan.

3.2.3 Example: Variables plan with non-negligible MU with laboratory bias

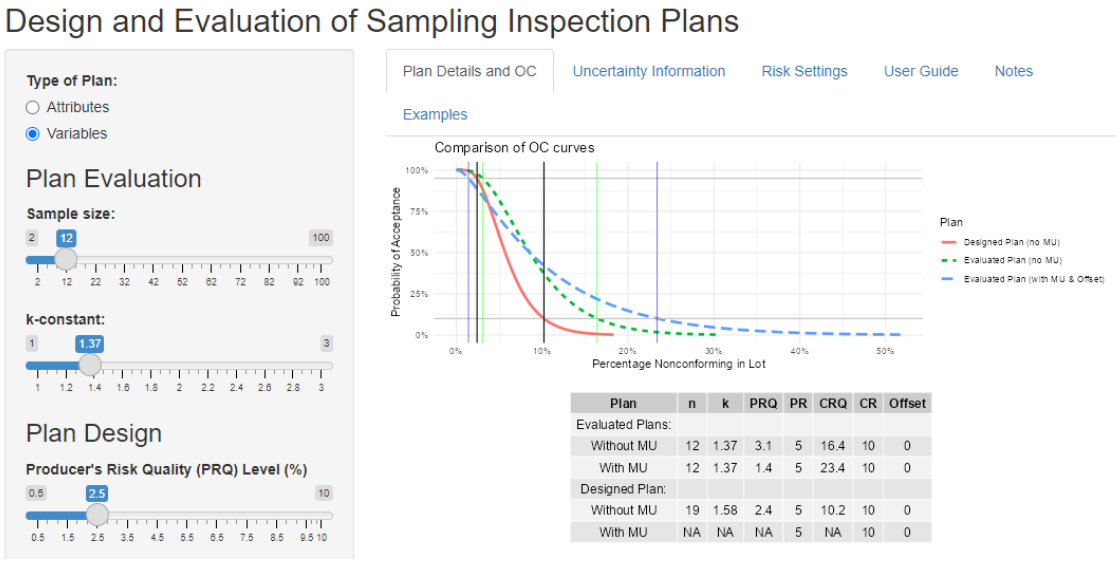
In this example it is assumed that the lot standard deviation of σ = 0.2 is known and that the between laboratory component of measurement uncertainty has a standard deviation of $\sigma_L = 0.08$ as in the original example (Section 3.2.1).

Figure 20 Entering uncertainty information - Non-negligible MU (inc. bias)



The NA (not available) values in last row of the table shows that a plan allowing control of the producer's and consumer's risks to the levels specified cannot be found.

Figure 21 Output - Non-negligible MU example



Plan	n	k	PRQ	PR	CRQ	CR	Offset
Evaluated Plans:							
Without MU	12	1.37	3.1	5	16.4	10	0
With MU	12	1.37	1.4	5	23.4	10	0
Designed Plan:							
Without MU	19	1.58	2.4	5	10.2	10	0
With MU	NA	NA	NA	5	NA	10	0

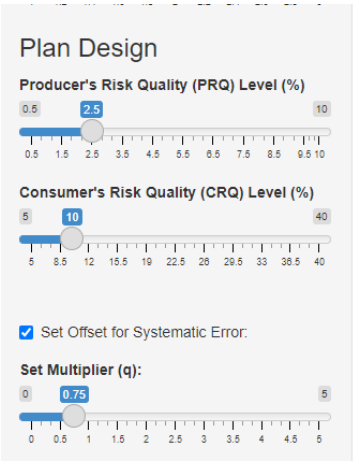
To find a plan it is necessary to introduce an offset.

Open the sub-window "Set Offset for Systematic Error" and move the slider to set a value of "q" until a plan is found. For example, as shown in the image below, if the multiplier is set to q = 0.75, then the offset in the acceptance criterion is then q*sigma_L = 0.75*0.08 = 0.06 so that with n=19 and k = 1.58 so the acceptance criterion would then become:

$$\bar{x} + k \cdot \sigma + 0.06 \leq USL$$

Note that while the consumer's risk quality level (CRQ) remains unchanged, the producer's risk quality is significantly reduced.

Figure 22 Plan details - non-negligible MU example



Plan	n	k	PRQ	PR	CRQ	CR	Offset
Evaluated Plans:							
Without MU	12	1.37	3.1	5	16.4	10	0
With MU	12	1.37	0.6	5	15.3	10	0.06
Designed Plan:							
Without MU	19	1.58	2.4	5	10.2	10	0
With MU & Offset	19	1.58	0.4	5	10	10	0.06

Help Information

These notes contain information to help users understand the results printed in the table below the plots.

Attributes Plans

Details of the Evaluated plan, evaluated from the specified sample size 'n' and the acceptance number 'c' and the designed plan, based on the specified Producer's and Consumer's risks are shown.

Variables Plans

Details of the Evaluated plans, evaluated from the specified sample size 'n' and the acceptability constant 'k' are shown. The OC curves for these plans are shown both without and with the effect of measurement uncertainty,

Details of the Designed, designed from the specified Producer's and Consumer's risks are shown. The OC

3.2.4 Example: Fractional Nonconformance plans

Suppose we have measurements from testing 15 samples from a lot to assess whether the lot conforms with the lower specification limit of L = 50. The measurement process is known to be normally distributed, with no laboratory bias and a standard deviation of $\sigma = 0.045$.

The following results were obtained:

50.01, 50.04, 50.07, 50.1, 50.15, 50.2, 50.29, 50.42, 50.45, 50.48, 50.55, 50.6, 50.8, 51.2, 51.3

The fractional nonconformance values for each sample can be calculated using Excel, using the formula:

$$fnc = NORMDIST(50, x, 0.045, TRUE)$$

where 'x' represents a single test result. This gives the following FNC values:

0.4121, 0.187, 0.0599, 0.0131, 0.0004, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0

The sum of these values is 0.6725, so if the acceptance limit Ac was 0.75 the lot would be accepted.

This example shows the principle behind the calculation, that can easily be extended to allow for measurement uncertainty distributions other than the normal distribution.

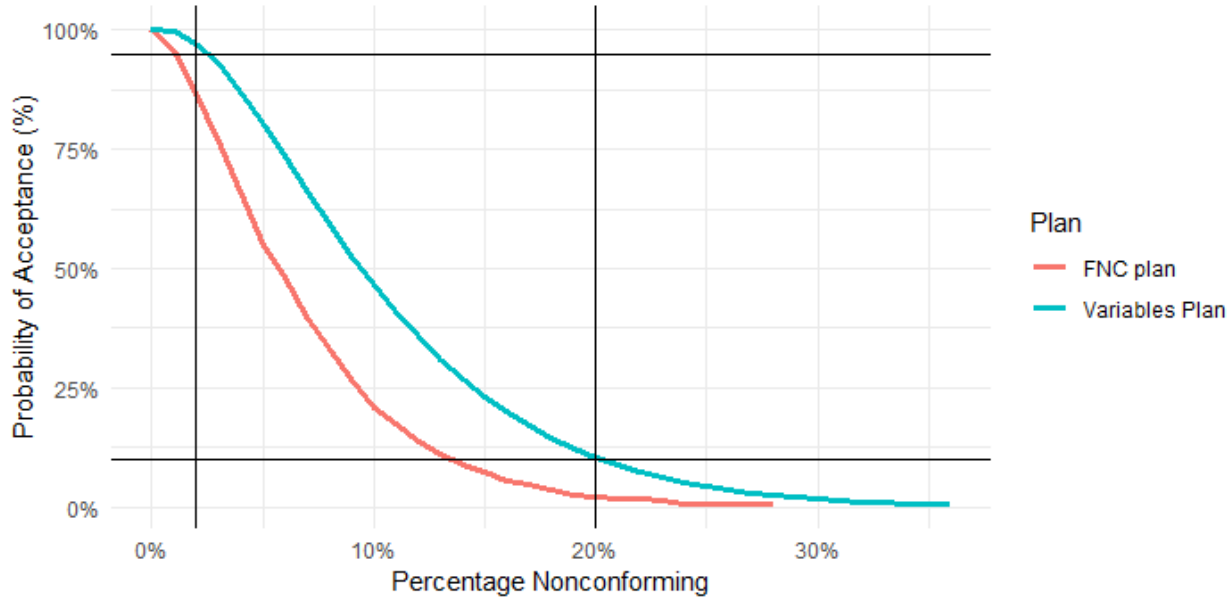
The following shows the OC curves for a variables plan and for a FNC plan, both with the same error-variance ratio.

Fractional nonconformance plans are an example of attribute-variables (attri-variables) plans where measurements are reclassified into another 'measure' of conformance for each sample, with the decision on acceptance of the lot made using the sum of these new measures.

The ICMSF sampling plans used in microbiological assessments for counted characteristics are another example of attri-variables plans as are attributes plans used with measurements that are classified as pass or fail with respect to a limit.

See also Section 4.4.

Figure 23 Fractional Nonconformance Plans



Plan	n	k_Ac	AQL	AQL.risk	LQL	LQL.risk
Variables Plan	15	1.20	2	5	20	10
FNC plan	15	0.75	1	5	14	10

3.3 Lots consisting of bulk materials

3.3.1 Example: Aflatoxin sampling plans due to Whitaker (2006) et al.

Refer to 5.6.5 Aflatoxin sampling plans

Shelled Almonds for further processing

Suppose the average concentration of aflatoxins in the lot was C = 8 µg/kg and ns = 20000, 20 kg @ 1000 shelled nuts per kg, were taken as a sample, and this sample was ground and well-mixed composite formed. If a subsample of 50g was taken and a single aliquot (n_a=1) tested, the standard deviation S representing the uncertainty of the average level would be:

$$S^{2=}\frac{7730 \times 5.759}{20 \times 1000}8^{1.561} + \frac{100 \times 0.170}{50}8^{1.646} + \frac{0.048}{1}8^2 = 70.67$$

Giving S = 8.41. The first component representing the sample-by-sample variation is much larger than the other two.

The maximum limit for shelled nuts for further processing is 20 µg/kg, based on an initial sample of 20kg shelled almonds and one laboratory determination.

At an average level of contamination of C=8 µg/kg, the variance $S^2 = 70.67$ and from the formula above, the value of k is worked out using:

$$70.67 = 8 + 8 \times 8/k$$

from which

$$k = \frac{64}{70.67 - 8} = 1.0212 \; and \; \frac{k}{C+k} = \frac{1.0212}{8 + 1.0212} = 0.1132$$

The probability of acceptance be calculated using Excel:

BETA.DIST(k/(C+k),k,maximum_Limit,TRUE) that is equivalent to the Negative Binomial distribution[6].

$$BETA.DIST\left(\frac{k}{C+k}, k, maximum.limit, TRUE\right) = BETA.DIST(0.1132, 1.0212, 20, TRUE) = 0.906$$

or 90.6%

[6] Although the negative binomial distribution function is available in Excel, it is not in a form suitable for these calculations.

Note that the probability of acceptance at the maximum limit C = 20µg/kg is 0.622, that shows again that the principle of offsets has been employed in the setting of limits to provide consumer protection.

The calculations of the probabilities of acceptance in the Mycotoxin S&T Guide appear approximate, the actual calculations are unknown but the differences from results calculated in other known ways are small enough so as not to matter.

Probabilities of Acceptance for Shelled Corn (n_s=3000, n_{ss}=50, n_a = 1)

Concentration (µg/g)	Variance	Mycotoxin S&T Guide (%)	Negative Binomial (R) (%)	Beta Distribution (R) (%)	Beta Distribution (Excel) (%)
0	0	100	100	100	100
5	72.76	94.07	94.29	94.29	94.29
10	148.01	84.9	85.3	85.3	85.3
20	302.74	61.53	62.23	62.23	62.23
30	461.41	38.87	39.8	39.8	39.80

3.3.2 Example: Plans based on the beta distribution

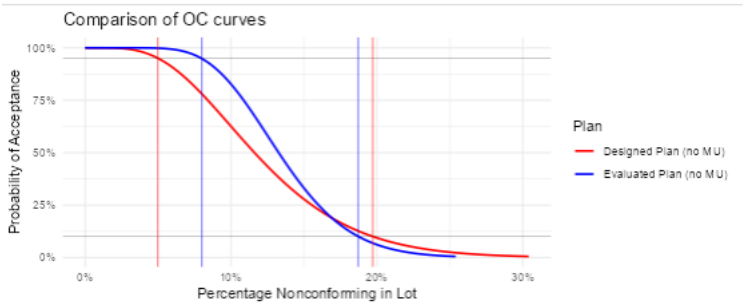
Plan for Capsaicin – based on Codex Standard 294-2023

Codex Standard CXS294 - 2023 for Gochujang contains a provision for capsaicin, that the levels should not be less than 10 mg/kg on a weight:weight (w/w) basis with lot acceptance decided using an attributes sampling plan with AQL = 6.5%, where a container is classed as nonconforming if the result from testing a sample taken from that container is less than the limit.

The number of samples will depend on the size of the lot but could be considerable, e.g. n=80 samples for a lot consisting of 1000 containers [packages]. However, capsaicin is tested using the HPLC method, so it is not feasible to perform more than relatively few tests on each lot.

As commented elsewhere, the use of attributes plans classifying measurements as attributes is inefficient and, for a lot size of 1000 containers the corresponding variables plan from ISO3951-1 (standard deviation unknown, negligible measurement uncertainty) with AQL = 6.5% is n=52, k=1.120, that has an AQL (PRQ) of 8%, with PR = 5%, and LQL (CRQ) of 18.7%, with CR = 10%.

Figure 24 Capsaicin example - ISO plans

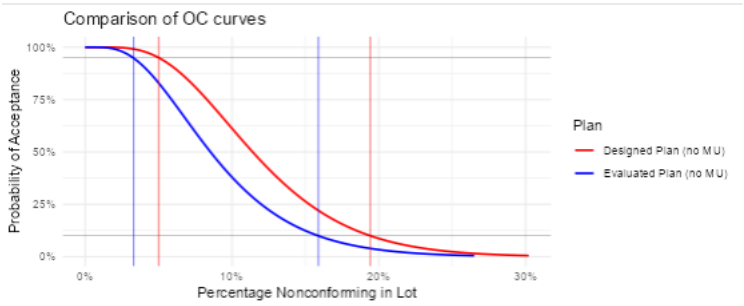


Plan	n	k	PRQ	PR	CRQ	CR	Offset
Evaluated Plan:							
Plan without MU	52	1.12	8	5	18.7	10	0
Designed Plan:							
Plan without MU	24	1.21	5	5	19.7	10	0

An alternative approach is to consider compliance of capsaicin levels in the overall lot rather than at the individual container level and, being a measured characteristic, it means that variables plans could be used.

Using the same consumer's and producer's risks as those for protein and moisture above (a producer's risk of 5% of rejecting lots containing 5% nonconforming product and a consumer's risk of 10% of accepting lots containing 20% nonconforming product) the resulting variables plan is (n=14, k=1.205) assuming the lot standard deviation is known. This plan can be modified to allow for non-negligible measurement uncertainty.

Figure 25 Capsaicin example - Variables plan



Plan	n	k	PRQ	PR	CRQ	CR	Offset
Evaluated Plan:							
Plan without MU	12	1.37	3.3	5	15.9	10	0
Designed Plan:							
Plan without MU	14	1.205	5	5	19.4	10	0

Further, capsaicin is a compositional characteristic, so that, if measurement uncertainty is negligible plans based on the beta distribution (refer to CXG 50-2004 Section 4.3.1) would be applicable. Use of these plans would mean that:

(1) a composite sample is formed from a requisite number of subsamples, that number being determined in the design of the plan based on specifications of allowable risks.

Acceptance of the lot would be determined by an acceptance criterion of the form:

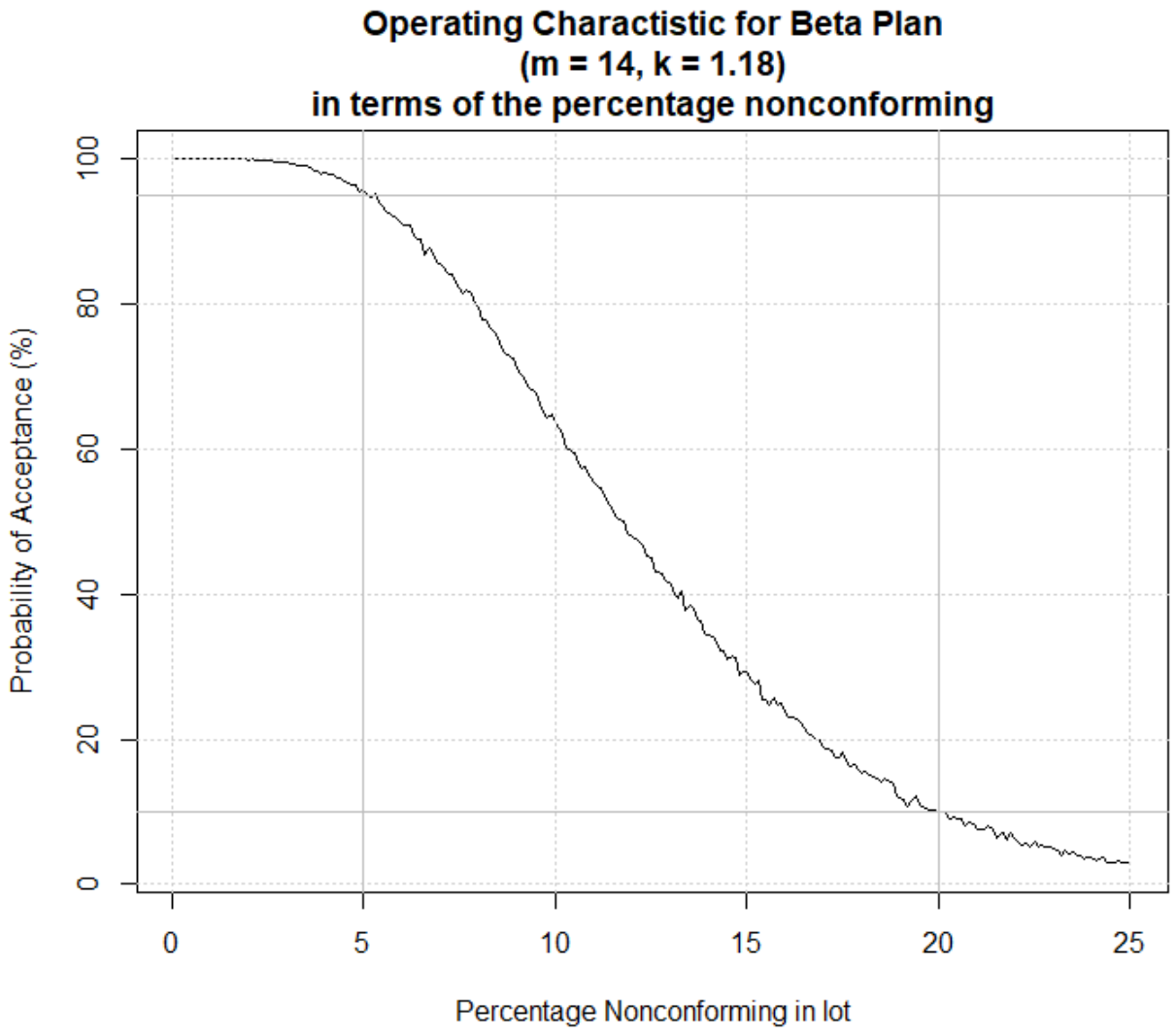
$P - k \times s \geq L$ where P is the test result or average test result and $s = \sqrt{P(1-P)/\theta}$, L is the minimum limit (10ppm) and k is the acceptability constant for the plan.

Historical data would first need to be analysed to estimate the precision parameter θ but a hypothetical value of the precision parameter of $\theta = 44 \times 10^6$ has been used in the following example.

Using those same consumer's and producer's risks the resulting plan is (m=14, k=1.18) i.e. a composite sample would be formed from 14 subsamples taken randomly from the lot, with the composite tested just once – the test result would then be the estimate of "P".

The Operating Characteristic for this plan is shown below.

Figure 26 Capsaicin example - Beta distribution plan



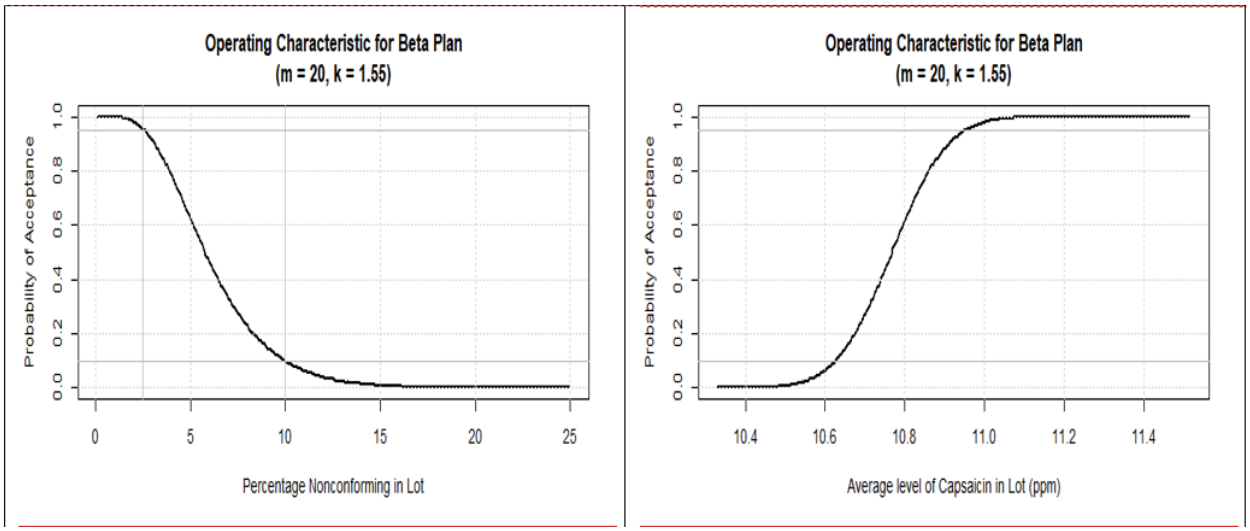
If however, we decided that capsaicin was a more critical characteristic for the product then we may wish to reduce the consumer's risk – instead of decreasing the chance of acceptance at the CRQ we can reduce the CRQ itself, to 10%, and also reducing the PRQ to 2.5%.

Consumer's Risk Quality level (CRQ)	
What percentage nonconforming would you allow in lots that you would want to reject most of the time?	10%
How often would you want to accept such lots (default = 10%)?	10%
Producer's Risk Quality level (PRQ)	
What percentage nonconforming would need to be present in lots that you would want to accept most of the time?	2.5%
How often would you want to reject such lots (default = 5%)?	5%

The corresponding sampling plan is (m=20, k=1.55) i.e. a composite sample would be formed from 20 subsamples randomly taken from the lot and the acceptance criterion would use a multiplier of the standard deviation of k=1.55.

The Operating Characteristics for this plan are shown below, in terms of both the percentage nonconforming and the average level of capsaicin in the lot.

Figure 27 OC curves for plan based on the beta distribution



3.4 Other sampling plans

3.4.1 Example: ISO sampling plans – 6.5% AQL

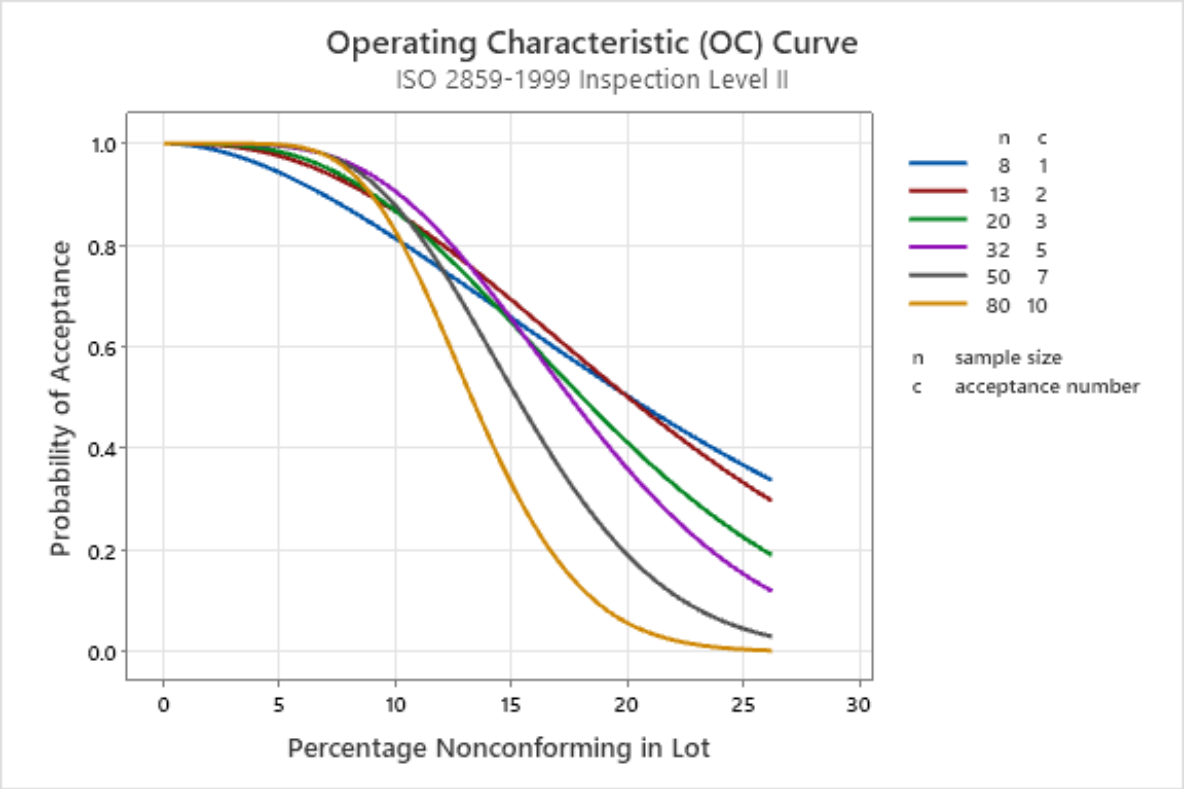
Refer to 3.2.2 Design of sampling plans, Table: Attributes plans from ISO 2859-1 for PRQ = 6.5 percent

A number of Codex standards contain sampling plans from the ISO standard ISO 2859[7] with an AQL of 6.5%, apparently because these plans were promulgated in the now defunct CODEX STAN 233, Codex sampling plans for prepackaged foods (AQL 6.5) (CODEX STAN 233-1969). While these plans might be suitable in some applications, users should first check that they will meet expectations around the control of producer's and consumer's risks before use. In particular, these plans might suffer from following problems:

- There could be poor control of the consumer's risk, that will vary according to the lot size.
- The plans should not be used by classifying variables data as attributes, except as a last resort, and not used in cases where measurement uncertainty is non-negligible.
- The ISO plans are intended to be used with switching rules; if not there will be no safeguards against deteriorating quality, nor any reward by way of reduced inspection for good quality.

[7] International standard ISO 2859-1: Sampling procedures for inspection by attributes - Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection.

Figure 28 OC curve ISO 2859 attributes plans



Refer: CXG 50 Appendix II ISO INSPECTION PLANS INDEXED BY PRODUCER'S RISK

Producer's and Consumer's Risk Quality Levels (PR = 5%, CR = 10%)

n	c	PRQ%	CRQ%
8	1	4.64	40.62
13	2	6.60	35.98
20	3	7.14	30.42
32	5	8.50	27.07
50	7	8.22	22.42
80	10	7.91	18.60

Conclusion

Whilst possibly facilitating trade, plans with low sample numbers do not provide high levels of consumer protection, that will vary according to the lot size.

Net weight

It appears Codex has not provided any guidance on sampling plans for net weight. However the need for pro visions relating to weight has possibly been superseded by the introduction of weight legislation using the Average Quantity System, based on the OIML International Recommendation R087 published by BIPM.

3.4.2 *Ad hoc* plans

Suppose, for example, that 4-6 samples are taken, formed into a composite sample and a single laboratory sample taken from the thoroughly mixed composite for analysis.

This is not the standard statistical approach to the design of sampling plans as the plan is not designed from specifications of allowable risks; therefore, we must evaluate it to check that it will control risks satisfactorily.

Four options have been evaluated:

(1) use of the single result in an (n=1, c=0) sampling plan for the assessment of compliance to an average level (see 3.4.2.1 Compliance of the average level scenario evaluation),

(2) use in an (n=1, c=0) attributes plan (see 3.4.2.2 Compliance of the average level scenario evaluation),

(3) use in a variables plan (see 3.4.2.3 Variable Plan scenario evaluation) and

(4) use in a plan based on the beta distribution if the characteristic is a compositional proportion and measurement uncertainty is negligible (see 3.4.2.4. Beta distribution plan scenario evaluation).

Notation:

U the upper specification limit,

σ (*sigma*) the assumed known value of the lot standard deviation (rather than an estimate of σ).

Alternatively, the error-variance ratio must be well known; in this case it refers to the ratio of the reproducibility variance to the lot variance.

$$Error - variance\ ratio = \frac{u^2}{\sigma^2}$$

u the assumed known standard deviation representing the standard measurement uncertainty.

The uncertainty of the average level $\bar{x}$ of the composite sample formed by taking n samples will be $\sigma/\sqrt{n}$ and the uncertainty variance of the measured value will be $\sqrt{u^2 + \frac{\sigma^2}{n}}$

The following examples reinforce the guidance in ISO2859 that ad hoc sampling methods are not recommended since they lead to uncalculated risks and often to unjustifiably high risks; further, there is no logical basis for either the acceptance or rejection of the product.

3.4.2.1 Compliance of the average level scenario evaluation

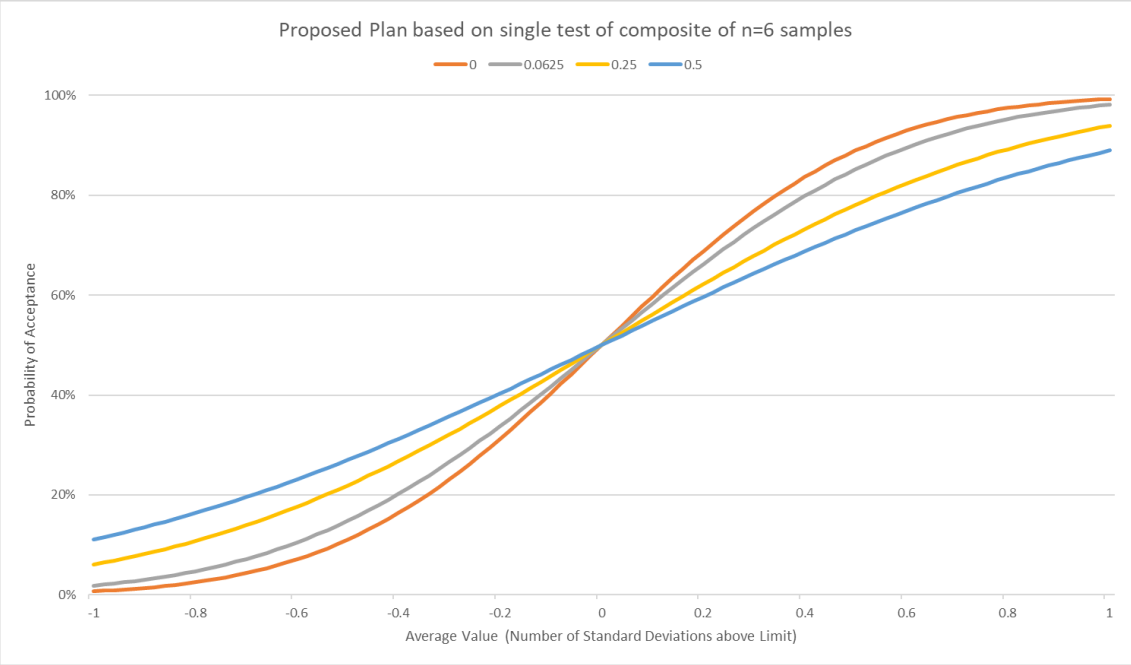
The probability of acceptance, of accepting a lot against a lower limit L in terms of the true average level μ in the lot is given by:

$$p.acc = NORMSDIST(k * sigma/SQRT(sigma\text{^}2/n + u\text{^}2))$$

Using Excel formula notation, where:

k*sigma is the offset from the limit.

Figure 29 OC curves - ad hoc plans - complinace of average level



The different lines show the OC curves for different values of the error-variance ratio.

Conclusion

This shows that the proposed plan is ineffective for assessing compliance of an average level – for example there is still a high chance of acceptance when the true average level is a reasonable number of standard

deviations below the limit; for example, around 20% chance of acceptance when the average level is 0.5 standard deviations below.

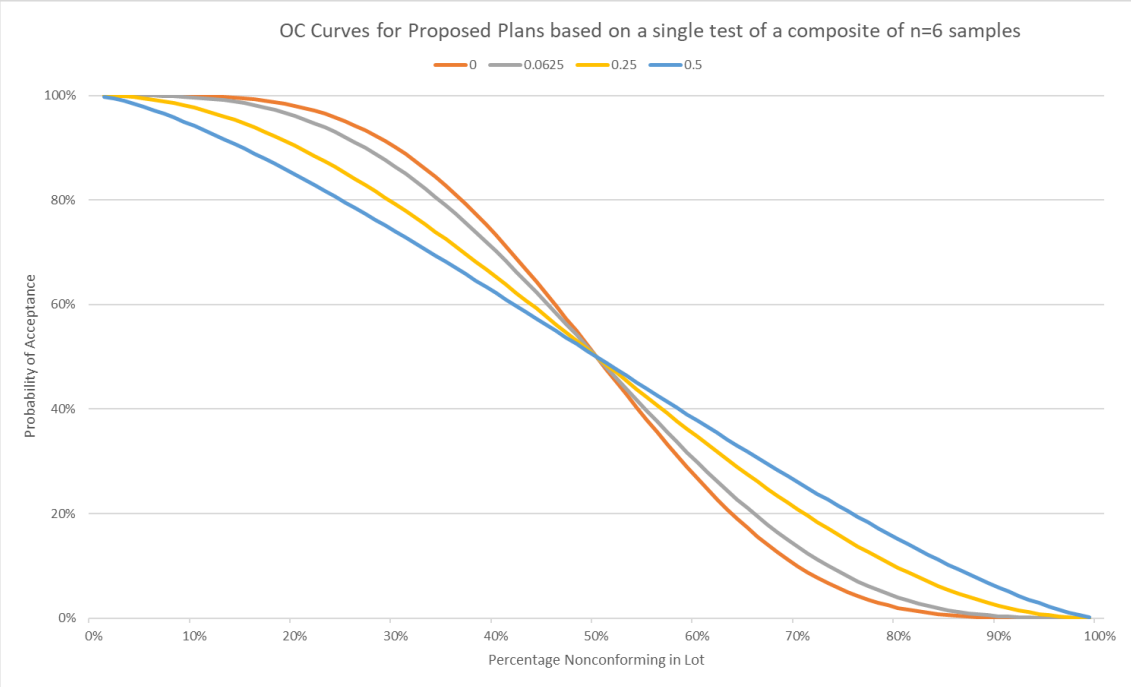
3.4.2.2 Attributes Plans

The probability that a single result will meet the upper limit is given by:

$$p.acc = NORMSDIST\big(NORMSINV(1 - NC) * sigma/SQRT(sigma^2/n + u^2)\big)$$

where NC is the percentage nonconforming in a lot.

Figure 30 OC curves - ad hoc attributes plans



Conclusion

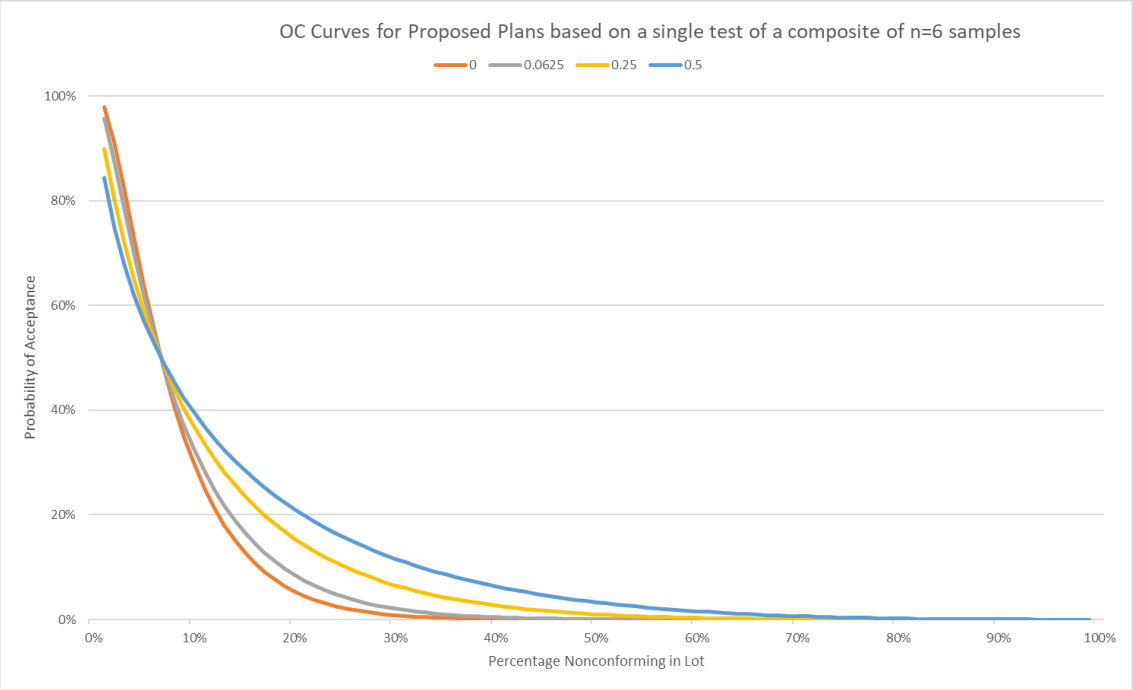
This plan also appears ineffective for assessments against upper or lower limits; this is not surprising since a composite sample represents the average level in a lot. Note that a single sample is also a representation of the average level of a lot.

Refer to section 4.1.2 for a discussion about (n=1, c=0) sampling plans.

3.4.2.3 Variables Plans

$$p.acc = NORMSDIST((NORMSINV(1 - NC) - k) * sigma/SQRT(sigma\text{^}2/n + u\text{^}2)))$$

Figure 31 OC curves - ad hoc variables plans



Conclusion

By using variables plans one can better control the risks of noncompliance by varying the value of k, the acceptability constant; the image shows the OC curves for k=1.5.

However, use by consumers of large k values for simplification or to reduce costs of testing does not seem to be a fair practice, considering that CXG50 is intended to apply mostly to commodity characteristics such as composition of 'commodity defects' and should serve to facilitate trade.

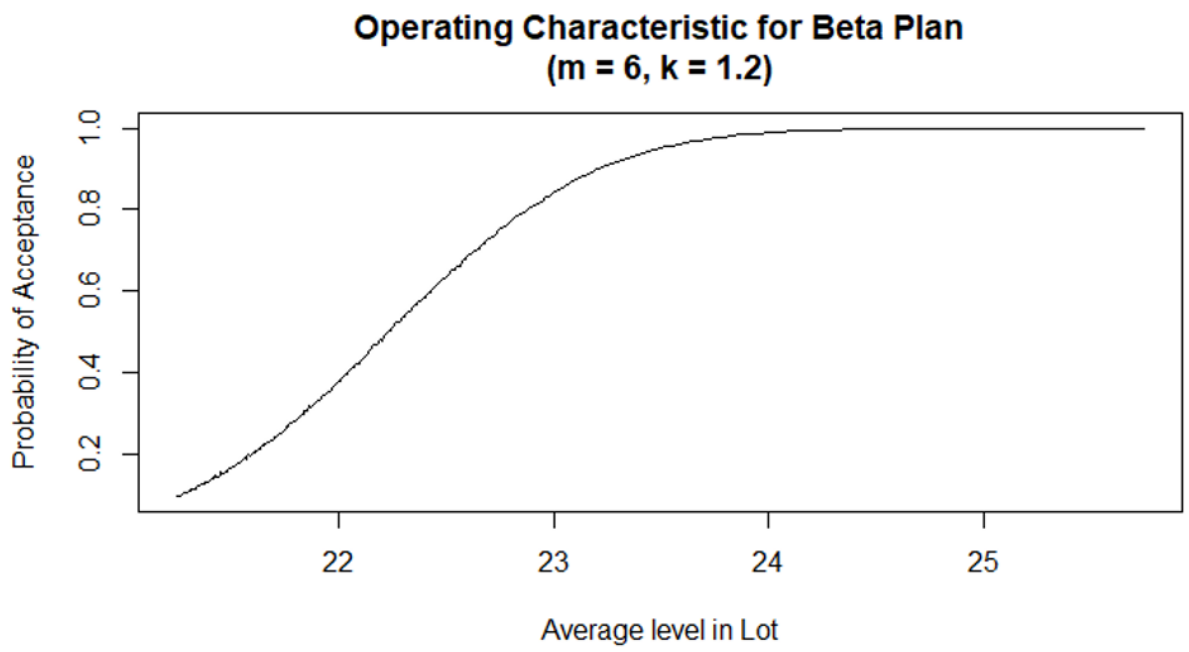
3.4.2.4 Beta distribution plans

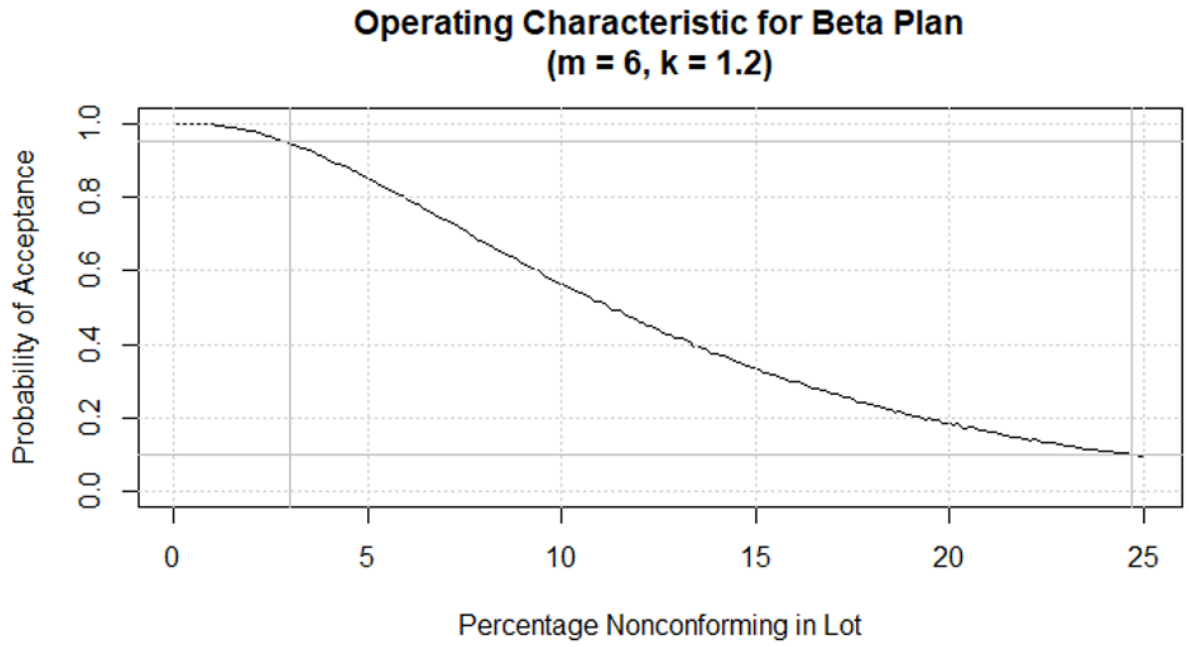
Use of a composite sample does not allow the use of attributes plans, and conventional variables plans won't be very useful unless one is assessing compliance against an average level (and maybe not even then).

The only possible classical solution seems to be the plans based on the beta distribution that requires the characteristic to be a compositional proportion and that measurement uncertainty is negligible.

These examples are based on a value of the theta parameter of θ = 500 and a minimum limit of L=20, (some previous work found that theta values for fat, protein and moisture in milk powders were between 700-3000).

Figure 32 OC curves for ad hoc beta plan





PART TWO

Background on sampling plans

Statistical appendix

4 Background to Acceptance Sampling Plans

4.1 Attributes plans

Two class attributes plans are based on the binomial distribution; for the plan (n, c) the probability of acceptance is given by:

$$Prob\ acceptance = \sum_{k=0}^{c} \binom{n}{k} p^k (1-p)^{n-k}$$

where p is the percentage nonconforming in the lot. This formula can be used to calculate the probability of acceptance for any level nonconforming p, to construct the operating characteristic.

This expression can be conveniently evaluated using the Excel function BINOM.DIST().

Example, the probability of accepting a lot in which p= 10% of the items are nonconforming, using the sampling plan (n=10, c=1) is given by.

BINOM.DIST(1,10,0.1,TRUE) = 0.736

or by the formula:

$$p_{acc} = \binom{10}{0} 0.9^{10} + \binom{10}{1} 0.9^9 \times 0.1 = 0.9^{10} + 10 \times 0.9^9 \times 0.1 = 0.736$$

However, if the level nonconforming varies between lots, this OC curve will not properly reflect the longer-term probability of acceptance with this plan; one solution due to Calvin [14] is to describe the variation in the level conforming by a beta distribution, in which case the long-term probability of acceptance will be given by a Polya distribution.

In general, if both the consumer's and producer's risks are specified in the design of the plan, as might be appropriate for non-food safety characteristics such as commodity defects, it is unlikely that the acceptance numbers, the c values, will be zero. It should be noted that rather large sample sizes (and large acceptance numbers) might be needed for plans where the operating ratio (CRQ/PRQ) is small.

Section 2.4.2 contains an Excel file that allows the calculation of 95% confidence intervals for the percentage nonconforming or the number of defects in a lot when nonconforming items have been found in the sample.

The lower 95% confidence limit (LCL) for the proportion of nonconforming items in the lot can be calculated using the Excel formula:

LCL = BETAINV(0.025, x, n-x+1),

and the upper 95% confidence limit by:

UCL = BETAINV(0.975, x+1, n-x)

where x is the number of nonconforming items observed in the sample and n is the total sample size.

Similarly, the lower 95% confidence limit (LCL) for the number of defects in the lot can be calculated using:

LCL = 2*GAMMA.INV(0.025, x, 0.5),

and the upper 95% confidence interval by:

UCL = 2*GAMMA.INV(0.975,x+1,0.5)

4.2 Variables plans

In the case of variables plans, once PR, CR, PRQ and CRQ have been specified, then the sample size n and the acceptance constant k can be calculated as follows:

$$k = \frac{z_{1-PR} \cdot z_{1-CRQ} - z_{1-PRQ} \cdot z_{CR}}{z_{1-PR} - z_{CR}}$$

where, for $0 < p < 1$, z_p denotes the one-sided quantile of a standard normal distribution, i.e.

$$\mathcal{P}(X \leq z_p) = p$$

for

$$X \sim \mathcal{N}(0,1).$$

In Excel, these quantiles can be calculated by means of the NORM.S.INV(p) function.

For the case that the lot standard deviation is known (σ method), the sample size can be determined as follows:

$$n = \left(\frac{z_{CR} - z_{1-PR}}{z_{1-CRQ} - z_{1-PRQ}} \right)^2$$

For the case that the lot standard deviation is not known (s method), the above expression for n must be multiplied by the factor $1 + \frac{k^2}{2}$.

The derivation of this concept is quite instructive and provided in Section 5.1.2.

4.2.1 Basis for calculations in App1

Firstly, when measurement uncertainty is negligible, the probabilities of acceptance for the variables sampling plans can be calculated using the formulas above, shown in terms of Excel functions as follows:

Known standard deviation 'sigma'

$$Prob.acc = NORMSDIST((NORMSINV(1 - theta) - k) * SQRT(n), TRUE)$$

Unknown standard deviation (sigma unknown, estimated from the inspection data):

$$Prob.acc = NORMDIST((NORMSINV(1 - theta) - k) * SQRT(n)/SQRT(1 + k * k/2))$$

where:

n is the number of samples,

k is the acceptability constant,

theta (θ) is the level nonconforming in the lot at which the probability of acceptance is to be calculated,

NORMSDIST() is the cumulative standard normal distribution function, and

NORMSINV() is the inverse of the standard normal distribution function.

There is no exact solution when the between laboratory standard deviation is non-negligible, so one must rely on an approximation. Wetherill [29] uses the following method based on a normal approximation for the negligible measurement uncertainty case.

$$p_{acc} = pr(\bar{x} + k \cdot s \leq U) = pr(\bar{x} + k \cdot s - \mu - k \cdot \sigma \leq U - \mu - k \cdot \sigma)$$

where k is the acceptability constant, U the upper specification limit and σ the lot standard deviation.

Making the substitution for $U - \mu = \sigma \cdot NORMSINV(1 - \theta)$, and standardising to a standard normal random variable Z~N(0,1) using the well-known normal approximations for the expected value and standard deviation (uncertainty) of a standard deviation:

$$E(s) = \sigma \; and \; var(s) = \frac{\sigma^2}{2(n-1)} \approx \frac{\sigma^2}{2n}$$

we get:

$$p_{acc} = pr\left(Z \leq \frac{NORMSINV(1-\theta) - k}{\sqrt{\frac{\sigma^2}{n} + \frac{k^2}{2n}}} \right) = NORMSDIST\left(\frac{NORMSINV(1-\theta) - k}{\sqrt{\frac{\sigma^2}{n} + \frac{k^2}{2n}}} \right)$$

This expression can be extended to allow for measurement uncertainty and any offsets used to allow for non-negligible between laboratory MU.

$$p_{acc} = pr\left(Z \leq \frac{NORMSINV(1-\theta) - k}{\sqrt{\sigma_b^2 + \frac{\sigma_r^2}{n} + \frac{\sigma^2}{n}}} \right) = NORMSDIST\left(\frac{NORMSINV(1-\theta) - k}{\sqrt{\sigma_b^2 + \frac{\sigma_r^2}{n} + \frac{\sigma^2}{n}}} \right)$$

In this expression the uncertainty of the calculated mean value depends on the repeatability standard deviation σ_r and the between laboratory standard deviation σ_b.

If an offset ($offset = \; q.\sigma_b$) is applied to allow for between laboratory measurement uncertainty, the probability of acceptance becomes.

$$p_{acc} = pr\left(Z \leq \frac{NORMSINV(1-\theta) - k - q.\sigma_b}{\sqrt{\sigma_b^2 + \frac{\sigma_r^2}{n} + \frac{\sigma^2}{n}}}\right) = NORMSDIST\left(\frac{NORMSINV(1-\theta) - k - q \cdot \sigma_b}{\sqrt{\sigma_b^2 + \frac{\sigma_r^2}{n} + \frac{\sigma^2}{n}}}\right)$$

where q is the multiplier of between laboratory standard deviation in the offset. Allowance for method biases can be made in the same way, subtracting the method bias from the numerator in the equation.

4.3 Calculating measurement uncertainty from precision estimates

The calculation of the probability of acceptance and the form of the acceptability criterion must take account of how the sampling and analytical components of measurement uncertainty are affected by any compositing, or by averaging of results performed as part of the overall sampling, subsampling, sample preparation and analytical procedures for the plan.

The following examples show the basic principles, where:

σ_r is the repeatability standard deviation

σ_L is the standard deviation representing the laboratory bias

σ_R is the reproducibility standard deviation

u is the standard measurement uncertainty

σ is the lot standard deviation

We distinguish the following 5 cases:

1. A single sample (increment) taken from a lot.

   $$u = \sigma_R = \sqrt{\sigma_L^2 + \sigma_r^2}$$

   This is the analytical component of measurement uncertainty.

2. A single taken from a lot, interpreting the result as the average level of the lot.

   $$u = \sqrt{\sigma^2 + \sigma_L^2 + \sigma_r^2}$$

3. n samples taken from lot, and tested, and the results averaged to provide an estimate of the average level

   $$u = \sqrt{\frac{\sigma^2}{n} + \sigma_L^2 + \frac{\sigma_r^2}{n}}$$

4. A composite of n subsamples is tested once to provide an estimate of the average level.

   $$u = \sqrt{\frac{\sigma^2}{n} + \sigma_L^2 + \sigma_r^2}$$

5. n samples taken from the lot, each is tested m times, and results averaged to estimate the average level

   $$u = \sqrt{\frac{\sigma^2}{nm} + \sigma_L^2 + \frac{\sigma_r^2}{nm}}$$

4.4 Combined attributes-variables plans

It is possible to modify the acceptance criterion for variables plans by including an additional requirement on the individual analytical results, typically that none of the results should exceed the specification limit. This leads to a combined attributes-variables plan.

This additional requirement will reduce the probability of acceptance, the decrease is obviously greater at higher levels nonconforming.

Refer also to CXG 50-2004 section 5.2.9 Fractional Nonconformance plans that are another type of combined attribute-variables plans.

4.5 Multi-stage plans

In multi-stage plans the inspection is carried out in several stages, most commonly two-stage plans are used. At each stage, a specified number of samples is taken and tested, although practically, a larger number of samples may be taken at the first stage in case they need to be tested at Stage 2:

- if the results meet the acceptance criterion for that stage, the lot is accepted without any further inspections needed.
- If the results meet the rejection criterion for that stage, the lot is rejected.
- If neither criterion is met, sampling continues to the next stage [if there is one].

The following example shows how a double attributes sampling plan operates. This example is based on a producer's risk of 5% at a quality level of 1% nonconforming and a consumer's risk with a 10% at a quality level of 5% nonconforming.

Stage 1:

n1 = 88 samples are taken at random from a lot.

- If at most one nonconforming item was found, then accept the lot.
- If four or more nonconforming items were found, then reject the lot.
- If two or three nonconforming items were found, proceed to Stage 2.

Stage 2:

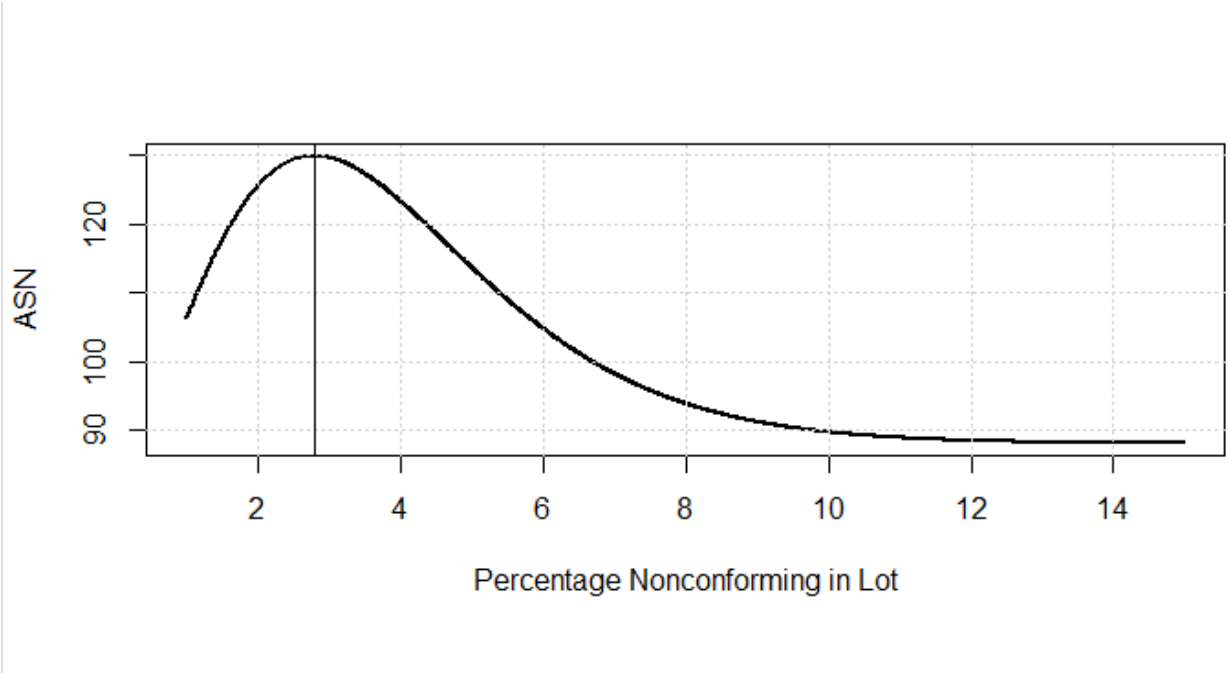
n2 = 88 additional samples are taken at random from a lot.

- If at most four nonconforming items were found in both stages, then accept the lot.
- If five or more nonconforming items were found in both stages, then reject the lot.

The main advantage of multi-stage plans is the reduction in the overall average sample size relative to the sample size for a single stage plan for the same control of producer's and consumer's risks; lots of very good quality lots are accepted, and lots of very poor quality are rejected, at the first stage. However, a disadvantage of multi-stage plans is the increased administrative and other costs and the possible delay making a final decision on the disposition of marginal lots.

The maximum Average Sample Number (ASN) for the double sampling plan is n1 + n2 = 130 approximately, at a quality level of about 2.8% nonconforming, but the ASN is considerably less at other levels nonconforming.

Figure 33: Two stage plan – expected average sample number (ASN) by level nonconforming



The corresponding single sample plan is (n=132, c=3).

4.6 Lots consisting of bulk materials

This section provides information on the design of plans for bulk materials, particularly relating to plans to assess compliance of the average level to a maximum or minimum limit that are often used for chemical contaminants. In particular, this section provides:

- The scope, some understanding, few basic properties of sampling plans for bulk materials, and 'motivation' for their use.
- Review of ISO 10725, Acceptance sampling plans and procedures for the inspection of bulk material, for assessment of the average level.
- Acceptance sampling for aflatoxins, in particular, the plans described in Whitaker's work, including explanation of tables from the *General standard for contaminants and toxins in food and feed* (CXS193-1995).

Detailed guidance is not provided on the following topics because of the statistical complexity involved; it is recommended that users seek assistance from a statistician:

- Characterizing the heterogeneity in bulk sampling, partitioning total heterogeneity in various components.
- The design of sampling plans for bulk materials to assess compliance against a minimum or maximum limits.

Note: plans for bulk materials are generally one-off i.e. applicable to a specific situation or a limited range of situations, so are not necessarily transferable to other matrices or characteristics.

Bicking (1970) defines the following process for the design of sampling plans for bulk materials:

1. State the problem for which an estimate of the average value is required.
2. Collect information on the relevant properties of the material (averages and components of variance of the properties)
3. Identify the components of variation in the overall sampling and testing process that might be relevant to the intended sampling plan options.
4. Estimate these components using a suitable statistical design (often 'hierarchical' designs are used)
5. Consider various approaches, taking account of cost, precision and difficulties.
6. Evaluate these plans in terms of the cost of sampling and testing, delay, supervisory time and convenience.
7. Calculate the standard deviations associated with the estimates of the average levels for these plans and their uncertainty (degrees of freedom).
8. Provisionally, select a plan from one of these approaches.
9. Reconsider the preceding steps.

The acceptance criterion will be of the form: $\bar{x} + k \cdot S \leq USL$ for upper specification limit USL for the average level, where S is the standard error (standard deviation) of the estimate $\bar{x}$ of the mean level and k is the multiplier[8] of the standard error in the acceptance criterion. Note that this multiplier is different from the acceptability constant used with variables plans used to control the percentage nonconforming.

If the cost associated with the initial sampling step was low the plan could be economized by taking more increments to improve the precision of the estimate of the average level.

4.6.1 Example: Variables Plans for homogeneous lots—negligible MU

As an alternative to the sampling plan discussed in section 4.2.1, the contents of the cans in the lot could be considered as a bulk material so that the assessment relates to the contents of the lot as a whole rather than to compliance at the can level. This approach would also allow plans for the beta distribution to be used, with the possible benefit of being able to carry out an inspection based on a single test of a composite sample, if the characteristic inspected was a compositional proportion.

[8] The multiplier is based on the Student's T-distribution for which the number of degrees of freedom has to be determined using a statistical procedure. See Schilling for an example and the CXG 54 information document for more information, but the details of this procedure are outside the scope of these guidelines.

If both producer's and consumer's risks are specified in the design of the sampling plan, the design process is the same as in section 4.2.1, noting that the sample size does not depend on the lot size if both risks are specified.

To avoid repetition, the full details of the plan in section 4.2.1 are not repeated here, but in summary that plan was based on the following assumptions:

The CRQ was chosen as 10% and the PRQ as 2.5%, with the CR and PR at 10% and 5% respectively. This means that the plan will have:

- A 10% chance of accepting a lot in which 10% of the product is nonconforming.
- A 5% chance of rejecting a lot in which 2.5% of the product is nonconforming.

The lot standard deviation was assumed to be $\sigma = 0.3$ and the measurement uncertainty was considered negligible, leading to the plan (n=43, k=1.59), i.e. 43 samples need to be taken from the lot and tested, with the lot is accepted with respect to milkfat provided the average and the standard deviation of the results meet the acceptance criterion:

$$\bar{x} - 1.59 \times s \geq 26$$

where:

- $\bar{x}$ is the average of the 43 individual results and 's' their standard deviation,

4.6.2 Example: Variables plan with non-negligible MU with no laboratory bias

Refer to section 4.2.2 – the process for the design of the sampling plan is the same.

4.6.3 Example: Variables plan with non-negligible MU with laboratory bias

Refer to section 4.2.3 – the process for the design of the sampling plan is the same as described in that section.

4.6.4 ISO 10725

This standard follows the work by Schilling & Neubauer and is discussed in their book [3], available on-line [5].

ISO10725 describes procedures for the design of sampling plans for the assessment of the average levels of lots, based on a three-component model:

- A number of increments are taken from the lot and combined to form composite samples.
- Test portions are taken from each of the well-mixed composite samples.
- Each test portion is tested a number of times.

As well as the variation of each component, the standard also allows for the actual (or relative) costs of each step to be taken into account to obtain cost optimal plans for specified levels of producer's and consumer's risks.

It is assumed that the standard deviations and the costs of each of the steps are known but the standard contains procedures to deal with situations where the costs or the standard deviations are not known.

4.6.5 Aflatoxin sampling plans

Introduction

The sampling plans for mycotoxins derived by Whitaker et al. are special cases of plans for bulk materials. Whitaker used 46 years of laboratory data, including some from contaminated lots, to derive Horwitz-type equations for the sampling, subsampling and analytical components of the total variation.

This method cannot be applied for the design of plans for new matrices or for new contaminants for which limited, or possibly unsuitable, historical data is available. In this case the classical approach described in Schilling, that underpins the ISO10725 standard, would be applied. A first step is to quantify the components of variation relevant to the intended sampling procedure using a suitable experimental design.

However, there are potential problems in that not every lot will be contaminated, and contamination might not be found in those lots that are actually contaminated, so that a considerable number of lots might be required for this exercise. Bayesian approaches might provide a way forward.

Following that, a sampling plan can be developed in terms of:

- the number of segments sampled
- the number of samples taken from each segment

- compositing and subsequent subsampling of those samples
- compositing of the subsamples
- the number of laboratory samples taken for testing
- analytical measurement uncertainty

The usual approach involves experimentation with the number of segments sampled, number of 'samples' taken at each stage and the number of results that are produced, the results from which are averaged. These numbers might also be chosen taking account of the cost involved in each operation. The statistical objective of the design process is the find the acceptability constant K in the acceptance criterion:

$$\bar{X} + t \cdot S \leq U$$

where:

U is the average test result

$\bar{X}$ is the average test result, that will be an estimate of the overall average level in the lot

S is the standard deviation of the estimate of the average level, usually referred to as the 'standard error'

t is the multiplier of the standard error in the acceptance criterion, a percentage point on the "t-distribution", obtained using a statistical procedure by taking account of the uncertainties of the components of the sampling and measurement variation.

There may be other considerations in the design of plans, not taken into account in Whitaker's work, such as:

- Whether one can make use of an assumed distribution for the characteristic in the lot considering that the behaviour of heterogenous materials cannot normally be explained in terms of a single standard deviation or distribution.
- Is it necessary to use discrete distributions to describe this behaviour given that the composite sampling will cause averaging?
- Does the use of cluster sampling need to be allowed for?

Other issues:

Designers of plans should consider what contamination scenarios they wish to detect, that is, the required probabilities of detection of detecting 'spikes' of contamination that contain certain levels of contamination and are of certain durations.

Compositing strategies should be developed to ensure that 'important' spikes of contamination are not averaged out to the extent that they cannot be found.

Aflatoxin sampling plans

The Horwitz type equations were derived for the three variance components (sample to sample, subsampling and the analytical components of variation) in terms of the average concentration of aflatoxin.

CXS 193 shows the breakdown of the total variation for aflatoxins in tree-nuts, for example, into components S_s^2, S_{sp}^2 and S_a^2, due to sampling, subsampling and testing respectively. It should be noted that provisions for aflatoxins are expressed in terms of the average levels in a lot; these plans employ large offsets between the limits and the levels at which the foods become unsafe to consume in order to provide consumer protection (refer CXG 50-2004 4.3.5 Plans for the average level in the lot).

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

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

The sampling plan is defined in terms of laboratory sample size n_s, test portion size n_{ss} and the number of aliquots n_a, the number of analytical samples taken from each subsample. The information in this table can be used to calculate the uncertainty of the estimated average value obtained using the sampling procedure and

thereby the probability of acceptance. For example, the variance of the estimate of the average level for almonds is given by:

$$S^{2=}\frac{7730 \times 5.759}{n_s}C^{1.561} + \frac{100 \times 0.170}{n_{ss}}C^{1.646} + \frac{0.048}{n_a}C^2$$

This equation implies the following sampling and testing procedure:

1. n_s samples are taken from the lot under consideration.
2. A composite sample is formed.
3. A laboratory sample of size n_{ss} grams is taken from that well-mixed composite.
4. n_a aliquots are taken from that subsample for testing.

In the FAO Mycotoxin plans this procedure can be carried out on more than a single sample, but the results for the different samples are not averaged but compared with the limit separately.

This criterion differs from the usual acceptance criterion for the assessment of compliance of the average level for bulk materials in general that would be of the form:

$$\bar{X} + t \times S \leq USL$$

where S is the uncertainty of the average level, t is the multiplier of the standard deviation in the criterion and USL is the upper limit for the mean.

This is a further example of the use of offsets that, in this case, allow simplification of the acceptance criterion.

CXS 193-1995 describes the operational details of the sampling and testing procedure:

1. A 20kg sample taken (1000 [shelled] almonds per kg) from a lots or part lot (sublot), with a 25-tonne limitation on lot size. These samples should be formed from many smaller increments, each no less than 200g. CXS 193-1995 provides guidance on the number of increments, in terms of sample size.
2. The entire sample is ground to a uniform particle size and thoroughly mixed.
3. A test portion of no less than n_{ss} = 50g is taken from the composite sample
4. A number (n_a) of aliquots is taken for testing.
5. The results from these n_a tests are averaged. However, it appears that CXS 193-1995 assumes only single tests are performed (n_a =1) and that usually one or two different samples might be tested with the lot accepted provided no result exceeds the limit. This leads to different probabilities of acceptance, depending on the number of samples that are taken.

Example – Shelled Almonds for further processing

Suppose the average concentration of aflatoxins in the lot was C = 8 µg/kg and ns = 20000, 20 kg @ 1000 shelled nuts per kg, were taken as a sample, and this sample was ground, and well-mixed composite formed. If a subsample of 50g was taken and a single aliquot (n_a=1) tested, the standard deviation S representing the uncertainty of the average level would be:

$$S^{2=}\frac{7730 \times 5.759}{20 \times 1000}8^{1.561} + \frac{100 \times 0.170}{50}8^{1.646} + \frac{0.048}{1}8^2 = 70.67$$

Giving S = 8.41. The first component representing the sample-by-sample variation is much larger than the other two.

Comments

The web-based FAO Mycotoxin Tool [2] for the evaluation of sampling plans is provided at http://tools.fstools.org/mycotoxins/.

This tool allows for only a single component of measurement uncertainty; there is no allowance for bias when multiple tests are performed. The tool allows users to select whether “within lab” or “among lab” variation is used, with the among laboratory variation equal to twice the within laboratory figure. The tables below show the within laboratory variance.

The sampling component is included using an assumed distribution, most often the negative binomial, a discrete distribution to allow contamination at the individual particle (e.g. grain) or sample level to be modelled - due to the small percentages (typically less than 1%) of contamination and the extreme distribution of contamination within lots very large sample sizes are needed to estimate the distribution.

The decision rule for Almonds for further processing in CXS 193 is that the lot is accepted "if the aflatoxin result is less than 15µg/kg in both samples…", so that each individual result is classified as pass or fail with respect to the limit. However, as the analytical component is small relative to the sampling component, this does not seem to matter.

To calculate probabilities of acceptance (and the OC curve) we need to know the distribution of the sample-to-sample variation within a bulk lot. As above, Whitaker assumed, mostly, that the sample-to-sample variation follows a negative binomial distribution.

The negative binomial distribution is used in situations where the variation is more extreme than the binomial; it is defined in terms of an average value and a variance.

$$Average = \mu;\ Variance = S^2 = \mu + \frac{\mu^2}{k}$$

where ***k*** is the dispersion factor that allows for the extra variation.

To work out the theoretical probability of acceptance at a concentration C of aflatoxin Whitaker used the 'method of moments', equating the theoretical concentration C to the mean and the estimate of S^2 to the variance, i.e.

$$\mu = C\ and\ S^2 = C + \frac{C^2}{k}$$

The second equation is be solved to determine k and the probability of acceptance calculated. This process must be repeated for each value of C, as S^2 depends on C.

Components of Variance for Aflatoxin Sampling Plans

Study #	Mycotoxin	Commodity	References	Variance			Mycotoxin Test Procedure				Distribution Among Sample Test Results
				Sampling (S^2_s)	Sample Preparation (S^2_{sp})	Analytical (Within Lab) (S^2_a)	Laboratory Sample Size (ns)	Comminuted Test Portion Size (nss)	Number of Aliquots (na)	Concentration (C)	
1	Aflatoxin	Shelled Peanuts	1, 2, 3, 34	$(10,644/ns)9.19C^{1.335}$	$(275/nss)0.294C^{1.729}$	$(1/na)0.083C^{1.664}$	Number of shelled kernels (1,952ker/kg)	Mass (g) Dry Comminution USDA mill powder	Number of aliquots quantified by HPLC from Ref 34	ng/g (ppb) aflatoxin total	Negative Binomial
2	Aflatoxin	Cottonseed	4, 5, 6, 34	$(43,200/ns)6.776C^{1.344}$	$(200/nss)0.180C^{1.3308}$	$(1/na)0.086C^{1.667}$	Number of seed (Hull removed) (19,031ker/kg)	Mass (g) Dry Comminution USDA mill powder	Number of aliquots quantified by HPLC from Ref 34	ng/g (ppb) aflatoxin either total or B1	Negative Binomial
3	Aflatoxin	Harvested Inshell Peanuts (Farmer's Stock)	7, 8, 9	$(3713/ns)37.607C^{1.161}$	$(100/nss)2.887C^{1.401}$	$(1/na)0.083C^{1.664}$	Number of inshell pods (882pods/kg)	Mass (g) Dry Comminution VCM Paste	Number of aliquots quantified by HPLC from Ref 34	ng/g (ppb) aflatoxin total	Negative Binomial
4	Aflatoxin	Shelled Corn	10, 11, 12	$(3,390/ns)11.36C^{0.98}$	$(50/nss)1.254C^{1.27}$	$(1/na)0.143C^{1.16}$	Number of shelled kernels (3,000ker/kg)	Mass (g) Dry Comminution Romer Powder	Number of aliquots quantified by HPLC	ng/g (ppb) aflatoxin total	Compound Gamma Used Negative Binomial
5	Aflatoxin	Shelled Almonds	13, 14, 15	$(7,730/ns)5.759C^{1.581}$	$(100/nss)0.170C^{1.646}$	$(1/na)0.0041C^{1.986}$	Number of shelled kernels (773ker/kg)	Mass (g) Dry Comminution VCM Paste	Number of aliquots quantified by HPLC	ng/g (ppb) aflatoxin either total or B1	Negative Binomial
6	Aflatoxin	Inshell Almonds	13, 14, 15	$(7,730/ns)5.759C^{1.581}$	$(100/nss)0.170C^{1.646}$	$(1/na)0.0041C^{1.986}$	Number of Inshell Nuts (309nuts/kg) Shell/ker Ratio = 60/40	Mass (g) Dry Comminution VCM Paste	Number of aliquots quantified by HPLC	ng/g (ppb) aflatoxin either total or B1	Negative Binomial
7	Aflatoxin	Shelled Hazelnuts	15, 16, 17	$(10,000/ns)4.291C^{1.609}$	$(50/nss)0.021C^{1.646}$	$(1/na)0.0028C^{1.990}$	Number of shelled kernels (1,000ker/kg)	Mass (g) Dry Comminution VCM Paste	Number of aliquots quantified by HPLC	ng/g (ppb) aflatoxin either total or B1	Negative Binomial
8	Aflatoxin	Inshell Hazelnuts	15, 16, 17	$(10,000/ns)4.291C^{1.609}$	$(50/nss)0.021C^{1.646}$	$(1/na)0.0028C^{1.990}$	Number of Inshell nuts (500Nuts/kg) Shell/Ker Ratio = 50/50	Mass (g) Dry Comminution VCM Paste	Number of aliquots quantified by HPLC	ng/g (ppb) aflatoxin either total or B1	Negative Binomial
9	Aflatoxin	Shelled Pistachios	15	$(8,000/ns)7.913C^{1.475}$	$(25/nss)2.334C^{1.622}$	$(1/na)0.0368C^{1.598}$	Number of Shelled Kernels (1,600ker/kg)	Mass (g) Dry Comminution VCM Paste	Number of aliquots quantified by HPLC	ng/g (ppb) aflatoxin either total or B1	Negative Binomial
10	Aflatoxin	Inshell Pistachios	15	$(8,000/ns)7.913C^{1.475}$	$(25/nss)2.334C^{1.622}$	$(1/na)0.0368C^{1.598}$	Number of Inshell Nuts (800nuts/kg) Shell/Ker Ratio = 50/50	Mass (g) Dry Comminution VCM Paste	Number of aliquots quantified by HPLC	ng/g (ppb) aflatoxin either total or B1	Negative Binomial

Study #	Mycotoxin	Commodity	References	Variance			Mycotoxin Test Procedure				Distribution Among Sample Test Results
				Sampling (S^2_s)	Sample Preparation (S^2_{sp})	Analytical (Within Lab) (S^2_a)	Laboratory Sample Size (ns)	Comminuted Test Portion Size (nss)	Number of Aliquots (na)	Concentration (C)	
11	Aflatoxin	Shelled Brazil Nuts	15	$(1{,}850/ns)4.862C^{1.889}$	$(50/nss)0.0306C^{0.632}$	$(1/na)0.0164C^{1.117}$	Number of Shelled Kernels (185ker/kg)	Mass (g) Slurry (Water/Ker 1/1) Comminution	Number of aliquots quantified by HPLC	ng/g (ppb) aflatoxin either total or B1	Negative Binomial
12	Aflatoxin	Inshelled Brazil Nuts	15	$(1{,}850/ns)4.862C^{1.889}$	$(50/nss)0.0306C^{0.632}$	$(1/na)0.0164C^{1.117}$	Number of Inshelled Nuts (93Nuts/kg) Shell/Ker Ratio=50/50	Mass (g) Slurry (Water/Ker 1/1) Comminution	Number of aliquots quantified by HPLC	ng/g (ppb) aflatoxin either total or B1	Negative Binomial
13	Aflatoxin	In Field Ear Corn	18	$(600/ns)8.919C^{2.230}$	$(50/nss)1.254C^{1.27}$	$(1/na)0.143C^{1.16}$	Number of shelled kernels per ear 200 g ker/ear (3,000ker/kg)	Mass (g) Dry Comminution Romer Powder	Number of aliquots quantified by HPLC	ng/g (ppb) aflatoxin B1	Negative Binomial
14	Aflatoxin	In Field Farmer's Stock Peanuts	19	$(116/ns)17.056C^{1.6686}$	$(100/nss)2.887C^{1.401}$	$(1/na)0.083C^{1.654}$	Number of inshell pods (882pods/kg)	Mass (g) Dry Comminution VCM Paste	Number of aliquots quantified by HPLC from Ref 34	ng/g (ppb) aflatoxin total	Negative Binomial
15	Aflatoxin	Powdered Ginger in Capsules	20	$(5/ns)0.138C^{1.0}$	No Test Portion, Entire Sample Extracted	$(1/na)0.0178C^{1.70}$	5 g Laboratory Sample is also the 5 g Test Portion	No Test Portion, Entire Sample Extracted	Number of aliquots quantified by HPLC	ng/g (ppb) aflatoxin total	Normal
16	Aflatoxin	Powdered Ginger in 1-Lb Bags	21	$(5/ns)4.218C^{1.0}$	No Test Portion, Entire Sample Extracted	$(1/na)0.00349C^{1.70}$	5 g Laboratory Sample is also the 5 g Test Portion	No Test Portion, Entire Sample Extracted	Number of aliquots quantified by HPLC	ng/g (ppb) aflatoxin total	Normal
17	Aflatoxin	Dried Figs	Not Published	$(590/ns)2.219C^{1.433}$	$(55/nss)0.012C^{1.465}$	$(1/na)0.006C^{1.368}$	Number of dried Figs (59 Figs/kg)	Mass (g) Slurry (Water/Ker 1/1) Comminution	Number of aliquots quantified by HPLC	ng/g (ppb) aflatoxin total	Negative Binomial
18	Fumonisin	Shelled Corn	22, 23, 24	$(3{,}390/ns)0.033C^{1.75}$	$(25/nss)0.011C^{1.59}$	$(1/na)0.014C^{1.44}$	Number of shelled kernels (3,000ker/kg)	Mass (g) Dry Comminution Romer Powder	Number of aliquots quantified by HPLC	ug/g (ppm) Fumonisin either B1, B2, B3 or total	Compound Gamma Used Lognormal
19	Deoxynivalenol (DON)	Shelled Corn	25	$(3{,}000/ns)0.202C^{1.923}$	$(50/nss)0.0193C^{1.140}$	$(1/na)0.0036C^{1.507}$	Number of shelled corn kernels (3,000ker/kg)	Mass (g) Dry Comminution Romer - 25 g	Number of aliquots quantified by Romer - Malone HPLC	ug/g (ppm) DON	Lognormal (not published)
20	Deoxynivalenol (DON)	Wheat	26	$(13{,}620/ns)0.026C^{0.833}$	$(25/nss)0.066C^{0.833}$	$(1/na)0.026C^{0.833}$	Number of raw wheat kernels (30,000ker/kg)	Mass (g) Dry Comminution Romer 25 g	Number of aliquots quantified by Romer FluoroQuant	ug/g (ppm) DON	Lognormal (not published)

Study #	Mycotoxin	Commodity	References	Variance			Mycotoxin Test Procedure				Distribution Among Sample Test Results
				Sampling (S^2_s)	Sample Preparation (S^2_{sp})	Analytical (Within Lab) (S^2_a)	Laboratory Sample Size (ns)	Comminuted Test Portion Size (nss)	Number of Aliquots (na)	Concentration (C)	
21	Deoxynivalenol (DON)	Barley	27	$(77{,}000/ns)0.0122C^{0.947}$	$(50/nss)0.003C^{1.956}$	$(1/na)0.0108C^{1.055}$	Number of raw barley kernels (30,800ker/kg)	Mass (g) Dry Comminution Romer 50 g	Number of aliquots quantified by Romer FluoroQuant	ug/g (ppm) DON	Lognormal (not published)
22	Ochratoxin A (OTA)	Green Coffee Beans	28, 29, 30	$(1{,}500/ns)1.350C^{1.090}$	$(25/nss)0.272C^{1.646}$	$(1/na)0.008C^{1.605}$	Number of beans (1,500ker/kg)	Mass (g) Dry Comminution VCM Paste	Number of aliquots quantified by HPLC	ng/g (ppb) OTA total	Lognormal
23	Ochratoxin A (OTA)	Powdered Ginger in Capsules	20	$(5/ns)0.108C^{1.0}$	No Test Portion, Entire Sample Extracted	$(1/na)0.00654C^{1.70}$	5 g Laboratory Sample is also the 5 g Test Portion	No Test Portion, Entire Sample Extracted	Number of aliquots quantified by HPLC	ng/g (ppb) OTA total	Normal
24	Ochratoxin A (OTA)	Powdered Ginger in 1-Lb Bags	21	$(5/ns)1.336C^{1.0}$	No Test Portion, Entire Sample Extracted	$(1/na)0.00146C^{1.70}$	5 g Laboratory Sample is also the 5 g Test Portion	No Test Portion, Entire Sample Extracted	Number of aliquots quantified by HPLC	ng/g (ppb) OTA total	Normal
25	Ochratoxin A (OTA)	Oats	Not Published	$(55{,}796/ns)1.440C^{1.278}$	$(100/nss)0.0074C^{1.638}$	$(1/na)0.0103C^{1.58}$	Number of raw oat kernels (27,898ker/kg)	Mass (g) Dry Comminution Retsch SR300 #20 Screen	Number of aliquots quantified by HPLC	ng/g (ppb) OTA total	Negative Binomial
26	Ochratoxin A (OTA)	Wheat	Not Published	$(60{,}180/ns)1.557C^{1.132}$	$(5/nss)0.207C^{1.152}$	$(1/na)0.0204C^{1.866}$	Number of raw wheat kernels (30,090ker/kg)	Mass (g) Dry Comminution Retsch SR300 #20 Screen	Number of aliquots quantified by HPLC	ng/g (ppb) OTA total	Negative Binomial
27	FAPAS among lab variability		31			$(1/na)0.0484C^{2.000}$					
28	Horwitz among lab variability (ppb)		32,33			$(1/na)0.2048C^{1.70}$					
29	Whitaker, Horwitz, Analytical Variances - TLC, Immuno, HPLC		34			Among Lab = 2*Within Lab					

Study 3 the sampling variance was calculated by subtracting analytical and sample prep variances from total variances for each of the three (2.26, 4.21, and 6.91 kg) sample sizes.

Studies 13 and 14 measured only total variance. Used sample prep and analytical variances from studies 4 and 3, respectively.

Study 28 analytical variance was determined for various methods, mycotoxins, and commodities using data base from Horwitz Ref 32

4.6.6 General plans for assessment against minimum or maximum levels

One approach, more suited to food safety than commercial characteristics on the grounds of fairness, is to use offsets and assess compliance of lots against the average level. This has the considerable advantage of simplicity.

However, these plans are also important in a commercial context where one might, for example, wish to provide assurance about the average level of protein in a lot of grain that is to be further processed, for example to make flour.

In general, however, the design of sampling plans for bulk materials to assess compliance against a minimum or maximum limit is difficult statistically and no information is included in this information document.

4.7 Plans for microbiological assessment

Plans used for the assessment of lots for microbiological characteristics, often referred to as *Microbiological Criteria*, frequently employ 2-class attributes plans that require n=5 samples to be taken. These plans are suitable only for characteristics where the measurements are counts and there are adequate offsets between the limits used in these plans and the levels at which foods are considered to be unsafe.

If the offsets are not adequate there could be a higher rate of acceptance of contaminated product. Testing of pathogens is usually carried out using detection tests that produce presence or absence outcomes; in this case there are no offsets between the limits (zero) and the levels at which foods become unsafe to consume. For this reason, the use of (n=5, c=0) plans for pathogens is inadvisable; this is also the reason why sampling plans for pathogenic characteristics require much larger sample numbers and a greater total amount of sample is tested. Use of a larger numbers of samples also provides some safeguard against potentially inhomogeneous contamination within lots. Some examples of microbiological criteria are given in the *Code of hygienic practice for powdered formulae for infants and young children* (CXC 66-2008) that contains the following microbiological criteria (see Codex definition at the end of the section) along with some points on the Operating Characteristic:

Microorganism	n	c	m	Class Plan
Cronobacter sp.	30	0	0/10g	2
Salmonella	60	0	0/25g	2

Points on the operating characteristic have been calculated by *Zweiterung et al*. [10] assuming a Poisson-lognormal distribution, being a Poisson distribution whose mean varies according to a lognormal distribution.

Cronobacter:

- At a mean concentration of 1 cfu/340g, the probability of detection is 95%, assuming a standard deviation [for the lognormal distribution] of sd = 0.8.
- At a mean concentration of 1 cfu/100g, the probability of detection is 99%, assuming a standard deviation of sd = 0.5.

Salmonella:

- At a mean concentration of 1 cfu/526g, the probability of detection is 95%, assuming a standard deviation of sd = 0.8.

4.7.1 3-class attribute plans

Refer CXG 50-2004 section 4.2.6

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

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

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

Lots are accepted provided:

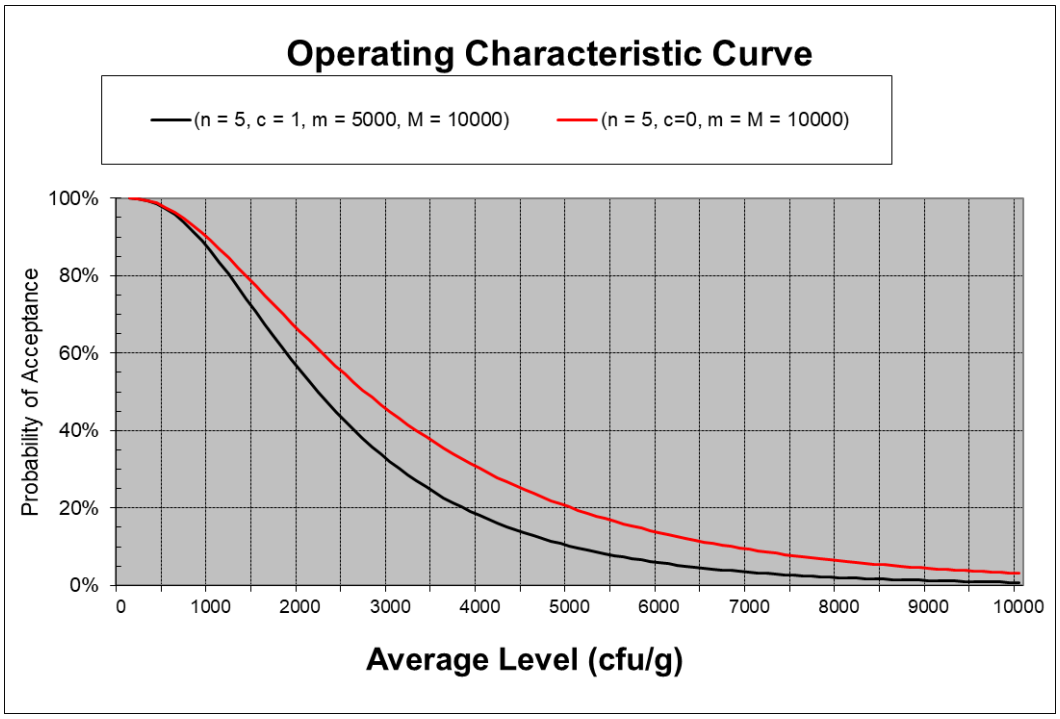
- None of the n samples is poor, having levels exceeding M
- At most c of the samples are marginal, with levels between m and M.

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

Evaluation of these plans generally requires an assumption about the underlying distribution of the identified characteristic, the lognormal distribution is commonly used for microbiological characteristics for counts occurring at higher levels, whereas the Poisson distribution is often used for counts at lower levels.

The following plot shows the operating characteristic curves for a two-class plan (n=5, c=0, m=10000) and a three-class plan (n=5, c=1, m=5000, M=10000); it shows that the three-class plan is more stringent despite allowing one result to be marginal.

Figure 34 OC curves - three class attributes plans



Although the plans mentioned in this section are used primarily in microbiological inspections, they are nevertheless useful in other applications such as those where acceptance is decided in terms of the total defects found in the sample, with the possibility that an item selected in the sample may contain more than one defect. One possible application of these plans is to the inspection of herbs and spices for insects or insect parts.

5 Statistical appendix

5.1 Background for the main (attributes & variables) sampling plans

5.1.1 Calculating acceptance probabilities – attributes plans

Attributes Plans are based on the binomial distribution (two-class plans) or the multinomial distribution, an extension of the binomial distribution, for three or more class plans.

The probability of acceptance for the two-class binomial model is given by:

$$prob\_acceptance = \sum_{k=0}^{c} \binom{n}{k} p^k (1-p)^{n-k}$$

where:

n is the sample size, the number or items or samples taken

c is the acceptance number, the maximum number of nonconforming items permitted for acceptance of the lot

p is the percentage nonconforming in the lot

For any given expression which depends on a variable k, the symbol

$$\sum_{k=value_1}^{value_n} \text{expression}(k)$$

means 'the sum of' the expression evaluated at

$$k = value_1,\ k = value_1 + 1,\ k = value_1 + 2,\ \ldots,\ k = value_n$$

For example

$$\sum_{k=1}^{5} k^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2$$

The symbol $\binom{n}{k}$ is the binomial coefficient, i.e. it is the number of ways of choosing k items from a total of n items. For example, $\binom{5}{1} = 5$ since there are 5 ways of choosing one item from 5 items, viz. Aaaaa, aAaaa, aaAaa, aaaAa and aaaaA, where A represents the item selected.

The design of an attributes sampling plan involves finding the values of the number of samples n and the acceptance number c from the probabilities of acceptance at two specified points on the operating characteristic curve. Typically, these points are chosen as the producer's and the consumer's risk quality levels.

In the case p = PRQ, the probability of acceptance is equal to "one minus the producer's risk"

$$Prob\ acceptance = 1 - PR = \sum_{k=0}^{c} \binom{n}{k} p^k (1-p)^{n-k}$$

and when p = CRQ, the probability of acceptance CR is equal to the consumer's risk

$$Prob\ acceptance = CR = \sum_{k=0}^{c} \binom{n}{k} p^k (1-p)^{n-k}$$

These two equations are usually solved iteratively in a statistical package or using a computer program following the algorithm due to Hailey [15]:

1. Start by setting $n = 0$ and $c = 0$
2. If the probability of acceptance at the CRQ exceeds the specified maximum allowable consumer's risk CR, then increase n by one, and go back to step 2.
3. If the probability of rejection at the PRQ exceeds the specified maximum allowable producer's risk PR, then increase c by one, and go to step 2.

Note that because n and c are integers, and can only increase in steps of one, the actual producer's and consumer's risks in the final plan might not be exactly equal to the producer's and consumer's risks specified in the design of the plan.

Calculation of confidence intervals

Section 3.1.2 discussed ZAN plans and their use in applications such as inspections for foreign matter. The final part of that section described the calculation of confidence intervals for the percentage nonconforming or the number of defects in a lot when at least one nonconforming item or defect has been found in the samples examined.

For the binomial case that relates to the percentage of defective items in the lot overall, the lower and upper limits are calculated using the Excel formulas:

$$LCL = BETA.INV(0.025, c, n - c + 1)$$

and

$$UCL = BETA.INV(0.975, c + 1, n - c)$$

where n is the number of items or samples examined and c is the number of nonconforming items found among those n items.

For the Poisson case that relates to the percentage of defective items in the lot overall, the lower and upper limits are calculated using the Excel formulas:

$$LCL = 2 * GAMMA.INV(0.025, c, 0.5)$$

and

$$UCL = 2 * GAMMA.INV(0.975, c + 1, 0.5)$$

where n is the number of items or samples examined and c is the number of defects found during the inspection.

5.1.2 Derivation of formulas for variables plans

The formulas for k and n are derived as follows for the case of a known lot standard deviation σ and an upper specification limit U.

We use the notation z_p to denote the one-sided quantile of a standard normal distribution, i.e.

$$\mathcal{P}(X \leq z_p) = p$$

for

$$X \sim \mathcal{N}(0,1).$$

The acceptance limit A is defined as

$$A = U - k\sigma$$

We thus have

$$\begin{aligned} U - A &= k\sigma \\ &= \mu_{PRQ} + z_{1-PRQ} \cdot \sigma - \left(\mu_{PRQ} + z_{1-PR} \cdot \frac{\sigma}{\sqrt{n}}\right) \quad \text{Eq. 1} \\ &= \mu_{CRQ} + z_{1-CRQ} \cdot \sigma - \left(\mu_{CRQ} + z_{CR} \cdot \frac{\sigma}{\sqrt{n}}\right) \quad \text{Eq. 2} \end{aligned}$$

The following example and figure illustrate these two equations. Consider the case that we are asked to design a plan with

$$PRQ = 6.5\%$$

$$PR = 5\%$$

$$CRQ = 26\%$$

$$CR = 10\%$$

The corresponding standard normal quantiles are:

$$z_{1-PRQ} = 1.514$$

$$z_{1-PR} = 1.645$$

$$z_{1-CRQ} = 0.643$$

$$z_{CR} = -1.282$$

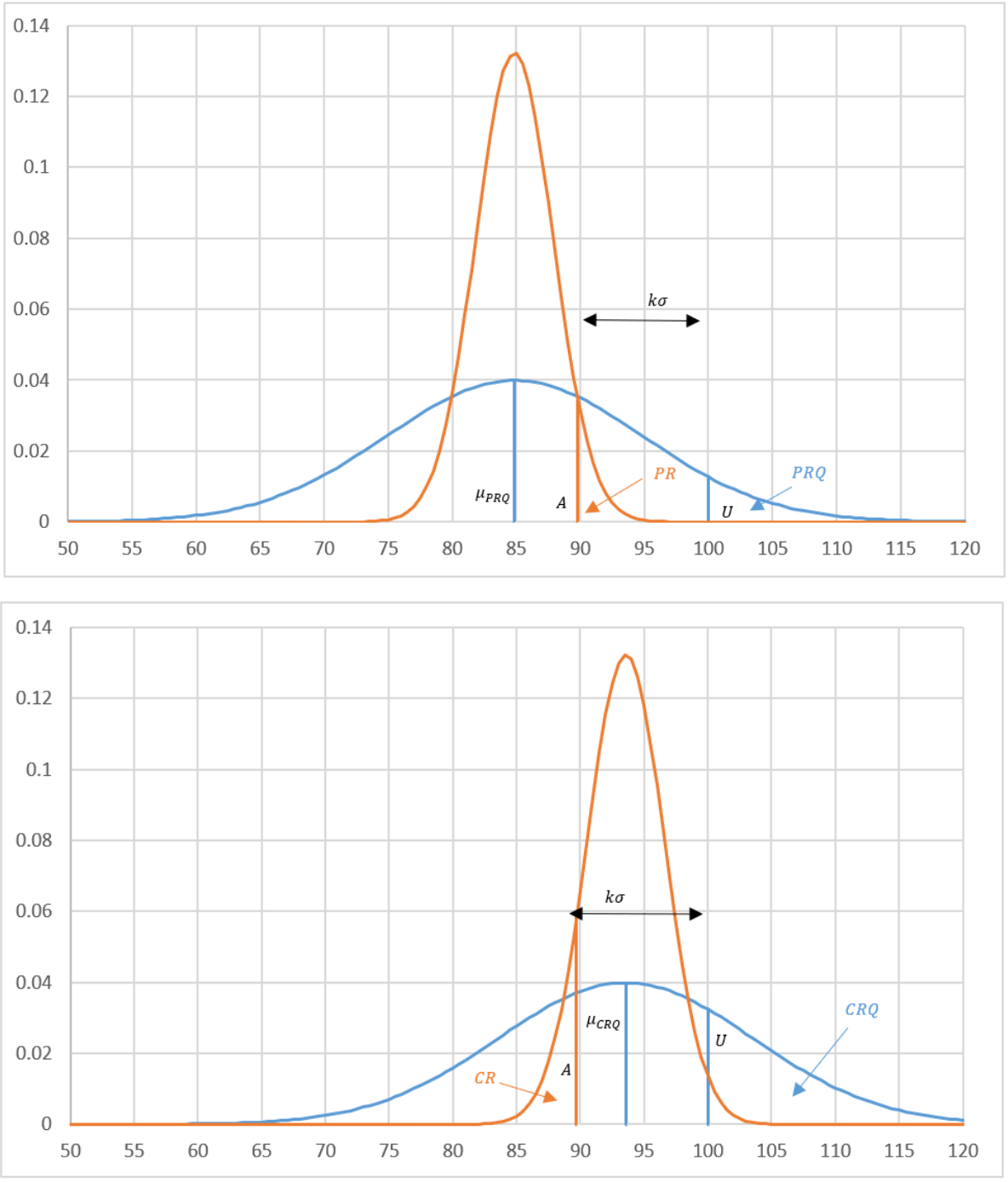
Applying the formulas for n and k (known σ), we obtain

$$n = 11.3$$

$$k = 1.025$$

This is illustrated in the following diagrams. We consider the situation that $U = 100$ (generic unit) and that the lot standard deviation is known with $\sigma = 10$. A lot with quality PRQ will have a mean value (across items) $\mu_{PRQ} \approx 85$. The sample size is $n = 11$. The acceptance limit (for the decision to accept or reject the lot) is calculated as $A = U - k\sigma \approx \mu_{PRQ} + z_{1-PR} \cdot \frac{\sigma}{\sqrt{n}} \approx 90$.

Figure 35: The blue curve represents the distribution of the property of interest in a lot with quality PRQ (top diagram) and in a lot with quality CRQ (bottom diagram) and lot standard deviation $\sigma = 10$. The red curve represents the statistical distribution of the arithmetic mean.



It follows from Equation 1 and Equation 2 that

$$\frac{1}{\sqrt{n}}(z_{CR} - z_{1-PR}) = z_{1-CRQ} - z_{1-PRQ}$$

And hence

$$\sqrt{n} = \frac{z_{1-PR} - z_{CR}}{z_{1-PRQ} - z_{1-CRQ}}$$

As far as k is concerned, it follows from Equation 1 and Equation 2 that

$$k = z_{1-PRQ} - \frac{z_{1-PR}}{\sqrt{n}}$$

$$k = z_{1-CRQ} - \frac{z_{CR}}{\sqrt{n}}$$

Hence, we have

$$\frac{k\sqrt{n}}{z_{1-PR}} = \frac{\sqrt{n} \cdot z_{1-PRQ}}{z_{1-PR}} - 1$$

and

$$\frac{k\sqrt{n}}{z_{CR}} = \frac{\sqrt{n} \cdot z_{1-CRQ}}{z_{CR}} - 1$$

It follows that

$$\frac{k\sqrt{n}}{z_{1-PR}} - \frac{k\sqrt{n}}{z_{CR}} = \frac{\sqrt{n} \cdot z_{1-PRQ}}{z_{1-PR}} - \frac{\sqrt{n} \cdot z_{1-CRQ}}{z_{CR}}$$

and thus

$$k \cdot \left(\frac{z_{CR} - z_{1-PR}}{z_{1-PR} \cdot z_{CR}}\right) = \frac{z_{1-PRQ} \cdot z_{CR} - z_{1-PR} \cdot z_{1-CRQ}}{z_{1-PR} \cdot z_{CR}}$$

From which we obtain

$$k = \frac{z_{1-PR} \cdot z_{1-CRQ} - z_{1-PRQ} \cdot z_{CR}}{z_{1-PR} - z_{CR}}$$

5.1.3 Within-item variability

In general, for a lot consisting of discrete items, there are two sources of variation: between-item variation and within-item variation. In “classical” acceptance sampling plans, there is a tacit assumption that within-item variation is negligible, and that one test result per item is thus sufficient. If non-negligible within-item variation is expected, then it may be necessary to modify the acceptance sampling plan. In particular, it may be necessary to correct the estimate of the lot standard deviation by “subtracting” the within-item component.

At low concentrations and for certain types of products (more specifically, when the presence or absence of the analyte is modelled via a discrete statistical distribution, as may be the case for sufficiently coarse-grained powders), within-item variation may be present even in the case of a “perfect” mixing process. This is due to an irreducible component of variation which remains even in the case of a perfectly homogenous item. This irreducible component is called the fundamental variability and is modelled via the Poisson distribution.

In the presence of fundamental variability, it may be necessary to apply specially developed models to distinguish between- from within-item variation. In particular, two different cases must be distinguished:

Case 1: the exact same quantity of analyte (corresponding to the property of interest, e.g. vitamin D in milk powder) is added separately to each item. Hence, there is no between-item variation, only within-item variation.

Case 2: the analyte is added to the mixing tank and then mixed with the pre-blend powder prior to filling the individual item recipients. In this case there is both between- and within-item variation.

Models for these two cases are discussed in Uhlig et al. (2025) [27].

5.2 Understanding ISO plans

The ISO standards apply a risk-based approach to the design of acceptance sampling plans. In inspection by attributes, it is the product AQL (PRQ) × sample size which informs the design of the plans. In inspection by variables, the plans indexed by AQL (PRQ) aim to achieve a producer's risk which depends on the lot size.

5.2.1 Attributes plans constructed in terms of unity value

In the ISO 2859-1 standard, the plans are constructed in such a manner as to have constant acceptance number values across the diagonals of the sampling plan tables. This section provides a short rationale for this approach.

ISO 2859-1, the AQL (PRQ) values and the sample size values are "approximate" geometric series. The following table shows a selection of sample size values along with the ratio between consecutive values.

Table 1: Sample size values from ISO 2859-1 as a geometric series

Sample size	Ratio between consecutive sample sizes
5	-
8	1.60
13	1.63
20	1.54
32	1.60
50	1.56

As can be seen, the ratio between two consecutive sample size values is always close to 1.6. The ratio between consecutive AQL (PRQ) values is also approximately 1.6, as shown in the following table.

Table 2: AQL (PRQ) values from ISO 2859-1 as a geometric series

AQL	Ratio between consecutive AQL values
0.010	-
0.015	1.50
0.025	1.67
0.040	1.60
0.065	1.63
0.100	1.54

As a result, the product AQL (PRQ) × sample size remains "near constant" across the diagonals of the sampling plan tables. This is illustrated in the following table, for a selection of AQL values.

Table 3: The product PRQ × sample size remains "near constant" across the diagonals of the sampling plan tables

Sample size	AQL (PRQ)										
	0.001	**0.0015**	**0.0025**	**0.004**	**0.0065**	**0.01**	**0.015**	**0.025**	**0.04**	**0.065**	**0.1**
2	0.002	0.003	0.005	0.008	0.013	0.02	0.03	0.05	0.08	0.13	0.2
3	0.003	0.005	0.008	0.012	0.020	0.03	0.05	0.08	0.12	0.20	0.3
5	0.005	0.008	0.013	0.020	0.033	0.05	0.08	0.13	0.20	0.33	0.5
8	0.008	0.012	0.020	0.032	0.052	0.08	0.12	0.20	0.32	0.52	0.8
13	0.013	0.020	0.033	0.052	0.085	0.13	0.20	0.33	0.52	0.85	1.3
20	0.020	0.030	0.050	0.080	0.130	0.20	0.30	0.50	0.80	1.30	2.0
32	0.032	0.048	0.080	0.128	0.208	0.32	0.48	0.80	1.28	2.08	3.2
50	0.050	0.075	0.125	0.200	0.325	0.50	0.75	1.25	2.00	3.25	5.0
80	0.080	0.120	0.200	0.320	0.520	0.80	1.20	2.00	3.20	5.20	8.0
125	0.125	0.188	0.313	0.500	0.813	1.25	1.88	3.13	5.00	8.13	12.5
200	0.200	0.300	0.500	0.800	1.300	2.00	3.00	5.00	8.00	13.00	20.0
315	0.315	0.473	0.788	1.260	2.048	3.15	4.73	7.88	12.60	20.48	31.5
500	0.500	0.750	1.250	2.000	3.250	5.00	7.50	12.50	20.00	32.50	50.0
800	0.800	1.200	2.000	3.200	5.200	8.00	12.00	20.00	32.00	52.00	80.0
1250	1.250	1.875	3.125	5.000	8.125	12.50	18.75	31.25	50.00	81.25	125.0
2000	2.000	3.000	5.000	8.000	13.000	20.00	30.00	50.00	80.00	130.00	200.0

The product AQL (PRQ) × sample size is called the *unity value* and can be understood as the number of expected nonconforming items in the sample for lot quality AQL. For example, for lot quality 1% percentage nonconforming and a sample size of 20 items, we can expect 0.2 nonconforming items. This is the rationale for having constant acceptance number values across diagonals in ISO 2859-1.

5.2.2 Variables plans constructed in terms of the producer's risk

The "philosophy" of ISO acceptance sampling plans for inspection by variables is as follows.

First, the ISO plans are designed in such a manner as to ensure either a high probability of acceptance at the acceptance quality limit (AQL) i.e. at the producer's risk quality level (PRQ) or a low probability of acceptance at the limiting quality (LQ) i.e. at the consumer's risk quality level (CRQ).

Secondly, the ISO plans indexed by AQL are constructed in such a way that the producer's risk decreases as the lot size increases. The following table, taken from the Mathematical and Statistical Principles underlying Military Standard 414 [30], the forerunner of the ISO 3951 standard, shows the producer's risk in terms of the sample size code letter (reflecting lot size):

Table 4: Lot size in the ISO standards

Sample size code letter	Producer's risk
B	0.11
C	0.10
D	0.10
E	0.10
F	0.10
G	0.09
H	0.08
I	0.07
J	0.06
K	0.06
L	0.05
M	0.05
N	0.04
O	0.03
P	0.02
Q	0.01

As can be seen the "target" PR of 5 % is only achieved from code letter L onwards. Indeed, the PR is better than 5 % from code letter N onwards, achieving 1 % for code letter Q.

In the plans in ISO 3951-2, the producer's risk remains "near constant" along the diagonals (from bottom left to top right).

The principle behind the ISO 3951-6 plans (which are indexed by LQ) is different: here the aim is to design plans whose OC curves correspond to the OC curves in ISO 2859-2.

5.3 Acceptance sampling versus conformity assessment

There is an extensive normative body of work on conformity assessment: the ISO 17000 series, JCGM 106, etc. The question thus arises to what extent this normative literature is relevant for acceptance sampling. In particular, the question arises whether conformity assessment procedures can be used in acceptance sampling.

In this section, the following abbreviations will be used:

- AS = acceptance sampling
- CA = conformity assessment

It is important to note that the question addressed here cannot currently be answered definitively one way or another. This section can thus be considered to provide basic orientation and considerations which may prove useful in untangling these various concepts in a given context.

5.3.1 Definitions

JGCM 106

In JCGM 106, conformity assessment is defined (definition 3.3.1) as

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

Note that this definition is so general as to allow lot inspection to fall under its scope. Indeed, in many cases a lot can be considered the product output of a process.

ISO 3534

In ISO 3534-2, in section 4 Inspection and general acceptance sampling, we find definition 4.1.1

> *Conformity evaluation*
>
> *Systematic examination of the extent to which an item/entity fulfills specific requirements.*

If conformity evaluation is taken as synonymous with conformity assessment, then the fact this definition is found in the section on acceptance sampling is an indication that the CA normative literature is indeed relevant for AS.

ISO 17025

It seems useful to recall the definition of a “decision rule” (for use in conformity assessment) in ISO 17025. (This definition highlights the central role which measurement uncertainty plays in conformity assessment.)

> Rule that describes how measurement uncertainty is accounted for when stating conformity with a specified requirement.

5.3.2 Positions in ISO standards

ISO 10576 (Guidelines for the evaluation of conformity with specified requirements)

On the other hand, the following paragraph from ISO 10576 would seem to indicate that a resounding no is the correct answer:

> *Because of the apparent similarity to acceptance sampling procedures, it is sometimes seen that acceptance sampling plans are used in conformity testing activities. Acceptance sampling and conformity testing activities both utilize elements of hypothesis testing (see e.g. ISO 2854). It is, however, important to realize that the objectives of the two activities are fundamentally different and in particular the two activities imply different approaches to the risk involved (see ISO 2854 and Holst).*

ISO 2859 and ISO 3951

In the standards from both series, the following phrase is found in the forward:

> *For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO’s adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: http://www.iso.org/iso/home/standards_development/resources-for-technical-work/foreword.htm*

This phrase seems to imply that there is a connection between acceptance sampling and conformity assessment.

5.3.3 Positions in the literature

Paper by Holst, Thyregod and Wilrich (On conformity testing and the use of two stage procedures)

This paper draws the following distinction:

- Acceptance sampling plans are used in the context of transactions between two parties and should provide unambiguous rules for accepting or rejecting the lot. Both parties are aware of the risks involved.
- On the other hand, in conformity testing, “it is crucial that the user can have confidence in a declaration of conformity”. Thus, when an item meets the conformity criterion, this means that “the test has demonstrated beyond any reasonable doubt that the entity conforms to the requirements.”

5.3.4 Discussion

The following distinctions between CA and AS seem to be clear:

- in CA, testing is performed on the basis of one single item, and measurement uncertainty is taken into account
- in AS, there are many cases where testing is performed on the basis of several items (sampled from the lot). In inspection by variables, it is not the conformity of each item which is determined, but rather, one test result is obtained per item and the decision to accept or reject the lot is made on the basis of calculations performed on these test results.
- The outcome of CA may be inconclusive whereas in AS there is always an acceptance or rejection outcome.

A related difference between the two is as follows:

- In CA, measurement uncertainty is taken into account in the decision rule. Thus, in CA, the focus is on the measurand (in the strict metrological sense).
- In AS, the rule for lot acceptance or rejection takes into account the lot standard deviation, which describes how the property of interest varies in the lot, rather than variation between *test results*, which may reflect other effects such as analytical uncertainty, effects due to the sampling procedure, etc. Thus, in AS, the acceptance rule is expressed in terms of the statistical properties of the lot.

The following points highlight conceptual similarities between conformity assessment and acceptance sampling:

- Acceptance sampling can be “re-interpreted” in such a way that the entire framework is formulated in terms of a “measurand” – thus achieving a common conceptual framework with conformity assessment. In this re-interpretation it is the statistical parameters of the lot (e.g. lot mean and lot standard deviation) which constitute the measurand. See Uhlig et al. (2022) [28].
- In the classical CA framework, conformity often requires the measurement uncertainty to be sufficiently low, e.g. in the case of a decision rule such as $y_m + U < USL$. Similarly, in AS, one could formulate requirements regarding sufficiently low specific producer or consumer risks.
- for both CA and AS, one can define both “parametric” and “specific” risks (see section 6.5.1).

5.4 The role of measurement uncertainty in acceptance sampling

The criterion for lot acceptance or rejection is often expressed in terms of statistical parameters such as lot mean and lot standard deviation. If the measurement uncertainty is non-negligible, then the estimates of these statistical parameters may be affected. Accordingly, in some cases, it may be appropriate to apply a correction for measurement uncertainty. Naturally, such a correction presupposes that a reliable estimate of measurement uncertainty is available.

The reader is referred to the *Guidelines on measurement uncertainty* (CXG 54-2004) and the information document for information about the estimation of measurement uncertainty; another key reference is ISO 5725 Parts 1 & 2.

The Eurachem Guide on *Measurement Uncertainty arising from sampling* provides guidance regarding the estimation of the sampling component of uncertainty. The duplicate method can also be applied to estimate the lot standard deviation. Such a method would be appropriate if within-item variation is expected. If an estimate of repeatability precision is available, then Hahn’s method (see Section 4.2.1 and Section 4.2.2) can be used to adjust the estimate of the lot standard deviation, avoiding the need to test items in duplicate. The Eurachem Guide also provides information regarding the use of control charts for the monitoring of consistency.

The following questions are discussed in this section:

- How is the measurand specified?
- Terminological clarification: sampling uncertainty versus acceptance sampling
- What effect does analytical uncertainty have on the producer’s risk?

5.4.1 Specifying the measurand

In determining measurement uncertainty, the first question is: *what is the measurand?*

The term *measurand* has a very specific definition in metrology. The full definition (which is based on the definition of another term, namely *quantity*) can be found in VIM. These definitions are rather technical. For our purposes here, it is sufficient to highlight 2 aspects of the definition of *measurand*.

In order to specify a measurand, it is necessary to define both

- the property of interest (e.g. analyte concentration)
- where this property of interest is being measured

For instance: measuring a given analyte concentration *in an individual item* and measuring the mean analyte concentration *in the lot* correspond to two different measurands.

It should also be noted that a measurand is by definition a property whose characterization is quantitative rather than qualitative.

The question which sources of uncertainty are relevant is answered by considering the definition of the measurand. For instance, if the measurand is defined in terms of the laboratory sample, then only analytical sources are relevant. If the measurand is defined in terms of a population/lot/container ("sampling target") from which the laboratory sample was obtained, then both sampling and analytical sources are relevant.

In connection with acceptance sampling, the concept of measurand can be understood in two different ways.

5.4.1.1 Classical definition of measurand

Insofar as test results are obtained (whether on the basis of discrete items, or on the basis of a composite sample), these test results involve the specification of a measurand. The question arises whether the measurand is specified in relation to the laboratory sample, or in relation to the lot. Two separate cases must be considered: lots consisting of discrete items and lots consisting of bulk material.

Lots consisting of discrete items

In the case of lots consisting of discrete items, acceptance is often based on a characterization of percentage nonconforming. The acceptance criterion is expressed in terms of the lot standard deviation (estimated from the item-specific test results) and the mean value across items. For a given item, the aim is to characterize the item-specific mean value—not the lot mean. Accordingly, the measurand is defined in relation to the laboratory sample, and only analytical sources of measurement uncertainty need be considered. *In particular, there is no sampling component of measurement uncertainty.*

Note: the mean value across the item-specific test results may be considered an estimate of the lot mean. Nonetheless, for lots consisting of discrete items, the measurand is the item-specific mean—not the lot mean.

Lots consisting of bulk material

By contrast, in the case of bulk materials, the aim is to obtain an estimate of the mean concentration in the lot. Accordingly, the measurand is specified in terms of the lot, *and both analytical and sampling sources of uncertainty apply.*

5.4.1.2 Reinterpretation of the concept of measurand for acceptance sampling

Insofar as acceptance is based on a criterion expressed in terms of the statistical parameters of the lot under inspection, it is useful to take a step back and to generalize the concept of measurand as follows:

- In acceptance sampling, the statistical parameters of the lot (e.g. lot mean, lot standard deviation) play the role of the measurand.
- The measurement uncertainty can then be considered to be reflected in the producer's and consumer's risks.
- This reinterpretation is particularly relevant in connection with Bayesian approaches to acceptance sampling.
  See Uhlig et al. (2022) [28].

5.4.2 Sampling uncertainty versus acceptance sampling (terminological clarification)

Sampling uncertainty

Sampling uncertainty is a component of measurement uncertainty.

If the measurand is specified in terms of a larger population such as a lot/container/area, then the laboratory sample must be considered the result of a sampling procedure which may contribute to the uncertainty of the test result. The larger population from which the laboratory sample was obtained is often referred to as the sampling target.

If the total measurement uncertainty is too large, it may be necessary to improve the sampling procedure.

If the measurand is specified in terms of the laboratory sample, there is no contribution to measurement uncertainty due to sampling.

Acceptance sampling

In acceptance sampling, the aim is not to obtain an estimate of measurement uncertainty. The only connection between acceptance sampling and measurement uncertainty is the possible effect of the latter on the calculation of statistical parameters such as lot mean value and lot standard deviation in terms of which the acceptance criterion is expressed.

Lot standard deviation versus sampling component of measurement uncertainty

The concept of the lot standard deviation may seem to be closely related to the concept of the sampling component of measurement uncertainty, namely as a measure of the variation of the property of interest within the lot (where the lot is interpreted as a sampling target). However, this similarity is merely superficial; at a more fundamental level, the two concepts must be carefully distinguished.

In order to clarify the distinction between the lot SD and the sampling component of measurement uncertainty, consider the following case: if all items in the lot are tested (one test result per item), there is no sampling component of the uncertainty of the lot mean. Nonetheless, variation between the items may be small or large.

Note regarding lots consisting of bulk material

For lots consisting of bulk material, the acceptance criterion is often expressed in terms of the lot mean (rather than in terms of the proportion non-conforming). The estimate of the lot mean can be obtained from a composite sample. If the acceptance criterion also involves the uncertainty of the lot mean, and the calculation of this uncertainty includes contributions reflecting the sampling procedure, then this procedure may be indistinguishable from the procedures typically applied in connection with conformity assessment. In such cases, there may thus be no useful distinction between acceptance sampling and conformity assessment. In particular, in such cases, measurement uncertainty takes on a completely different role: rather than a “nuisance parameter” which must be corrected for (if non-negligible), it now plays a central role.

5.4.3 Effects of analytical and sampling uncertainty in acceptance sampling

In this section, we consider the case that the lot consists of discrete items. The criterion for lot acceptance is thus expressed in terms of the percentage nonconforming.

Notation: let n denote the sample size (i.e. the sample consists of n items). For item i, the corresponding test result is denoted x_i. Proposed model for test result x_i:

$$x_i = \mu_i + B + e_i$$

where

μ_i is the “true” mean value for item i. The “true” lot standard deviation σ characterizes the variation of the μ_i in the lot. If sigma is unknown, it is estimated on the basis of the x_i values obtained from the items in the sample.

B is the bias (systematic effect). First and foremost, this term reflects laboratory bias or analytical method bias, but there may be other contributions to bias, e.g. from the sampling procedure.

e_i is the random effect for item i. First and foremost, this term reflects analytical measurement uncertainty (repeatability effects), but there may be contributions from sampling uncertainty.

For lots consisting of discrete items, the acceptance criterion typically has the following form:

$$\bar{x} + ks \leq U \text{ (or } \bar{x} \leq A \coloneqq U - ks)$$

where

$\bar{x}$ is the mean value across the item-specific test results $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

s is the standard deviation across the item-specific test results $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

Analytical uncertainty will manifest itself in the term $e_i \sim \mathcal{N}(0, u_{analytical})$ and will always **inflate** the estimate of the lot standard deviation s:

$$s^2 = \sigma^2 + u_{analytical}^2$$

where

σ is the “true” lot standard deviation

$u_{analytical}$ is the analytical measurement uncertainty

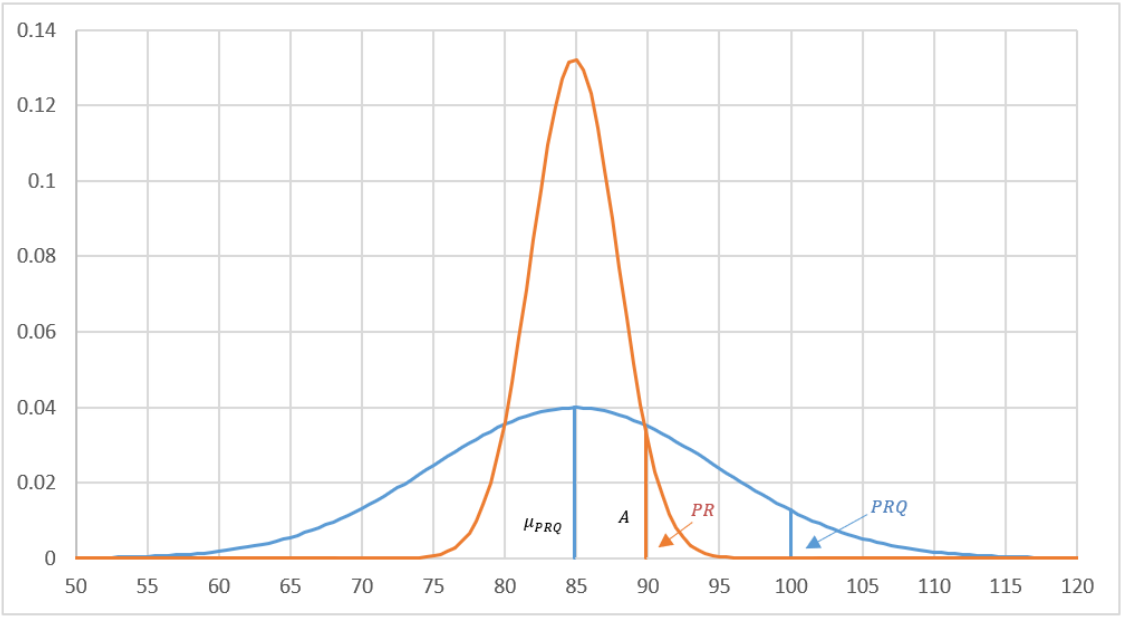
If left uncorrected, the presence of analytical uncertainty will increase producer and consumer risks.

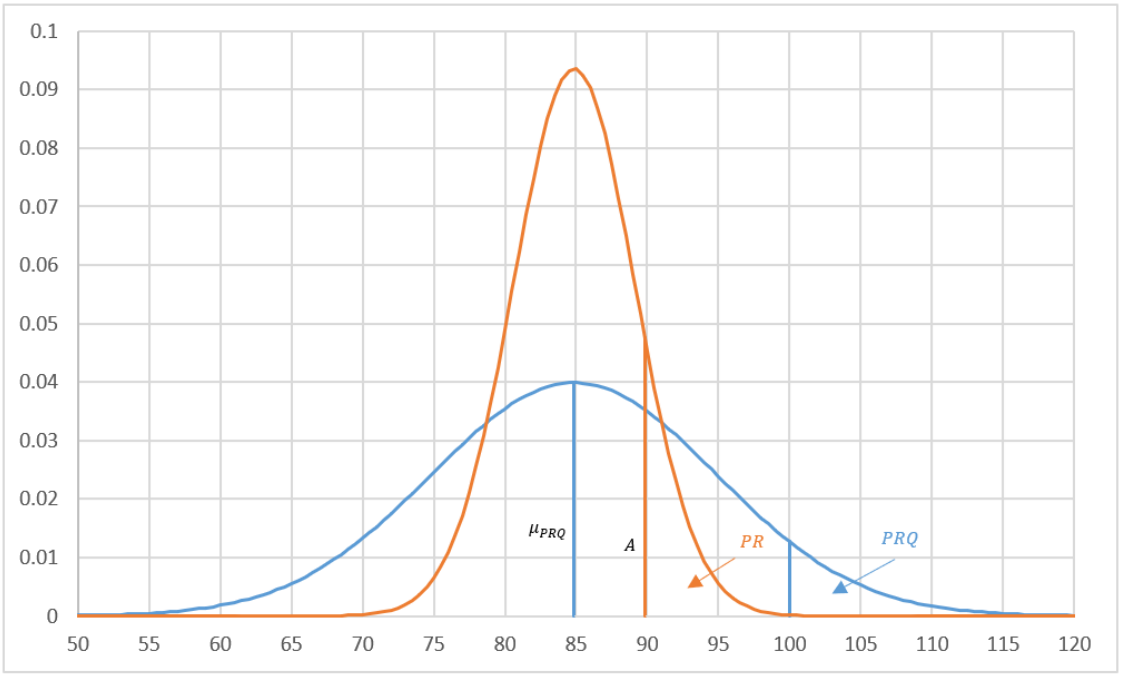
This is illustrated on the basis of two scenarios.

Scenario 1: known σ, increase in producer risk

In this scenario, the upper specification limit is $U = 100$ and the lot standard deviation $\sigma = 10$ is known. The following acceptance sampling plan is applied:$n = 11$, $k = 1.025$ ($A = U - k \cdot \sigma = 90$). However, the analytical uncertainty is nonnegligible, with $u_{analytical} = 10$ (the analytical uncertainty is the equal to the lot SD, and hence can be considered considerable). As a result of the analytical uncertainty, for a lot with quality $PRQ = 6.5\%$, the PR is over 11% (instead of 5%) due to the inflated variation of the x_i.

Figure 36: The blue curve represents the distribution of the property of interest in the lot. The area of the blue curve above $U = 100$ is $PRQ = 6.5\%$. The red curve represents the statistical distribution of the arithmetic mean. In the top diagram, there is no analytical uncertainty, so the PR is 5%. In the bottom diagram, the analytical uncertainty is non-negligible, resulting in an increase in PR to over 11%.

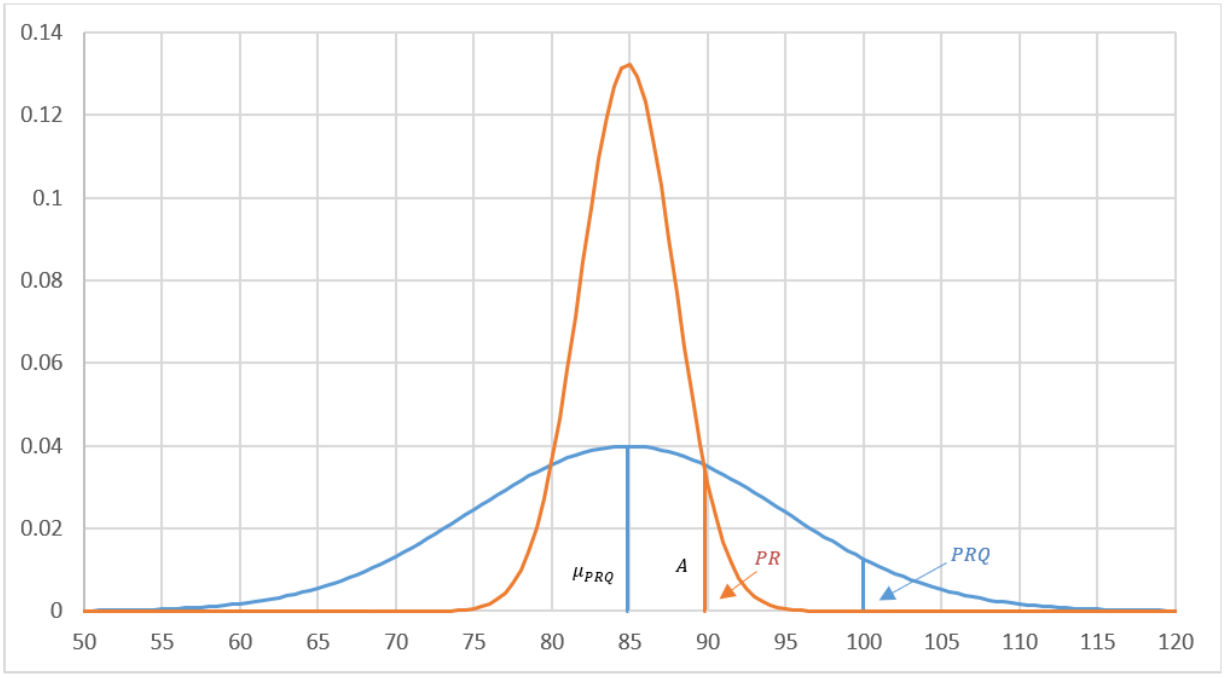


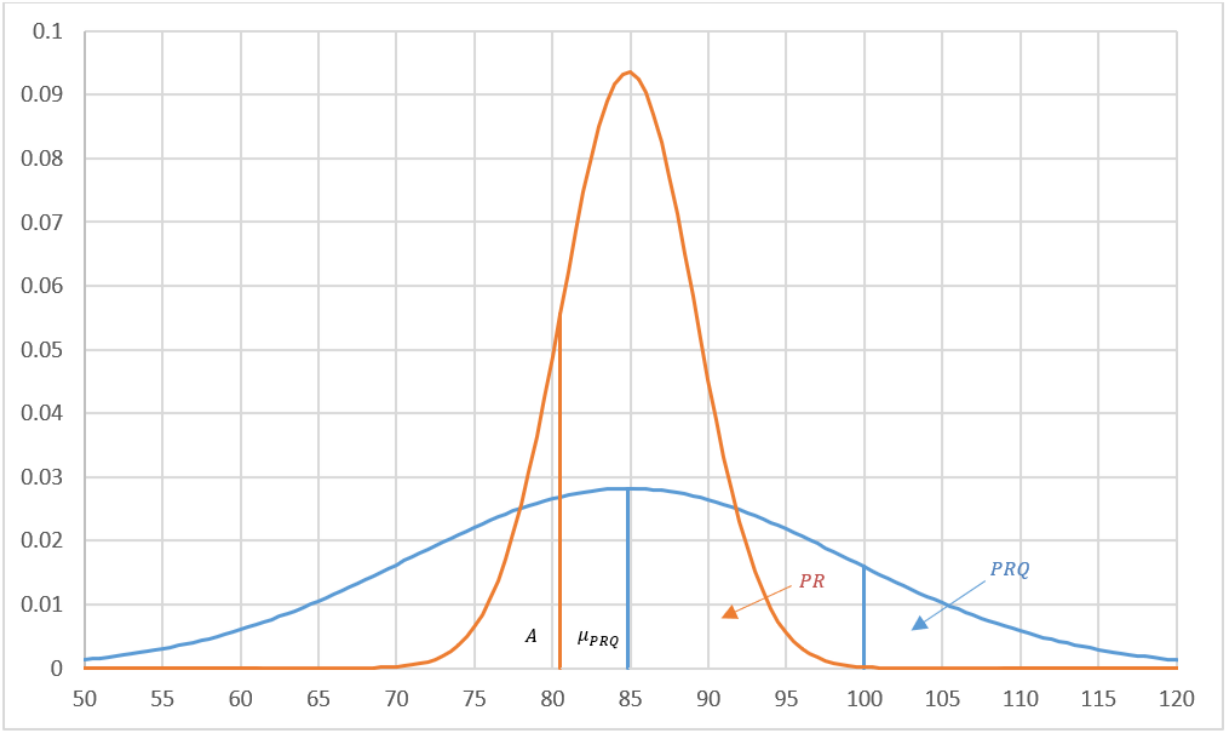


Scenario 2

In this scenario, the upper specification limit is $U = 100$ and the lot standard deviation is unknown and estimated from the x_i. At this point, various things can happen. For instance, the producer could notice that the lot quality is now 8.5 % percentage nonconforming instead of 6.5 %. If this discrepancy is ignored, and the same plan as originally contemplated is applied (in particular: $k = 1.025$), the acceptance limit is now 80.5 (instead of 90) due to the inflated estimate s, and the PR is now over 85 %.

Figure 37: The blue curve represents the distribution of the property of interest in the lot. The area of the blue curve above $U = 100$ is $PRQ = 6.5\%$. The red curve represents the statistical distribution of the arithmetic mean. In the top diagram, there is no analytical uncertainty, so the PR is 5%. In the bottom diagram, the analytical uncertainty is non-negligible, resulting in an increase in a much wider blue curve and a distorted value for A, resulting in a PR of nearly 85 %.





5.5 Bayesian plans

It is often the case that prior information regarding lot quality is available. For instance, the consumer may have previously purchased lots from the producer of the lot currently under inspection. It is thus sensible to ask the following question: is it possible to propose a Bayesian framework for the design of acceptance sampling plans which mobilizes prior information in order to achieve a reduction in sample size? In the following, this question will be discussed in relation to *inspection by attributes*.

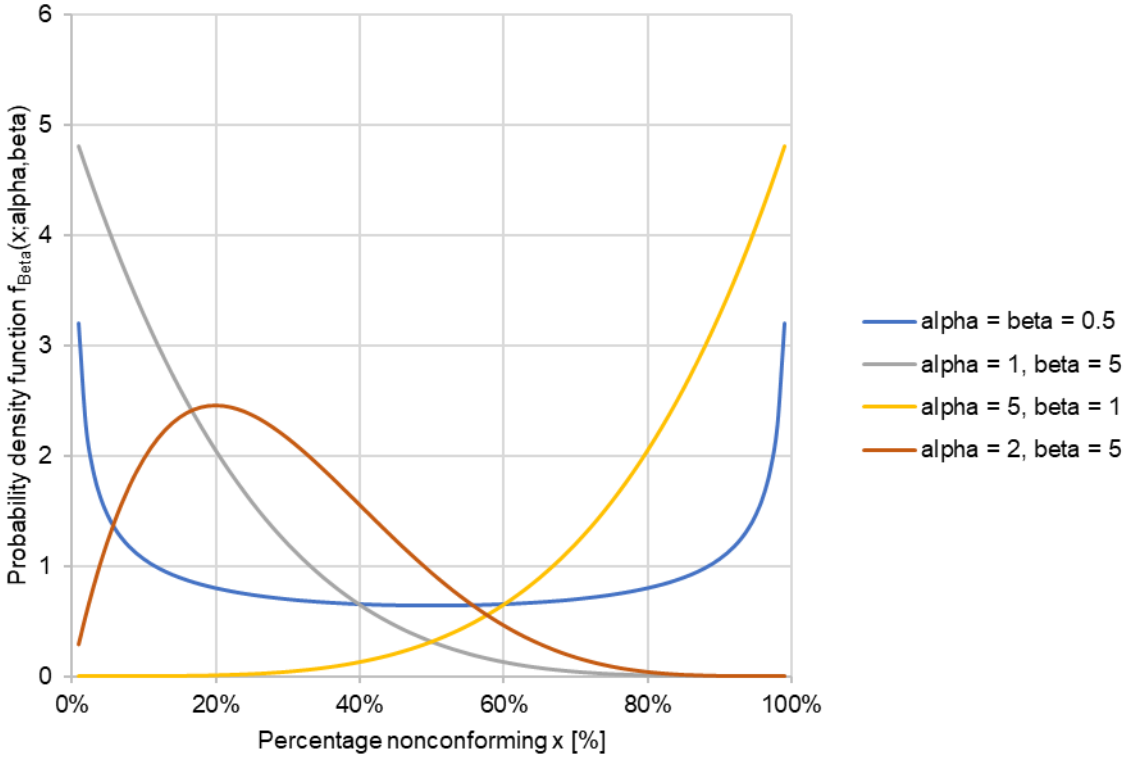
5.5.1 Prior distributions

The consumer's or producer's prior information regarding the percentage nonconforming x is encapsulated in the prior distribution. In the case of inspection by attributes, the simplest assumption regarding the prior for x is that it follows a beta distribution.

The beta family of distributions is generated via two hyperparameters α and β. For a given choice of α and β, the corresponding beta distribution is denoted Beta(α, β) and the probability density function is denoted $f_{Beta}(x; \alpha, \beta)$.

The following diagram shows different beta distributions. As can be seen, this family of distributions is quite versatile, allowing very different curves to be mapped via the choice of α and β.

Figure 38 Different beta distributions



Note: The case that no prior information is available can be represented by the choice $\alpha = \beta = 0.5$.

Once items from a lot have been tested, the prior distribution can be updated to obtain a posterior distribution.

If the prior is a beta distribution, the posterior is also a beta distribution and the posterior hyperparameters α_1 and β_1 are obtained from the prior hyperparameters α_0 and β_0 and from the number of nonconforming items y as follows:

$$\alpha_1 = \alpha_0 + y$$

$$\beta_1 = \beta_0 + n - y$$

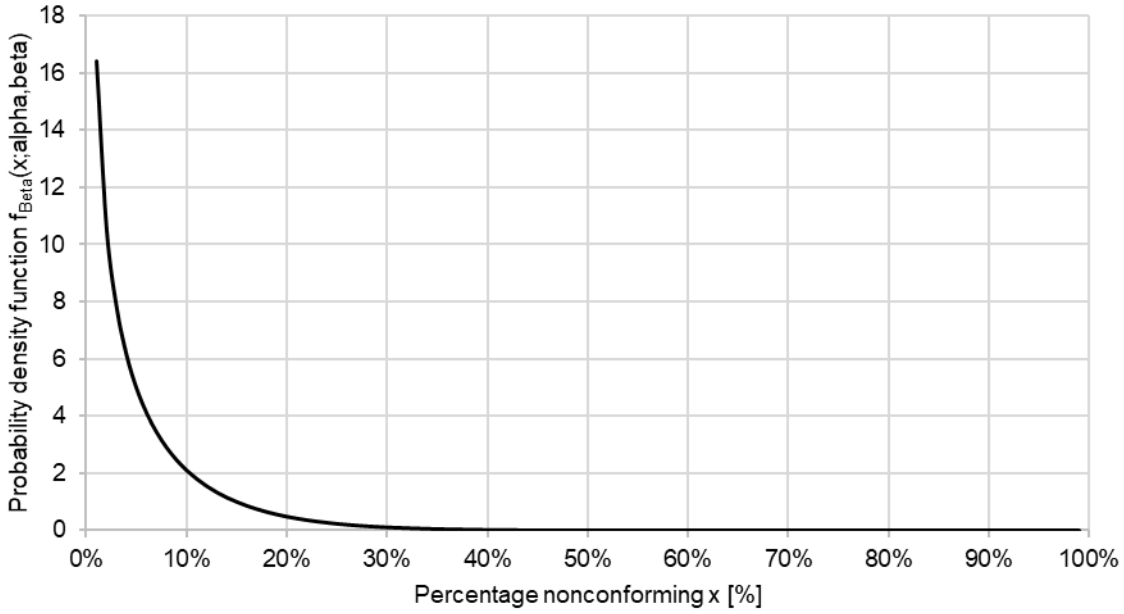
For example, if 10 items from a given lot have been inspected and all 10 have been found to be conforming (i.e. $y = 0$), the $\alpha_0 = \beta_0 = 0.5$ prior is updated as follows:

$$\alpha_1 = 0.5 + 0 = 0.5$$

$$\beta_1 = 0.5 + 10 - 0 = 10.5$$

The following diagrams shows the beta distribution corresponding to $\alpha = 0.5$ and $\beta = 10.5$.

Figure 39 Probability density for Beta distribution α=0.5, ß= 10.5



As can be seen, the Beta(0.5,10.5) density function falls steeply in the range $x = 0\%$ to $x = 30\%$ and is near 0 for all percentage nonconforming values above $x = 30\%$. This means that—on the basis of the prior Beta(0.5,0.5) and the test outcome $y = 0$ (out of 10 tests)—it is now expected that the percentage nonconforming will be no greater than 30%.

5.5.2 Conformance probability approach

In this section, an approach is presented for specifying acceptance sampling plans on the basis of the concept of conformance probability from JCGM 106. The approach described in the ISO 2859 and ISO 3951 standards asks the following question: given a certain quality level (expressed e.g. as percentage nonconforming) what is the probability that the lot is accepted? By contrast, the conformance probability approach asks the following question: given that a lot is accepted, what is the probability that it is actually conforming? Insofar as the conformance probability approach starts from lot acceptance or rejection (i.e. information which is known), this approach can be considered more pragmatic. In the conformance probability approach, probabilities and risks are calculated via the Bayesian approach. The starting point is a prior distribution, which encapsulates all available knowledge regarding the property of interest prior to lot inspection. Once tests have been performed, the prior distribution is updated on the basis of the test results. The updated distribution is called the posterior distribution.

Definition of conformance probability

The definition of conformance probability found in JCGM 106 can be adapted to lot inspection and acceptance sampling as follows.

Conformance probability is the probability that the lot quality actually lies in the conformance region $\mathcal{C}$. This probability is calculated on the basis of the *posterior* distribution.

As can be seen from the definition, a conformance region for lot quality must be specified. This is a clear departure from the approach described in the ISO 2859 and ISO 3951 standards in which, even though the plans are indexed in terms of certain quality levels considered “good” (PRQ, AQL) or “poor” (CRQ, LQ), conformance regions for lot quality are not specified.

The conformance region $\mathcal{C}$ can be specified via an upper limit for the percentage nonconforming. This upper limit is denoted $x_{\mathcal{C}}$.

Definition of parametric and specific consumer’s and producer’s risks

In the ISO 2859 and ISO 3951 standards., the consumer and producer risks are defined as follows:

- Producer risk = probability that a lot of good quality (e.g. PRQ or AQL) will not be accepted
- Consumer risk = probability that a lot of poor quality (e.g. CRQ or LQ) will be accepted

In the calculation of the ISO risks, lot quality is treated as the parameter of a statistical distribution. For this reason, the ISO risks are also called “parametric” risks.

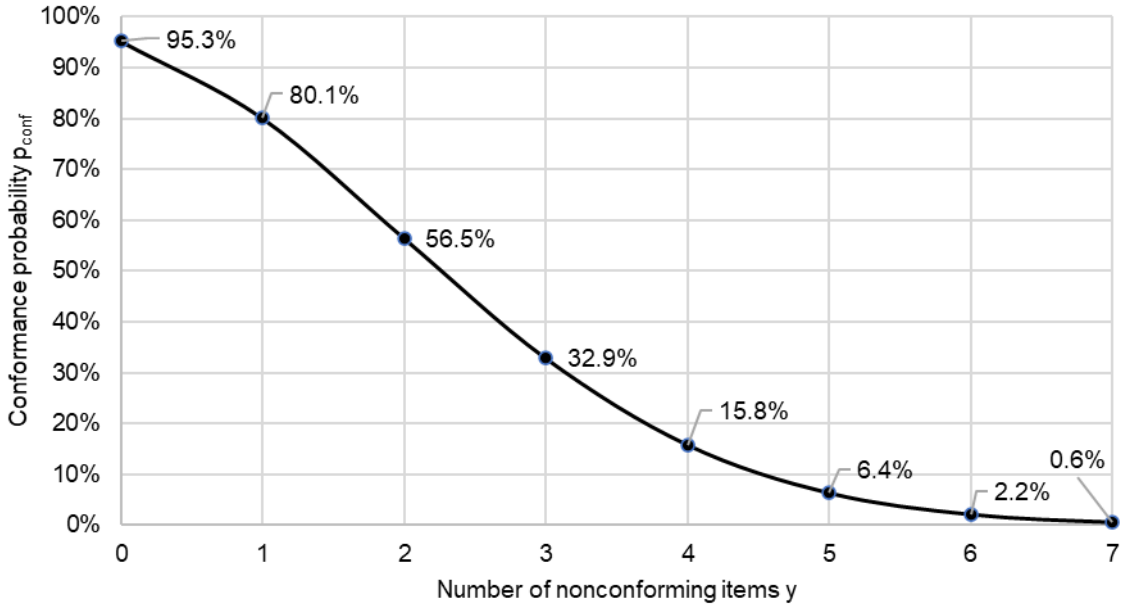
In JCGM 106, specific consumer and producer risks are defined by "going in the opposite direction" as compared to the ISO definitions:

- Specific producer risk = probability that a lot *which has not been accepted* actually has conforming quality
- Specific consumer risk = probability that a lot *which has been accepted* actually has nonconforming quality

Conformance probability curves

The concept of conformance probability can be used to derive acceptance sampling plans. For a given sample size n, the conformance probability is calculated for each possible testing outcome y (e.g. number of nonconforming items in the sample). An acceptance sampling plan can then be determined by requiring that the specific consumer risk (the complement of the conformance probability) should be less than a certain threshold (e.g. 5 %). This procedure is illustrated in the following diagram which shows the conformance probability curve for a lot conformity region specified via the upper limit $x_{\mathcal{C}} = 10\%$ for the percentage nonconforming, the prior Beta (1,9) and the sample size $n = 20$.

Figure 40: Conformance probability curve for upper limit of the lot conformity region $x_{\mathcal{C}} = 10\%$, prior Beta(1,9) and sample size $n = 20$



As can be seen, the conformance probability is greater than 95 % at $y = 0$. Accordingly, the plan $n = 20$, $c = 0$ is a plan which allows a specific consumer risk less than 5 %.

Consumer and producer risks

The Bayesian framework from JCGM 106 allows the definition of various risks. These risks address different questions, as summarized in the following tables.

Table 5: Bayesian producer's risks – notation and interpretation

Risk	Notation	Definition
Specific PR (evaluated for a specific test outcome y resulting in rejection)	SPR(y)	How likely is it that a lot is conforming, given that it is rejected?
Conditional PR (conditioned on the lot quality x)	CPR_x	How likely is it that a lot is rejected, given that it is conforming?
Conditional PR (conditioned on all test outcomes y resulting in rejection)	CPR_y	How likely is it that a lot is conforming, given that it is rejected?

Global PR	GPR	How likely is it that a lot is both conforming and rejected?
Global probability of rejection	GP_{rej}	How likely is it that a lot is rejected – no matter whether it is conforming or not?

Table 6: Bayesian consumer's risks – notation and interpretation

Risk	Notation	Interpretation
Specific CR (evaluated for a specific test outcome y resulting in acceptance)	SCR(y)	How likely is it that a lot is nonconforming, given that it is accepted?
Conditional CR (conditioned on the lot quality x)	CCR_x	How likely is it that a lot is accepted, given that it is nonconforming?
Conditional CR (conditioned on all test outcomes y resulting in acceptance)	CCR_y	How likely is it that a lot is nonconforming, given that it is accepted?
Global CR	GCR	How likely is it that a lot is both nonconforming and accepted?
Global probability of acceptance	GP_{acc}	How likely is it that a lot is accepted – no matter whether it is conforming or not?

More information regarding the conformance probability approach can be found in Uhlig et al. (2024) [21].

5.5.3 Utility approach

5.5.3.1 Definition

The definition of utility is as follows:

Utility for a lot which has been **accepted**	=	benefit associated with an accepted lot (under the assumption that all items are conforming) i.e. returns minus expenditures
		minus damages associated with nonconforming items in an accepted lot
		minus testing and sampling costs
Utility for a lot which has been **rejected**	=	**minus** testing and sampling costs

Depending on the context and the consumer, benefits may reflect profits from commercial sales, tax income, long-term environmental considerations, laying the foundations of new business opportunities, etc.; while damages may reflect commercial losses, negative health impacts, costs associated with recalling a lot and a tarnished public image.

5.5.3.2 The consumer

In acceptance sampling, the consumer is defined as the party which accepts or rejects the producer's lot. The consumer may be a retailer purchasing commodities at a wholesale market, a manufacturer acquiring parts or a customs officer or food safety agent checking that legal limits for contaminants are not exceeded before admitting a lot at the border, etc... The following table provides different interpretations of benefit and damages for three different types of consumers.

Table 7: The meaning of benefit and damages for different types of consumers

	Retailer	Manufacturer	State
Benefit associated with conforming items	Profit = *returns* minus *expenditures* of conforming items *Returns* = income from selling the items in the lot *Expenditures* = the lot's purchase price, transport costs, staff remuneration, retail outlet overhead, taxes	Profit from conforming items depends on aspects like cost efficiency, innovation, secure supply, flexibility, and sustainability	Benefit from conforming items could include: Positive impact on trade balance Benefits for the end-user Employment impact Tax revenue Innovation Positive environmental impact Positive health impact
Damages associated with nonconforming items	Income loss associated with nonconforming items, e.g. with items which were not sold, contractual penalties or legal fines, costs associated with containing and recalling items	Income loss from nonconforming items, associated with disruptions in the manufacturing process, quality issues and supply chain challenges	Damages from nonconforming items could include Negative health impact Loss in GDP Negative employment impact Financial losses (e.g. tax, extra expenditures) Negative environmental impact

5.5.3.3 Mathematical expression

The definition of utility makes it possible to apply a simple criterion in the design of acceptance sampling plans: select the plan which maximizes utility.

In order to express the utility function in mathematical terms, the following notation will be used:

B = benefit associated with one conforming item in an accepted lot

D = damages associated with one nonconforming item in an accepted lot (damages is understood here in a very general sense, see below. This parameter could also be called "losses".)

T = the testing and sampling costs per item

All costs are expressed in terms B (the benefit associated with a conforming item in an accepted lot). In other words, the common unit in which all the terms in the utility function are expressed is B. For example, the damages associated with a nonconforming item D could be 10 B; and the testing and sampling costs per item T could be 5 B. The values for D and T expressed in terms of B will be referred to the *cost structure*.

The quantity B thus has a dual role: on the one hand, it denotes the item-specific benefit and on the other hand, it functions as a unit in which to express D (damages) and T (testing costs). The values for D and T are thus *relative* to B. The advantages of expressing D and T in relation to B are twofold:

- on the one hand, it may be easier to provide values for D and T in relation to B rather than monetary values
- on the other hand, the fact that these values are relative highlights that it is the relation between the costs and benefits rather than the absolute or monetary values themselves which play a central role in the utility approach

Specifying the parameter B point of view of a state (the consumer is e.g. a food safety agency or a customs officer) is not always straightforward. Indeed, from the point of view of the state, some of the positive aspects of a successful commercial transaction may be intangible or difficult to assign a precise monetary value to. In such cases, as a rule of thumb, it is suggested to use the lot's purchase price.

The damages (or losses) parameter D can have different meanings. If there are no other costs and if there are no re-purposing options for items which are not sold, the value $D = B$ corresponds to the case that the only cost associated with a nonconforming item is the *loss* in income caused its not being sold. A value such as $D = 1.5\,B$ corresponds to the scenario that, in addition to losses, there are additional costs associated with nonconforming items, such as costs associated with waste management. A value such as $D = 10\,B$ could reflect additional costs associated with environmental pollution, damage to the reputation of the retailer (dissatisfied customers taking their business elsewhere, tarnished public image) etc. Higher values such as $D = 25\,B$ or $D = 100\,B$ reflect substantial additional costs such as those associated with recalling a lot or healthcare costs. It may be difficult to quantify healthcare costs associated with nonconforming items (e.g. hospitalization costs due to the ingestion of contaminated meat). One approach could be to introduce an auxiliary unit such as W = wages corresponding to a day's work for an "average worker." For example, if a brief stay in a hospital is quantified as 5 W and if the profit (benefit B) associated with the item which caused the health issue (e.g. contaminated meat) is 0.1 W, then D is calculated as 50 B.

Finally, we also introduce the following notation:

N	Number of items in the lot
M	Number of nonconforming items in the lot
n	Sample size (i.e. number of items in the sample)

The utility function is defined as follows:

$$\mathrm{U}(N, B, M, D, T, n) = \begin{cases} B \cdot N - D \cdot M - T \cdot n, \text{ if the lot is accepted} \\ -T \cdot n, \text{ if the lot is rejected} \end{cases}$$

In the case of an accepted lot, the utility function can be rewritten as

$$\mathrm{U}(N, B, M, D, T, n) = N \cdot \left(B - D \cdot \frac{M}{N} - T \cdot \frac{n}{N}\right)$$

It should be noted that M is an integer which, in general, remains unknown. The aim of the acceptance sampling procedure is to obtain an estimate of U by means of a suitable (nonbiased) estimator for M (or for the proportion of nonconforming items $\frac{M}{N}$), taking all prior information into account.

5.5.3.4 Example

In order to illustrate the concept of utility as well as the coefficients for the benefit and damages discussed in the previous section, we consider the case where the utility is simply a given value rather than the expected value of a random variable.

Consider the scenario where a retailer has accepted and purchased a lot of 2000 apples (400 kg) at a wholesale price of 600 €. The retailer expects to sell all apples at an average retail price of 4 € per kg. Total sales (if all apples are sold) for the lot will thus be 1600 €. Transport costs were 20 €. The retailer paid a salesperson 100 € for the day's work.

The benefit for the entire lot (under the assumption that all the apples are sold) for the retailer is thus

$$1600\ € - 600\ € - 20\ € - 100\ € = 880\ €$$[9]

A total of 100 apples (20 kg) had blemishes which resulted in being discarded (i.e. not sold). The retailer concludes that the damages caused by nonconforming items are 80 €.

Benefit, damages and testing costs for the example

Utility for a lot which has been **accepted**	=	benefit associated with an accepted lot (under the assumption that all items are conforming) i.e. returns minus expenditures	880 €
		minus damages associated with nonconforming items in an accepted lot	**minus** 80 €.
		minus sampling and testing costs	**minus** 0 €.

[9] For the sake of simplicity, neither overhead costs (associated with the retail outlet) nor taxes are included here.

Accordingly, in this example, the utility for the lot is **800 €.** Now, let us calculate the parameters B and D (we already know that $T = 0$). The benefit per conforming item (apple without blemishes) B expressed in Euros is

$$B = \frac{880\ €}{2000\ \text{apples in the lot}} = 0.44\ € \text{ per apple}$$

The damages associated with nonconforming items are

$$D = \frac{80\ €}{100\ \text{apples with blemishes}} = 0.8\ € \text{ per apple}$$

If the damages coefficient is expressed in terms of *B*, we have

$$D = \frac{0.8\ € \text{ per apple}}{0.44\ € \text{ per apple}} = 1.82\ B \text{ per apple}$$

The following table summarizes the example.

Table 8: Summary of example of lot of 1000 apples

Lot size	$N = 2000$
Sample size	$n = 0$
Acceptance number	c Not applicable
Sampling and testing costs	$T = 0\ B$
Damages per nonconforming item	$D = 1.82\ B$
Conversion of benefit per conforming item	$B = 0.44\ €$
Proportion of nonconforming items	x_0 = 5 %
Utility	$B \cdot N - D \cdot M - T \cdot n$ $= 0.44\ € \cdot 2000 - 1.82 \cdot 0.88\ € \cdot 100$ $= 800\ €$

Note 1

If the retailer purchases lemons instead of apples and if the lemons are displayed for purchase at the retail outlet during 10 days, then the retailer will worry about mold spread via contamination with neighboring spoilt lemons. If the lemons are sold individually and if the retailer observes that, on average, one spoilt lemon contaminates 5 neighboring lemons in the time span of 10 days, this state of affairs can be taken into consideration by multiplying the D coefficient by 5.

Note 2

If an intermediate distributor is involved in the purchase of a lot of 500 crates of 2000 apples (10^6 apples), it is likely that the lot will be inspected prior to acceptance. In such a situation, T will be nonzero.

Note 3

If the aim is to use the utility in order to determine the acceptance sampling plan (sample size n and acceptance number c), then the damages coefficient can be specified via the mean value of the prior distribution for the percentage nonconforming.

In order to illustrate this point, we start with the observation that, in the example above, at least 900 apples must be sold in order for the retailer to break even:

Expenditures (Purchase price of the lot + transport costs + sales person remuneration)	720 €
Retail price of apple	0.8 €
Income from the sale of 900 apples (180 kg)	720 €

Accordingly, 1 maximum of 1100 nonconforming items (1100 apples with blemishes leading to discarding) can be tolerated in order to break even. This corresponds to a maximum of 55% for the percentage nonconforming.

If the retailer works with a prior distribution for the percentage nonconforming whose mean value is 10% (i.e. well below the maximum of 55%), then *D* can be determined as follows. First, we note that the percentage nonconforming corresponds to $\frac{M}{N}$. In the absence of testing costs, the utility is simply

$$U = N \cdot B - D \cdot M$$

Setting $U = 0$ (i.e. lot acceptance in the sense of breaking even), we obtain the following expression for the damages coefficient

$$D = \frac{N}{M} \cdot B$$

If D is expressed in terms of B, then this simplifies to

$$D = \frac{N}{M}$$

For $\frac{M}{N} = 0.1$ (the prior which the retailer is working with), we thus obtain $D = 10$.

Note 4

It is important to understand the impact which D has on the utility. In particular, in the absence of sampling and testing costs, the value $D = 1\ B$ means that one single sold item is sufficient to have a positive utility.

Indeed, for $M = N - 1$ and $D = 1\ B$ we have

$$U = N \cdot B - D \cdot M = N - M = N - (N - 1) = 1$$

5.5.3.5 The role of the prior

In the example from the previous section, the lot has already been accepted and purchased and the proportion of nonconforming items x_0 can be empirically determined by 100% inspection of the lot. In other words, in such a case, x_0 is an empirically determined value rather than a random variable (hence the notation with the subscript). However, the aim of the utility approach is to determine sample size and acceptance number—in other words to design the acceptance sampling plan—prior to lot inspection. For this reason, in the following, the empirically determined x_0 value will be replaced with a random variable X whose distribution as called the *prior*. This gives a slightly different meaning to the utility: instead of an *empirical* value corresponding to a given lot, it is now an *expected* value, in the probabilistic sense. The expected utility for a given plan (sample size n and acceptance number c) calculated on the basis of the prior distribution is denoted $u_{prior}(n, c)$.

5.5.3.6 Utility curves

The following diagrams show utility curves:

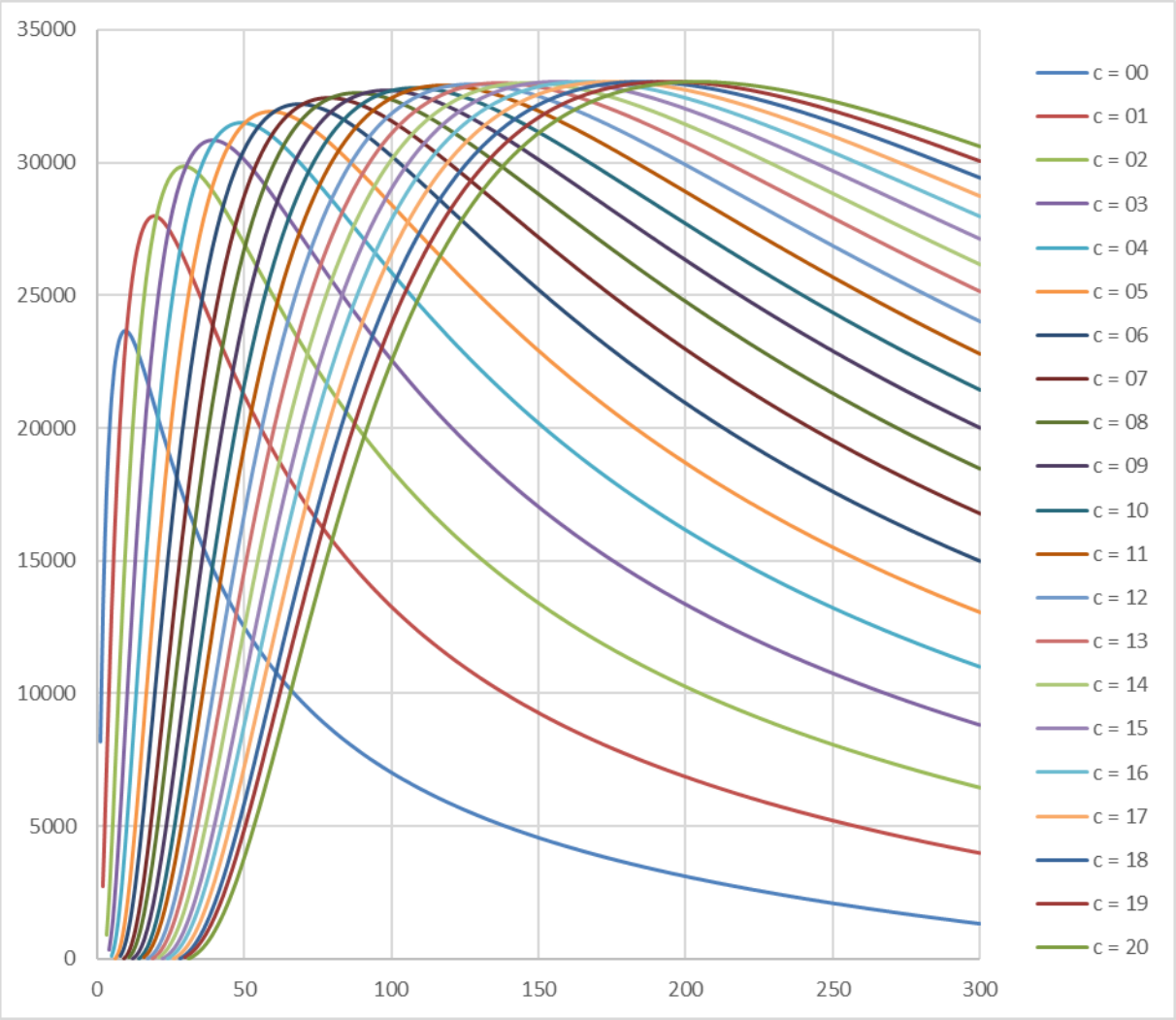
- for the prior Beta(1,9) with mean value 10% percentage nonconforming
- for the prior Beta(0.5,0.5) with mean value 50% percentage nonconforming (“non-informative” prior)

The x-axis shows the sample size n and the y-axis shows the expected utility $u_{prior}(n, c)$. There is a separate curve for each acceptance number c.

As can be seen, the utility values for the more optimistic prior are greater (maximum around 33000 B) than for the non-informative prior (maximum around 12600 B).

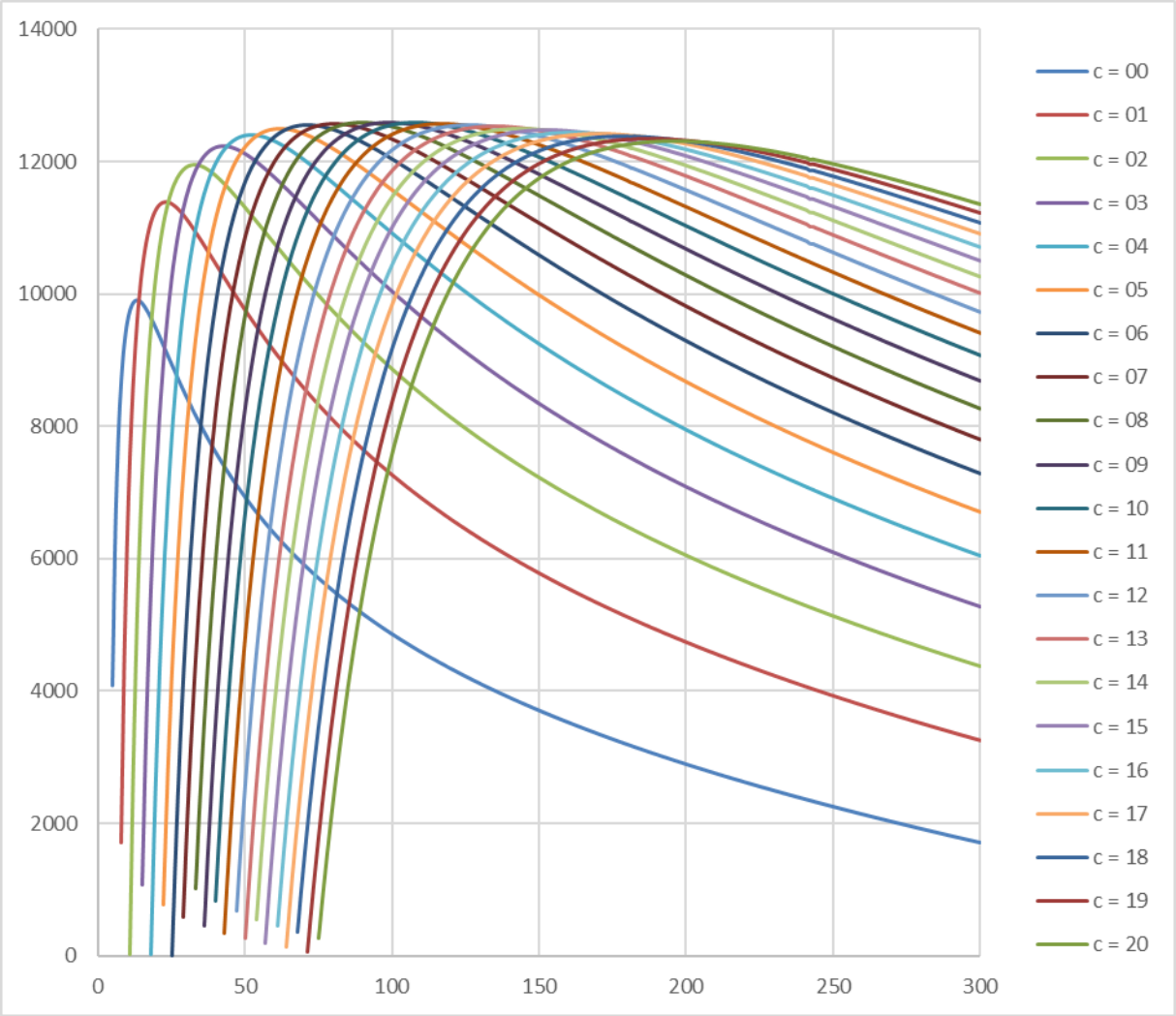
The following diagram shows utility curves for the prior Beta(1,9), lot size N = 100000 and cost structure $D = 10\ B$ and $T = 5\ B$. The optimum plan is $n = 175$, $c = 17$ (utility = 33043 B).

Figure 41: Utility curves for different acceptance number c values as a function of sample size n for the prior Beta(1,9), lot size N = 100000 and the cost structure D = 10 B and T = 5 B



The following diagram shows the utility curves for the Jeffreys prior Beta(0.5,0.5), which represents the case that no prior information is available (noninformative prior). As in the previous diagram, the lot size N is 100000 and the cost structure is $D = 10\ B$ and $T = 5\ B$. The optimum plan is $n = 99$, $c = 9$ (utility = 12592 B).

Figure 42: Utility curves for different acceptance number c values as a function of sample size n for the prior Beta(0.5,0.5), lot size N = 100000 and the cost structure D = 10 B and T = 5 B



As can be seen in the above diagrams, the utility curves appear to plateau around the maximum value. This motivates the following procedure: we propose to consider plans whose utility lies within 10% of the maximum utility. This will allow a considerable reduction in sample size while maintaining a negligible impact on utility.

For instance, the maximum utility for the Beta(1,9) prior is 30457 B. Subtracting 10% of this maximum value, we obtain 29739 B. The $c = 2$ curve exceeds this value at $n = 27$. Hence the $n = 175$, $c = 17$ plan can be replaced with $n = 27$, $c = 2$.

5.5.3.7 Standard plans

This section provides tables with standard plans for the practitioner. It is assumed that there is no inspection error and that prior information is available in the form of results from previous testing performed on items from the producer of the lot currently under inspection. It is assumed that these previous tests have been performed recently—say, within the previous 2 or 3 months. In the tables, the results from previous testing are represented via the pair (n_0, y_0) with n_0 denoting the number of items previously tested and y_0 denoting the number of nonconforming items from these recent tests**.**

For a given pair (n_0, y_0), the plans are calculated on the basis of a Beta(α, β) prior for the percentage nonconforming where

$$\alpha = y_0 + 0.5$$

$$\beta = n_0 - y_0 + 0.5$$

The pair $(n_0 = 0, y_0 = 0)$ stands for the absence of any prior information. This scenario is represented as the first row in each of the tables. The corresponding standard plans are based on the noninformative Jeffreys prior, i.e. the Beta(0.5,0.5) distribution.

The standard plans can be used to answer two questions.

Question 1: Given prior information, a cost structure and a lot size, which acceptance sampling plan should be applied?

Question 2: Given a cost structure and lot size, what type of prior information is required in order to achieve an acceptance sampling plan whose workload is acceptable (or indeed, for it to be worthwhile to perform acceptance sampling at all)

In order to illustrate the use of the standard plans to answer these two questions, consider the following examples.

Example 1
A lot with size $N = 100\,000$ packages of grains was purchased at a price of 100 000 GCU (where GCU stands for generic currency unit). The retailer who purchased the lot intends to sell each package at 6 GCU. A simplified calculation (i.e. without taking overhead costs etc. into consideration) thus yields a benefit per item of $B = 5$ GCU. Say that sampling and testing costs are 25 GCU per package. The corresponding parameter is thus $T = 5\,B$. Finally, if the damages associated with a nonconforming item consist in a dissatisfied customer who may no longer return to the retailer's outlet, the D parameter can be specified as $D = 10\,B$.
Answer to question 1: Given this background, if the retailer has recently purchased another lot and tested 20 packages, none of which were nonconforming, then the plan $n = 33$, $c = 1$ can be applied.
Answer to question 2: If the retailer would like to test, say, no more than 10 packages, then prior information based on at least 50 tested packages would be required.

Example 2
A lot of canned salmon is inspected following the item conformity criteria set out in Section 8 of CXS 3-1981. Say the lot size is $N = 1000$ items (cans) and that the sampling and testing costs per item are $T = 5\,B$ (as in Example 1). Finally, the damages associated with a nonconforming item consist in a dissatisfied customer who may no longer go to the retailer's outlet, and the D parameter can be specified as $D = 10\,B$ (as in Example 1).
Answer to question 1: Given this background, if the consumer has data from 5 recently tested cans from the same supplier, one of which was nonconforming, the decision would be to reject without testing.
Answer to question 2: The consumer would need to test an additional three cans prior to investing resources in inspecting a lot of 1000 cans. With prior information $n_0 = 8$, $y_0 = 1$, the acceptance sampling plan would be (10,0).

Notation

The following notation is used in the tables:

n_0	Previous testing / prior information The number of items tested prior to the current lot inspection
y_0	Previous testing / prior information The number of nonconforming items
N	Lot size
D	Damages/losses/costs per nonconforming item (expressed in terms of the benefit per item B)
T	Sampling and testing costs per item (expressed in terms of the benefit per item B)
a	Accept without testing
r	Reject without testing
(n, c)	Acceptance sampling plan with sample size n acceptance number c

Table 9: Standard plans for testing and sampling costs $T = 5\ B$

Prior test results		N	1000					10000					100000				
		D	1.5	3	10	30	100	1.5	3	10	30	100	1.5	3	10	30	100
		T	5					5					5				
n_0	y_0	T/N	0.500%					0.050%					0.005%				
0	0		(1,0)	(2,0)	(7,0)	r	r	(2,1)	(5,1)	(15,1)	(25,0)	r	(2,1)	(7,2)	(21,1)	(50,1)	(93,0)
1	0		a	(2,0)	(7,0)	r	r	a	(4,1)	(15,1)	(27,0)	(78,0)	a	(4,1)	(24,2)	(53,1)	(110,0)
2	0		a	(1,0)	(6,0)	r	r	a	(3,1)	(14,1)	(28,0)	(77,0)	a	(3,1)	(20,1)	(54,1)	(119,0)
3	0		a	(1,0)	(5,0)	(20,0)	r	a	(1,0)	(13,1)	(28,0)	(76,0)	a	(3,1)	(18,1)	(54,1)	(131,0)
4	0		a	(1,0)	(5,0)	(19,0)	r	a	(1,0)	(12,1)	(28,0)	(76,0)	a	(1,0)	(16,1)	(53,1)	(147,1)
5	0		a	(1,0)	(4,0)	(18,0)	r	a	(1,0)	(11,1)	(28,0)	(76,0)	a	(1,0)	(14,1)	(52,1)	(147,1)
8	0		a	(1,0)	(3,0)	(15,0)	r	a	(1,0)	(7,0)	(26,0)	(74,0)	a	(1,0)	(10,1)	(49,1)	(144,1)
10	0		a	(1,0)	(3,0)	(14,0)	r	a	(1,0)	(5,0)	(24,0)	(73,0)	a	(1,0)	(9,1)	(46,1)	(143,1)
13	0		a	(1,0)	(2,0)	(12,0)	r	a	(1,0)	(4,0)	(21,0)	(71,0)	a	(1,0)	(7,1)	(42,1)	(140,1)
20	0		a	(1,0)	(2,0)	(8,0)	r	a	(1,0)	(3,0)	(15,0)	(65,0)	a	(1,0)	(3,0)	(33,1)	(133,1)
30	0		a	(1,0)	(2,0)	(6,0)	r	a	(1,0)	(3,0)	(9,0)	(57,0)	a	(1,0)	(3,0)	(22,1)	(122,1)
50	0		a	(1,0)	(2,0)	(5,0)	(22,0)	a	(1,0)	(2,0)	(7,0)	(41,0)	a	(1,0)	(3,0)	(9,0)	(80,0)
80	0		a	(1,0)	(2,0)	(4,0)	(14,0)	a	(1,0)	(2,0)	(6,0)	(23,0)	a	(1,0)	(3,0)	(7,0)	(50,0)
100	0		a	(1,0)	(2,0)	(4,0)	(12,0)	a	(1,0)	(2,0)	(6,0)	(18,0)	a	(1,0)	(3,0)	(7,0)	(34,0)
1	1		(2,1)	(3,0)	r	r	r	(4,2)	(7,1)	(11,0)	r	r	(6,3)	(12,3)	(22,1)	(36,0)	(87,0)
2	1		a	(5,1)	r	r	r	(2,1)	(9,2)	(16,0)	r	r	(3,2)	(15,4)	(33,2)	(59,1)	r
3	1		a	(4,1)	r	r	r	a	(8,2)	(21,1)	r	r	a	(14,4)	(41,3)	(64,1)	r
4	1		a	(3,1)	r	r	r	a	(7,2)	(22,1)	r	r	a	(10,3)	(43,3)	(69,1)	r
5	1		a	(3,1)	r	r	r	a	(6,2)	(22,1)	r	r	a	(8,3)	(46,3)	(72,1)	r
8	1		a	(1,0)	(10,0)	r	r	a	(2,1)	(23,1)	r	r	a	(3,1)	(48,4)	(96,2)	r
10	1		a	(1,0)	(8,0)	r	r	a	(2,1)	(25,2)	(45,0)	r	a	(2,1)	(45,4)	(98,2)	r
13	1		a	(1,0)	(6,0)	r	r	a	(1,0)	(19,1)	(42,0)	r	a	(2,1)	(37,3)	(102,2)	(159,0)
20	1		a	(1,0)	(3,0)	r	r	a	(1,0)	(10,1)	(58,1)	r	a	(1,0)	(20,2)	(106,2)	(159,0)
30	1		a	(1,0)	(3,0)	r	r	a	(1,0)	(6,1)	(49,1)	r	a	(1,0)	(7,1)	(97,2)	(156,0)
50	1		a	(1,0)	(2,0)	(9,0)	r	a	(1,0)	(4,0)	(30,1)	r	a	(1,0)	(5,1)	(64,2)	(147,0)
80	1		a	(1,0)	(2,0)	(6,0)	r	a	(1,0)	(3,0)	(10,0)	r	a	(1,0)	(3,0)	(20,1)	(180,1)
100	1		a	(1,0)	(2,0)	(5,0)	r	a	(1,0)	(3,0)	(8,0)	(87,0)	a	(1,0)	(3,0)	(15,1)	(167,1)
8	2		a	(3,1)	r	r	r	a	(6,2)	(25,1)	r	r	a	(12,4)	(48,3)	(68,1)	r
10	2		a	(1,0)	r	r	r	a	(3,1)	(25,1)	r	r	a	(6,2)	(56,4)	(74,1)	r
13	2		a	(1,0)	r	r	r	a	(2,1)	(26,1)	r	r	a	(3,1)	(62,5)	(80,1)	r
20	2		a	(1,0)	(8,0)	r	r	a	(2,1)	(26,2)	r	r	a	(2,1)	(56,5)	(109,2)	r
30	2		a	(1,0)	(3,0)	r	r	a	(1,0)	(12,1)	r	r	a	(2,1)	(30,3)	(133,3)	r
50	2		a	(1,0)	(3,0)	r	r	a	(1,0)	(6,1)	(59,1)	r	a	(1,0)	(7,1)	(122,3)	r
80	2		a	(1,0)	(2,0)	(10,0)	r	a	(1,0)	(5,1)	(32,1)	r	a	(1,0)	(6,1)	(87,3)	r
100	2		a	(1,0)	(2,0)	(7,0)	r	a	(1,0)	(3,0)	(19,1)	r	a	(1,0)	(5,1)	(46,2)	(186,0)
20	3		a	(1,0)	r	r	r	a	(2,1)	(27,1)	r	r	a	(2,1)	(70,5)	(81,1)	r
30	3		a	(1,0)	(8,0)	r	r	a	(2,1)	(27,2)	r	r	a	(2,1)	(66,6)	(92,1)	r
50	3		a	(1,0)	(3,0)	r	r	a	(1,0)	(7,1)	r	r	a	(1,0)	(15,2)	(143,3)	r
80	3		a	(1,0)	(2,0)	r	r	a	(1,0)	(5,1)	(60,1)	r	a	(1,0)	(6,1)	(127,3)	r
100	3		a	(1,0)	(2,0)	r	r	a	(1,0)	(5,1)	(41,1)	r	a	(1,0)	(6,1)	(107,3)	r
50	4		a	(1,0)	(4,0)	r	r	a	(1,0)	(17,2)	r	r	a	(2,1)	(46,5)	(100,1)	r
80	4		a	(1,0)	(3,0)	r	r	a	(1,0)	(6,1)	r	r	a	(1,0)	(8,1)	(147,3)	r
100	4		a	(1,0)	(3,0)	r	r	a	(1,0)	(5,1)	r	r	a	(1,0)	(6,1)	(142,3)	r

Table 10: Standard plans for testing and sampling costs $T = 25\ B$

Prior test results		N	1000					10000					100000				
		D	1.5	3	10	30	100	1.5	3	10	30	100	1.5	3	10	30	100
		T	25					25					25				
n_0	y_0	T/N	2.500%					0.250%					0.025%				
0	0		(1,0)	(2,0)	r	r	r	(2,1)	(4,1)	(8,0)	r	r	(2,1)	(5,1)	(16,1)	(29,0)	(78,0)
1	0		a	(1,0)	r	r	r	a	(2,0)	(8,0)	(22,0)	r	a	(4,1)	(16,1)	(34,0)	(80,0)
2	0		a	(1,0)	(5,0)	r	r	a	(1,0)	(7,0)	(21,0)	r	a	(3,1)	(15,1)	(44,1)	(81,0)
3	0		a	(1,0)	(4,0)	r	r	a	(1,0)	(7,0)	(21,0)	r	a	(1,0)	(14,1)	(43,1)	(81,0)
4	0		a	(1,0)	(4,0)	r	r	a	(1,0)	(6,0)	(20,0)	r	a	(1,0)	(13,1)	(43,1)	(81,0)
5	0		a	(1,0)	(3,0)	r	r	a	(1,0)	(5,0)	(19,0)	r	a	(1,0)	(12,1)	(42,1)	(81,0)
8	0		a	(1,0)	(2,0)	r	r	a	(1,0)	(4,0)	(17,0)	r	a	(1,0)	(9,1)	(39,1)	(80,0)
10	0		a	(1,0)	(2,0)	r	r	a	(1,0)	(3,0)	(15,0)	r	a	(1,0)	(7,1)	(37,1)	(80,0)
13	0		a	(1,0)	(2,0)	r	r	a	(1,0)	(3,0)	(13,0)	r	a	(1,0)	(4,0)	(34,1)	(80,0)
20	0		a	(1,0)	(2,0)	(5,0)	r	a	(1,0)	(2,0)	(9,0)	(55,0)	a	(1,0)	(3,0)	(22,0)	(77,0)
30	0		a	(1,0)	(2,0)	(4,0)	r	a	(1,0)	(2,0)	(6,0)	(44,0)	a	(1,0)	(3,0)	(11,0)	(70,0)
50	0		a	(1,0)	(2,0)	(4,0)	r	a	(1,0)	(2,0)	(5,0)	(28,0)	a	(1,0)	(3,0)	(7,0)	(53,0)
80	0		a	(1,0)	(2,0)	(3,0)	r	a	(1,0)	(2,0)	(5,0)	(16,0)	a	(1,0)	(2,0)	(6,0)	(30,0)
100	0		a	(1,0)	(2,0)	(3,0)	(9,0)	a	(1,0)	(2,0)	(5,0)	(14,0)	a	(1,0)	(2,0)	(6,0)	(22,0)
1	1		(1,0)	r	r	r	r	(2,1)	(4,0)	r	r	r	(4,2)	(9,2)	(15,0)	(29,0)	r
2	1		a	(2,0)	r	r	r	(2,1)	(5,1)	r	r	r	(2,1)	(11,3)	(22,1)	(36,0)	r
3	1		a	(2,0)	r	r	r	a	(5,1)	r	r	r	a	(10,3)	(24,1)	(40,0)	r
4	1		a	(1,0)	r	r	r	a	(4,1)	(13,0)	r	r	a	(8,2)	(30,2)	(42,0)	r
5	1		a	(1,0)	r	r	r	a	(3,1)	(12,0)	r	r	a	(6,2)	(30,2)	(43,0)	r
8	1		a	(1,0)	r	r	r	a	(1,0)	(11,0)	r	r	a	(2,1)	(29,2)	(45,0)	r
10	1		a	(1,0)	r	r	r	a	(1,0)	(10,0)	r	r	a	(2,1)	(28,2)	(66,1)	r
13	1		a	(1,0)	r	r	r	a	(1,0)	(9,0)	r	r	a	(2,1)	(24,2)	(65,1)	r
20	1		a	(1,0)	(2,0)	r	r	a	(1,0)	(4,0)	r	r	a	(1,0)	(12,1)	(63,1)	r
30	1		a	(1,0)	(2,0)	r	r	a	(1,0)	(3,0)	(25,0)	r	a	(1,0)	(6,1)	(58,1)	r
50	1		a	(1,0)	(2,0)	r	r	a	(1,0)	(2,0)	(11,0)	r	a	(1,0)	(5,1)	(37,1)	r
80	1		a	(1,0)	(2,0)	(4,0)	r	a	(1,0)	(2,0)	(6,0)	r	a	(1,0)	(3,0)	(16,1)	(111,0)
100	1		a	(1,0)	(2,0)	(4,0)	r	a	(1,0)	(2,0)	(6,0)	r	a	(1,0)	(3,0)	(11,0)	(111,0)
8	2		a	(1,0)	r	r	r	a	(3,1)	r	r	r	a	(9,3)	(27,1)	r	r
10	2		a	(1,0)	r	r	r	a	(2,1)	r	r	r	a	(5,2)	(34,2)	r	r
13	2		a	(1,0)	r	r	r	a	(2,1)	r	r	r	a	(2,1)	(35,2)	r	r
20	2		a	(1,0)	r	r	r	a	(1,0)	(9,0)	r	r	a	(2,1)	(35,3)	r	r
30	2		a	(1,0)	(3,0)	r	r	a	(1,0)	(5,0)	r	r	a	(1,0)	(18,2)	(78,1)	r
50	2		a	(1,0)	(2,0)	r	r	a	(1,0)	(3,0)	r	r	a	(1,0)	(6,1)	(68,1)	r
80	2		a	(1,0)	(2,0)	r	r	a	(1,0)	(3,0)	(11,0)	r	a	(1,0)	(5,1)	(41,1)	r
100	2		a	(1,0)	(2,0)	(5,0)	r	a	(1,0)	(2,0)	(7,0)	r	a	(1,0)	(5,1)	(22,1)	r
20	3		a	(1,0)	r	r	r	a	(2,1)	r	r	r	a	(2,1)	(37,2)	r	r
30	3		a	(1,0)	r	r	r	a	(1,0)	(9,0)	r	r	a	(2,1)	(37,3)	r	r
50	3		a	a	(2,0)	r	r	a	(1,0)	(4,0)	r	r	a	(1,0)	(9,1)	r	r
80	3		a	(1,0)	(2,0)	r	r	a	(1,0)	(3,0)	r	r	a	(1,0)	(6,1)	(69,1)	r
100	3		a	(1,0)	(2,0)	r	r	a	(1,0)	(3,0)	(16,0)	r	a	(1,0)	(5,1)	(55,1)	r
50	4		a	(1,0)	(3,0)	r	r	a	(1,0)	(5,0)	r	r	a	(2,1)	(21,2)	r	r
80	4		a	(1,0)	(2,0)	r	r	a	(1,0)	(5,1)	r	r	a	(1,0)	(6,1)	r	r
100	4		a	(1,0)	(2,0)	r	r	a	(1,0)	(3,0)	r	r	a	(1,0)	(6,1)	(73,1)	r

General notes

Note 1

As the lot size N increases, the sample size also increases, but at the same time, it becomes "easier" to justify investing resources in lot inspection—i.e. plans are available at lower n_0 values. This is due to the increase in benefit due to the increased lot size.

Note 2

For a given cost structure and lot size, the sample size is less for $T = 25\ B$ than for $T = 5\ B$. On the other hand, a higher n_0 value is required prior to committing resources to acceptance sampling for $T = 25\ B$.

Note 3

The 'r' entries (highlighted in red in the table) indicate that the lot should be rejected without testing. This should be interpreted as follows: given the testing costs and the damages associated with nonconforming items, it only makes sense to invest resources in lot inspection given a minimum level of confidence in the lot quality. This minimum level of confidence is codified via the (n_0, y_0) pair. For example, for a relatively small lot ($N = 1000$) and a very high value for D ($D = 100$, reflecting for example a health hazard in connection with nonconforming items), it only makes sense to conduct acceptance sampling if the prior information is based on at least 50 previously tested items with zero nonconforming results.

If the consumer rejects the lot without testing, the producer has the following options:

1. Increase the lot size N. As seen in the standard plans, a larger lot size translates to higher income for the consumer, thus lowering the threshold for investing resources in acceptance sampling.
2. Decrease the purchase price of the lot. This is tantamount to increasing the parameter B, which, in turn, will result in lower values for D and T. This will lower the threshold for investing resources in acceptance sampling.
3. Develop test methods which are cheaper to apply. This option will result in a decrease in the T parameter, thus lowering the threshold for investing resources in acceptance sampling. An important caveat here is that the performance of the new method must be at least as good as that of the original method.

Technical notes

Note 1

The standard plans were calculated via a hierarchical Bayesian model which "mixes" two priors: the "actual" prior corresponding to the prior information and the "non-informative" Jeffreys prior. The latter's influence is increased the further the testing outcome deviates from the "actual" prior. See Uhlig et al. (2025) [25]. In addition, the 10% approximation (see discussion at the end of the previous section) is applied.

Note 2

We define the ratio $x_0 = B/D$ (or $1/D$ if D is expressed in terms of B). The probability that the percentage nonconforming exceeds x_0 is an interesting pendant to the consumer's risk in the ISO 2859 series of standards. See the discussion in Hald [23].

Note 3

The sample size increases with n_0 as long as the mean value of the beta distribution corresponding to (n_0, y_0) is less than the ratio $x_0 = B/D$ (or $1/D$ if D is expressed in terms of B). When this mean value is greater than x_0, the sample size decreases as n_0 increases. This can be explained as follows: if the mean value is less than x_0, it is unlikely that utility will be positive, and the natural tendency of the model is to resist investing resources in lot inspection. This resistance decreases as the prior becomes more optimistic. By contrast, if the mean value is greater than x_0, then it is likely that the utility will be positive, and the natural tendency of the model is to invest resources in lot inspection. As the prior is becomes more optimistic, fewer resources are required.

Note 4

The standard plan for $N = 1000$, $D = 3$, $T = 25$, $n_0 = 50$, $y_0 = 3$ is "accept" rather than (1,0). This anomaly may possibly be related to rounding issues.

More information regarding this approach can also be found in Uhlig et al. (2025) [25].

5.5.3.8 Broader view of utility

As can be seen in the standard plans, the acceptance sampling plan depends very much on the lot size. Indeed, a large lot size translates to an increase in total benefit, thus impacting the calculation of utility. For example, for the cost structure $D = 30\ B$ and $T = 25\ B$, and for the prior information $(n_0 = 20, y_0 = 1)$, there is an acceptance plan (i.e. it makes sense to invest resources in lot inspection) only for the lot size $N = 100\ 000$. Indeed, for the lot size $N = 1000$ and the lot size $N = 10\ 000$, the utility approach results in the decision to reject the lot without testing. These considerations show the extent to which the acceptance samplings plans (including the decision to reject without testing) reflect the cost structure internalized in the parameters of the utility model.

There are three takeaways from this for the consumer.

The first is that the consumer can perform some preliminary calculations and inform the producer prior to the lot being shipped that, given the cost structure, the transaction is only commercially viable for a minimum lot size.

The second is that, in certain circumstances, it could lie in the interest of the consumer to "pretend" that the lot size is greater than it actually is in order to achieve a viable acceptance sampling plan. For instance, consider the case that an importing country is initiating commercial relations with a new supplier and that the first lot—intended as a trial—is smaller in size than subsequent "routine" lots would be.

The third and last takeaway is that an indispensable condition for achieving plans which successfully balance the producer's and the consumer's interests is transparency. For instance, the producer must be able to ascertain whether the consumer intends to reject without testing or to apply an acceptance sampling plan *prior to shipping the lot*. Indeed, this last take away leads directly to the question whether it is possible to combine utility functions representing the consumer's and the producer's perspectives so as to achieve a broader notion of utility representing, as it were, a win-win situation for both parties to the transaction. This is discussed in the following section.

5.5.3.9 Adversarial approach

The standard plans from the previous section were calculated on the basis of a utility concept which reflects the consumer's perspective. However, the utility approach can be extended to include the producer's perspective. This will be briefly described in this section.

If the consumer reaches the decision to reject without testing, then the producer has two possible courses of action.

1. Apply one of the options listed under **Note 3** in the **General notes,** at the end of Section 5.5.3.7
2. Offer to pay for testing and sampling costs. In such a case, the sampling plan can be calculated from the point of view of the producer. This is where the *adversarial* approach is applied. The aim of the adversarial approach is to calculate an acceptance sampling plan (sample size and acceptance number) to maximize the producer's utility. The latter in turn, takes into account the consumer's decision to accept or reject the lot. In other words, the adversarial approach consists in combining the consumer's and the producer's utility functions.

The adversarial approach was first described in articles by Lindley and Singpurwalla [24]. For further information regarding the adversarial approach, the reader is referred to Uhlig et al. (2025) [26].

5.5.4 Bayesian plans: glossary of terms

Percentage nonconforming	The proportion of nonconforming items in the lot, expressed as a percentage. In the ISO 2859 and ISO 3951 standards, this is called percentage nonconforming.
Prior distribution (short: prior)	Statistical distribution which encapsulates information regarding the lot quality which is available prior to the lot inspection
Posterior distribution (short: posterior)	Statistical distribution which combines the prior and the testing outcomes
Hyperparameter	Parameter of the prior or posterior distribution
Beta distribution	Typical choice of prior for the percentage nonconforming in the case of inspection by attributes
Parametric risk (“classical” or “ISO” risk)	The producer’s and consumer’s risks as defined in the ISO 2859 and ISO 3951 standards. These risks start from a given quality level and calculate the corresponding probability of acceptance or rejection. In other words, the quality level is treated as the parameter of a statistical distribution.
Bayesian risk	These risks are calculated via a prior distribution for the parameter which characterizes the lot quality. For example, in the case of inspection by attributes, the risks are calculated via a prior distribution for the percentage nonconforming. See Section 5.5.2.
Lot conformity	In the case of inspection by attributes: a lot is conforming if the percentage nonconforming lies within a conformance region $\mathcal{C}$ for lot quality specified via an upper limit $x_{\mathcal{C}}$ for the percentage nonconforming (e.g. upper limit $x_{\mathcal{C}} = 10\%$). Lot conformity must be carefully distinguished from item conformity.
Conformance probability	The probability that a lot is conforming, given a test outcome.
Specific consumer risk	A type of Bayesian risk. The probability that a lot is nonconforming given a test outcome resulting in lot acceptance.
Global producer risk	A type of Bayesian risk. The probability that lot is both conforming and rejected.
Utility	A value reflecting benefits and costs associated with an acceptance sampling plan and a lot.
Expected prior utility	Insofar as the utility is calculated from a prior for the percentage nonconforming, the utility is a random variable. The expected value of the utility is called the expected prior utility.

5.6 References

[1] FAO Mycotoxin Tool and Mycotoxin S&T User Guide

[2] The web-based FAO Mycotoxin Tool for the evaluation of sampling plans is provided at http://tools.fstools.org/mycotoxins. It is not a secure site but contains more detail than allowed for in the formulae and there is an accompanying user guide.

[3] Edward G. Schilling & Dean V. Neubauer (2017); Acceptance Sampling in Quality Control, (3rd edition) Chapman & Hall. ISBN 9781498733571

[4] https://www.crcpress.com/downloads/K26314/CRC%20book%20website%20files.zip

[5] https://www.crcpress.com/Acceptance-Sampling-in-Quality-ControlThird-Edition/Schilling-Neubauer/p/book/9781498733571

[6] https://www.amazon.com/Acceptance-Sampling-Quality-Control-Statistics-dp-1498733573/dp/1498733573/ref=dp_ob_title_bk. Second edition:

[7] https://www.amazon.com/Acceptance-Sampling-Quality-Control-Statistics/dp/1584889527

[8] Link to copy of version 2 on Google Books (see Chapter 9): https://www.google.co.nz/books/edition/Acceptance_Sampling_in_Quality_Control_S/Ryv1EGIkAaYC?hl=en&gbpv=1&dq=schilling+%26+neubauer&printsec=frontcover

[9] Bicking, C.A. 1970. ASTM E-105-58 and ASTM E-300-69 standards for the sampling of bulk materials, Journal of Quality Technology, 2(3): 165-173.

[10] Zweiterung et al. Relating Microbiological Criteria to Food Safety Objectives and Performance Objectives, Food Control (20) 2009, 967-979

[11] ISO 10725:2000 Acceptance sampling plans and procedures for the inspection of bulk materials

[12] International Standards Organisation

[13] ASQC Basic References in Quality. Volume 7: How and When to Perform Bayesian Acceptance

[14] Calvin, T.W. (1990) How and When to Perform Bayesian Acceptance Sampling. American Society for Quality Control. ISBN 0-87389-0097-3,

[15] Hailey, W.A. (1980) Minimum Sample Size Single Sampling Plans: A Computerized Approach Journal of Quality Technology 12, 4, pp230-235

[16] MH Ramsey, SLR Ellison and P Rolston (eds.) Eurachem/EUROLAB/CITAC/NORDTest/AMC Guide: Measurement uncertainty arising from sampling: a guide to methods and approaches. Second Edition Eurachem (2019). ISBN (978-0-948926-8). Available from http://www.eurachem.org

[17] General Standard for Contaminants and Toxins in Food and Feed (CXS 193-1995)

[18] Code of Hygienic Practice for Powdered Formulae for Infants and Young Children (CXC 66 – 2008)

[19] ISO 5725-1:2023: Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions

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

[21] Uhlig, S.; Colson, B.; Kissling, R.; Ellis, S.; Hicks, M.; Vandenbemden, J.; Pennecchi, F.; Göb, R.; Gowik, P. Acceptance Sampling Plans Based on Conformance Probability—Inspection of Lots and Processes by Attributes. Preprints 2024, 2024092228. doi.org/10.20944/preprints202409.2228.v1

[22] Whitaker T. (2006) Sampling food for Mycotoxins, Food Additives and Contaminants, January 2006, 23(1) p50–61

[23] Hald A (1981). Statistical Theory of Sampling Inspection by Attributes. Academic Press Inc, London New York.

[24] Lindley D, Singpurwalla N. Adversarial life testing. J. R. Statist. Soc. B (1993) 55, No. 4, pp. 837-847

[25] Uhlig et al. (2025) Acceptance sampling based on utility

[26] Uhlig et al. (2025) Acceptance sampling based on utility: adversarial approach

[27] Uhlig et al. (2025) Within-item variability

[28] Uhlig et al. (2022) Application of the Bayesian conformity assessment framework from JCGM 106:2012 for lot inspection on the basis of single items. doi:10.20944/preprints202206.0032.v1

[29] Wetherill G. Barrie (1977) Sampling Inspection and Quality Control (2nd edition) Chapman & Hall ASIN B00L6YW5PW

[30] Defense Technical Information Center (1958) Mathematical and statistical principles underlying military standard 414 (sampling procedures and tables for inspection by variables for percent defective).

Appendix II

Summary of the EWG consultation on the information document for the *General guidelines on sampling* (CXG 50-2004)

Introduction

1. New Zealand and Germany worked closely to review and update the information document based on the discussion and comments provided to CCMAS43. The information document was provided to the electronic working group (EWG) in December 2024. The EWG was advised that this version included a range of updates such as:

   - moving some of the section on Attributes plans "theory" into the section on Examples to improve understanding;
   - an updated section on variables plans to include examples based on Codex provisions;
   - formulas for evaluation of variables plans (negligible MU);
   - an example of a plan used in a real-life dispute situation that is not part of a Codex standard. However, an example of a designed attributes plan has been added;
   - commentary regarding Whitaker's plans and the approach to new matrices or contaminants, in response to questions raised, noting that the information document does provide an overview of sampling plans for bulk materials; and
   - a rewritten section on Bayesian approaches, noting that this area is being discussed at various technical committees (e.g. ISO TC 69 WG10) outside of CCMAS and that with a wider focus on the Bayesian approach for sampling, this may need to be considered in any updates of this information document.

2. Following the EWG consultation, New Zealand and Germany reviewed the comments received and made further updates to the information document. A summary of the responses is set out in the following table:

No.	Summary of the comment	Response	Section
Australia			
1	The 'CXG 50 Information Document December 2024' document is well progressed, we welcome the elaboration of the additional case studies, these are very useful. Particularly the discussion on the now defunct CODEX STAN 233 plus Net weight sampling plans based on the OIML International Recommendation R 87 "Quantity of product in prepackages", International Organization for Legal Metrology, Paris.	Noted – thank you	
2	We note the issue of how CCMAS should document sampling plans when endorsed in CXS-234, may not have sufficient detail, especially if the sampling is based on 'provisions' as well as 'commodity categories' or a 'hybrid' which is not specifically stated. But possibly this format is now redundant, and CCMAS may have to specify if the provision in the commodity is in principle an 'attribute' or 'variable', 'the distributed characteristics', and a starting point 'maximum allowable producer's and consumer's risks for the inspection of a provision', whether a plan is expect to have 'negligible' or 'non-negligible MU' or 'bias' (based on the endorsed methods), plus suggest 'Method of Sampling' (if available and 'fit-for-purpose') but allowing 'specific ways to economize sample numbers'. This way	Noted – thank you	

No.	Summary of the comment	Response	Section
	CCMAS is providing some initial guidance as a starting point, facilitate the use of sampling App(s) to develop plans and accept that there is more than one option for the sampling plan that could be used.		
3	Section 2 Introduction – The previous version had two parts, now there are three. As the section 1 is a 'Content', it is good to show a sectional breakdown of what had changed in this latest version (in blue), but looking forward, this detail may have no purpose. Suggestion is to potentially simplify the 'Introduction' to: Part 1 contains general information relating to the design of sampling plans, including examples (Section 3 - 4.4). Part 2 contains more background on sampling plans including a statistical appendix (section 5 - 5.7). Part 3 Statistical appendix (Section 6 - 6.5) plus Section 6.6 References Retaining the text 'Note: Some Excel formulas in the text (and the two Excel files provided) use the English style, with decimal points and comma separators'		
4	Amending the reference as follows 'OIML International Recommendation R087 (OIML R 87) - Quantity of products in prepackages', where we expect the relevant section in OIML R 87 is '4.2.1 Metrological requirements when an inspection lot is sampled'	Agreed - done	Section 3.4.1
5	Replacing 'label weight quantity, e.g. weight,'(two instances); plus replace 'inspection of weights of quantities by weights,' to be consistent with OIML R 87.	Agreed - done	Section 3.4.1
6	Section 3.3.1 Remove empty text box.	Agreed - done	
7	Section 4.1.2 offers an Excel file 'pexact.xlsx' with formulas in Annex. But there appears to be no file or Annex. Can these be included?	Agreed - text included in section 5.1.2	Section 5.1.2
8	Section 4.1.2 'Example: Inspection for Foreign Matter' there is the text 'The number of samples n can be calculated directly using the formula given in section 2.5.1:'. As with many of the section references in the former version 'black' text, these have not been updated after the inclusion of the new version 'blue' text. We suggest the formula is now at the top of the page 14 under the same 'section 4.1.2:'.	Agreed - done	Section 3.4.1

No.	Summary of the comment	Response	Section
9	Amend section 4.2.1.1 first line page 16, 'The (n=1, c=0) plans have been extended to include allowance for MU [diagram CXG 54-2004 See CXG 54-2004 Figure 1 where MU uncertainty bars are included for the decision process]'. The tables and figures in this document would be more informative if they all had numbering and captions.	Agreed	Section 3.1.2 And elsewhere, done for figures
10	Amend as follows Section 4.2.1 Example: Variables plan with negligible MU, 'Fat in Milk powder' step 3. '….manufacturing process is in a state of statistical control.'	Agreed - done	Section 3.2.1
11	Amend as follows 'Section 4.3.1 Example: Aflatoxin sampling plans due to Whitaker et al.(2006)'. Also, inclusion of related reference in section 6.6, 'Whitaker T. (2006) Sampling food for Mycotoxins, Food Additives and Contaminants, January 2006, 23(1) p50–61'.	Agreed - done	Section 4.6.5
12	Amend as follows Section 4.3.2 'Plan for Capsaicin – based on Codex Standard 294-20023. Also amend text 'Codex Standard CXS 294 - 20023 for Gochujang…..'. Also we note previously in CXS 294R:2009 the Quality Factors was (a) Capsaicin not less than 10.0 ppm (w/w). Suggest what is now in the CXS 294:2023 is incorrect and should appear as either '10 µg/mLg (w/w)' or preferably '10 µmg/mLkg (w/w), or alternatively the 'unconventional' 10 µg/mL (w/wv).	Agreed - done, using '10mg/kg'	Section 3.3.2
13	Section 4.3.2 figure at top of page 34. We do not believe the plan parameters in figure are correct. They should be m=13, k=1.20, not (m = 20, k = 1.55). Also, it may be better to keep the graph X-axis and Y-axis unit convention as it is provided in App1, i.e. not a proportion but a percentage value as axis naming suggests.	Agreed – corrected, the plan is actually (m=14, k=1.18) for the first example and (m=20, k=1.55) for the second	Section 3.3.2
14	Section 4.3.2 last paragraph page 34 states 'The corresponding sampling plan is (m=20, k=1.55) i.e. a composite sample would be formed from 18 subsamples randomly taken…' shouldn't this be 20 subsamples?	Agreed - corrected	Section 3.3.2
15	We believe the table under 'From CXG 50 Appendix II ISO INSPECTION PLANS INDEXED BY PRODUCER'S RISK' on page 36-37 replicates the table on page 18-19. Suggestion removal of table on pages 36-37 and referencing the Table # 'From CXG 50 Appendix II ISO INSPECTION PLANS INDEXED BY	Agreed - done	NA

No.	Summary of the comment	Response	Section
	PRODUCER RISK QUALITY (AQL) page 18-19:'		
16	Section 4.4.2 third paragraph. We have assumed the following options and subsections are linked (if correct it would be better for the reader that this is formalized), e.g. amending as follows. 'Four options have been evaluated: use of the single result in an (n=1, c=0) sampling plan for the assessment of compliance to an average level (see 4.4.2.1 Compliance of the average level scenario evaluation), use in an (n=1, c=0) attributes plan (see 4.4.2.2 Attribute plan scenario evaluation), use in a variables plan (see 4.4.2.3 Variable Plan scenario evaluation), and use in a plan based on the beta distribution if the characteristic is a compositional proportion and measurement uncertainty is negligible (see 4.4.2.4. Beta distribution plan scenario evaluation)'.	Agreed - done	Section 3.4.2, 4.4
17	Then amend the following subsections accordingly e.g. '4.4.2.1 Compliance of the average level scenario evaluation'	Agreed - done	Section 3.4.2.1
18	Is the reference in section '5.4 Combined attributes-variables plans' correct? As 'Refer CXG 50-2004 section 5.2.9.' is titled 'Fractional nonconformance'?	Agreed – done Additional note added.	Section 3.2.4
19	Section 5.6 'Lots consisting of bulk materials' page 46, last paragraph on page. Suggest the following amendment to highlight that this is a reference. "Bicking (1970) defines the following process for the design of sampling plans for bulk materials:"	Agreed - done	Section 4.6
20	Section 5.6, second last paragraph in section. Suggest the following amendment. "…level, where S is the standard error (standard deviation) of the estimate $\bar{x}$ of the mean level and k is the acceptability constant'.	Agreed – reworded but using different wording as this is different to the usual variables plans	Section 4.6.5
21	Section 5.6.5 'Aflatoxin sampling plans - Introduction', first paragraph, second sentence. Suggest following amendment 'Whitaker used data from 46 years of laboratory data, ….'	Agreed - done	Section 4.6.5
22	Section 6.1 'Background for the main (attributes & variables) sampling plans' last paragraph on page 57. We doubt this reference 'Section 5.1.1' is correct, should it be Section 4.1.2?	Corrected.	Section 3.1.2

No.	Summary of the comment	Response	Section
23	Section 6.2.2 'Attributes plans constructed in terms of OR and unity value' page 63, last paragraph, third sentence, 'For a given OR, refer to a table such as the following (Cameron). Is 'Cameron' a literature reference? If so, add a (year date) and provide details in section 6.6 'References'.	This entire section has been rewritten and simplified. The reference to Cameron has been removed.	Section 5.2
24	Section 6.4.3.1 'Lots consisting of discrete items - Scenario 1: known $\boldsymbol{\sigma}$, increase in producer risk' second sentence appears as 'The same acceptance sampling plan as in section Error! Reference source not found.2 is applied (n = 11, k = 1.025 with A = 90)'. The highlighted text requires link to be amended.	Corrected.	Section 5.4.3
25	Section 6.4.3.1 page 73 has text in brackets '[This part on uncertainty associated with the bias is under development]'. Has this been completed in Section 6.5.3. 'How to obtain a prior distribution'? In which case this text can be removed.	This text has been removed and the title of the section has been improved.	Section 5.4.3
Canada			
1	Canada thanks New Zealand and Germany for all of their work on the issue of sampling plan development and, in particular, the information document. There is an incredible amount of information shared in it.	Noted – thank you	
2	There are some concepts (such as utility functions) that may be familiar to those with experience in Bayesian statistics, but may be conceptually difficult to comprehend for someone with less experience. Perhaps including a glossary of terms would be beneficial.	Agreed – glossary of terms for Bayesian statistics included	Section 5.5.4
3	It is suggested that the examples taken from the Codex standards (e.g., attributes plans CXS 3-1981, CXS 207-1999, CXS 294-2003) be placed inside a box, so the reader immediately understands that they are not a continuation of the text.	Noted	Sections 3.1.1-3.1.3, 3.2.1 – 3.2.4, 3.3.1 & 3.3.2, 3.4.1 & 3.4.2
4	One of the most important things to clarify is the difference between plan evaluation (determine how successful a current plan is by estimating the risk) and design (choose plan based on obtaining desired risks).	Agreed	Section 2.3
5	Slightly different results from the app were obtained when testing the app and comparing it against the text in the document (e.g., P. 25, when using the	Agreed – yes, see point 11	Section 3.2.1

No.	Summary of the comment	Response	Section
	app, obtained n=43 instead of n=44 provided in the example). Was the app updated?		
6	It was noted that although the introduction states that “the document consists of three parts:” Part 1 and 2 are in bold but, Part 3 is not identified.	Agreed – changed	NA
7	P. 3: Correct section number: 5.6.3 to 5.6.5 – sampling plans for aflatoxins from Whitaker et al.	Agreed	Refers to section 4.6.5
8	P. 4, 3,1,2: Suggest rearranging sentence to read: "These two risks are generally referred to as the consumer’s risk and the producer’s risk, respectively."	Agreed	Section 2.1.2
9	P. 6: Suggest changing the decision diamond where it reads “OK?” to “Acceptable risks?” on the flow chart, to improve clarity.	Agreed	Section 2.2.2
10	P. 11: Suggest changing explanation in step 1 from “…as either PASS or having excessive…” to “…as either PASS or FAIL when compared against a reference powder” to improve clarity.	Agreed	Section 3.1.1
11	P. 14 (bottom): There is no section 2.5.1 (Section 2 is a intro/content summary, so any mention in the text of a Section 2.XX needs to be revised)	Agreed Reference changed to “see above”	Section 3.1.2
12	P. 21 Section 4.2.1: What are the units of the standard deviations? (percentage points?). Standard deviations have the same ‘units’ as the original data.	Agreed	Section 3.2.1
13	P. 22: Apps description: It appears the app has been updated since the document was written, as we get different results between the document when inputting the parameters in the app (the Workshop examples appear to reflect the updated app, though). Also, it appears that the numbering of the - Evaluation and Design Plans (1 and 2, respectively) have been removed. Suggest removing the numbering in the screenshots as well.	Agreed Yes, the app has been updated.	Section 3.2.1 and elsewhere
14	Also, in general, it is not obvious what the difference is between design and evaluated plans. It could be clarified that the aim of a "design" is to find the required sample size n and k, while the "evaluation" estimates the risks CR and PR. For example, at the end of P. 22, please clarify that one is "designing" a plan by finding n and k, so the "evaluation plan" is not relevant to the example.	Agreed	Refer section 2.3
15	P. 24: In the screenshot, suggest setting PRQ to 2.5% and CRQ to 10.0% to get	Agreed – screenshot aligned to text	Section 3.3.2

No.	Summary of the comment	Response	Section
	the same results and also to match the text.		
16	P. 24 (under Step 3): What is a "spinner"? Is this referring to the arrow buttons?	Text altered and image included	Section 2.5.1
17	[refer to earlier question] P. 25: For the design plan, n=43 is obtained with MU+offset, which is reflected in the Workshop example, instead of n=44 as in the screenshot. If this has been updated, please revise text in the document.	Yes, the app has been updated, the reason for the difference is statistical, and explains why the two versions of the same plan have different sample numbers, n=43 and n=44, in the last two rows of the table below the OC curve.  The results in the 'without MU' row are based on an exact calculation (assuming the data is normally distributed) but the number of samples in the last row 'with MU & Offset' are based on an approximation, as there is no exact calculation when the between-laboratory standard deviation is allowed for (even if it is zero). A note could be included in the document somewhere, but the explanation is rather technical and not in keeping with the content in the rest of the document that we've tried to keep on a relatively simple level statistically. The basis for the calculations used in the app has been documented but we haven't considered whether it should be included of if so, how.	Section 4.6 and section 4.6.3 in particular Section 3.2.1
18	P. 32 Section 4.3.2 (first paragraph): Suggest define w/w?	Agreed	Section 3.2.2
19	P. 40 Section 4.4.2.3 in the p.acc: NC should not be a percentage	Agreed	Section 3.4.2.3
20	P. 47: Please complete the sentence "The acceptance criteria will be of the form: $\bar{x} + k \cdot S \leq USL$ for upper specification limit USL for the average level, where S is the standard error (standard deviation) of the estimate $\bar{x}$ of the mean level and k is . . ."	Agreed Reworded	Section 4.6.5
21	P. 48, 5.6.5: Suggest rearranging sentence to read: "Whitaker used 46 years of laboratory data, including some from contaminated lots, to derive Horwitz-type equations for the sampling, subsampling,	Agreed	Section 4.6.5

Plan	n	k	PRQ	PR	CRQ	CR	Offset
Evaluated Plans:							
Without MU	12	1.37	2.1	5	20.4	10	0
With MU	12	1.37	2.1	5	20.4	10	0
Designed Plan:							
Without MU	43	1.59	2.4	5	9.7	10	0
With MU & Offset	44	1.59	2.4	5	9.7	10	0

No.	Summary of the comment	Response	Section
	and analytical components of the total variation."		
22	P. 50: The FAO Mycotoxin Tool user guide explicitly indicates that the "among lab" analytical variance is 2x the within lab analytical variance. Suggest a change to reflect this, rather than identify this value based on a common assumption.	Agreed Text has been amended	Section 4.6.5
23	P. 75, Table 1: What is difference between specific PR and conditional PR? The definitions are the same. Likewise, what is difference between specific CR and conditional CR in Table 2? It appears specific PR refers to a unique value of y (in rejection region), whereas condition PR is across all possible outcomes of y that are rejected. Please clarify this.	Correct. The text has been changed to clarify.	Section 5.5.2
24	P. 79 Section 6.5.3 (first paragraph): Suggest replacing 'bias' with 'influence' to read ". . . unduly influence the analysis" because bias has a different statistical implication.	Text has been removed.	
25	P. 82: k doesn't seem to be defined for Equation 2 and 3, is k the index for the integration (# of non-conforming units, from 0 to c)? If so, suggest another letter, or something like c', as this is an attributes plan, and the "k" appears to be different than the one in the variables plan.	Agreed. Text has been simplified by removing the equations.	Section 5.5.3.9
26	Also, the descriptions of both p_conf and p_nonconf suggest they are posterior distributions, but according to the formulas it seems that they are joint distributions (that is, p_conf+p+nonconft is the probability of acceptance). Please clarify. If the interpretation is correct, please describe p_conf as "probability that a given item in a lot is both accepted and conforming"? Likewise for p_nonconf (lot if both accepted+nonconforming).	This is correct. However, we have changed the utility function approach. These expressions no longer play a role.	Section 5.5.3.9
27	Some of the examples in the Algorithms section 6.5.4 are a little difficult to follow, as some quantities (such as the Utility in Table 5 on p.82) appear to need the use of numerical or computational methods, and require experience with statistical software. Perhaps mention the need for additional steps.	Agreed. The explanation of this approach has been rewritten and simplified.	Section 5.5.3.9
28	Table 5- The headings for Benefit and Utility are both [B_conf], please clarify what is meant. It appears this is due to different reasons, B_conf is the variable name for Benefit (equal to 1), whereas Utility is expressed in B_conf as a unit. This implies that Utility is proportional to B_conf, is this true? Please confirm.	Agreed. This has been clarified.	Section 5.5.3.3

No.	Summary of the comment	Response	Section
29	P. 84 Equation 6: Should U_c(0|x) be U_cons(0|x) for consistency?	Correct. However, this text has been removed.	Section 5.5.3.9
30	P. 85 Phase 1 description: A Beta(1,10) consumer prior is mentioned in the text, but in Table 7, as well as the Phase 2 section, it says Beta (1,6). Please clarify.	This text has been removed.	Section 5.5.3.9
31	P. 85: Please remove blank first column of Table 7, if there is no content there.	This table has been removed.	Section 5.5.3.9
Japan			
1	Japan appreciates the preparation of this information document by New Zealand and Germany as the Chair and co-Chair of EWG. Japan would like to recall that this information document is to serve as a supplement document to CXG 50, as agreed by CCMAS40. Thus, Japan generally supports the sampling plan in the document within the scope of CXG 50 because the parts will help users to understand CXG 50.	Noted – thank you	
2	However, we have concern about some part of this document containing information beyond CXG 50, such as Bayesian sampling. Such information is included neither in CXG 50 nor any other Codex guidelines. CCMAS has never discussed this issue thoroughly at the plenary, either. Japan expresses strong concern to include such new concept as a part of Codex recommendation without any discussion at the plenary, even though it is an “unofficial” format (i.e. information document).	We would like to thank Japan for this comment which will allow us to shed some light on several important issues. We agree with Japan that introducing a new concept in CCMAS guidelines requires justification and thorough discussion. The reason why we have introduced Bayesian concepts in the discussion is that an urgent need for plans with low sample size was expressed in CCMAS42, and that many delegations expressed strong interest in Bayesian plans in the workshop held in CCMAS43. We believe that Bayesian methods represent a pragmatic “real-world” approach to acceptance sampling which allows—under certain circumstances—a considerable reduction in sample size, and which is also in line with Section 3.2.2 of CXG 50 that states that “in selecting a sampling plan, it should be ensured that producers are not exposed to unreasonable costs in terms of sampling and testing, loss of yields, […]” We believe this is the right time to discuss Bayesian plans in the plenary of CCMAS since	

No.	Summary of the comment	Response	Section
		such plans are currently being finalized in ISO committees and working groups. In particular, a technical report on the application of Bayesian methods to acceptance sampling will soon be published by ISO TC 69 SC 5 WG 10. In the current draft of the CXG 50 information document, the Bayesian plans are described in a statistical annex. They are thus separated from the main body of the text and described in a “for information purposes” spirit. There is no imperative to apply Bayesian methods.	
3	With regard to Bayesian sampling, it could potentially be applied during the commodity production phase at the manufacturing stage, where it is feasible to predict a prior distribution of a characteristics of the lot. However, does this mean that Bayesian sampling will not only be intended for continuous series lots? In our view, Bayesian sampling is not suitable for import/export inspections of food, where only isolated lots are handled, as clearly stated in the scope of CXG 50.	We agree with Japan that the application of Bayesian methods in acceptance sampling is only possible if there is a similarity between the current lot and previous lots. Such a similarity can also exist with isolated lots and can be empirically checked. Our understanding is that “isolated lot inspection” means that the lot is inspected rather than produced in isolation. Hence, it is perfectly reasonable to assume that prior to a new lot inspection, the consumer has past experience with the supplier of the new lot and has information from previous tests on products from this supplier. Indeed, this point is made in Section 1.2 (Scope) of CXG 50: “The term ‘isolated’ means that the inspection of each lot is done in isolation, without considering the outcome of the inspection of adjacent lots or, for example, other lots from the same producer. This does not mean that information from previous inspections cannot be used; in particular, there are cases where the lot standard deviation may be known from the inspection of previous lots.”	

Appendix III

LIST OF PARTICIPANTS

CHAIR

New Zealand

Susan Morris
Ministry for Primary Industries - New Zealand

Roger Kissling
Fonterra - New Zealand

CO-CHAIR

Germany

Petra Gowik
BVL - The Federal Office of Consumer Protection and Food Safety - Germany

Steffen Uhlig
QuoData - Germany

Bertrand Colson
QuoData – Germany

MEMBER NATIONS AND MEMBER ORGANIZATIONS
ÉTATS MEMBRES ET ORGANISATIONS MEMBRES
ESTADOS MIEMBROS Y ORGANIZACIONES MIEMBROS

AUSTRALIA - AUSTRALIE

Richard Coghlan
Senior Technical Expert, Analytical Services Branch, NMI-Australia

BRAZIL - BRÉSIL - BRASIL

Ligia Lindner Schreiner
Health Regulation Expert, Brazilian Health Regulatory Agency – Anvisa

Ana Claudia Marquim Firmo De Araujo
Health Regulation Expert, Brazilian Health Regulatory Agency – Anvisa

CANADA - CANADÁ

Thea Rawn
Research Scientist, Chemical Contaminant Section, Health Canada

ECUADOR – ÉQUATEUR

Rosa Chalon
Analista, ARCSA, Ecuador

EGYPT - ÉGYPTE - EGIPTO

Mariam Barsoum Onsy
Food Standards Specialist, Egyptian Organization for Standardization & Quality (EOS)

EUROPEAN UNION - UNION EUROPÉENNE - UNIÓN EUROPEA

Franz Ulberth
Scientific Expert, European Commission, European Union

FRANCE - FRANCIA

Laurent Guillier
Statistician, French Agency for Food, Environmental and Occupational Health & Safety (ANSES)

HUNGARY - HONGRIE - HUNGRÍA

Attila Nagy
Chairman of CCMAS
Krisztina Bakó-Frányó
Codex Contact Point of Hungary

JAPAN - JAPON - JAPÓN

Hidetaka Kobayashi
Coordinator, Risk and Crisis Management, Ministry of Agriculture, Forestry and Fisheries of Japan

Takahiro Mori
Associate Director, Ministry of Agriculture, Forestry and Fisheries of Japan

Takahiro Watanabe
Section Chief, Division of food safety information, National Institute of Health Sciences

Yuusuke Miyaaki
Assistant Director, Min of Health, Labour and Welfare

Kazuko Fukushima
Director, Office of Import Food Safety, Min of Health, Labour and Welfare

NIGERIA - NIGÉRIA

Ibrahim Yahaya
Codex Contact Person SDD Nigeria

PARAGUAY

Mauricio Rebello

PHILIPPINES – FILIPINAS

Lourdes Timario
Supervising Science Research Specialist, Chairperson, NCO Sub-Committee on Methods of Analysis and Sampling (SCMAS). Food Development Center, Dept of Agriculture

Christmasita Oblepias
Food-Drug Regulation Officer IV, Co-Chairperson, NCO SCMAS, Food and Drug Administration. Department of Health

REPUBLIC OF KOREA - RÉPUBLIQUE DE CORÉE - REPÚBLICA DE COREA

Kim Youngjun
Codex Researcher, Ministry of Food and Drug safety

Korea Codex Contact Point
Quarantine Policy Division, Ministry of Agriculture, Food and Rural Affairs (MAFRA)

Kiseon Hwang
CODEX/SPS Researcher, Ministry of Agriculture, Food and Rural Affairs

SAUDI ARABIA - ARABIE SAOUDITE - ARABIA SAUDITA

Nimah M Baqadir

Abdulaziz A Al Qaud
Senior Product Registration Support Expert, Saudi Food and Drug Authority, Kingdom of Saudi Arabia

Mubarak M Al-Garaiwi
Senior Scientific Evaluation Expert, Saudi Food and Drug Authority, Kingdom of Saudi Arabia

Abdullah A Al Sayari
Section Head of Hormones and antibiotics, Saudi Food and Drug Authority, Kingdom of Saudi Arabia

Mohrah A Alenazi
Lab Expert, Saudi Food and Drug Authority, Kingdom of Saudi Arabia

SINGAPORE - SINGAPOUR - SINGAPUR

Ken Lee
Branch Head, Singapore Food Agency

Ivan Ng
Senior Scientist, Singapore Food Agency

THAILAND - THAÏLANDE - TAILANDIA

Chitrlada Booncharoen
Standards Officer, National Bureau of Agricultural Commodity and Food Standards, Ministry of Agriculture and Cooperatives, Thailand

Kittiporn Pinke Phuangsukaw
Standards Officer, National Bureau of Agricultural Commodity and Food Standards, Ministry of Agriculture and Cooperatives, Thailand

Rungrassamee Mahakhaphong
Standards Officer, National Bureau of Agricultural Commodity and Food Standards, Ministry of Agriculture and Cooperatives, Thailand

URUGUAY

Laura Flores
Laboratorio Tecnologico del Uruguay

UNITED STATES OF AMERICA ÉTATS-UNIS D'AMÉRIQUE - ESTADOS UNIDOS DE AMÉRICA

Patrick Gray
Research Chemist, Center for Food Safety and Applied Nutrition, Office of Regulatory Science, US Food & Drug

Timothy Norden
Chief Scientist, Agricultural Marketing Service – Technology and Science Division, US Dept of Agriculture