CODEX ALIMENTARIUS COMMISSION E

Food and Agriculture Organization of the United Nations

World Health Organization

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

Agenda Item 5 **CX/MAS 24/43/7**
March 2024

JOINT FAO/WHO FOOD STANDARDS PROGRAMME

CODEX COMMITTEE ON METHODS OF ANALYSIS AND SAMPLING

43rd Session
13 – 18 May 2024

INFORMATION DOCUMENT *THE GENERAL GUIDELINES ON SAMPLING* (CXG 50–2004)

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

Codex members and Observers wishing to submit comments on this document should do so as instructed in CL 2024/16-MAS available on the Codex webpage/Circular Letters:
http://www.fao.org/fao-who-codexalimentarius/resources/circular-letters/en/

Introduction

1. The 42nd Session of the Committee on Methods of Analysis and Sampling (CCMAS42) agreed to re-establish the EWG, chaired by New Zealand and co-chaired by Germany, working in English, to continue working on the information document namely the e-book with the sampling plans applications for consideration by CCMAS43.

2. It was also advised that this supporting information document containing the sampling plan apps would be part of the next phase of the work to facilitate the understanding and implementation of the revised Guidelines. The information document would contain more detailed examples on measurement uncertainty and some practical examples of sampling plans, amongst others.

3. It was anticipated that the information document would be completed during 2023-2024 and presented to CCMAS43 for finalization. Following the EWG process, completion of the information document is going to take more time. This report therefore includes the information document to date (Appendix I), along with a proposal to extend finalization of the work to 2025.

EWG Registration and consultation

4. EWG registration on the information document was sent out in June 2023. Registrations included 27 member countries and as well as an observer organisation. The list of participants is in Appendix II.

5. A first round of consultation with the EWG was undertaken in July 2023. We, the chairs of the EWG. sent out a proposed structure for the information document and asked for feedback on this. We noted this to be an overview of the information document and its purpose. We set out the aim to provide specific examples relating to CCMAS questions and/or Codex standards to be used and to provide guidance for CCMAS and/or other committees responsible for creating or modifying, revising sampling plans. We also noted the possible discussion of the current relationship between CCMAS guidance documents and ISO standards. Excellent comments were received from Canada, Brazil, Australia, Uruguay and Singapore. There was strong support for examples for specific scenarios. There was also support for presenting a first draft of the information document to the CCMAS43 as well as proposals to hold an exemplary seminar in the margins of CCMAS. In addition, on-line tutorials were proposed to help with the use of the app and development of sampling plans.

6. A second round of consultation with the EWG was undertaken in December 2023. A draft of the information document as well as a link to the sampling plan app and an excel version were provided. We considered it important to note that the draft document is long and that there are parts that are statistically complex. On that basis we summarized the content at the start. Detailed comments were received from Japan, Canada and Australia.

7. Comments from these consultations resulted in updates to the information document. The content of the draft document covered a lot of information delegations want included. However, important points were raised:

- Whether related and relevant new science that is happening in parallel should be included in the information document? There are active discussions on Bayesian approaches in other standards

development organisations, given the way this approach can reduce the number of samples and therefore costs, and this is relevant to CXG 50.

- Additional content, for example more focus on real-life practical examples based on commodity standards. While we have included some examples in the draft information document, we can provide more, but to do this, we need to have a better understanding of what country delegations want and then time to develop these specific examples.
- To have both current and a forward-looking information document. This takes into account both relevant practical examples as well as additional theoretical background.
- There was also EWG comment on the need to review CXS 234 references to sampling plans.

CONCLUSION

8. The EWG has prepared a draft information document taking up discussions to date, however, several questions have been raised which will need to be addressed in order to complete the information document. It is proposed to extend the timeframe for completion of this work by one year to 2025.

RECOMMENDATIONS

9. CCMAS is invited to:

a. consider the draft information document (Appendix I) and the questions raised in para. 7 above:

b. re-establish the EWG to further develop the information document for finalisation by CCMAS44.

c. Consider a review of the sampling plans contained in CXS 234 as part of the further work of the EWG (see b) above).

Appendix I

Proposed draft Information Document for the *General Guidelines on Sampling* (CXG 50-2004)

(For comment through CL 2024/16-MAS)

1 Introduction

The purpose of this document is to provide additional information on the sampling plans referred to in CXG 50, including background and examples for each of the main types of sampling plan. A link to the app1 for the design and evaluation of these sampling plans is included.

The document consists of two parts:

Part 1 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 when measurement uncertainty is negligible. An Excel version of this app is also provided.

Section 3 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 4.1 to 4.5 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 4.6 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 4.7 contains information about other sampling plans, including 3-class attribute plans, used for microbiological assessments.

Part 2 contains statistical background detail including some information about current developments in Bayesian sampling plans:

Section 5 is a statistical annex that describes the derivation of attributes and variables plans when measurement uncertainty is negligible and references for plans discussed in the document.

Section 5.4 discusses measurement uncertainty and its role in acceptance sampling.

Section 5.5 discusses options for Bayesian plans that allow control of specific risks, based say, on historical experience. These plans offer considerable potential for controlling risks using economical levels of sampling and testing. Some examples are given.

Note:

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

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 of the following:

- 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 risk 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 referred to generally 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.

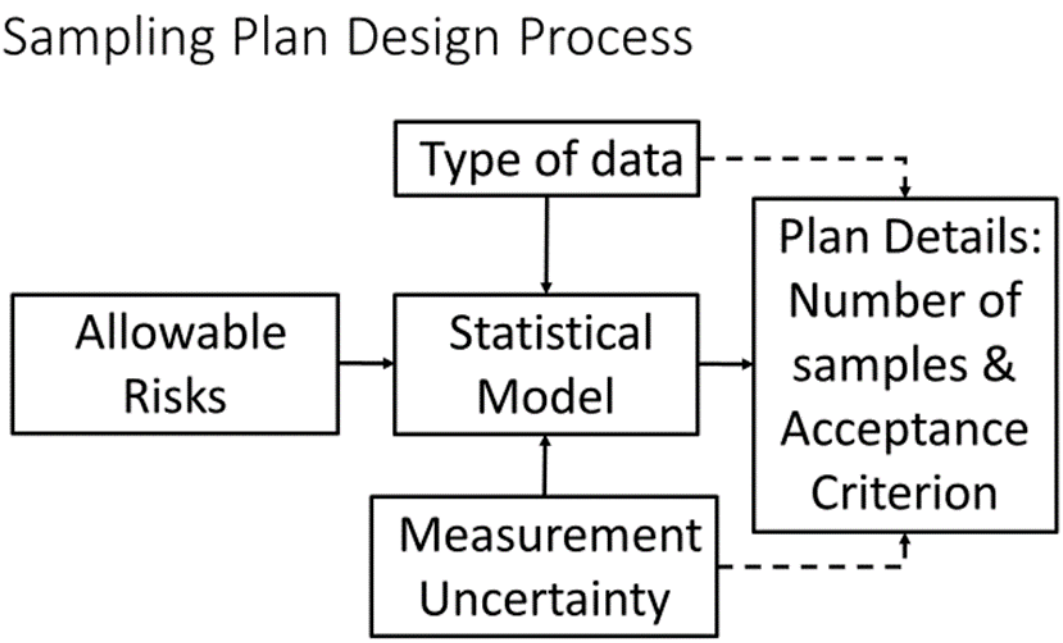
This is the traditional approach to the design of sampling plans, in this approach risks are considered non-specifically "if a lot contained x% of nonconforming product, then the probability of acceptance would be "y".

However, it is also possible to consider risks in a specific sense, using a Bayesian approach employing a prior distribution based on historical inspections of a characteristic in a product.

The use of Bayesian methods has the potential to develop sampling plans that control risks using economical levels of testing. Refer section 5.5 for more details.

2.2 Design of sampling plans

2.2.1 Overview of the 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).

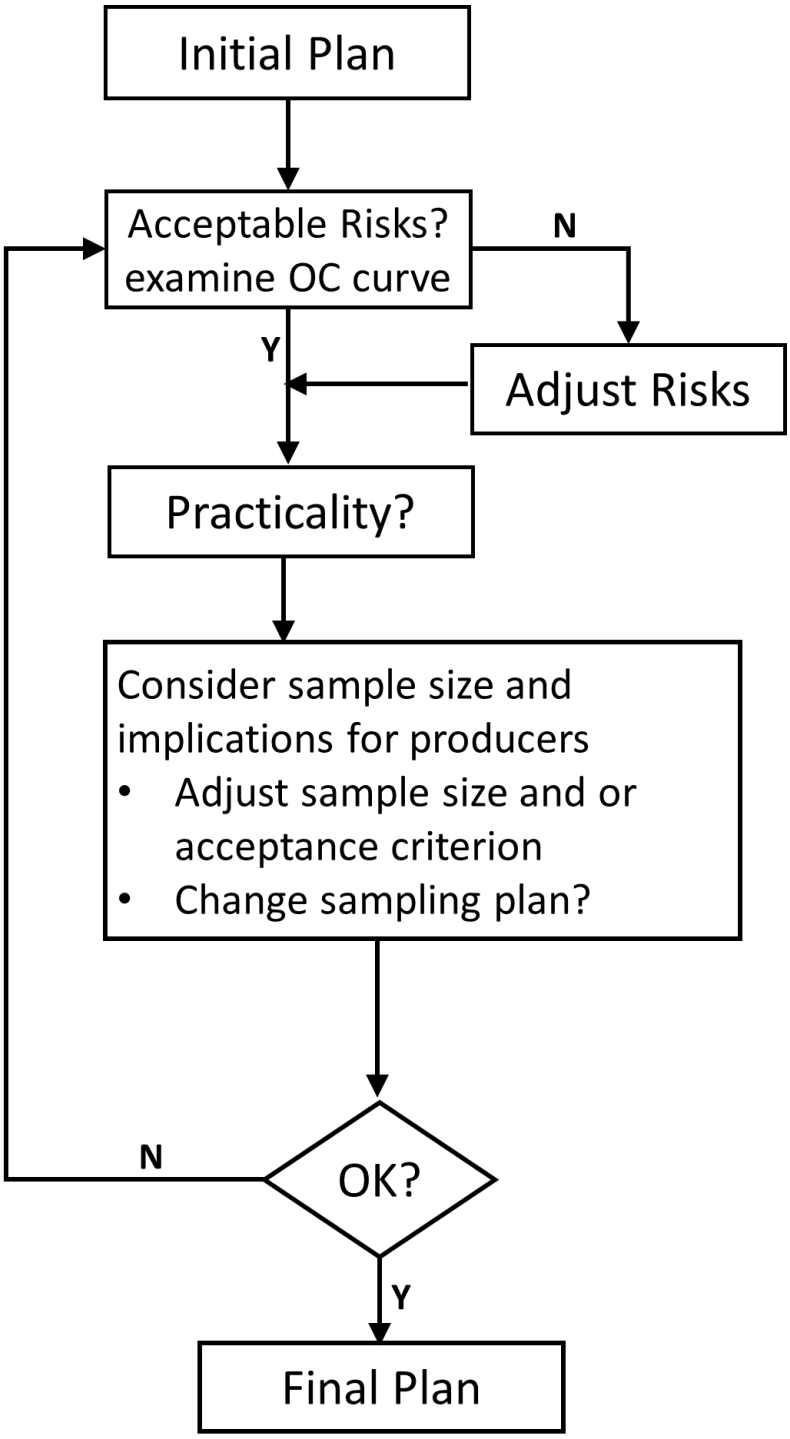
Specification of the allowable risks is a key input, the producer's risk and the consumer's risk might both be specified or only one of those risks might be specified. In the ISO plans only the consumer's risk would be specified for the inspection of isolated lots of incoming goods whereas plans based 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.



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 percent

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 percent

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 percent (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 parameters.

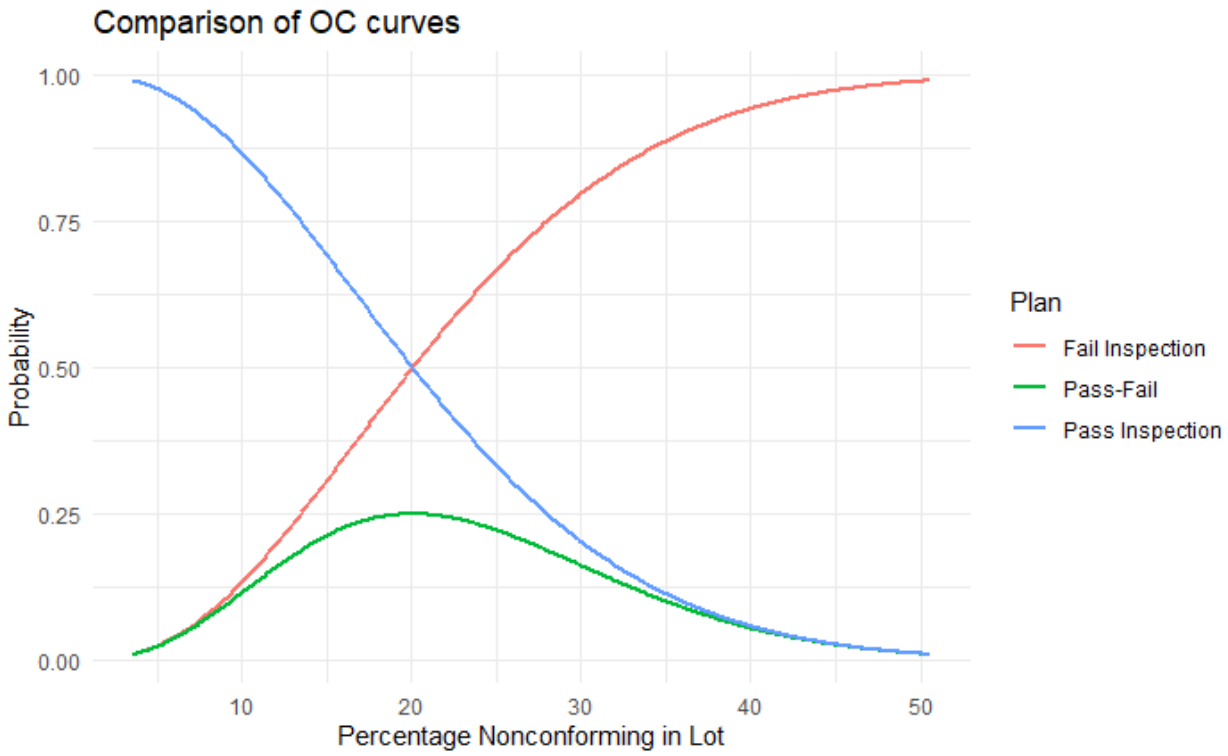
In the interests of fairness, designers of plans should also take account of the measures that the producer may have to take to ensure compliance, remembering that in Codex, CXG 50-2004 will be applied mainly to compositional characteristics and commodity defects (refer CXG 50-2004 3.2.2 Fairness).

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 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 test the product; that should be considered in the design of plans to ensure fairness.



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 BIPM International Recommendation R087 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 weight should be at most 0.5%.

Consumer's Risk

The probability of accepting a lot whose true mean weight is less than the label 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 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 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.

2.3.1 Description of apps

An Excel version of App 1 has been included with this Information Document as well as a version of App1 – see the following section.

Sampling plan design and evaluation tool.xls

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

An updated version of this app has been deployed on the shinyapps.io server:
https://codex-testing.shinyapps.io/codex-testing-SamplingPlan/

App3 demonstrates variables plan for averages. There are different parameters that can be selected. These include whether the standard deviation is known, whether the specification limit is upper or lower, and what this particular specification limit is. If the standard deviation is known, its value is entered. The sample size and *k*-constant are also selected, along with the producer's and consumer's risks. The OC curves will be different depending on whether the standard deviation was known or not, and these curves can be compared.

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

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

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

3 Case studies (examples for specific scenarios)

Note: among other things, these examples illustrate the use of the apps and there is step-by-step design process described in CXG 50-2004 Appendix 1.

3.1 Attributes plans

3.1.1 Example: Attributes plan with c>0

Browning in Milk powder

- A customer has 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 having excessive browning 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%.

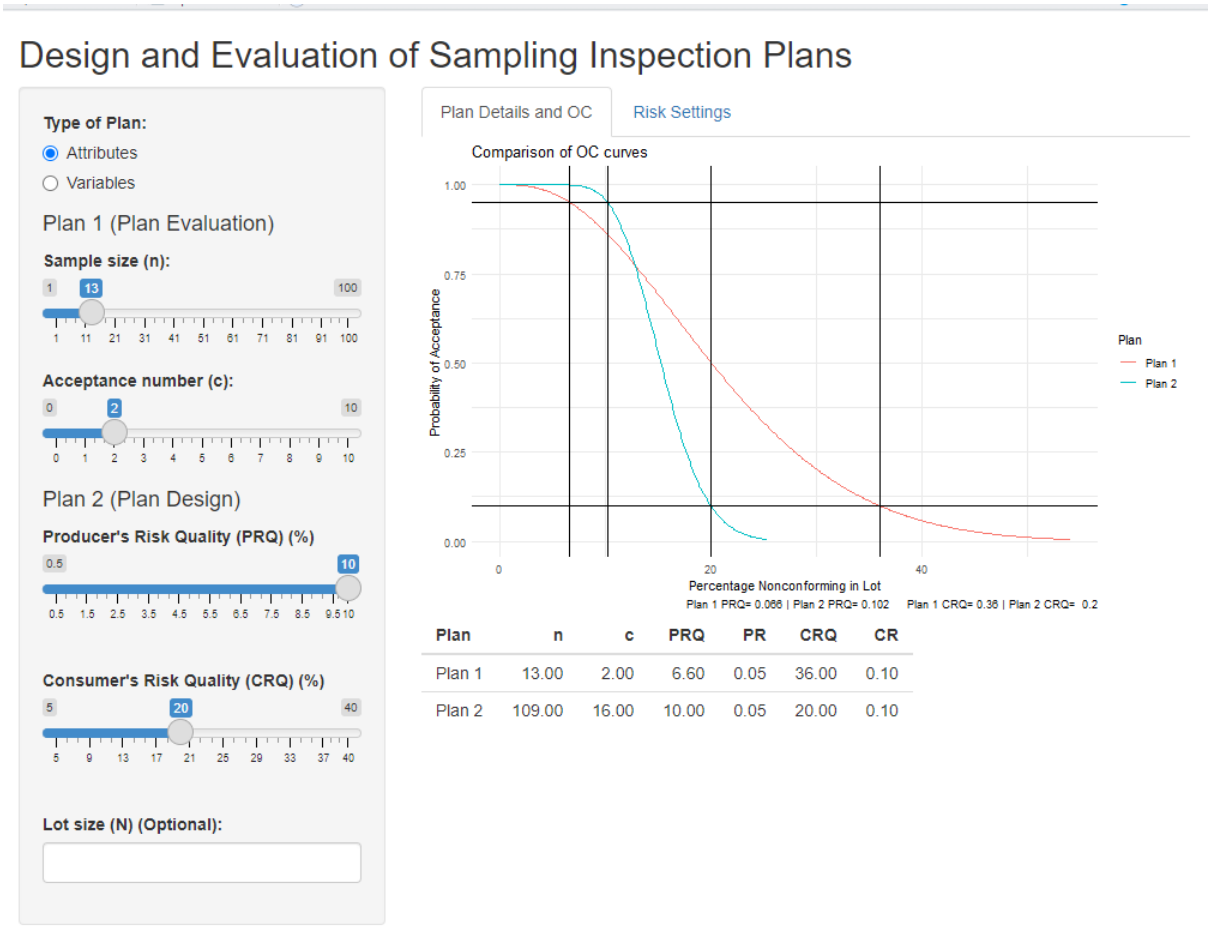
4. Set producer's risk quality level (PRQ):

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:

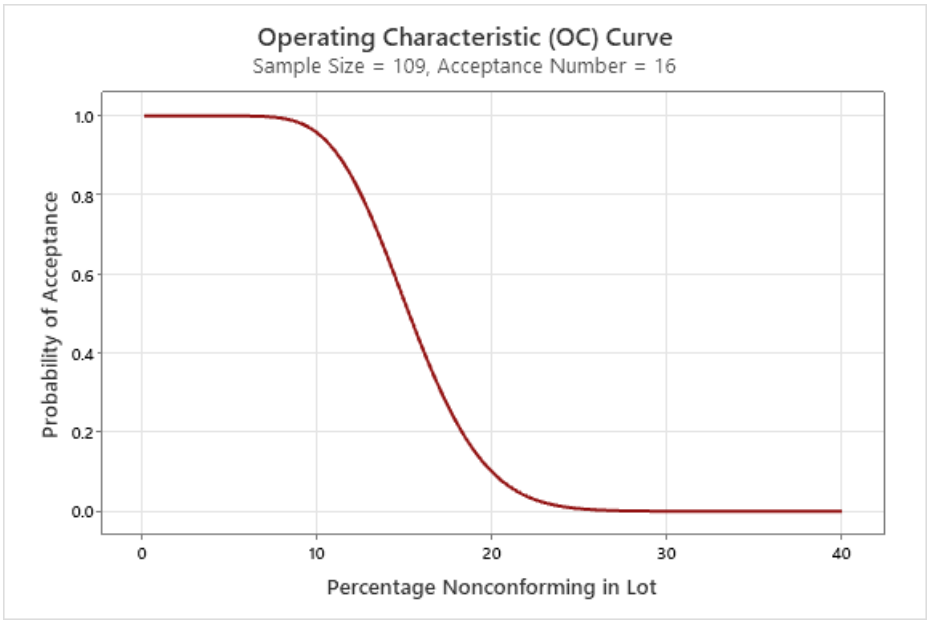
PRQ	**CRQ**	**n**	**c**
5%	20%	38	4
10%	20%	109	16
15%	20%	500	88

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.



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.



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

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 in section 2.5.1.:

$CR = (1 - CRQ)^n$ 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 (as also mentioned in section 2.5.1).

3.1.3 Example: Attributes plan based on ISO 2859-2

A commercial farming operation supplies eggs to retailers for sale. A retailer, the consumer, enters a contract to purchase eggs on condition that only a small proportion of the cartons of one dozen (12) eggs in a lot contain any broken or cracked eggs.

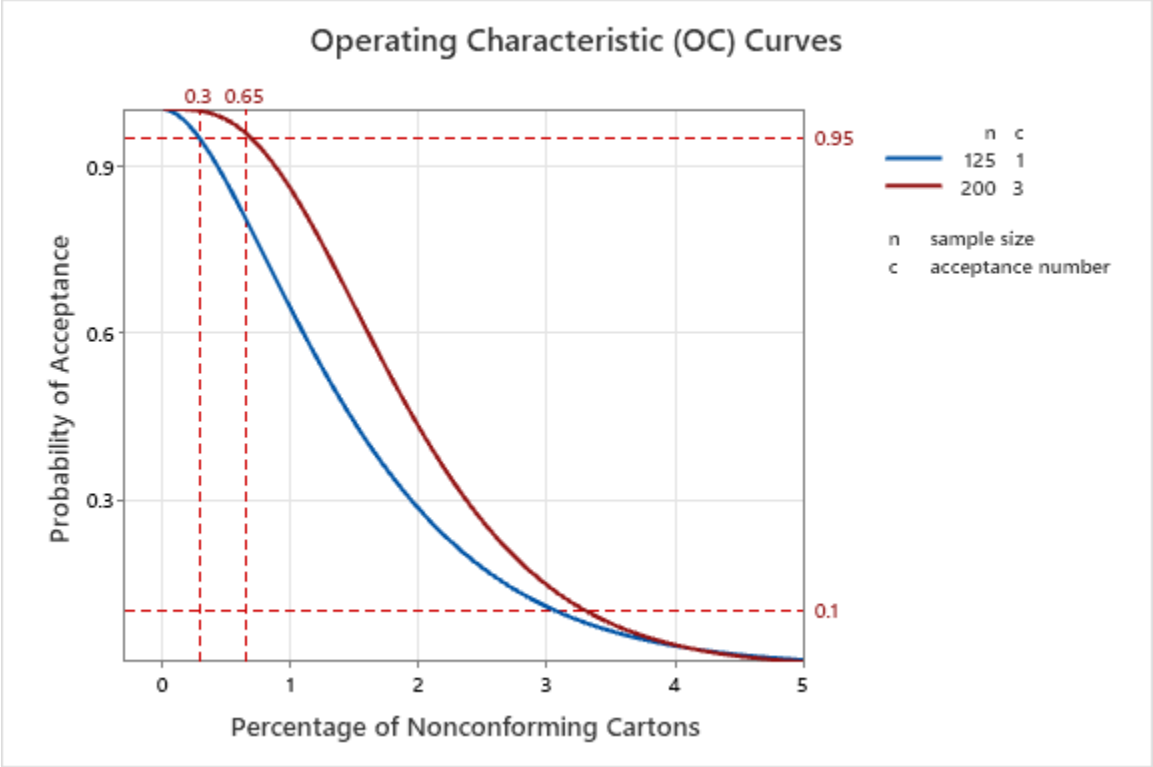
It was agreed that the retailer could carry out an inspection of each lot of eggs sold to the retailer using sampling, and that the sampling plan should have a Consumer's Risk Quality Level (LQL) of 3.15%; i.e. there is a 10% chance of accepting a lot in which 3.15% of the cartons containing cracked or broken eggs.

The proposed lot size is N = 1250 cartons. Using the tables in ISO2859-2, for normal inspection (inspection level II) of a lot of size of N = 1250 cartons, with CRQ (LQ or LQL) of 3.15% gives the sampling plan (n=125, c=1). This plan has a PRQ (AQL) of 0.40%.

Lot sizes for inspection levels					ISO 2859-1 Single sampling plan (normal inspection)			Sample size code letter
S-1 to S-3	S-4	I	II	III	AQL	~~A~~ *n*	Ac	
126 or over	126 or over	126 to 35 000	126 to 3 200	126 to 1 200	0,40	125	1	K
		35 001 to 150 000	3 201 to 10 000	1 201 to 3 200	0,65	200	3	L
		150 001 to 500 000	10 001 to 35 000	3 201 to 10 000	0,65	315	5	M
		500 001 or over	35 001 or over	10 001 or over	1,00	500	10	N

If, however, the lot size was N=5000, the resulting sampling plan would be (n=200, Ac = 3) with an AQL of 0.65%. This shows that when using the ISO plans the sampling requirements for larger lots is more economical, one sample per 10 cartons in the first case, one per 25 in the second.

The Operating Characteristic curves for these two sampling plans are shown below:



3.2 Variables plans

3.2.1 Example: Variables plan with negligible MU

Moisture in Milk powder

The provision for Moisture in the *Standard for milk powders and cream powder* (CXS 207-1999) states that Moisture in whole milk powder should not exceed a maximum of 5%.

Key steps in the step-by-step design process

1. Attributes or variables?

Moisture is a measured parameter so measurements are variables data.

2. Does the provision relate to the average value or to the entire distribution?

The provision specifies a maximum 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 the characteristic [moisture] is normally distributed that is a reasonable assumption if the manufacturing process is in a state of statistical control. However, since moisture is a compositional proportion, the beta distribution might provide a better description of the behaviour of moisture with the lot [provided measurement uncertainty was negligible] as well as being more economical (refer section 3.3.2).

4. Is measurement uncertainty negligible or non-negligible?

In this example measurement uncertainty is considered negligible.

5. Specify the stringency required for the sampling plan.

Consumer's Risk Quality (CRQ):

What percentage nonconforming would you 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)

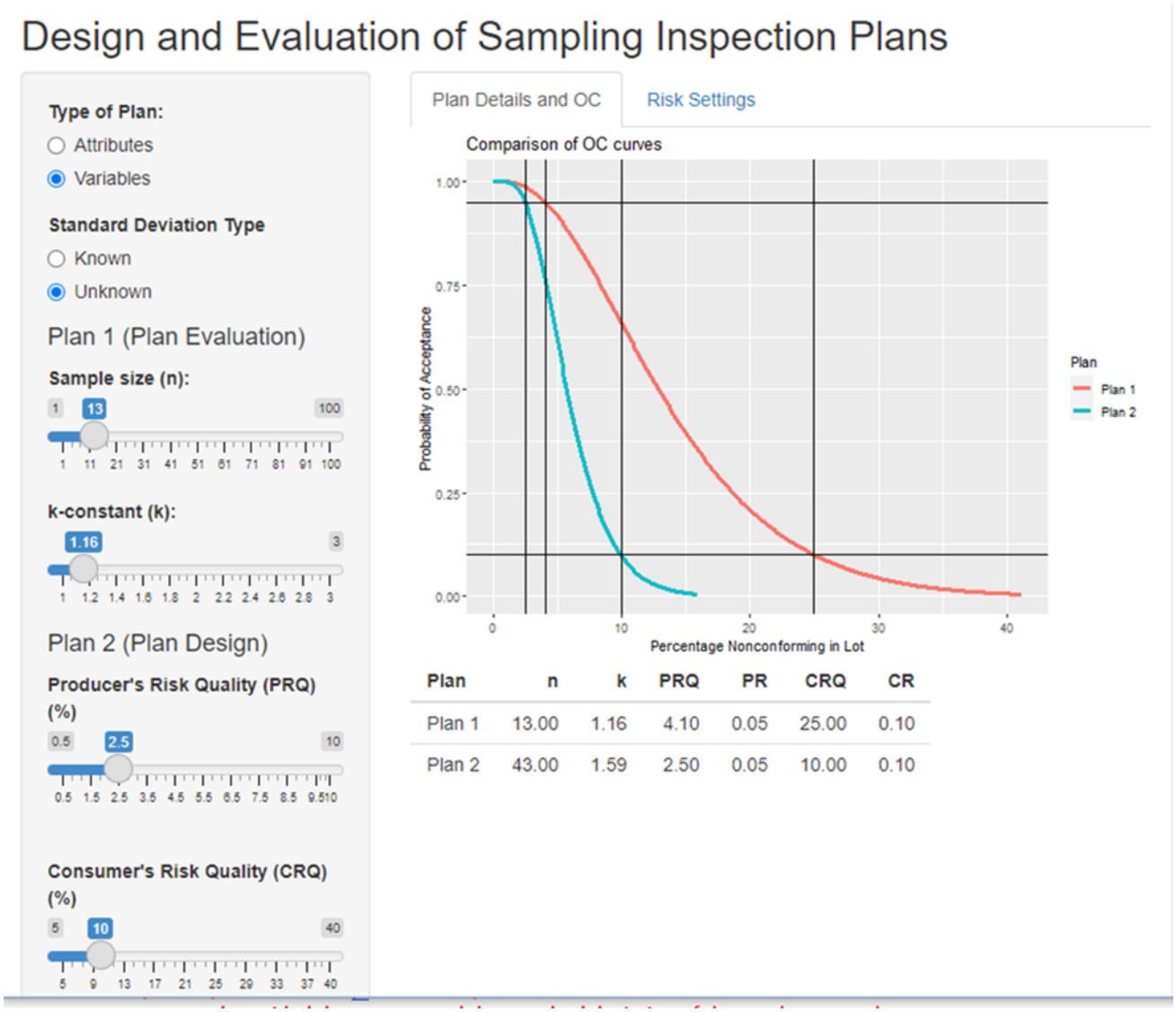
What percentage nonconforming would you allow in lots that you would want to accept most of the time?

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,



The plan required to control risks to the specified levels is (n=43, c=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:

$$\bar{x} + \mathrm{k} \times s \leq 5$$

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. moisture of 5% on a weight/weight basis.

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

Refer CXG 50-2004 Section 5.2.5

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.

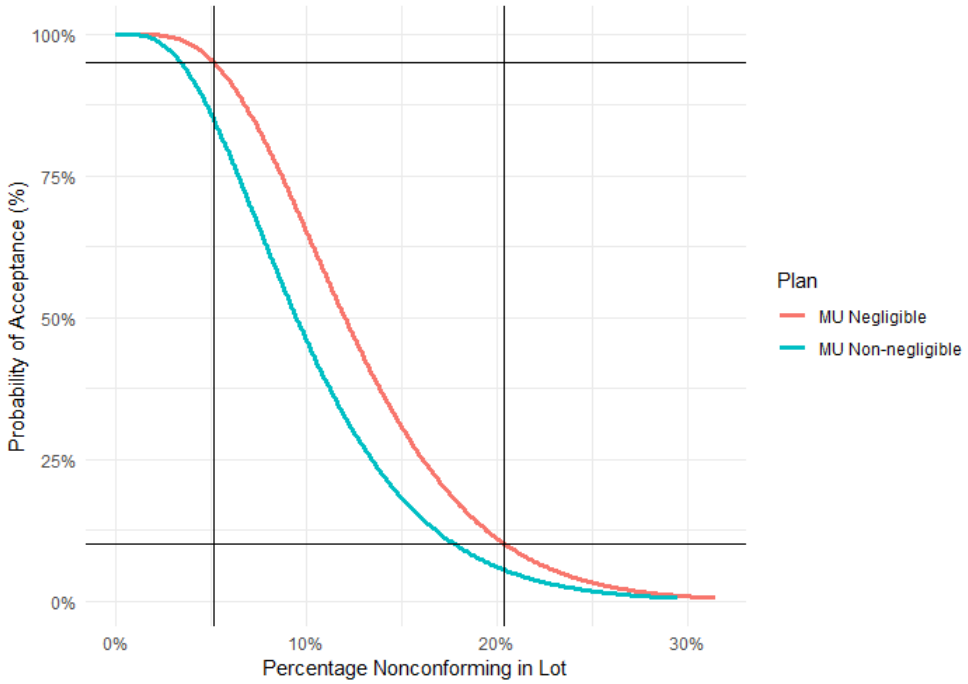
If the assessment of compliance proceeds in the usual way, the mean value of the results is m=9.90, the standard deviation s = 0.12, so that the acceptance criterion has a value of 9.90 + 1.19*0.12 = 10.04, and the lot should not be accepted.

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.



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

See examples in the *Guidelines on measurement uncertainty* (CXG 54-2004) Information document Section 10.

3.2.4 Example: Fractional Nonconformance plans

Refer CXG 50-2004 Section 5.2.7

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:

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

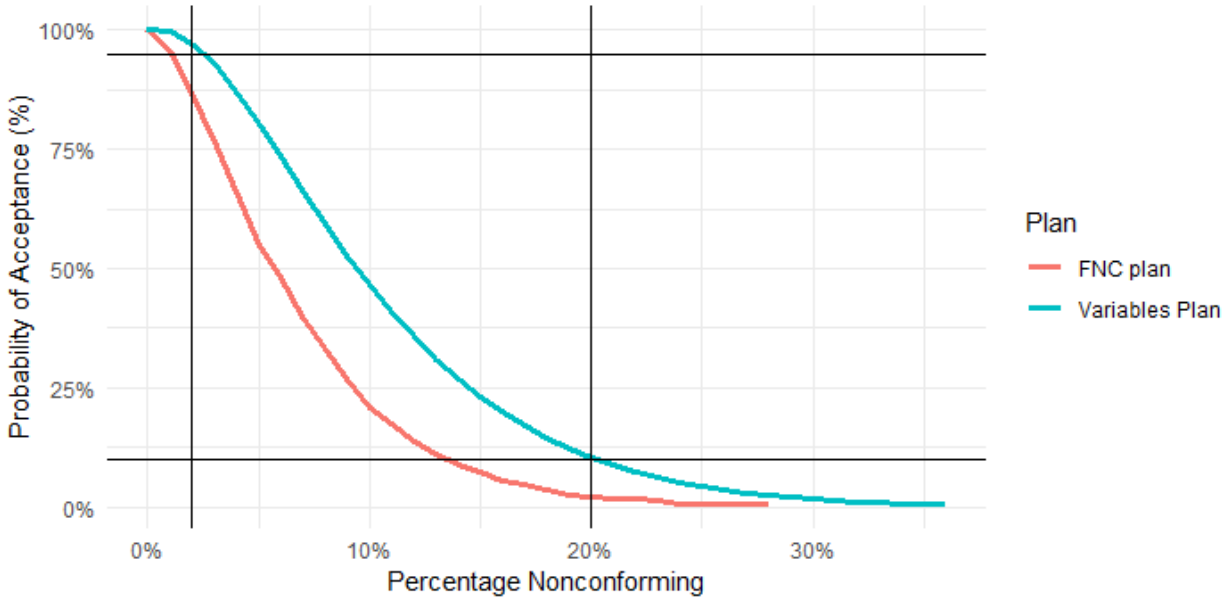
where 'x' represents a single test result. This gives the following FNC values:

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.



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 et al.

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 20µg/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[1].

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

or 90.6%

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 – Hypothetical Example

For capsaicin it is not feasible to perform more than relatively few tests on each lot.

As above, if measurement error 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:

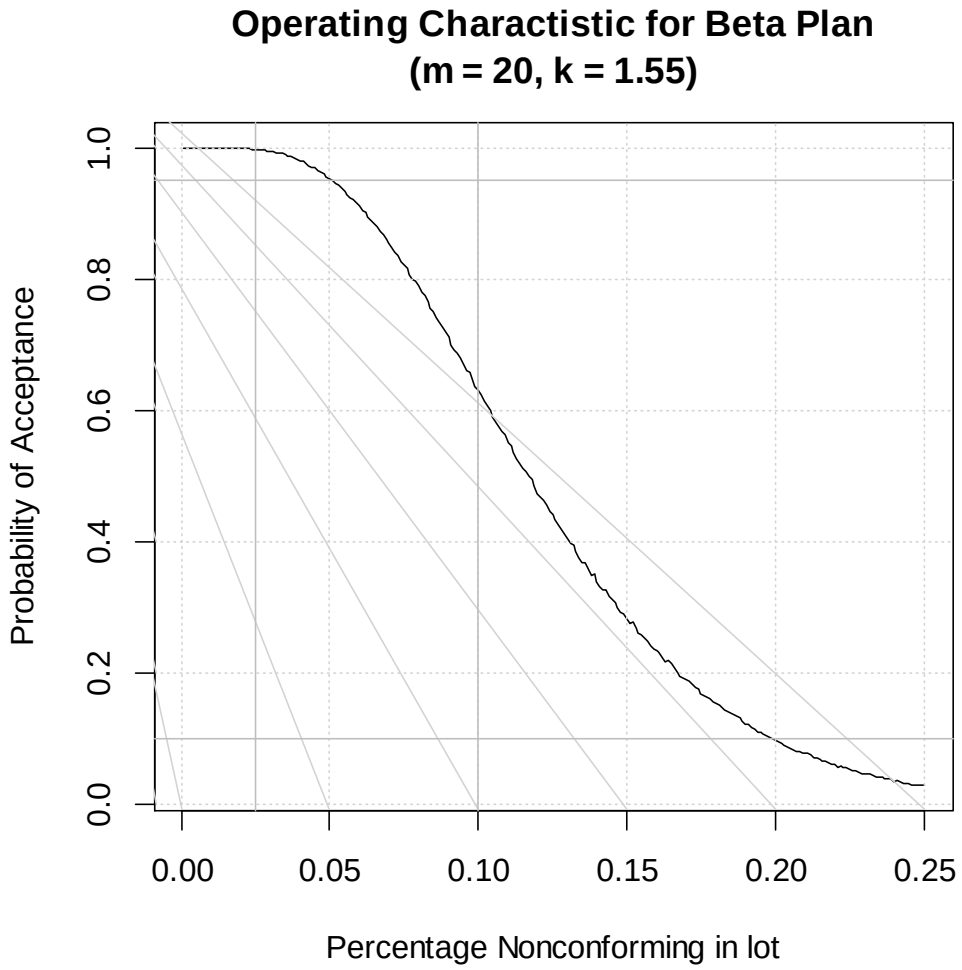
$P - k \times s \geq L$ where P is the test result or average test result and $s = \sqrt{P(1-P)/\theta}$, L is the minimum limit (10ppm) and k is the acceptability constant for the plan.

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.

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

Using the same consumer's and producer's risk as those for protein and moisture above (a 5% chance of rejecting lots containing 5% nonconforming product and a 10% chance of accepting lots containing 20% nonconforming product) the resulting plan is (m=13, k=1.20) i.e. a composite sample would be formed from 13 subsamples randomly taken from the lot and the composite would be tested once to produce the estimate of "P".

A sampling plan can be derived using App10. The Operating Characteristic for this plan is shown below.



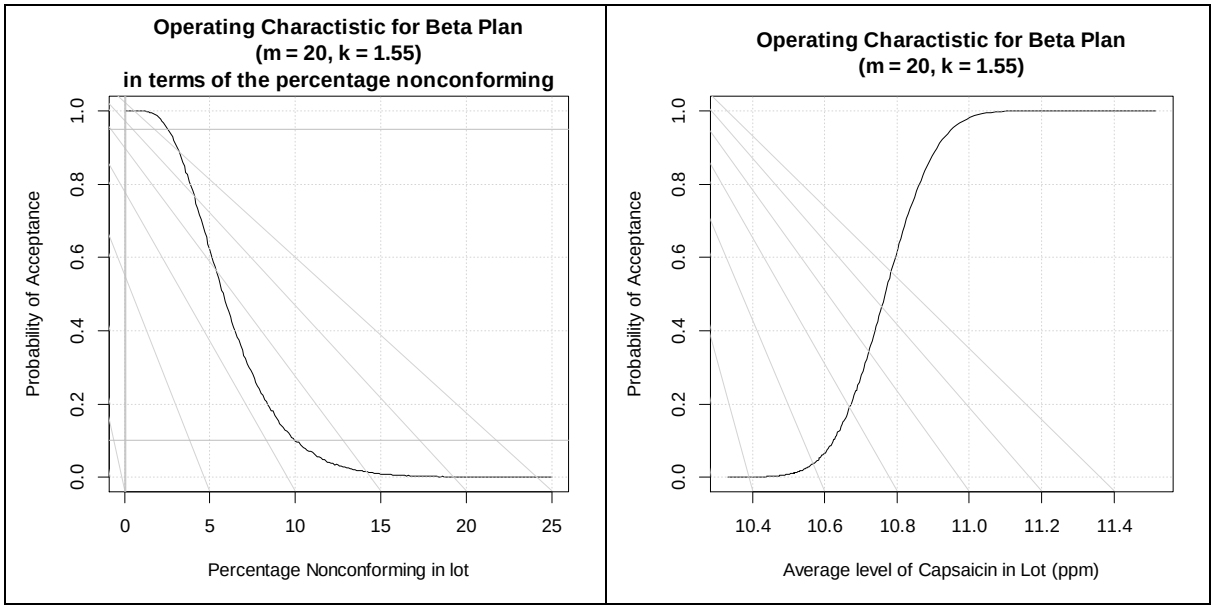
If however, we decided that capsaicin was a more critical parameter for the product then we may wish to reduce the consumer's risk – instead of decreasing the chance of acceptance at the CRQ we can reduce the CRQ itself, to 10%, and also reducing the PRQ to 2.5%.

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 18 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 in a lot and by the average level of capsaicin.



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 proportion 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 (refer to References) 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.

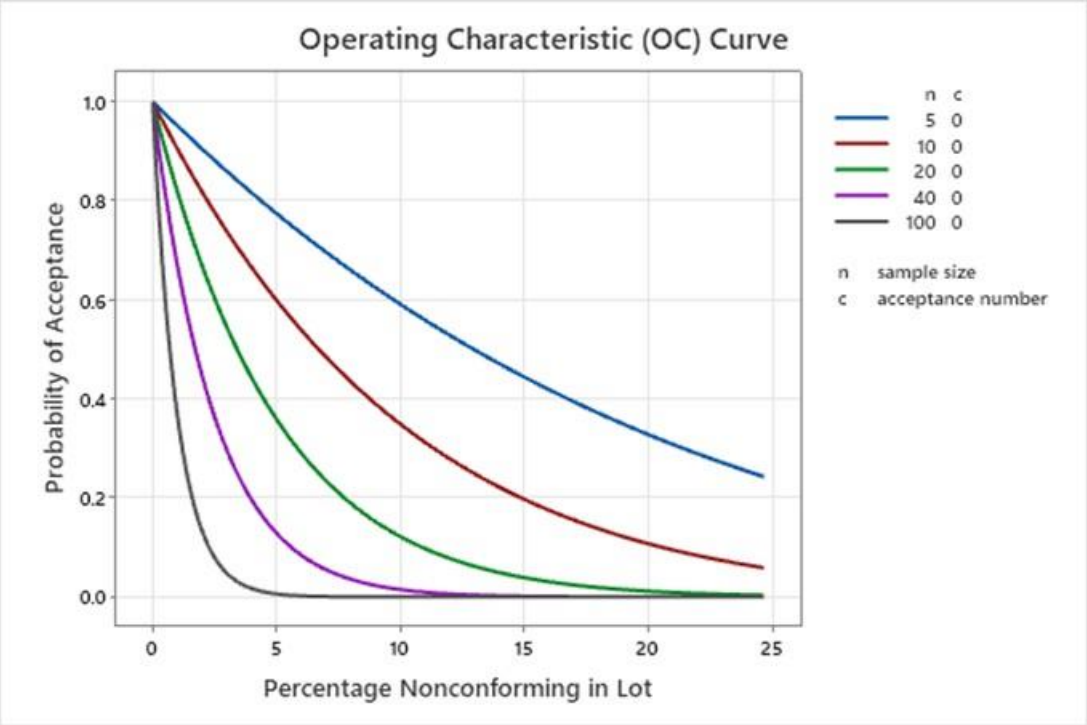
4.1.1 Zero-acceptance number plans

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

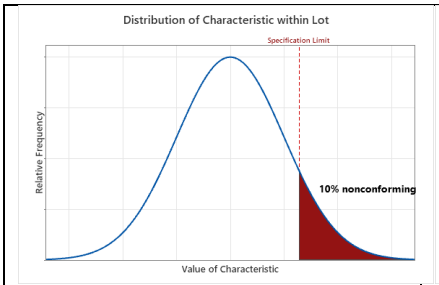
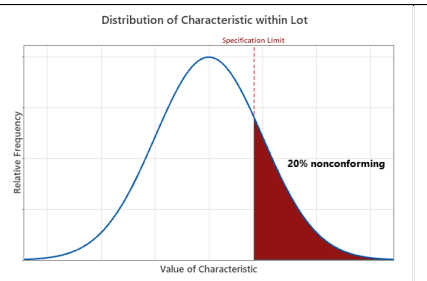
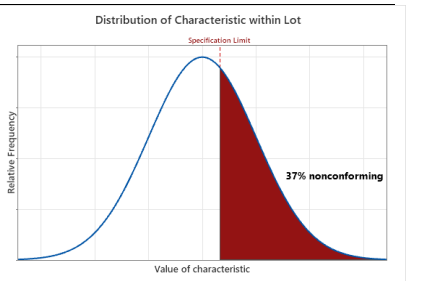
Zero-acceptance number (ZAN) plans are a special case of two-class plans in which the acceptance numbers are set to c = 0. They are used in more critical situations such as for pathogens or foreign matter where only CR is considered directly and acceptance of lots demands that nonconforming items are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

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:



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.

		
10% nonconforming, 59% of lots are accepted	20% nonconforming, 33% of lots are accepted	37% nonconforming, 10% of lots are accepted

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

[Excel file, formulas in annex]

pexact.xlsx

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.

4.1.2 Why the value of c in attribute plans need not be zero

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.

4.1.3 (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.

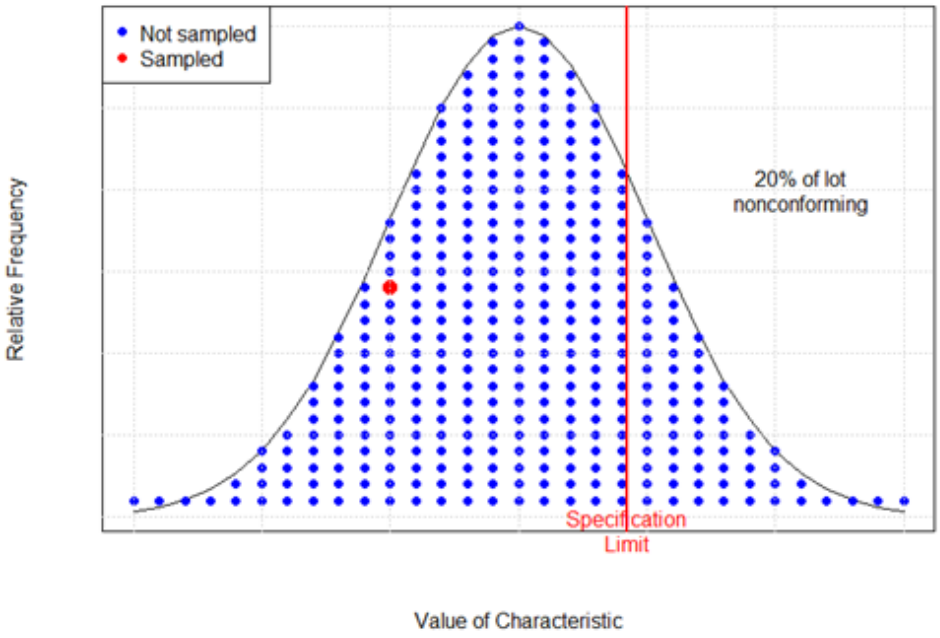
For contaminants these plans rely on an assumption of homogeneity and possibly also on the usually large offsets between the decision limits used in those plans and food safety levels, so that allowance for MU may not be necessary. However, there is a considerable risk incorrectly accepting a noncompliant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

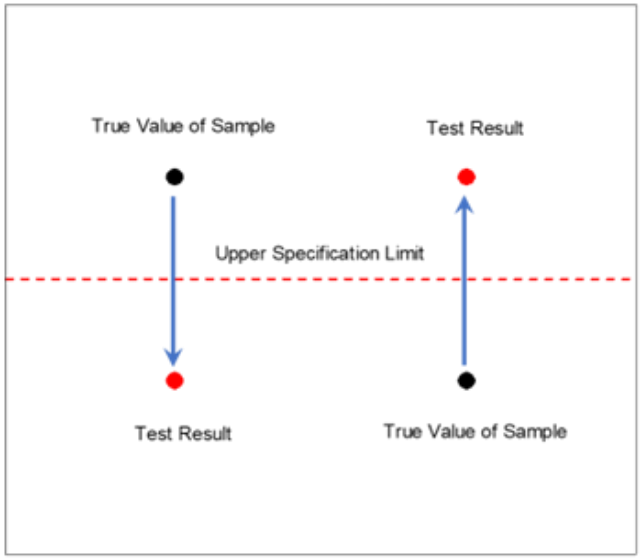
- Attempting to minimize the cost of testing,
- Ignoring of the principles of sampling,

- 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 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.



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



The (n=1, c=0) plans have been extended to include allowance for MU [diagram CXG 54-2004].

These plans, however, cannot simply be extended to lot inspection by including sampling components in the MU while allowing PR and CR to be controlled to specified levels. Given that a single result is an estimate of the mean of a lot, this 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 the case studies section 3.3.2.

Section 5.7 discusses Bayesian sampling plans that allow plans controlling producer's and consumer's risks to be designed while requiring only small sample sizes.

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 Annex A.2.

[These formulas are implemented in App 1]

4.2.1 Within-item variability

High variability between test results obtained from different test portions taken from the same item may also have an impact on the estimate of the lot standard deviation, if the latter is calculated on the basis of test results from a sample of n items.

In theory, variation within items should not be allowed to affect lot assessment (since item compliance is defined in terms of an item's mean value). However, variation within items will impact the test results and thus the assessment of the lot. In other words: within-item variation may inflate observed variation between item-specific test results, thus increasing the probability of lot rejection.

Accordingly, in acceptance sampling, variation within items plays a similar role to measurement uncertainty. Indeed, the question may be raised whether within-item variation should be considered a component of measurement uncertainty. Strictly speaking – insofar as an item corresponds to the laboratory sample – within-item variation should be subsumed under *analytical* measurement uncertainty. However, in connection with acceptance sampling, it is expedient to consider analytical measurement uncertainty and within-item variation as two separate sources of random variation. The reason is that any available estimate of analytical measurement uncertainty may or may not reflect the within-item variation actually observed in the lot under consideration.

In this section, two models, first described in Uhlig (2024), will be discussed. These two models can be used to analyze the relationship between within- and between-item variability, and thus allow an estimate of the lot standard deviation which is corrected for any between-item variability.

In the first model, it is assumed that the production process is item-oriented in the sense that – on account of the production process – the true lot standard deviation (variability between items) can be expected to be near zero, even if there is considerable within-item variability. This is the case if a given volume of the substance under consideration is added separately to each item. In the model, it is assumed that if there are m possible test portions in a given item, the m corresponding test results are correlated: if the concentration is relatively high in one of the test portions, then the concentration in the others must "make up" for this by being low.

In the second model, it is assumed that the production process is batch-oriented in the sense that it both within- and between-item may be observed. This is the case if a given volume of the substance under consideration is added to bulk material (the "batch") from which the items are subsequently extracted. In this model, the concentration within item i is no longer constrained as in model 1. Rather, it is "free" to reflect the sum of its constituent test portions.

The two models are described and then illustrated on the basis of examples. Furthermore, it will be shown how to obtain an estimate of both between-item variability which is corrected for within-item variability.

4.3 The role of measurement uncertainty in acceptance sampling

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 the total measurement uncertainty are affected by any compositing or 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

1. A single sample 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 sample taken from a lot, representing 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

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

Refer to Section 5.4 for more information about the role of measurement uncertainty in acceptance sampling.

The *Guidelines on measurement uncertainty* (CXG 54-2004) and the information document contain more information about the estimation of measurement uncertainty; one of the key references is ISO 5725 Parts 1 & 2.

The lot standard deviation can be estimated using the duplicates method. Alternatively, Hahn's method (refer section 3.2.2) can be used to adjust observed standard deviations for repeatability type uncertainty, avoiding the need to test samples in duplicate. Either approach could be applied to estimates of lot standard deviations obtained from the analysis of data from a series of lots; these standard deviations could be considered as known if the within lot variation was consistent across the lots, thereby enabling a possible reduction in the number of samples required by the sampling plan.

The Eurachem Guide on measurement uncertainty arising from sampling contains information on the estimation process and the use of control charts for monitoring of consistency.

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 CXG 50-2004 section 5.2.9.

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.

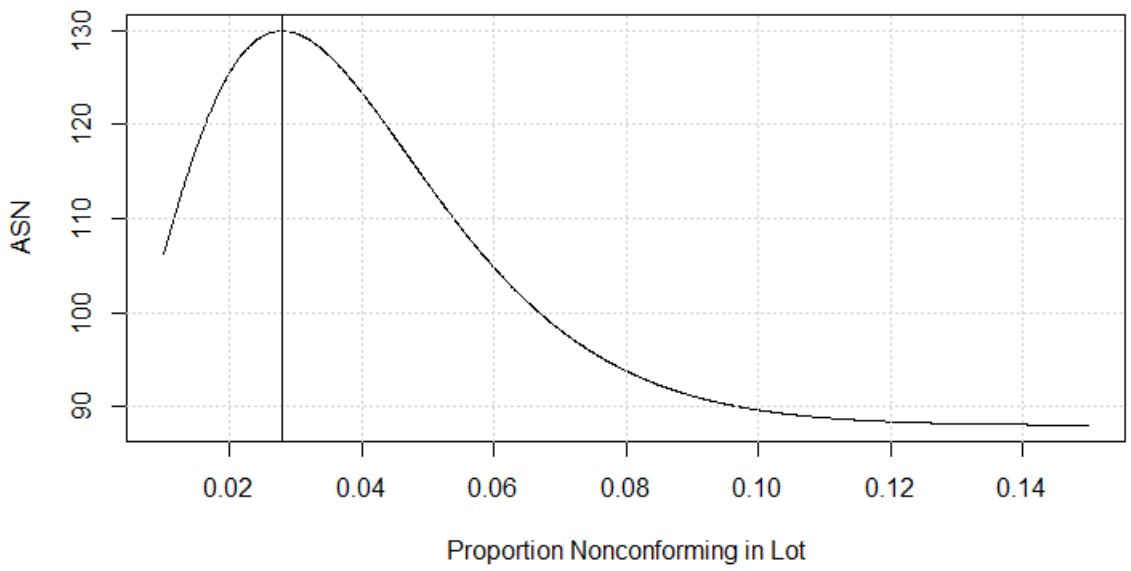
Stage 2:

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 is considerably less at other levels 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 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.

4.6.1 Sampling plans for the average level in bulk materials

It is possible that the chosen plan could be economized if the cost associated with the initial sampling step was low then more increments could be taken to improve the precision of the estimate of the average level.

The acceptance criteria will be of the form: $\bar{x} + k \cdot S \leq USL$ for upper specification limit USL for the average level.

4.6.2 ISO 10725

[This standard follows the work by Schilling and discussed in his book, available on-line]

See references section.

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.

4.6.3 Sampling plans for Aflatoxins (CXS 193-1995)

Introduction

Whitaker et al. analyzed 46 years of laboratory data to obtain estimates of the sample-to-sample, the subsampling and the analytical components of variation for each of the lots in the data where significant contamination was found. Following this analysis Horwitz type equations were derived for each of these three variance components 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 for the evaluation of sampling plans is provided at http://tools.fstools.org/mycotoxins/.

This formula 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 "between lab" between laboratory variation is used, assuming that the between laboratory variation is twice the within laboratory figure (a common assumption). 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	$(7{,}730/ns)5.759C^{1.581}$	$(100/nss)0.170C^{1.846}$	$(1/na)0.0041C^{1.986}$	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	$(7{,}730/ns)5.759C^{1.581}$	$(100/nss)0.170C^{1.846}$	$(1/na)0.0041C^{1.986}$	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

4.6.4 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. 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.

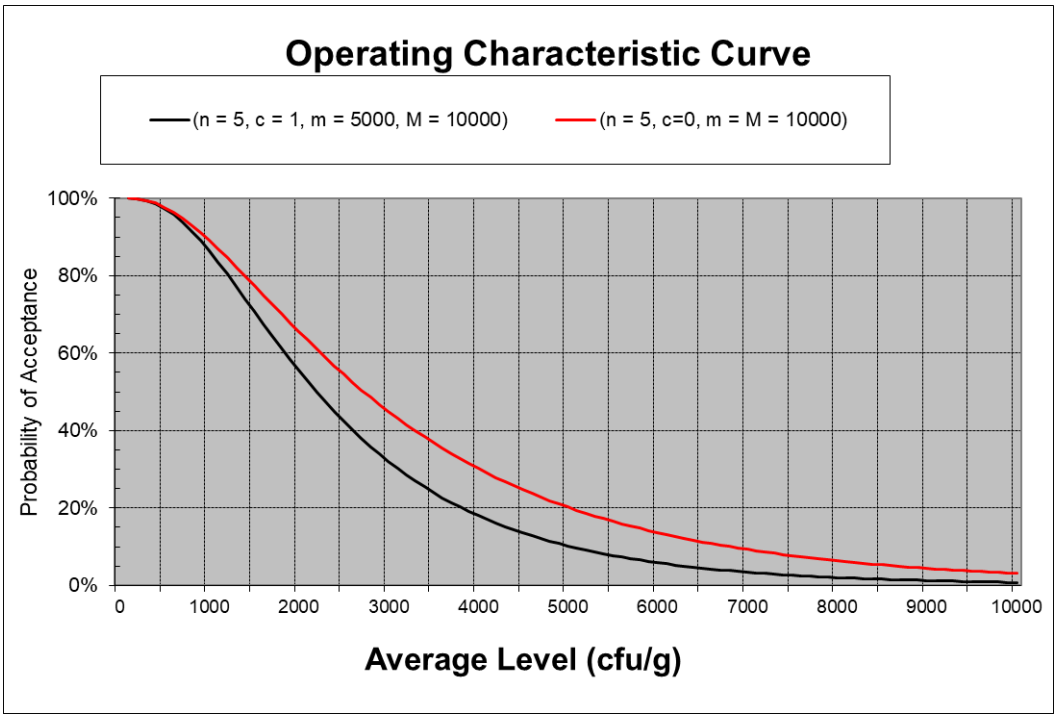
Lots are accepted provided:

- 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 parameters 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.



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, as 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 proportion nonconforming in the lot

The symbol ∑ means 'the sum of' the expression evaluated at values of k from zero to the acceptance number 'c' and $\binom{n}{k}$ is the binomial coefficient, it is the number of ways of choosing k items from a total of n item. 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 "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:

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 4.1.1 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 proportion 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 [English version] 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 [English version] 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:

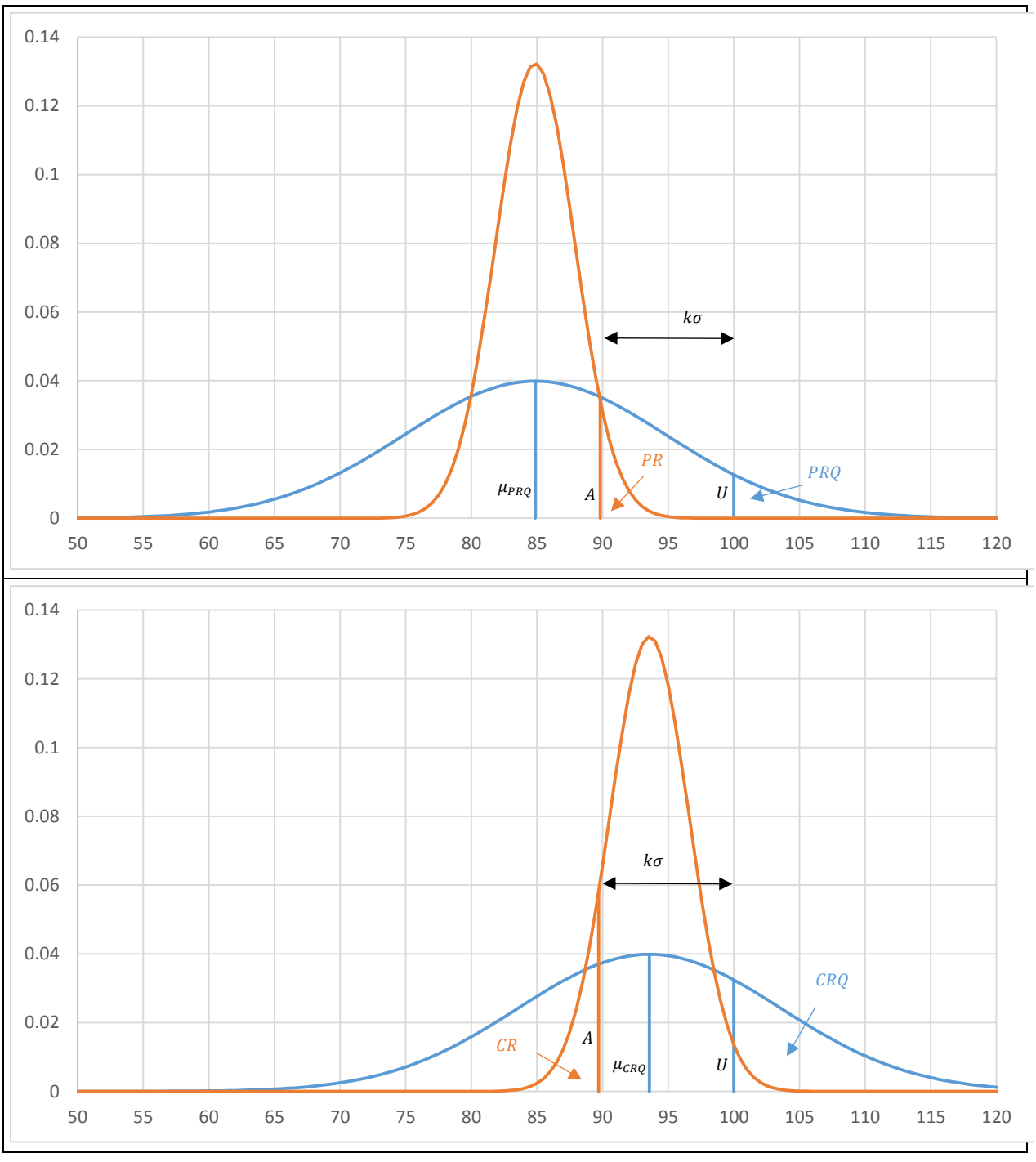
$$z_{1-PRQ} = 1.514$$
$$z_{1-PR} = 1.645$$
$$z_{1-CRQ} = 0.643$$
$$z_{CR} = -1.282$$

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 1: 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

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

5.2 Understanding ISO plans

5.2.1 Variables plans constructed in terms of producer risk

The “philosophy” or “rationale” of ISO variables plans is as follows.

First, ISO variables plans take into consideration either PRQ (plans “indexed by AQL”) or CRQ (plans “indexed by LQ”), but not both at the same time.

Secondly, the ISO plans indexed by AQL are constructed in such a way, that the producer risk decreases as the lot size increases. Taken from the Mathematical and Statistical Principles underlying Military Standard 414, the forerunner of the ISO3951 standard, the following table shows the producer’s risk in terms of the sample size code letter that determines the sample size based on the lot size:

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. (Recall that the code letter reflects lot size.)

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

Table N.2 — Producer's risk (in percent) for normal inspection: σ-method

Code letter	Acceptance quality limit (in percent nonconforming)															
	0,01	0,015	0,025	0,04	0,065	0,10	0,15	0,25	0,40	0,65	1,0	1,5	2,5	4,0	6,5	10,0
B	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	3,57	2,96	6,72
C	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	7,17	3,59	6,06	4,54
D	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	6,33	3,89	6,37	4,81	2,86
E	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	7,17	3,94	6,29	4,62	2,81	2,74
F	↓	↓	↓	↓	↓	↓	↓	↓	↓	7,65	4,32	5,42	4,66	1,89	2,80	0,865
G	↓	↓	↓	↓	↓	↓	↓	↓	7,44	4,96	6,87	4,66	2,04	2,09	1,41	1,15
H	↓	↓	↓	↓	↓	↓	↓	7,47	4,68	7,35	5,48	1,98	1,86	1,26	1,38	0,871
J	↓	↓	↓	↓	↓	↓	6,69	4,70	7,38	6,40	2,56	1,59	1,05	1,25	1,32	1,24
K	↓	↓	↓	↓	↓	7,32	4,16	7,17	5,56	2,74	2,10	0,572	1,08	1,06	1,58	0,602
L	↓	↓	↓	↓	7,64	4,82	6,30	5,80	2,67	2,48	0,788	0,854	1,09	1,39	1,07	↑
M	↓	↓	↓	7,52	5,16	7,26	5,29	2,56	2,19	0,933	1,17	0,682	1,28	0,829	↑	↑
N	↓	↓	7,30	5,02	7,95	5,82	2,04	2,12	0,844	1,36	1,07	0,808	0,774	↑	↑	↑
P	↓	6,70	4,77	7,55	6,30	2,64	1,82	0,832	1,23	1,42	1,46	0,481	↑	↑	↑	↑
Q	7,06	4,16	7,25	5,85	2,84	2,26	0,578	1,02	1,07	1,69	0,776	↑	↑	↑	↑	↑
R	4,89	6,71	5,76	2,73	2,68	0,830	0,738	1,04	1,43	1,20	↑	↑	↑	↑	↑	↑

NOTE The producer's risk is the probability of not accepting a given lot when the process fraction nonconforming is equal to the AQL.

Note: 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.

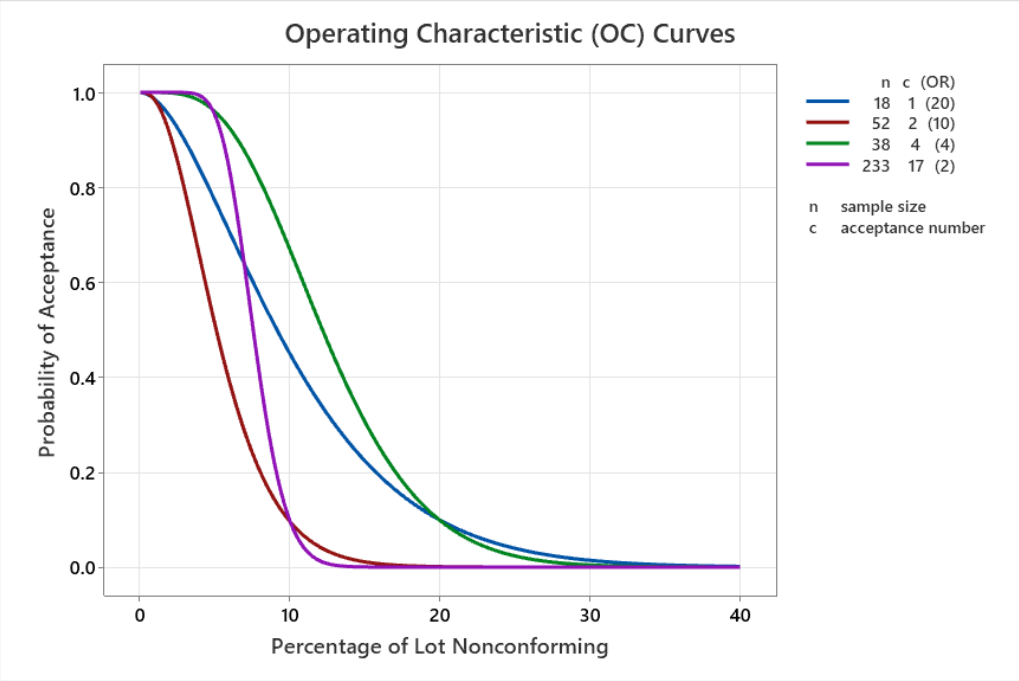
code	1e-04	0.00015	0.00025	4e-04	0.00065	0.001	0.0015	0.0025	0.004	0.0065	0.01	0.015	0.025	0.04	0.065	0.1
B	7.04%	6.68%	7.29%	7.51%	7.62%	7.31%	6.69%	7.45%	7.43%	7.66%	7.18%	6.32%	7.17%	3.56%	2.96%	6.71%
C	7.04%	6.68%	7.29%	7.51%	7.62%	7.31%	6.69%	7.45%	7.43%	7.66%	7.18%	6.32%	7.17%	3.58%	6.06%	4.54%
D	7.04%	6.68%	7.29%	7.51%	7.62%	7.31%	6.69%	7.45%	7.43%	7.66%	7.18%	6.32%	3.89%	6.36%	4.81%	2.86%
E	7.04%	6.68%	7.29%	7.51%	7.62%	7.31%	6.69%	7.45%	7.43%	7.66%	7.18%	3.95%	6.30%	4.63%	2.80%	2.73%
F	7.04%	6.68%	7.29%	7.51%	7.62%	7.31%	6.69%	7.45%	7.43%	7.66%	4.32%	5.42%	4.65%	1.89%	2.80%	0.87%
G	7.04%	6.68%	7.29%	7.51%	7.62%	7.31%	6.69%	7.45%	7.43%	4.95%	6.85%	4.65%	2.05%	2.09%	1.42%	1.14%
H	7.04%	6.68%	7.29%	7.51%	7.62%	7.31%	6.69%	7.45%	4.68%	7.33%	5.47%	1.97%	1.86%	1.26%	1.38%	0.87%
J	7.04%	6.68%	7.29%	7.51%	7.62%	7.31%	6.69%	4.70%	7.38%	6.41%	2.56%	1.58%	1.05%	1.25%	1.32%	1.25%
K	7.04%	6.68%	7.29%	7.51%	7.62%	7.31%	4.17%	7.16%	5.55%	2.74%	2.09%	0.57%	1.07%	1.06%	1.57%	0.61%
L	7.04%	6.68%	7.29%	7.51%	7.62%	4.82%	6.32%	5.80%	2.67%	2.48%	0.79%	0.85%	1.09%	1.40%	1.07%	0.61%
M	7.04%	6.68%	7.29%	7.51%	5.14%	7.26%	5.28%	2.57%	2.19%	0.94%	1.16%	0.68%	1.28%	0.83%	1.07%	0.61%
N	7.04%	6.68%	7.29%	5.04%	7.93%	5.80%	2.04%	2.13%	0.84%	1.37%	1.08%	0.82%	0.77%	0.83%	1.07%	0.61%
P	7.04%	6.68%	4.48%	7.53%	6.30%	2.66%	1.81%	0.83%	1.22%	1.42%	1.48%	0.48%	0.77%	0.83%	1.07%	0.61%
Q	7.04%	4.16%	7.26%	5.83%	2.84%	2.27%	0.58%	1.02%	1.07%	1.67%	0.78%	0.48%	0.77%	0.83%	1.07%	0.61%
R	4.88%	6.74%	5.76%	2.73%	2.68%	0.83%	0.74%	1.04%	1.41%	1.19%	0.78%	0.48%	0.77%	0.83%	1.07%	0.61%

5.2.2 Attributes plans constructed in terms of OR and unity value

The operating ratio (OR) is defined as the ratio $\frac{CRQ}{PRQ}$. It is thus an immediate measure of the quality of a sampling plan: the lower the OR, the steeper the operating characteristic (OC) curve – i.e. the sooner the OC curve dips below the CR as the lot quality worsens – and thus, the better the sampling plan.

The following plot shows the OC curves for a selection of sampling plans with different operating ratios:

PRQ	CRQ	n	c	OR
1%	20%	18	1	20
1%	10%	52	2	10
5%	20%	38	4	4
5%	10%	233	17	2



Several questions arise:

- Can sampling plans be designed for a given OR?
- What is the relationship between OR and sample size?
- What is the OR of ISO sampling plans?

In the case of attributes plans, one possible approach is to calculate sample size by means of unity values. The unity value is defined as sample size multiplied by PRQ. The unity value approach is as follows (this approach is described in Schilling, Grubbs, etc.). For a given OR, refer to a table such as the following (Cameron). Choose the OR value equal or less than the desired OR. The sampling plan can then be obtained directly from the table: the acceptance number c is provided in the table and the sample size is obtained by dividing the unity value by PRQ.

Table: Acceptance number and unity value for a given OR and for PR = 5 % and CR = 10 %. The unity values are calculated on the basis of the binomial distribution.

OR	c	Unity value = $n \cdot PRQ$
44.890	0	0.052
10.946	1	0.355
6.509	2	0.818
4.890	3	1.366
4.057	4	1.970
3.549	5	2.613
3.206	6	3.286
2.957	7	3.981

For instance, for a specified PRQ of 6.5 % and a specified CRQ of 25 % (and for PR = 5 % and CR = 10 %), then the OR is calculated as

$$\frac{CRQ}{PRQ} = \frac{25}{6.5} \cong 3.85$$

Accordingly, we choose the plan corresponding to $OR = 3.549$, yielding $c = 5$ and a unity value of 2.613. The sample size is calculated as $n = \frac{unity\ value}{PRQ} = \frac{2.613}{0.065} \cong 40$.

It should be noted that the OR (or unity value) depends only on the choice of c.

Unity values computed on the basis of the binomial distribution remain near constant across a large range of PRQ.

Table: Unity values for different values of PRQ

	PRQ				
c	0.0001	0.01	0.1	0.2	0.3
0	0.052	0.051	0.001	0.001	0.001
1	0.356	0.351	0.301	0.401	0.301
2	0.818	0.821	0.801	0.801	0.901
3	1.367	1.371	1.401	1.401	1.501
4	1.971	1.981	2.001	2.001	2.1
5	2.614	2.621	2.701	2.801	3.001
6	3.286	3.291	3.401	3.401	3.601
7	3.981	3.991	4.101	4.201	4.2
8	4.696	4.71	4.8	5.001	5.1
9	5.426	5.441	5.601	5.8	0.001

The following table shows OR values for the sampling plans from ISO 2859-1. As can be seen, the OR values are constant across the diagonals. This due to the fact that in this standard both AQL and sample size are geometric series. As the lot size increases and the PRQ decreases, the product remains near constant. Equivalently: the values for c remain constant across the diagonals.

n	0.0100	0.0150	0.0250	0.0400	0.0650	0.1000	0.1500	0.2500
2	68.4	45.6	27.4	17.1	10.5	9.5	6.3	3.8
3	53.6	35.7	21.4	13.4	8.2	8.0	5.4	3.9
5	36.9	24.6	14.8	9.2	9.0	5.8	5.0	3.6
8	25.0	16.7	10.0	10.2	6.3	5.4	4.4	3.4
13	16.2	10.8	10.7	6.7	5.5	4.4	4.0	2.9
20	10.9	12.1	7.2	6.1	4.7	4.1	3.5	2.6
32	11.6	7.7	6.3	4.9	4.2	3.4	2.9	2.3
50	7.6	6.9	5.2	4.4	3.4	2.9	2.5	2.1
80	6.5	5.4	4.5	3.6	2.9	2.4	2.2	1.3
125	5.3	4.9	3.7	3.0	2.4	2.2	1.5	0.9
200	4.6	3.9	3.0	2.5	2.1	1.4	0.9	0.6
315	3.7	3.2	2.5	2.2	1.4	0.9	0.6	0.4
500	3.1	2.7	2.2	1.4	0.9	0.6	0.4	0.2
800	2.5	2.3	1.4	0.9	0.5	0.4	0.2	0.1
1250	2.3	1.5	0.9	0.6	0.3	0.2	0.2	0.1
2000	1.4	0.9	0.6	0.4	0.2	0.1	0.1	0.1

As is to be expected given the discussion in the above sections, unity values also remain near constant across the diagonals.

n	0.0100	0.0150	0.0250	0.0400	0.0650	0.1000	0.1500	0.2500
2	0.02	0.03	0.05	0.08	0.13	0.20	0.30	0.50
3	0.03	0.05	0.08	0.12	0.20	0.30	0.45	0.75
5	0.05	0.08	0.13	0.20	0.33	0.50	0.75	1.25
8	0.08	0.12	0.20	0.32	0.52	0.80	1.20	2.00
13	0.13	0.20	0.33	0.52	0.85	1.30	1.95	3.25
20	0.20	0.30	0.50	0.80	1.30	2.00	3.00	5.00
32	0.32	0.48	0.80	1.28	2.08	3.20	4.80	8.00
50	0.50	0.75	1.25	2.00	3.25	5.00	7.50	12.50
80	0.80	1.20	2.00	3.20	5.20	8.00	12.00	20.00
125	1.25	1.88	3.13	5.00	8.13	12.50	18.75	31.25
200	2.00	3.00	5.00	8.00	13.00	20.00	30.00	50.00
315	3.15	4.73	7.88	12.60	20.48	31.50	47.25	78.75
500	5.00	7.50	12.50	20.00	32.50	50.00	75.00	125.00
800	8.00	12.00	20.00	32.00	52.00	80.00	120.00	200.00
1250	12.50	18.75	31.25	50.00	81.25	125.00	187.50	312.50
2000	20.00	30.00	50.00	80.00	130.00	200.00	300.00	500.00

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 should motivate us to answer the question posed above with a resounding yes.

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 distinction between CA and AS seems 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.

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, see section 5.3.4).

  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 definition of a "decision rule" (for use in conformity assessment) in ISO 17025 is as follows:

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

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, the measurand consists of the relevant summary statistics (e.g. lot mean and lot standard deviation) for the lot; and the MU of the acceptance or rejection decision is the specific producer or consumer risk as defined in JCGM 106. See the discussion in section 5.3.4.

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 5.5.1).

Finally, if acceptance sampling consists in applying a decision rule formulated in terms of a criterion for the lot mean and if only one test result is obtained from a unique composite sample obtained from the lot, then there may not be any meaningful distinction between conformity assessment and acceptance sampling, see the note in section 5.4.2 Hence, in clarifying the relationship between conformity assessment and acceptance sampling, the distinction between plans for the proportion nonconforming and plans for the lot mean should be borne in mind. This may be equivalent to distinguishing plans where the criterion is formulated in terms of a measurand in the strict metrological sense and plans where the criterion relates to the proportion nonconforming.

Note: Many of the issues discussed here are closely related to the discussion on measurement uncertainty in section 5.4.

5.4 The role of measurement uncertainty in acceptance sampling

Refer CXG 54-2004 and the information document.

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 following questions are discussed in this section:

- How is the measurand specified?
- Terminological clarification: sampling uncertainty versus acceptance sampling
- How can it be ensured that the measurement uncertainty estimate is reliable?
- How can a lot statistical parameter be corrected for measurement uncertainty?

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 proportion 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 question is 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 measurand is now considered as the relevant summary statistics, e.g. lot mean, lot standard deviation.
- The measurement uncertainty can then be considered the specific producer and consumer risks.
- This reinterpretation is particularly relevant in connection with Bayesian approaches to acceptance sampling, see Section 5.5.

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.

The lot standard deviation may seem to be related to sampling 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 or conceptual level, there is no connection between the two concepts. Indeed, even if there were *no measurement uncertainty* (the measurand is specified in terms of the items so there is no sampling uncertainty component, the analytical uncertainty is very small and thus negligible analytical, and there no bias), one would nonetheless need to take the lot standard deviation into account in the criterion for lot acceptance or rejection. The lot standard deviation is thus not a component of measurement uncertainty and must thus be carefully distinguished from sampling uncertainty.

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 the lot quality expressed as a 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 entire procedure may be indistinguishable from the conformity assessment of an individual sample where the measurand is specified as the mean concentration in the lot. In such cases, there may thus be no useful distinction between acceptance sampling and conformity assessment. In such cases, the lot standard deviation can be interpreted as a component of sampling uncertainty. Moreover, in such cases, the role of measurement uncertainty in acceptance sampling is completely different: rather than a “nuisance parameter” which must be corrected for, it now plays a central role.

5.4.3 Effects of analytical uncertainty versus effects of sampling uncertainty in AS

5.4.3.1 Lots consisting of discrete items

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:

$\bar{x} + ks \leq U$ (or, using the notation from Annex 4.1.2: $\bar{x} \leq A \coloneqq U - ks$)

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 $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

Effect of analytical uncertainty

Analytical uncertainty will manifest itself in the term e_i – i.e. $e_i \sim \mathcal{N}(0, u_{analytical})$ – 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

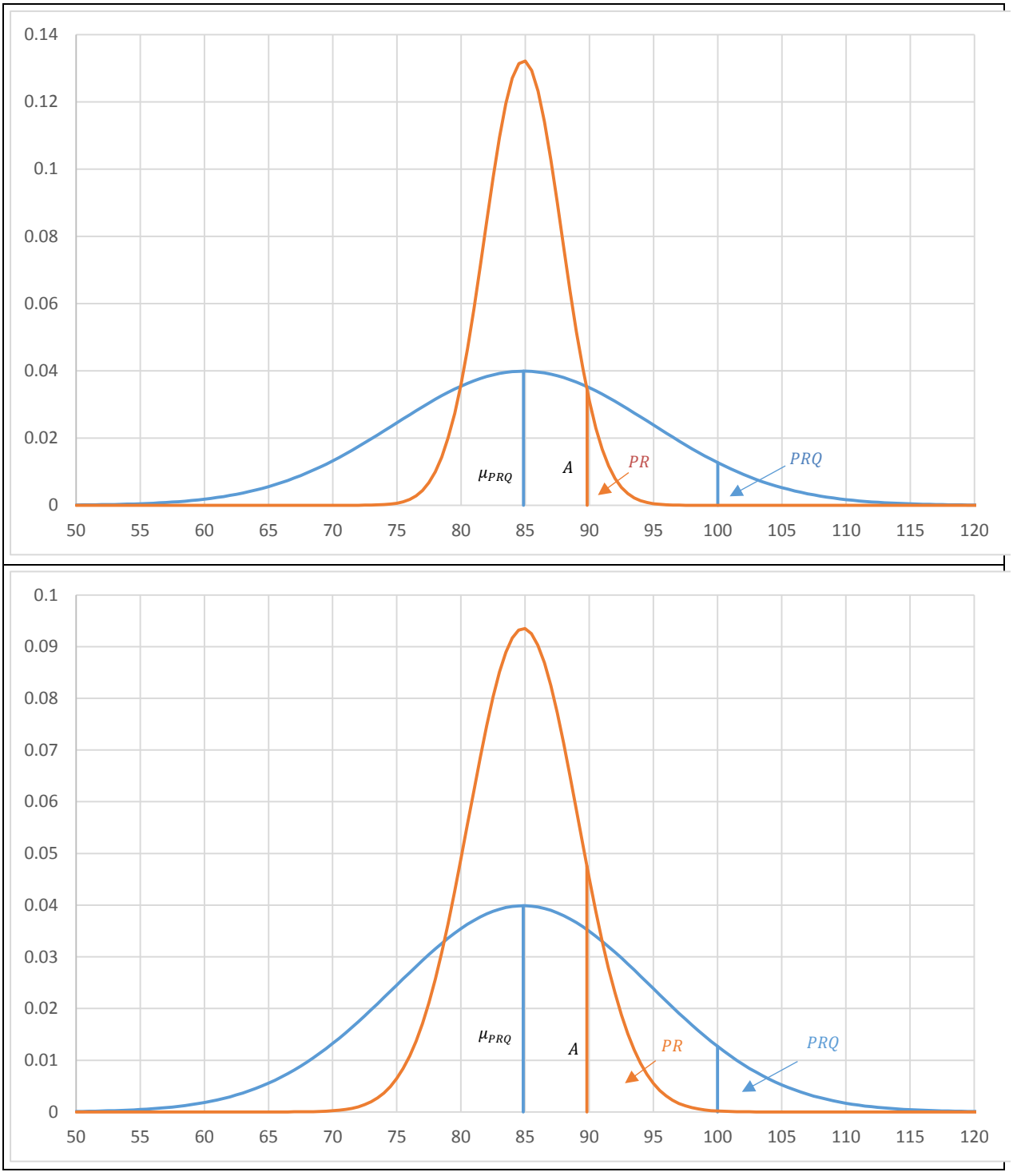
$u_{analytical}$ is the analytical measurement uncertainty

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 lot standard deviation $\sigma = 10$ is known. The same acceptance sampling plan as in Annex **Error! Reference source not found.**2 is applied ($n = 11$, $k = 1.025$ with $A = 90$). However, the analytical uncertainty is nonnegligible, with $u_{analytical} = 10$ ($= \sigma$). 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.

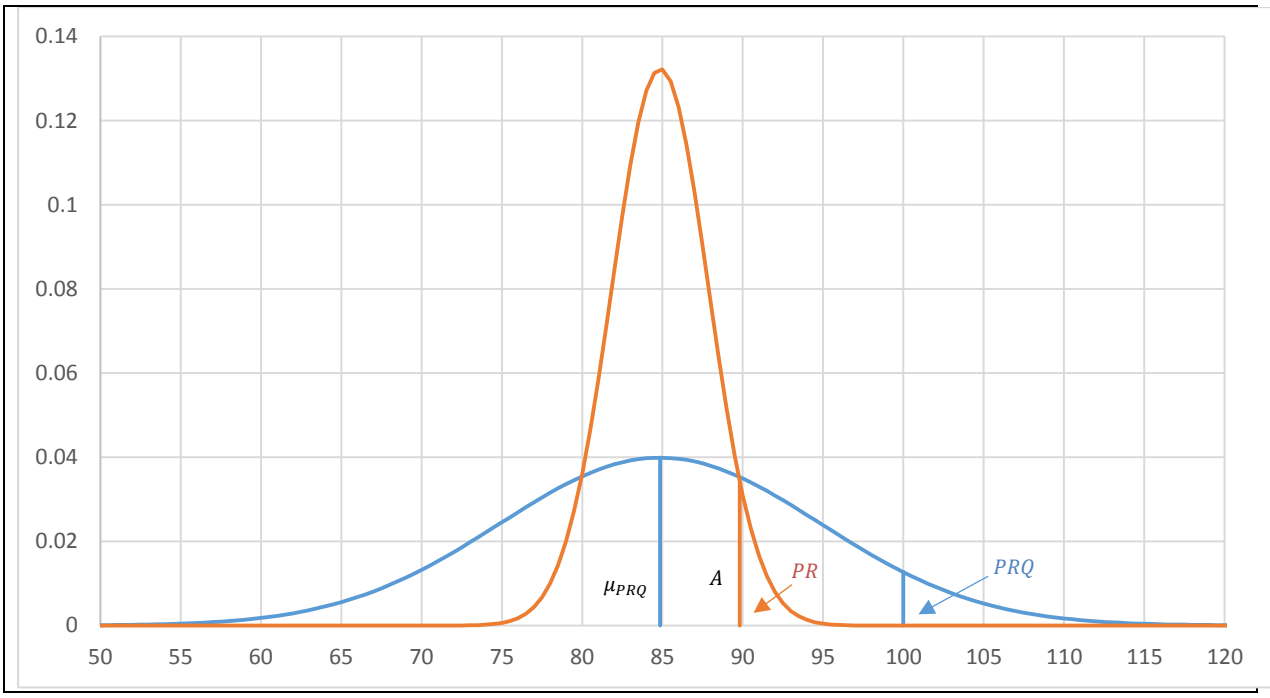
Figure 2: The blue curve represents the distribution of the property of interest in the lot and the red curve represents the statistical distribution of the arithmetic mean. In the top diagram, there is no analytical uncertainty, so the PR is 5%. In the bottom diagram, the analytical uncertainty is nonnegligible, resulting in an increase in PR to over 11%.

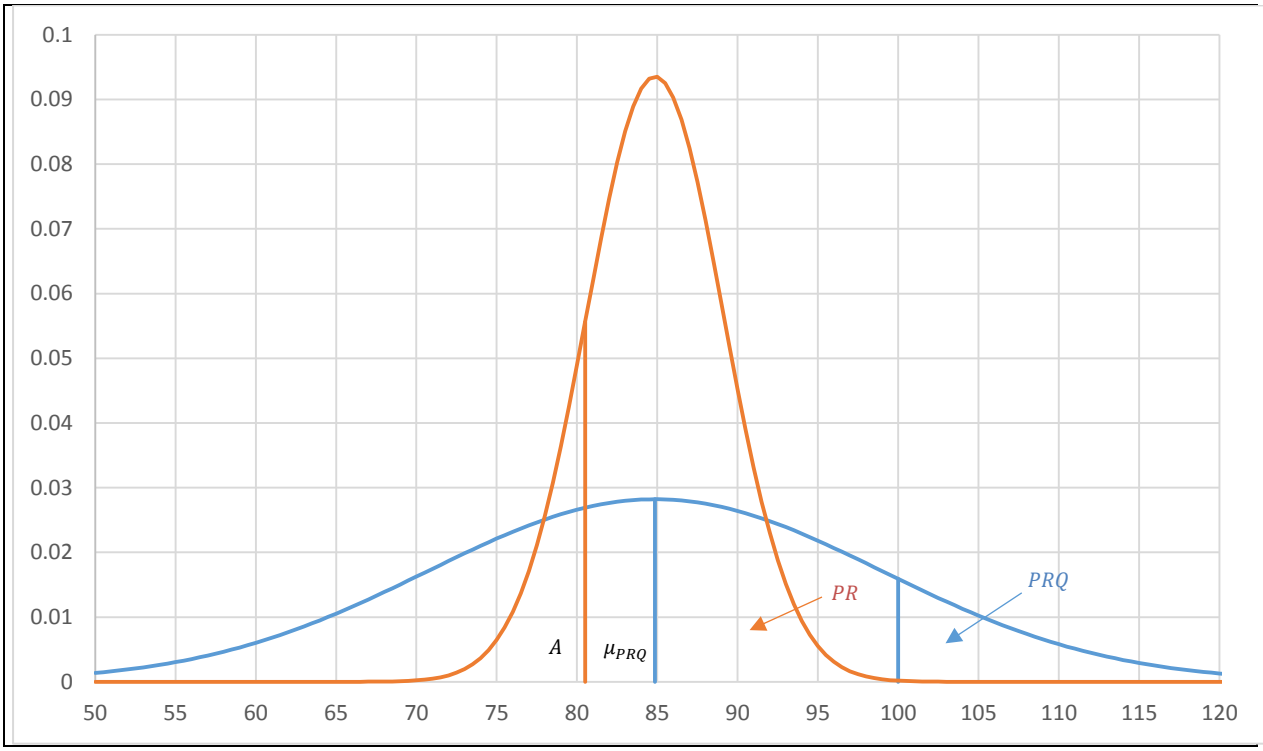


Scenario 2

In this scenario, 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 % proportion nonconforming instead of 6.5 %. Thus, requirements for the producer to achieve quality PRQ = 6.5 % would now be more demanding. 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 *3*: The blue curve represents the distribution of the property of interest in the lot and 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 nonnegligible, resulting in an increase in a much wider blue curve and a distorted value for A, resulting in a PR of nearly 85 %.





[This part on uncertainty associated with the bias is under development]

5.5 Bayesian plans

5.5.1 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. More specifically, it will be shown how to derive an acceptance sampling plan on the basis of the concepts:

- conformance probability
- specific consumer and producer risks
- specific OC curve

Definition of conformance probability

The conformance probability is defined in JCGM 106 as the probability that the measurand Y lies within the tolerance range $\mathcal{C}$ – i.e. conforms to the specification – given a measured value y_m. This probability is calculated on the basis of the *posterior* distribution, i.e. it is calculated on the basis of an updated distribution function for the measurand. The term *posterior distribution* comes from the Bayesian approach to statistics. In the Bayesian framework, one starts from a prior distribution which is then updated on the basis of test results:

- Prior to testing: information regarding the measurand is encapsulated in the prior distribution function.
- Testing: one or several test results are obtained
- Updating the prior: the posterior distribution function is obtained by taking into account both the prior distribution and the test results.
- Conformance probability: calculated on the basis of the posterior distribution.

In JCGM 106, the measurand is thought of as a random variable Y with realizations y, and the posterior distribution function for Y given y_m is denoted $g(y|y_m)$. Accordingly, the conformance probability p_c can be obtained as follows:

$$p_c = \mathcal{P}(Y \in \mathcal{C}|y_m) = \int_{\mathcal{C}} g(y|y_m)\,\mathrm{d}y$$

Definition of specific consumer and producer risks

In addition to the tolerance or conformity range $\mathcal{C}$, which is used for the calculation of the conformance probability p_c, the concept of acceptance range $\mathcal{A}$ from JCGM 106 will be required for the definitions of specific consumer and producer risks. Whereas the notion of conformity concerns the measurand Y, the question of acceptance is related to the test result y_m. Conformity thus concerns the true state of affairs or "ground truth" – which is represented by the measurand – whereas acceptance is decided on the basis of a test result taken as a (hopefully reliable) proxy of the measurand.

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

- Consumer risk = probability that a lot of nonconforming quality will be accepted
- Producer risk = probability that a lot of conformity quality will not be accepted

In the context of acceptance sampling, the term conforming quality – corresponding, in JCGM 106, to the measurand Y lying within the conformity range $\mathcal{C}$ – can be understood as corresponding to quality = PRQ.

Similarly, in the context of acceptance sampling, the term nonconforming quality – corresponding, in JCGM 106, to the measurand Y lying outside the conformity range $\mathcal{C}$ – can be understood as corresponding to quality = CRQ.

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 nonconforming quality
- Specific consumer risk = probability that a lot *which has been accepted* actually has conforming quality

Terminological note: the ISO risks are also referred to as "parametric" risks.

The difference between the specific and parametric risks is similar to the difference between precision and measurement uncertainty. In the case of the parametric risks (precision), the focus is the characterization of the performance of a sampling plan (method). In the case of the specific risks (measurement uncertainty), the focus is the characterization of a given lot (test item) on the basis of a test result.

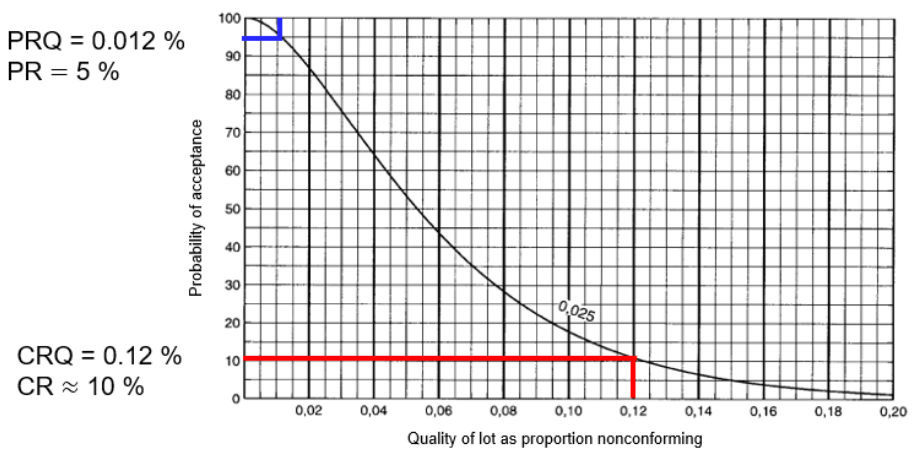
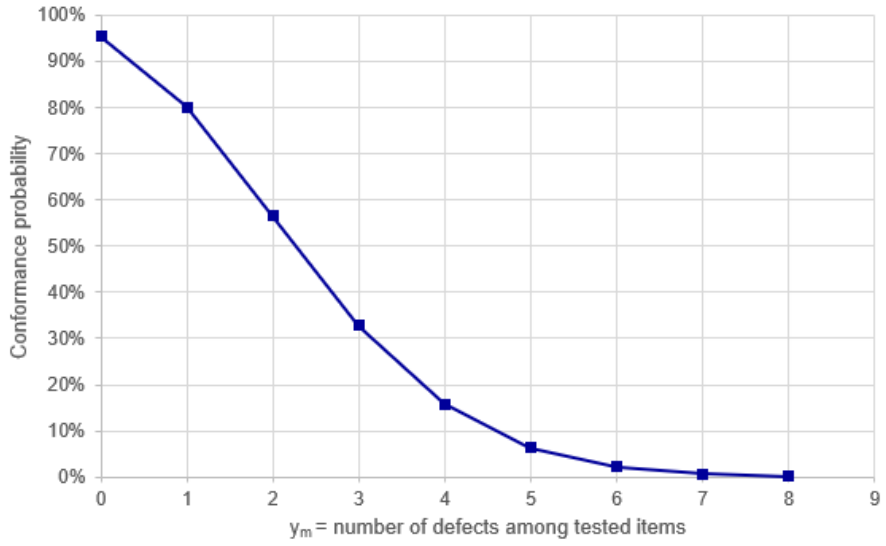
The different types of risks (parametric and specific) complement one another. A framework for acceptance sampling should thus be proposed which includes both types of risk.

Specific OC curves

In the ISO standards, the OC curve plots the parametric probability of acceptance against the quality level, expressed as proportion nonconforming.

In analogy to the distinction between parametric and specific risks it is proposed here to replace these "parametric" OC curves with "specific" OC curves, where the conformance probability is plotted against y_m.

Note: If the producer and consumer start from different priors, their specific OC curves will be different.

Parametric OC curve	
Specific OC curve	

Procedure for the specification of acceptance sampling plans

In the ISO standards, the producer risk is often specified as 5 % and the consumer risk as 10 %. Here we will simply choose a level of risk α (e.g. α = 5 %) for both types of risk.

In the context of acceptance sampling, depending on whether inspection by variables or by attributes is performed, the individual test result y_m is re-interpreted.

- in the case of a variables plan, as a vector of test results: $y_m = y_{m,1}, \ldots, y_{m,n}$, where n denotes the sample size
- or, in the case of a lot consisting of bulk material, as a single test result
- or, in the case of an attributes plan, as the number of defects, i.e. the sum across the test results where each $y_{m,i}$ is either 0 (conforming) or 1 (nonconforming)

Definition of the acceptance range $\mathcal{A}$

Once level of risk α has been specified, the acceptance range can be defined as follows.

$\mathcal{A}$ = test results y_m such that the conformance probability is greater than $1 - \alpha$.

In the case α = 5 %, this means that if the test result y_m is such that the conformance probability is less than 95 %, then the item is not accepted. In other words, the specific consumer risk is no greater than 5 %.

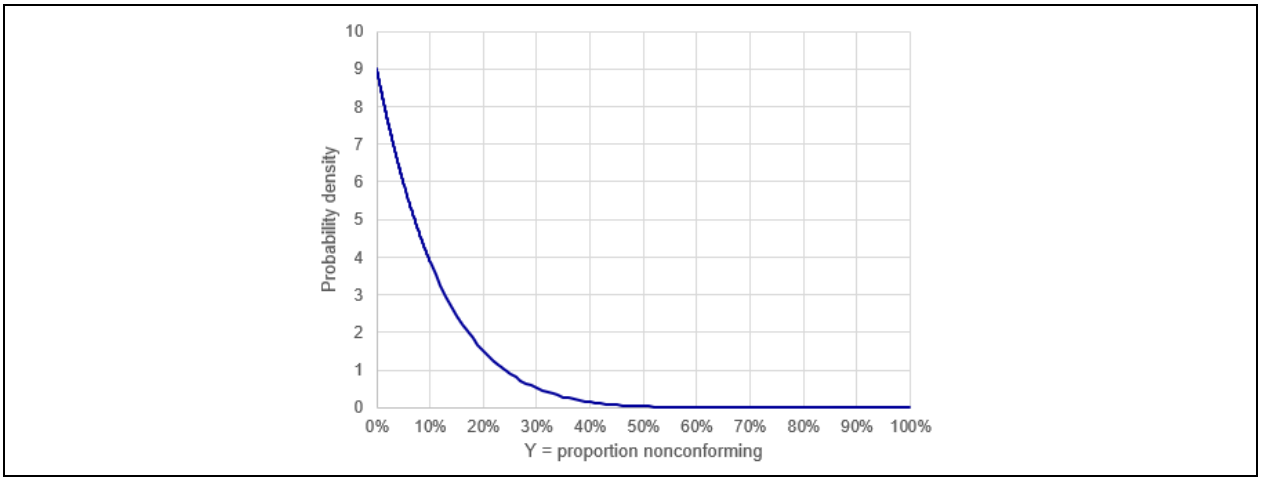
Example

In order to illustrate the above, consider the following example.

An attributes plan for a given lot must be proposed. The conformity range is

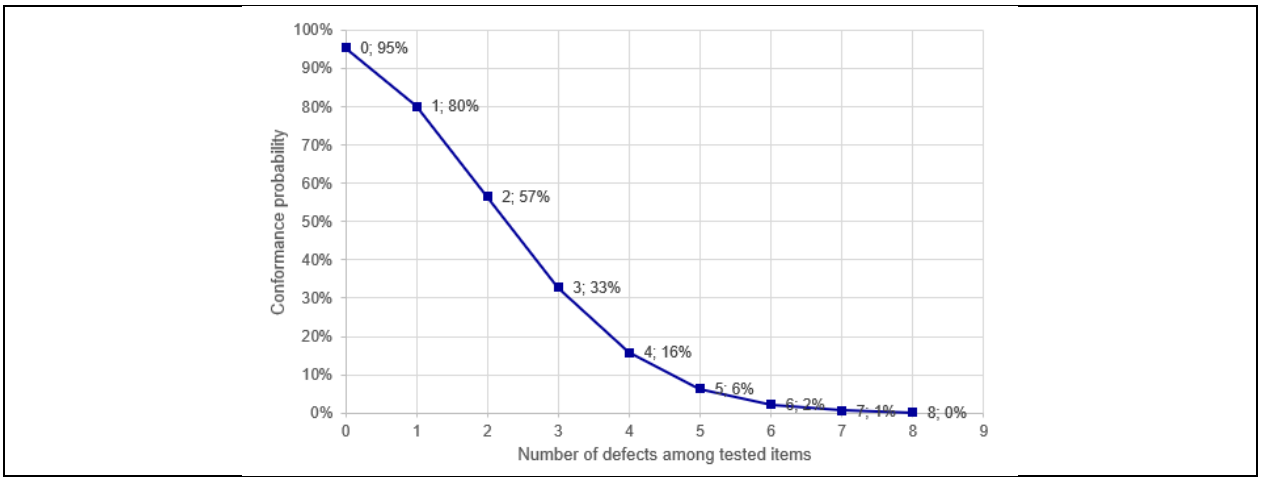
$\mathcal{C}$: Y = proportion nonconforming < 10 %

The consumer applies a prior for the proportion nonconforming which reflects their interest to exercise a conservative approach regarding the producer's quality claims.



With this prior, the conformance probability is 38.7 %.

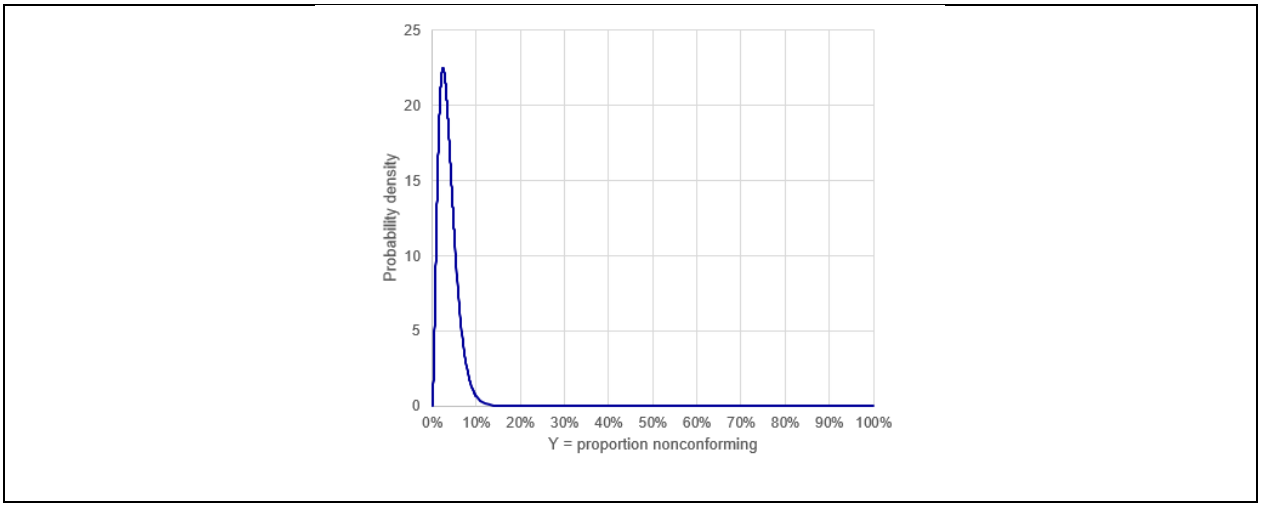
For an acceptance sampling plan with sample size n = 20, the specific OC is as follows:



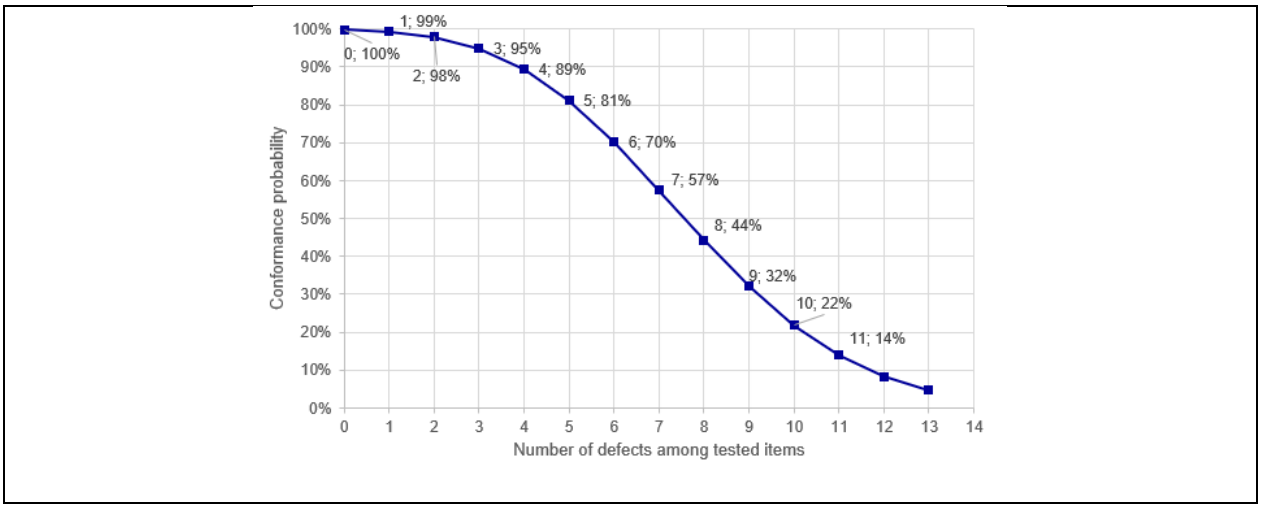
As can be seen, with acceptance number c=1, the conformance probability is already less than 95 %. Thus, the consumer defines

$\mathcal{A}$: c=0

An important question is how the producer views the proposed plan. The producer typically has a much better understanding of the process than the consumer. We consider the case that the producer's prior is as follows:



The specific OC curve based on this prior is as follows:



As can be seen, with the c=0 plan specified by the consumer, the specific producer risk

- for 1 defect is 99 %
- for 2 defects is 98 %
- etc.

In general, the producer will be interested in the *global* risk of rejection, i.e. the probability of conformance across all possible outcomes resulting in lot rejection. In this example, this global producer's risk is 48 %.

Of course, these plans depend on the sample size. The following table provides an overview for various values of n:

Plan	**Acceptance number**	**Specific consumer risk** Calculated on the basis of		**Global producer risk** Calculated on the basis of	
n	**c**	Consumer's prior	Producer's prior	Consumer's prior	Producer's prior
20	**0**	5%	0.7%	69%	48%
37	**1**	5%	0.6%	64%	37%
52	**2**	5%	0.6%	61%	29%

206	**14**	5%	0.5%	52%	9%

Comparison with the "starting point" attributes plans provided in section 2.2,2

PRQ = 6.5 %

p_max = 32.5 %

Scenario 1

The consumer prior is a Beta distribution with parameters (4,36)

Scenario 2

The consumer prior is a Beta distribution with parameters (4,50)

					Scenario 1	Scenario 2
c	n	PR	CRQ	Global PR	Global CR	Global CR
0	2	12.6%	68.4%	12.8%	0.011%	0.000116%
0	3	18.3%	53.6%	18.2%	0.007%	0.000077%
1	5	3.7%	58.4%	3.8%	0.012%	0.000119%
1	8	9.1%	40.6%	9.3%	0.005%	0.000049%
2	13	4.8%	36.0%	4.7%	0.003%	0.000034%
3	20	3.7%	30.4%	3.7%	0.001%	0.000015%
5	32	1.6%	27.1%	1.6%	0.000%	0.000005%
7	50	1.5%	22.4%	1.5%	0.000%	0.000000%

PRQ = 1.5 %

p_max = 7.5 %

Scenario 1

The consumer prior is a Beta distribution with parameters (4,36)

Scenario 2

The consumer prior is a Beta distribution with parameters (4,50)

					Scenario 1	Scenario 2
c	n	PR	CRQ	Global PR	Global CR	Global CR
0	8	11.4%	25.0%	11.4%	24.4%	18.1%
0	13	17.8%	16.2%	17.9%	13.5%	10.6%
1	20	3.6%	18.1%	3.6%	20.5%	16.1%
1	32	8.3%	11.6%	8.2%	7.8%	6.8%
2	50	3.9%	10.3%	3.9%	5.7%	5.2%

5.5.2 Adversarial Bayesian acceptance sampling

5.5.2.1 General description of the approach

In this approach, the consumer and producer each have a prior distribution for the property of interest in the lot under inspection. In addition, for each, a utility function reflects the various interests and risks at stake in the consumer's accepting or rejecting the lot. The procedure for acceptance sampling is articulated in two phases. In the first phase, the consumer either accepts or rejects the lot based on their own opinion of the producer's lots' quality, encapsulated in the consumer's prior. More specifically, the consumer accepts the lot if

$$\int U_c(A|\vartheta) > \int U_c(R|\vartheta)$$

where U_c denotes the consumer's utility function which takes two arguments, one which is acceptance (A) or rejection (R). The other argument ϑ denotes the mean value of the property of interest in the lot. This mean value is unknown, so we integrate over all possible values, taking the consumer's prior as the density.

If the consumer accepts the lot, the acceptance sampling procedure is over. If the consumer rejects the lot, we go to phase 2, which consists in the producer paying for sampling and testing. The question for the producer is to determine an appropriate sampling size. This calculation is performed as follows: for a given sample size, and a given set of test results, the consumer can update the prior to obtain a posterior distribution. The acceptance rule is then the same as above, but with the posterior distribution as density. The sample size is then chosen so as to maximize the integral

$$\int_{\boldsymbol{y_n}} U_p(X|\vartheta)$$

where U_p denotes the producer's utility function, X denotes either acceptance or rejection (based on the consumer's posterior assessment of lot quality), the density function is the producer's posterior distribution and integration is performed across all possible testing outcomes $\boldsymbol{y_n}$.

5.5.2.2 Public sector agents

In order to illustrate this approach, the following example is considered.

The consumer is a governmental food safety agency which specifies acceptance rules for imports. Accordingly, its utility function reflects

- tax revenue accruing from lot sales – in particular, revenue from value added tax (VAT)
- costs associated with health hazards such as healthcare costs, costs associated with any recall measures, etc.

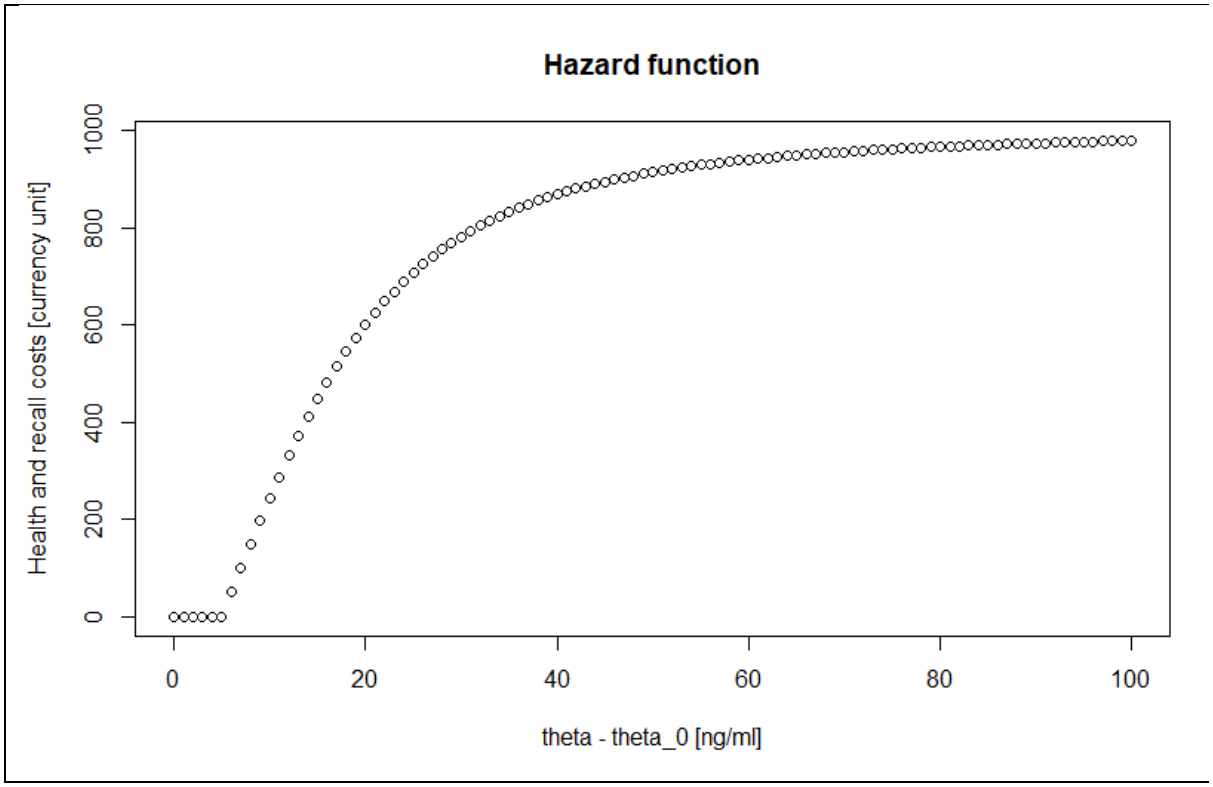
Accordingly, U_c can take the following form

$$U_c = X \cdot p_{VAT} \cdot (S_c - S_p) - hazard\ costs$$

where S_c denotes the income generated by selling the entire lot on the consumer's market after lot acceptance and S_p denotes the producer's price for selling the lot to the consumer.

It should be noted that the consumer's utility depends on whether the consumer accepts or rejects the lot ($X = 1$ or $X = 0$, respectively).

The hazard costs themselves can be represented via an appropriate function of the difference between the lot mean ϑ and the upper specification limit ϑ_0 (*legal threshold*), for example:



The producer's utility can take the following form:

$$U_p = X \cdot S_p - sampling\ and\ testing\ costs$$

where S_p denotes the producer's price for selling the lot to the consumer.

It should be noted that the producer's utility depends on whether the consumer accepts or rejects the lot ($X = 1$ or $X = 0$) and on the sample size for testing.

The two utility functions are summarized as follows:

Consumer utility function

Profits from sales

$$U_c = X \cdot p_{VAT} \cdot (S_c - S_p) - hazard\ costs$$

If the lot is accepted by the consumer ($X = 1$), tax revenue is generated

Healthcare and/or recall measures

Producer utility function

$$U_p = X \cdot S_p - sampling\ and\ testing\ costs$$

If the lot is accepted by the consumer ($X = 1$), sales revenue is generated

Sampling and testing costs depend on the sample size n

Numerical example

Consider the following scenario.

- The legal threshold for the property of interest is $\vartheta_0 = 100$ ng/ml.
- The producer sells the lot at $S_p = 200$ currency units.
- Selling all items in the lot on the consumer's market will generate revenue of $S_c = 500$ currency units. The consumer's profits will thus be $S_c - S_p = 300$ currency units.
- VAT in the consumer's country is $p_{VAT} = 20$ %.

The consumer's prior for the mean lot value is

$$\mu_{0,c} = 110 \text{ ng/ml}$$

$$\sigma_{0,c} = 1 \text{ ng/ml}$$

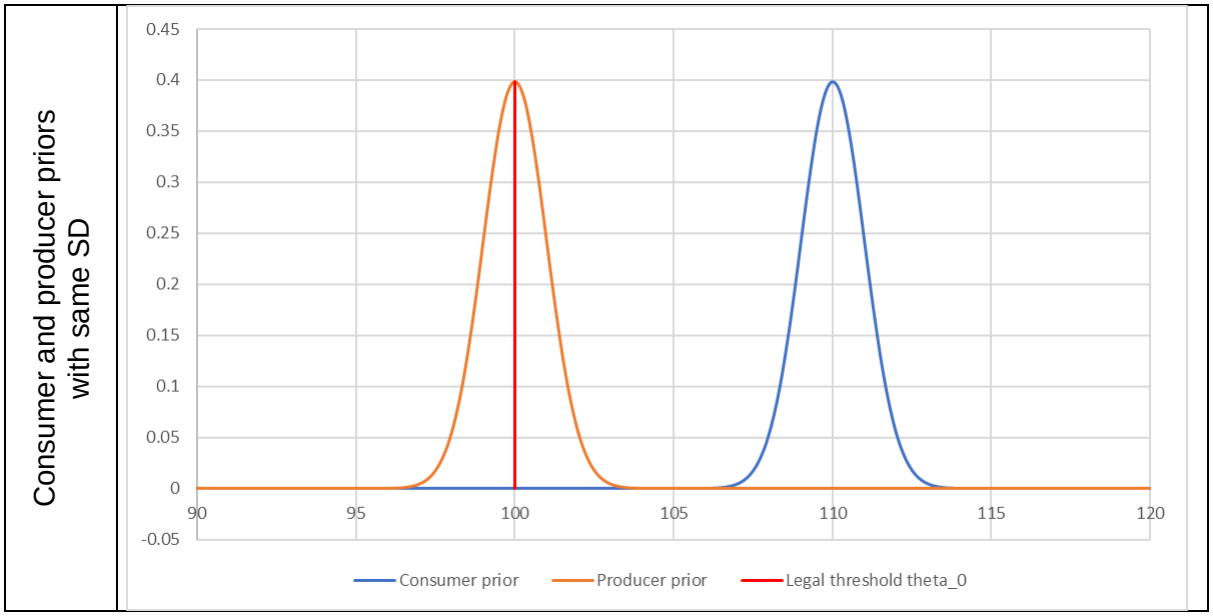
The producer's prior for the mean lot value is

$$\mu_{0,p} = 100 \text{ ng/ml}$$

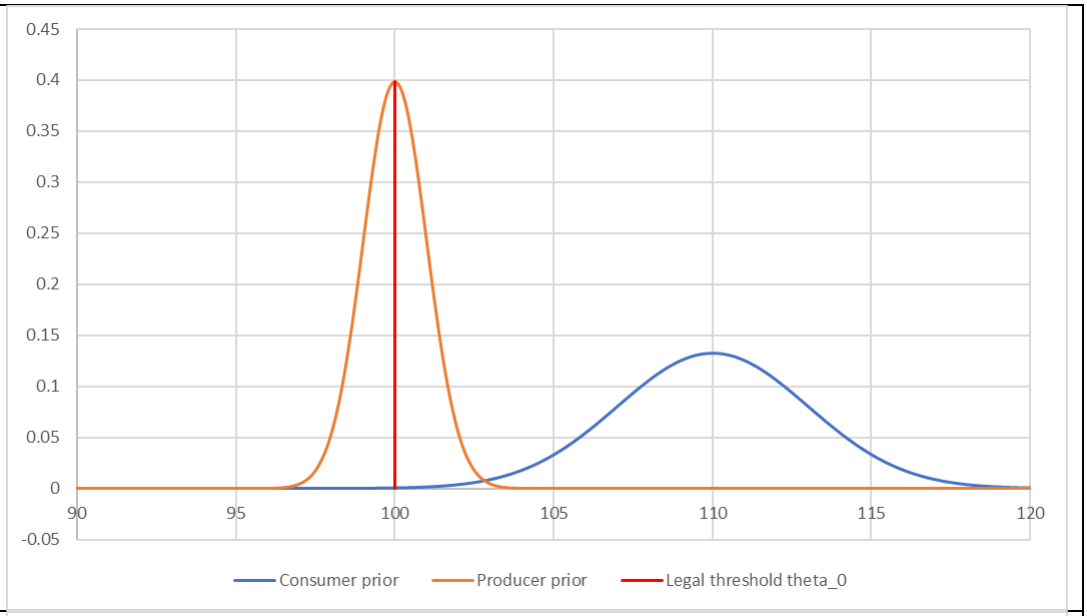
$$\sigma_{0,p} = 1 \text{ ng/ml}$$

Note that in this example, the two priors have the same standard deviation. It can be expected, however, that the producer's SD – given his greater familiarity with the lot – will be less than the consumer SD. Moreover, the producer's mean may typically lie lower than the threshold, rather than coincide with it.

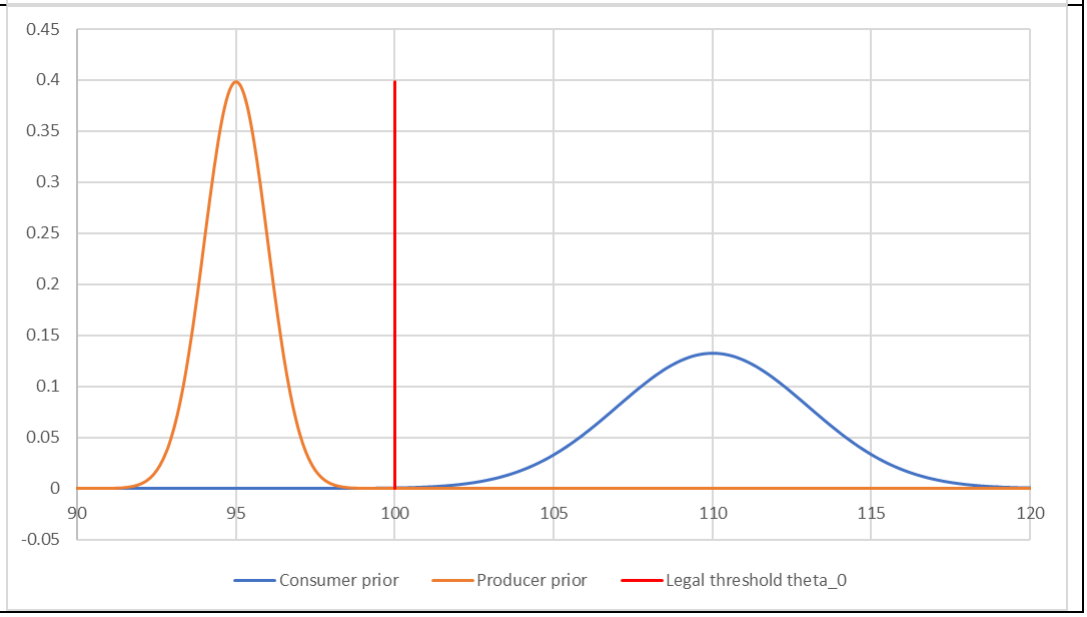
These various cases are illustrated in the following diagram.



Producer prior's SD is less than the consumer's



Producer prior's mean is less than the threshold

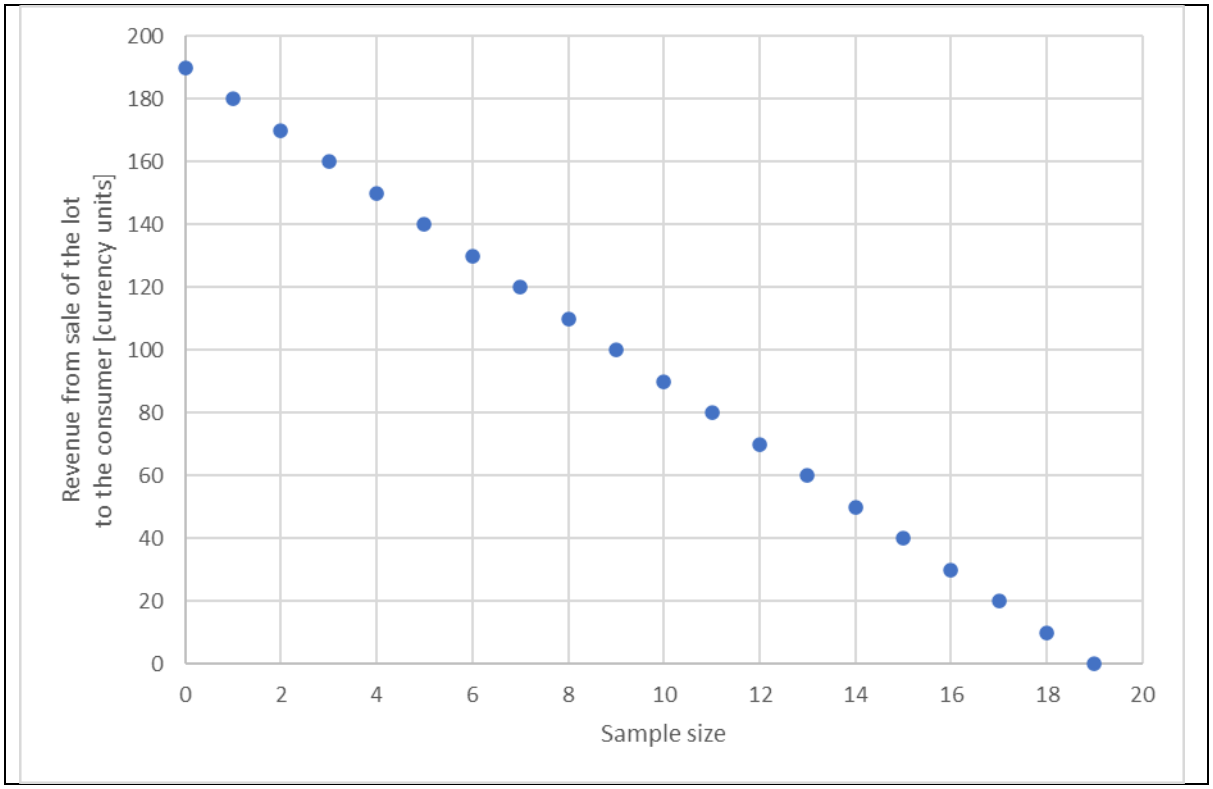


The lot standard deviation is considered by both consumer and producer to be

$$\sigma_{lot} = 3 \text{ ng/ml}$$

Note: this is the simplest case. In reality, it may be necessary to introduce priors for σ_{lot}. An appropriate distributional assumption is normal-inverse-gamma.

Sampling and testing costs are expensive, namely: 10 + (sample size x 10) currency units. In other words, for a sample size of 19, the producer no longer generates any revenue by selling the lot to the consumer. This may be the case if sample + testing costs are very high in comparison to the retail price of an item.



In this scenario, the consumer's decision in phase 1 is to reject the lot.
In phase 2, the producer calculates a sample size of 7 to be optimal.

5.5.2.3 Private sector agents

A second example will now be discussed. In this example, both parties to the transaction are private-sector companies.

The consumer's utility function U_c now takes the following form

$$U_c = X \cdot \left[(S_c - S_p) \cdot \left(1 - p_{fail} \cdot p_{inspection}\right) - penalty \cdot p_{fail} \cdot p_{inspection}\right] - overhead$$

while the producer's utility function U_p remains the same as above

$$U_p = X \cdot S_p - sampling\ and\ testing\ costs$$

Numerical example

Consider the following scenario.

- The legal threshold for the property of interest is $\vartheta_0 = 100$ ng/ml.
- The consumer's overhead is 50 currency units.
- The producer sells the lot at $S_p = 200$ currency units.
- Selling all items in the lot on the consumer's market will generate revenue of $S_c = 500$ currency units.
- The probability of inspection is $p_{inspection} = 0.2$ (i.e. 1 lot out of 5 is inspected, on average)
- The probability of failing inspection p_{fail} is a function of the difference $\vartheta - \vartheta_0$
- The penalty for failed inspection is 2500 currency units.

The consumer's prior for the mean lot value is

$$\mu_{0,c} = 110 \text{ ng/ml}$$
$$\sigma_{0,c} = 1 \text{ ng/ml}$$

The producer's prior for the mean lot value is

$$\mu_{0,p} = 90 \text{ ng/ml}$$
$$\sigma_{0,p} = 1 \text{ ng/ml}$$

The lot standard deviation is considered by both consumer and producer to be

$$\sigma_{lot} = 5 \text{ ng/ml}$$

As above, sampling and testing costs are expensive, namely: 10 + (sample size x 10) currency units. In other words, for a sample size of 19, the producer no longer generates any revenue by selling the lot.

The probability of failing inspection is defined as $0.2 \cdot (\vartheta - \vartheta_0)$, capped at 1.

This means that $p_{fail} = 100\%$ for $\vartheta - \vartheta_0 \geq 5$.

In this scenario, the consumer's decision in phase 1 is to reject the lot.
In phase 2, the producer calculates a sample size of 4 to be optimal.

5.5.3 Special "hybrid" scenarios with very small sample sizes

In this section, special cases are considered. These special cases can be described as follows:

Aspect 1

Due to external constraints, acceptance/rejection is based on the test results from only a few items. In this paper we will consider the cases *n* = 1, 2, 3, 4, where *n* denotes the sample size (number of items taken for the inspection of a given lot).

Aspect 2

The property of interest *X* is distributed as follows:

• for many (the vast majority of) items, we have *X* = 0

• for some (very small) proportion of items *p*, *X* follows a distribution such as normal or lognormal

These two aspects make it impossible to apply ISO 3951 [1] (inspection by variables) or ISO 2859 [2] (inspection by attributes). Indeed, the special type of lot inspection under consideration here is "hybrid" in the sense that it couples aspects of both attributes and variables inspection. Determining the proportion of items with $X > 0$ would be the focus in inspection by attributes, while determining the distribution of the $X > 0$ values would be the focus in inspection by variables.

The special type of lot inspection considered here is no theoretical construct; on the contrary, it reflects many real-life lot inspection situations. For instance, in the case of imported products of animal origin, lot inspection may involve testing for veterinary drug residues. In such a situation, *p* represents the proportion of animals having been treated for a disease.

While this special type of inspection clearly falls under the broad category of lot inspection, the focus is quite different from that of the above-mentioned ISO standards. Indeed, since the sample size is known in advance (determined as it is by external constraints), the main question is not determining sampling size, but rather determining consumer risk (CR) and producer risk (PR). In particular, acceptance sampling plans will be assessed by means of the specific consumer and producer risks discussed in Section 5.5.1. Moreover, an alternative framework for the understanding lot inspection will be presented, where the measurand is no longer the (quantitative) property of interest which is determined by performing testing on the basis individual items in variables inspection. Rather, the measurand is now considered to be a description of the statistical properties of the lot, see Section 5.4.1. This means that the prior and posterior distributions will relate to these statistical parameters rather than to the mean concentration per item or in the lot.

The following questions will be discussed:

1. Can JCGM 106 [3] be applied not only in connection with the conformity of a single item, but also in connection with lot inspection?

2. Calculations of risks with a posterior distribution of the mean.

3. Calculations of risks with a posterior distribution of both mean and standard deviation.

5.6 Validation of apps

[This section is still to be completed]

5.7 References

FAO Mycotoxin Tool and Mycotoxin S&T User Guide
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.
Edward G. Schilling & Dean V. Neubauer (2017); Acceptance Sampling in Quality Control, (3rd edition) Chapman & Hall
ISBN 9781498733571
https://www.crcpress.com/downloads/K26314/CRC%20book%20website%20files.zip
https://www.crcpress.com/Acceptance-Sampling-in-Quality-ControlThird-Edition/Schilling-Neubauer/p/book/9781498733571
https://www.amazon.com/Acceptance-Sampling-Quality-Control-Statistics-dp-1498733573/dp/1498733573/ref=dp_ob_title_bk
Second edition:
https://www.amazon.com/Acceptance-Sampling-Quality-Control-Statistics/dp/1584889527
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
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.
Zweiterung et al. Relating Microbiological Criteria to Food Safety Objectives and Performance Objectives, Food Control (20) 2009, 967-979
ISO 10725:2000 Acceptance sampling plans and procedures for the inspection of bulk materials International Standards Organisation
ASQC Basic References in Quality. Volume 7: How and When to Perform Bayesian Acceptance Sampling.Thomas W Calvin (1990)

Hailey, W.A. (1980) Minimum Sample Size Single Sampling Plans: A Computerized Approach Journal of Quality Technology **12**, 4, pp230-235
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
General Standard for Contaminants and Toxins in Food and Feed (CXS 193-1995)
Code of Hygienic Practice for Powdered Formulae for Infants and Young Children (CXC 66 – 2008)
ISO 5725-1:2023
Accuracy (trueness and precision) of measurement methods and results — Part 1: General principles and definitions
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

Appendix II

List of Participants

Chair
Susan Morris
Ministry for Primary Industries - New Zealand

Roger Kissling
Fonterra - New Zealand

Fonterra - NZ

Co-chair
Petra Gowick
BVL - The Federal Office of Consumer Protection and Food Safety - Germany
Steffen Ulrig
Quo data - Germany
Bertrand Colson
Quo data - Germany

Argentina

Sara Valdez
Coordinadora de Laboratorio- SENASA

Jonatan Pietronave
Jefe de Área de laboratorio – INAL

Juan Pablo Maseda

Analista de laboratorio – INAL

Australia

Mark Lewin
Technical and Development Manager, Certified Reference Values, NMI-Australia
Australia

Susan Maddalena
Assistant Director - Residue Chemistry and Laboratory Performance Evaluation Section, National Residue Survey

Richard Coghlan
Senior Technical Expert, Analytical Services Branch, NMI-Australia

Brazil

Ligia Lindner Schreiner

Ana Claudia Marquim Firmo De Araujo
Health Regulation Expert, Brazilian Health Regulatory Agency – Anvisa

Canada

Thea Rawn
Research Scientist, Chemical Contaminant Section, Health Canada

China

Jing Xiao
Researcher, China

Wei Wang

Associate Researcher, China

Yuzhe Li
Associate Researcher, China

Wei Yu

Associate Researcher, China

Costa Rica

Karla Rojas Arrieta
Coordinator of the National CCMAS Committee, National Animal Health Service, Ministry of Agriculture and Livestock

Melina Flores Rodriguez
Technical Advisor of the National Codex Contact Point
Ministry of Economy, Industry and Commerce

Egypt

Mariam Barsoum Onsy
Food Standards Specialist, Egyptian Organization for Standardization & Quality (EOS)
Ministry of Trade and Industry

European Union

Franz Ulberth
Scientific Expert,European Commission, European Union

Judit Krommer
Policy Officer, European Commision, Eurpoean Union

Irene Seipelt

Codex Contact point, European union

Guyana

Norissa King
Assistant Scientific Officer, Goverrntment Analyst Food and Drug Department

Hungary

Attila Nagy

Chairman of CCMAS

Krisztina Bakó-Fányó

Codex Contact Point of Hungary

India

Ahammed Shabeer TP
Senior Scientist and Quality Manager, National Referral Laboratory, ICAR-NRCG, Pune, Maharashtra, India

Lalitha Gowda
Retired Chief Scientist, CFTRI, Mysore

Vinay Tyagi
Joint Director, National Accreditation Board for Testing and Calibration Laboratories (NABL)

Rini Narayan
Joint Director, National Accreditation Board for Testing and Calibration Laboratories (NABL)

Codex Contact Point
National Codex Contact Point, Food Safety and Standards Authority of India, Ministry of Health and Family Welfare

Indonesia

Tantu Yulianti
Laboratory accreditation assessor

Ratri Alfitasari

Standardization Analyst

Jamaica

Georgette Thompson
Senior Veterinary Biochemical Analyst, Veterinary Services Division- Ministry of Agriculture & Fisheries

Japan

Hidetaka Kobayashi
Coordinator, Risk and Crisis Management, Ministry of Agriculture, Forestry and Fisheries of Japan

Takanori Ukena
Director, Ministry of Agriculture, Forestry and Fisheries of Japan

Takahiro Mori
Associate Director, Ministry of Agriculture, Forestry and Fisheries of Japan

Tadashi Izawa
Deputy Director, Food Inspection and Safety Division, Pharmaceutical Safety and Environmental Health Bureau, Ministry of Health, Labour and Welfare

Takahiro Watanabe
Section Chief, Division of food safety information, National Institute of Health Sciences

Korea

YoungjunKim
Researcher, Ministry of Food and Drug Safety, Republic of Korea

MAFRA Korea Codex Contact Point
Quarantine Policy Division, Ministry of Agriculture, Food and Rural Affairs (MAFRA)

Kwang Kisoen

SPS Researcher, MAFRA

Malaysia

Azalina Binti Othman
Ab Rahman, Head of Section, Food Quality Division, Chemistry Department of Malaysia

Morocco

Mounir Rahlaoui
Head of Microbiology Laboratory (MOROCCO FOODEX)

Bouchra Messaoudi
Engineer, Service of Standardization and Codex at the National Food Safety Office (ONSSA)

Nigeria

Arannilewa Comfort
Assistant Director, Standards Organisation of Nigeria

Panama

Joseph Gallardo

Codex Contact Point, Panama

Peru

Jenny Tupac
Standardization specialist at the National Quality Institute, Peru

Saudi Arabia

Abdulaziz AAl Qaud
Product Registration Support Expert, Food Sector, Saudi Food and Drug Authority, Kingdom of Saudi Arabia

Mamdouh AAL-Jabri
Monitoring Expert, Food Sector, Saudi Food and Drug Authority, Kingdom of Saudi Arabia

Abdullah IAl Tamim
Section Head of Toxins & organic contaminants, Labs & Research Sector,Saudi Food and Drug Authority, Kingdom of Saudi Arabia

Senegal

Astou Ndiaye
Chef Unité "Chimie" Laboratoire d'Analyses et de Contrôle

Mathilde Cabral Ndior

Chef du laboratoire « Centre Anti-Poison »

Diégane Diéne Thiaré

Lecturer au Cheikh Anta Diop University - Dakar

Singapore

Ken Lee

Branch Head, Singapore Food Agency

Thailand

Chanchai Jaengsawang
Advisor, Department of Medical Sciences, Thailand

Paveena Pinkeaw
Standard Officer, National Bureau of Agricultural Commodity and Food Standards

Rungrassamee Mahakhaphong
Standard Officer, National Bureau of Agricultural Commodity and Food Standards

Uganda

Arthur Mukanga

Uganda National Bureau of Standards

Uruguay

Laura Flores
LATU

Pedro Friedrich

Codex Contact Point

USA

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

Heather (US CCP) Selig
International Issues Analyst, US Dept of Agriculture

OIV International Orgaisation for vine and Wine

Jean-ClaudeRuf
Scientific Director, Intnernational Orgainstion of Vine and Wine