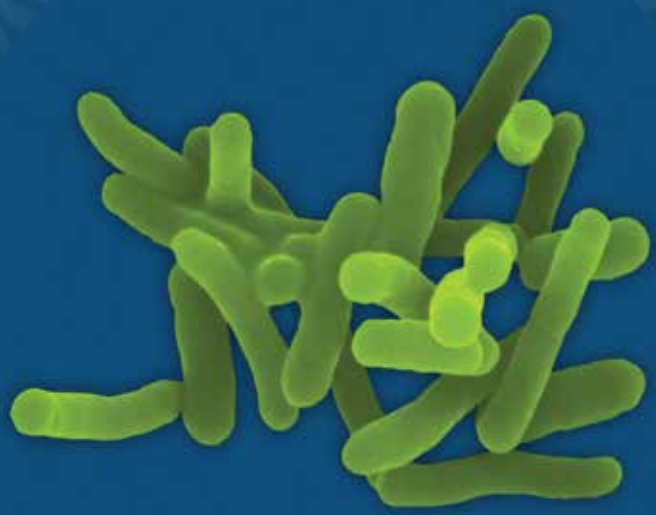
ISSN 1726-5274

Food and Agriculture Organization of the United Nations

World Health Organization

Statistical Aspects of Microbiological Criteria Related to Foods

A RISK MANAGERS GUIDE



24

MICROBIOLOGICAL RISK ASSESSMENT SERIES

For further information on the joint FAO/WHO activities on microbiological risk assessment and related areas, please contact

Office of Food Safety
Agriculture and Consumer Protection Department
Food and Agriculture Organization of the United Nations
Viale delle Terme di Caracalla
00153 Rome, Italy

Email: jemra@fao.org
Website: www.fao.org/food/food-safety-quality

Or

Department of Food Safety and Zoonoses
World Health Organization
20 Avenue Appia
1211 Geneva 27, Switzerland

Email: jemra@who.int
Website: www.who.int/foodsafety

Cover design
Food and Agriculture Organization of the United Nations and World Health Organization

Cover picture

24

MICROBIOLOGICAL RISK ASSESSMENT SERIES

Statistical Aspects of Microbiological Criteria Related to Foods

A RISK MANAGERS GUIDE

Food and Agriculture Organization of the United Nations
World Health Organization

Rome, 2016

WHO Library Cataloguing-in-Publication Data:
Statistical aspects of microbiological criteria related to foods: a risk manager's guide.

(Microbiological risk assessment series, 24)

1.Food Microbiology – statistics and numerical data. 2.Models, Statistical. 3.Risk Management. 4.Food Contamination. I.World Health Organization. II.Food and Agriculture Organization of the United Nations.

ISBN 978 92 4 156531 8 (WHO) (NLM classification: QW 85)
ISBN 978-92-5-108516-5 (FAO)
ISSN 1726-5274

Recommended citation:
FAO/WHO [Food and Agriculture Organization of the United Nations/World Health Organization]. 2016. Statistical Aspects of Microbiological Criteria Related to Foods. A Risk Managers Guide. *Microbiological Risk Assessment Series*, no 24. Rome. 120pp

Contents

ANNEXES

Figures

Examples

Preface

Microbiological Criteria (MC) have been used in food production and the food regulatory context for many years. While the food-specific aspects of MC are well understood, the mathematical and statistical aspects of MC are less well understood, which hinders the consistent and appropriate application of MC in the food industry. The Codex Committee on Food Hygiene (CCFH) recognized this issue, and at its 44th Session (12-16 November 2012) requested FAO and WHO to provide technical assistance on the mathematical and statistical aspects of MC.

In line with the Codex "Principles and Guidelines for the Establishment and Application of Microbiological Criteria Related to Foods" this document is targeted towards national regulators and food business operators (FBOs) who have a role in the establishment and implementation of MC for foods. However, other sectors, such as academia, that play a role in supporting government and the food industry may also find this material of value, for example, as a basis for training and education of future food safety professionals on this subject. Recognizing that knowledge in the area of statistics, or that access to such expertise, may be limited this document is intended to explain some of the basic concepts related to the mathematical and statistical description of microorganisms in food and sampling. Subsequently, issues related to the establishment and application of MC in two main areas are covered; decision-making about the safety of a particular food lot and assessing the ability of a process or system to produce safe food.

To make this document useful and accessible to a wide audience, the mathematical and statistical details, especially formulae, have been kept to a minimum, although the details are provided in the Annexes for those interested. In addition, a companion spreadsheet containing a collection of tools has been developed and used to illustrate some of the concepts discussed in this document. FAO/WHO strongly encourage the reader to use these tools to explore the sampling plans by changing the various parameters and observing the changes in the probability of lot acceptance.

Two versions of the companion spreadsheet – a Microsoft Excel version and a LibreOffice Calc version – are available from the FAO website (see Annex 3). Both spreadsheets contain "macros" and these need to be given permission to run for some calculations to work. In addition, Microsoft Excel functions are used in this document to show how some of the calculations can be performed, although the same functions also work in LibreOffice. However, some caution is in order, because all software programs can

contain bugs. In certain cases, there can be numerical problems (Almiron *et al*. 2010; McCullough and Yalta 2013), so it is always wise to verify results using another software program.

While the spreadsheet tools are intended to be used in conjunction with this document, FAO/WHO have also developed web-based tools that perform many of the same calculations and also include some additional functionality. In addition, the International Commission on Microbiological Specifications for Foods (ICMSF) has produced several tools for developing and evaluating sampling plans. FAO/WHO encourage readers to explore the various resources available (see Annex 2).

Lastly, a series of English-language videos has been prepared to accompany the text and examples (see Annex 3). These are aimed at providing complementary information and illustrating the use of the spreadsheet tools. These videos can be found at FAO's YouTube channel and links to specific videos are provided at appropriate places in the text.

Acknowledgements

The Food and Agriculture Organization of the United Nations and the World Health Organization would like to express their appreciation to all those who contributed to the preparation of this document through the provision of their time, expertise and other relevant information. Special appreciation is extended to all the members of the Expert Panel for their dedication to this project, to Dr Wayne Anderson for his expert chairing of the Panel and to Dr Andreas Kiermeier for all his work in expanding the text in accordance with the guidance provided by the Panel, developing the companion spreadsheets and multimedia materials and patiently and knowledgeably incorporating the feedback from the various rounds of reviews. All contributors are listed on the following pages.

We would also like to thank random.org for permission to use their tools in the video demonstrations. Thanks also to Sophie O'Conner for her support during the physical meeting of the Expert Panel and Roberto Sciotti for finalization of the videos.

Final editing for language was by Thorgeir Lawrence and layout by Joanne Morgante.

Contributors

EXPERT PANEL

Wayne Anderson, Director Food Science and Standards, Food Safety Authority, Ireland

Robert L. Buchanan, Director, Center for Food Safety and Security Systems, University of Maryland, United States of America

Darrell Donohue, Professor and Chair, College of Engineering, Michigan State University, United States of America

Moses Gathura Gichia, Department of Veterinary Services, Nairobi, Kenya

Sebastian Hielm, Ministerial Adviser, Ministry of Agriculture and Forestry, Finland

Andreas Kiermeier, Director, Statistical Process Improvement Consulting and Training Pty Ltd, Australia

Aditya Kumar Jain, QA Manager, National Dairy Development Board, India

Greg Paoli, Principal Risk Scientist, Risk Sciences International, Canada

Fernando Perez Rodriguez, Department of Food Science and Technology, University of Cordoba, Spain

Hajime Toyofuku, Chief, Department of International Health and Collaboration, National Institute of Public Health, Japan

Michael S. Williams, Senior Risk Analyst, Office of Public Health Science, Food Safety and Inspection Service, United States of America

Marcel Zwietering, Food Microbiology Laboratory, Wageningen University, The Netherlands

EXTERNAL REVIEWERS

John Bassett, John Bassett Consulting, United Kingdom

Catherine Bessy, Food Safety and Quality Unit, Food and Agriculture Organisation of the United Nations, Italy - Organization

Ian Jenson, Meat and Livestock Australia, Australia

Bobby Krishna, Food Control Department Dubai Municipality, United Arab Emirates

Anna Lammerding, AML Consulting, Canada

Maarten Nauta, Technical University of Denmark, Denmark

Tom Ross, University of Tasmania, Australia

Moez Sanaa, French Agency for Food, Environmental and Occupational Health & Safety, France

Donald Schaffner, Rutgers University, United States of America

Ousmain Toure, Institut National de Recherche en Santé Publique (INRSP) and Université des Sciences, des Techniques et des Technologies de Bamako (USTTB), Mali

Mieke Uyttendaele, Ghent University, Belgium

FAO/WHO SECRETARIAT

Sarah Cahill, Food Safety Officer, Food Safety and Quality Unit, Agriculture and Consumer Protection Department, FAO, Rome, Italy

Marisa Caipo, Food Safety Officer, FAO Regional Office for Latin America and the Caribbean, Santiago, Chile

Mina Kojima, Technical Officer, Department of Food Safety and Zoonoses, WHO, Geneva, Switzerland

Rei Nakagawa, Technical Officer, Department of Food Safety and Zoonoses, WHO, Geneva, Switzerland

DECLARATIONS OF INTEREST

All participants completed a Declaration of Interests form in advance of the meeting and were given the opportunity to update them in the course of the meeting. Based on the information provided, none of the experts were considered to present any potential conflict of interest.

Key terms and definitions

Acceptance number: The acceptance number (*c*) indicates the maximum number of unacceptable analytical units (two-class sampling plans) or marginally acceptable analytical units (three-class sampling plans) that can be tolerated in a sample while still accepting the lot. CAC/GL 50 (CAC 2004)

Acceptance sampling: Lots that are tested using a pre-specified sampling plan for the purpose of accepting/rejecting the lot.

Analytical unit: A single unit of food, from which a predetermined *analytical unit amount* is removed and tested for microorganisms. All or part of the sample unit may be used as the analytical unit.

Analytical unit amount: The relevant quantity – mass, volume or area – of the food product that is being tested, in each analytical unit. The analytical unit amount (*w*) is less than or equal to the sample unit amount.

Analytical unit detection probability: The proportion of analytical units that contain the target microorganism or contain the target organism above a predetermined microbiological limit, assuming the microbiological test is 100% specific and sensitive. The analytical unit detection probability is an estimate of the prevalence (of food units containing the target organism in a lot) and depends on the analytical unit amount, i.e. how much of the food is tested. The term *detection probability* is also used for brevity.

Batch: See *Lot*.

Consumer's risk: The probability of accepting a lot, P_1(accept), at a pre-specified analytical unit detection probability (p_1) or mean concentration (μ_1). The values of p_1 or μ_1 are usually chosen to indicate 'unacceptable' lots which, according to the customer or consumer, should only be accepted infrequently. See also 'Consumer's Risk' in CAC/GL 50 (CAC 2004).

Consumer's risk point: A pre-specified combination of analytical unit detection probability (p_1) or mean concentration (μ_1) and probability of acceptance, P_1(accept), i.e. consumer's risk. A suitable sampling plan will achieve a P(accept) that *is less than or equal* to P_1(accept). If a Producer's Risk Point is also specified, then p_1 (or μ_1) must be greater than p_0 (or μ_0). See also 'Consumer's Risk' in CAC/GL 50 (CAC 2004).

Detection probability: See *Analytical unit detection probability*.

Food Safety Control System: The combination of control measures that, when taken as whole, ensures that food is safe for its intended use. CAC/GL 69 (CAC 2008)

Hygiene indicator: A microorganism that is used as an indicator for hygiene of the production process.

Lot: A lot is a predefined quantity of food product, produced under similar, or uniform, conditions so that the units in the lot are similar in their microbiological status. See also CAC/GL 50 (CAC 2004).

Lot-by-lot testing: Under lot-by-lot testing every lot is tested using a pre-specified sampling plan for the purpose of accepting/rejecting each lot.

Negative: When the target organism is not detected in the analytical unit, then the analytical unit is often referred to as 'negative.'

P(accept): Probability of accepting a lot

P(reject): Probability of rejecting a lot; P(accept) + P(reject) = 100%

Positive: When the target organism is detected in the analytical unit, then the analytical unit is often referred to as 'positive.'

Prevalence: The percentage of units of the food lot that contain the target microorganism or contain the target organism above a predetermined microbiological limit.

Producer's risk: The probability of rejecting a lot, P_0(reject), at a pre-specified analytical unit detection probability (p_0) or mean concentration (μ_0). The values of p_0 or μ_0 are usually chosen to indicating 'acceptable' lots which, according to the supplier or producer, should only be rejected infrequently. Note that P_0(reject) is the complement of the probability of acceptance at this analytical unit detection probability or mean concentration, i.e. P_0(reject) = 100% - P_0(accept). See also 'Producer's Risk' in CAC/GL 50 (CAC 2004).

Producer's risk point: A user-specified combination of analytical unit detection probability (p_0) or mean concentration (μ_0) and probability of rejection, P_0(reject), i.e. producer's risk. A suitable sampling plan will achieve a P(reject) that is *less than or equal* to P_0(reject). The value of p_0 (or μ_0) must be less than those specified for p_1 (or μ_1) under the Consumer's Risk Point. See also 'Producer's Risk' in CAC/GL 50 (CAC 2004).

Sample: A subset of units from the lot or production process, selected in some predetermined manner.

Sample size: The number of sample units (n) in the sample.

Sample unit: A single unit of food of a predetermined sample unit amount. All or part of the Sample Unit may be used as the analytical unit.

Sample unit amount: The relevant quantity – mass, volume or area – of the food product that is being sampled, in each sample unit.

Sampling plan: A sampling plan is a scheme that defines the number of sample units to collect, the amount of food that constitutes a sample unit, the size of the analytical units tested, and the number of marginal and/or non-acceptable items allowed in a sample to evaluate the compliance status of a lot. See CAC/GL 50 (CAC, 2004)

Target (micro-)organism: The microorganism of interest.

Test-and-Hold: An application of acceptance sampling, where control over a lot is maintained, for example, through storage or tracking, until the test results are available and a decision about the lot's acceptability can be made.

Verification: The application of methods, procedures, tests and other evaluations, in addition to monitoring, to determine whether a control measure is or has been operating as intended. See CAC/GL 69 (CAC, 2008)

Zero acceptance number sampling plan: A sampling plan where the acceptance number (*c*) equals zero. This type of sampling plan is used primarily with two-class presence-absence sampling plans, but can also be used with two-class concentration-based sampling plans.

Zero tolerance: See *zero acceptance number sampling plan*.

List of abbreviations and mathematical symbols

APC: Aerobic Plate Count(s)

AOAC: Association of Analytical Communities

c: The acceptance number indicates the maximum number of unacceptable (two-class sampling plans) or marginally acceptable (three-class sampling plans) analytical units that can be tolerated without rejecting the lot or signalling that a process is out of control.

CCFH: Codex Committee on Food Hygiene

CCP: Critical Control Point

CFU: Colony forming unit(s)

FBO: Food Business Operator(s)

GHP: Good Hygienic Practice(s)

GMP: Good Manufacturing Practice(s)

HACCP: Hazard Analysis and Critical Control Point

ICMSF: International Commission on Microbiological Specifications for Foods

ISO: International Organization for Standardization

k: The critical value used for calculation of the probability of acceptance in variables sampling plans calculated from the sample size and consumer (or producer) risk point.

μ: Mean of a statistical distribution

m: The microbiological limit that differentiates acceptable from unacceptable microbial concentrations (two-class concentration-based and variables sampling plans) or acceptable from marginally acceptable microbial concentrations (three-class sampling plans).

M: The microbiological limit that differentiates marginally acceptable from unacceptable microbial concentrations (three-class sampling plans).

MC: Microbiological Criterion [or Criteria]

MPN: Most Probable Number

n: The sample size, i.e. the number of sample units that comprise a sample.

OC: Operating Characteristic

OCAP: Out-of-Control Action Plan

SD: Standard Deviation, also commonly denoted in mathematical equations by the Greek symbol σ

***w*:** The analytical unit amount, i.e. the size – mass, volume or area – of the analytical unit

Executive Summary

Microbiological Criteria (MC) related to foods have been used for many years. However, the mathematical and statistical aspects of MC are often not well understood. The Codex Committee on Food Hygiene (CCFH) recognized this issue, and, at its 44th Session (12-16 November 2012) requested FAO and WHO to provide technical assistance to enable it to develop an Annex to CAC/GL 21 "The Principles and Guidelines for the Establishment and Application of Microbiological Criteria Related to Foods" (CAC, 2013a). In this context, the CCFH requested FAO and WHO to provide inputs on issues such as the development and interpretation of operating characteristic (OC) curves, the effect of assumptions on the distribution and standard deviation (SD) of microorganisms in food in the elaboration of MC, as well as practical issues such as the establishment of the length of a moving window.

In response, FAO/WHO convened an expert meeting on the topic in Rome (8-10 October 2013) to establish the scope, structure and main content of a guidance document on these issues. It was agreed that this document should consist of a series of questions and answers. The questions, and key pieces of text, were developed using the collective experience and expertise of the expert panel. After the meeting, the answers to the questions, examples, companion spreadsheets and multimedia materials, were further developed.

A particular aim of this document is to illustrate the important mathematical and statistical aspects, but without the equations and mathematical details, which are relegated to the Annexes. It is hoped that the resulting document and support materials make this subject more accessible to a broad audience, including food business operators (FBOs), quality assurance managers, food safety-policy makers and risk managers.

The document is divided into three parts. Part 1 contains the basic concepts related to microorganisms in food and sampling, an understanding of which are required before embarking on the remainder of the document. This includes why the $\log_{10}$ transformation is used when dealing with microbiological data and why care needs to be taken when converting back to the arithmetic scale; random sampling and the alternatives to it; the importance of data to determine the statistical distribution to describe the microbial concentrations in the food; and a brief introduction to various sampling plans.

Part 2 is concerned with how to make decisions about individual lots. Information about how to define a lot for sampling and testing purposes is presented along

with the importance of lot independence. The OC curve and probability of acceptance are introduced. The sampling plans introduced in Part 1 are then examined in detail with respect to how the probability of acceptance is affected by some of the parameters that are specified in MC.

Part 3 deals with making decisions about process verification and control. The importance of process verification is outlined. A brief introduction to statistical process control is given and the moving windows approach is discussed in detail.

In developing this document, several areas were identified as falling outside of the scope of this document and consequently these are not addressed here. These areas include basic statistical concepts, such as calculating means and SDs; more advanced information on sampling, such as sequential and multiple sampling plans; the broader managerial aspects and approaches related to the establishment of MC, including risk-based MC; developing process control studies and national baseline surveys; and linking MC with Performance Objectives or Food Safety Objectives. Some of these areas are already adequately covered in other texts while others will require further development in the future.

Introduction

Microbiological Criteria (MC) have been used in food production and the food regulatory context for many years and Codex defines *Microbiological Criteria* as follows (CAC, 2013a):

> A ***microbiological criterion*** is a risk management metric which indicates the acceptability of a food, or the performance of either a process or a food safety control system following the outcome of sampling and testing for microorganisms, their toxins/metabolites or markers associated with pathogenicity or other traits at a specified point of the food chain.

In addition, according to Codex (CAC, 2013a), the components of MC for foods include

- The purpose of the microbiological criterion;
- The food, process or food safety control system to which the microbiological criterion applies;
- The specified point in the food chain where the microbiological criterion applies;
- The microorganism(s) and the reason for its selection;
- The microbiological limits (m, M) or other limits (e.g. a level of risk);
- A sampling plan defining the number of sample units to be taken (n), the size of the analytical unit and where appropriate, the acceptance number (c);
- Depending on its purpose, an indication of the statistical performance of the sampling plan; and
- Analytical methods and their performance parameters.

There are a number of approaches to developing MC. These range from developing MC based on empirical knowledge related to Good Hygienic Practices (GHPs), to using scientific knowledge of food safety control systems such as through HACCP, or by conducting a risk assessment. A range of examples of the application of MC have also been produced for Codex and these have since been published in *Food Control*[1].

While the food-specific aspects of MC are well understood, the mathematical and statistical aspects of MC, including sampling plans and statistical distributions, are less well understood, which hinders the consistent and appropriate application of MC in the food industry.

The Codex Committee on Food Hygiene (CCFH) recognized this issue, and, at its 44th Session (12-16 November 2012) requested FAO and WHO to provide technical assistance to enable it to develop an Annex to "The Principles and Guidelines for the Establishment and Application of Microbiological Criteria Related to Foods" (CAC, 2013a). The Annex concerns statistical and mathematical considerations for the elaboration of MC. The following terms of reference were provided by CCFH:

- How to develop and interpret operating characteristic (OC) curves;
- The impact of assumptions about the distribution and standard deviation (SD) of microorganisms in a food;
- How to establish the length of a moving window; and
- Any other relevant aspects.

In considering this request, FAO and WHO noted the wealth of information that is available on this issue and considered that there was no need to develop another textbook but rather a document that focused on those aspects relevant to the implementation of the Codex texts. It was further considered that any document developed needed to be understandable to an audience with limited statistical training, but who, once introduced to the issues in a clear and jargon-free manner could begin to quickly grasp the important elements. Thus, the challenge was to take the vast amount of information that exists on this topic and translate it into a concise document which a food safety professional in any country could understand.

In response FAO/WHO convened an Expert Meeting on the topic in Rome (8-10 October 2013). The expert panel decided that the framework for this document should consist of a series of questions and answers and that the document should be conversational and informal in style. The questions, and key pieces of text, were

[1] Volume 58, December 2015. Further details are available on the Food Control webpage http://www.journals.elsevier.com/food-control/).

developed during the meeting using the panel's collective experience and expertise. After the meeting, the text, examples, companion spreadsheet and multimedia materials were further developed, before the draft was reviewed by external peers from academia, regulatory agencies and the food industry.

ABOUT THIS DOCUMENT

Given the background of this document and its linkages with various Codex texts, it is recommended that this document be read in conjunction with key Codex texts, including:

- Principles and Guidelines for the Establishment and Application of Microbiological Criteria Related to Foods (CAC, 2013a)
- Guidelines for Food Import Control Systems (CAC, 2006)
- General Guidelines on Sampling (CAC, 2004)
- Principles and Guidelines for the Conduct of Microbiological Risk Management (CAC, 2008a)
- Guidelines for the Validation of Food Safety Control Measures (CAC, 2008b)
- Codex Alimentarius Commission, Procedural Manual (CAC, 2013b)

In developing this document, decisions had to be taken regarding what materials should and should not be included. Aspects that were excluded and the reasons for doing so are summarised below:

1. **Basic statistical concepts:** Information on sample means, SDs, confidence intervals, and descriptions of statistical distributions were excluded as it was considered that these are already covered adequately in existing texts, e.g. Moore, McCabe and Craig. (2012), Bower (2013) or some of the web-based resources provided in Annex 2.
2. **More advanced information on sampling:** More advanced sampling plans, such as sequential and multiple sampling plans, can offer benefits in requiring fewer sample units, on average. However, they are usually not feasible when testing for microorganisms because of the time lag between collecting the sample units and obtaining a test result. Information about such sampling plans can, for example, be found in Montgomery (2012) and Schilling and Neubauer (2009).
3. **Approaches for establishing Microbiological Criteria:** As the focus of this document is on the statistical aspects of MC, it does not address the broader issues of the role of MC or the risk management scenarios in which they might be established. Information on establishing MC has been covered elsewhere, e.g. the relevant Codex guidelines (CAC, 2013a), International Commission on Microbiological Specifications for Foods (ICMSF, 2002), and are outside

the scope of this document. In particular, this should be kept in mind when reading the final section for each sampling plan entitled "Putting it all together …" where we deal solely with the statistical aspects of establishing sampling plans.

4. **Risk-based Microbiological Criteria.** There are a number of approaches to developing MC, including risk assessment (which results in risk-based MC). However, the most appropriate approach will depend on the specific circumstance. For example, performing a risk assessment can be very time consuming and resource intensive, and is probably more appropriate for national regulatory agencies than individual food business operators (FBOs).
5. **Link between Microbiological Criteria and Performance Objectives or Food Safety Objectives.** While it was recognized that linking MC to risk-based metrics such as performance objectives or food safety objectives, is an important feature of risk-based management, these aspects were not part of the request from CCFH. Furthermore, it was recognized that inclusion of these aspects would require considerable additional work and should be treated separately in a future document.
6. **Process control study versus National baseline survey.** While the application of MC and baseline surveys both involve sampling, they have different aims. MC and the associated sampling plans are used to accept or reject lots and they can be used to assist in process verification. In contrast, baseline surveys are used to estimate the extent of microbiological contamination (prevalence, mean and SD) in a particular food product or commodity, taking into account different sources of variability, such as country of origin, supply chains, etc. This can be used to collect the necessary information needed for developing MC. Thus, while some of the basic statistical concepts are the same, the application of these concepts to a baseline surveys is not addressed here.

In discussing the amount of food that constitutes an individual sample unit the term *sample unit amount* is used throughout this document. Similarly, the amount of food that is used for the microbiological test is referred to as the *analytical unit amount*. These terms are used without loss of generality, and apply to sample units that are based on weight (per g), volume (per ml) or area (per cm^2). When reference is made to a particular example the more specific term, e.g. *sample unit weight* and *analytical unit weight*, may be used.

Finally, it should be noted that percentages have been used to denote probabilities. This approach is not strictly correct in a mathematical sense (where probabilities are expressed as proportions between 0 and 1), but it is hoped that this approach assists the readability and understanding of the materials for our target audience.

PART 1

Basic concepts related to microorganisms in food and sampling

1.1 WHY DO WE UNDERTAKE SAMPLING AND MICROBIOLOGICAL TESTING ON FOOD?

The presence of certain microorganisms in foods can affect public health and the quality of foods consumed. For this reason, the sampling and testing of foods for a variety of microorganisms is a common part of most food safety and quality systems. The main uses of microbiological sampling in the food industry are as follows:

- *Compliance testing*: Public health regulatory agencies often sample and test food product on the market for compliance with national food safety standards, e.g. *Listeria monocytogenes* in some ready-to-eat (RTE) foods. Food manufacturers sometimes also use "test and hold" to demonstrate compliance prior to releasing for product into the market place.
- *Import and export certification testing*: Food safety regulatory agencies may require food products to be sampled and tested for pathogens of public health interest prior to import/export, e.g. *Escherichia coli* O157 in beef trim prior to export to the United States of America.
- *Commercial supply agreements*: Commercial agreements often include microbial specifications that the supplier needs to meet. The supplier may need to demonstrate compliance by sampling and testing product prior to shipping, while customers may randomly sample and test product to check compliance.
- *Process Verification*: Food manufacturers can use sampling and testing to demonstrate that a food production process is in control and operating as intended.

1.2 WHAT DO WE NEED TO REMEMBER ABOUT THE CHARACTERISTICS OF MICROBIOLOGICAL POPULATIONS IN FOOD?

Despite the different reasons for microbiological testing of food there are a number of common factors that are integral to understanding the effectiveness of testing programmes.

The ability to detect microorganisms of interest is relatively easy when the degree of contamination of the food is high. However, as the level, or concentration, of the microorganisms drops it becomes increasingly difficult to detect them despite the fact that they are present. This reflects the particulate nature of microorganisms which means that at very low concentrations there is a distinct possibility that a microorganism will not be present in a given sample (Figure 1). This is different from chemical contaminants, which are generally assumed to be more uniformly distributed through the food product, or at least in the analytical unit after homogenization during laboratory sample preparation.

Consequently, in many situations only a proportion of all units in a food lot will contain the target microorganism. In the food safety, epidemiology and public health areas this phenomenon (sometimes called a *defect rate* in engineering terms) is commonly known as *prevalence*, while *proportion defective* or *percentage non-conforming* are terms often used in acceptance sampling and statistical process control. We use the term "prevalence" to denote the actual, unknown proportion of food units in a lot that are contaminated, usually with a pathogen, and the term "analytical unit detection probability", or simply "detection probability", to indicate the estimate obtained from the sample, using a particular microbiological test and analytical unit amount (see also "1.2.6 What is random sampling and what are the alternatives?" and Figure 4).

As indicated above, detecting microorganisms becomes more difficult the lower the prevalence. For example, if you have a prevalence of 50% (i.e. 1 in 2 food units contains the target organism), then you would be highly confident of selecting a contaminated unit of food (and thus detecting the microorganisms) even when you sample just three or four units. Conversely, if you had a prevalence of only 1% (i.e. 1 in 100 food units contains the microorganisms), then you would have to sample many more units of food to have a similar level of confidence, e.g. about 300 units are needed to achieve a 95% probability of detection if only 1% of food units are contaminated. In other words, as a process becomes more controlled and the degree of contamination of the food decreases, it becomes harder to find the microorganism by sampling. Therefore, it is necessary to increase the number and/

or size of the analytical units to detect the target organism, if present, or to have high confidence that the organism is in fact not present or is present only at very low levels, e.g. to verify compliance. This becomes a limiting factor in the direct detection of pathogenic microorganisms and hence the lack of utility in using sampling as way to 'control' the safety of food.

For this reason, when pathogen contamination levels are low, an alternate approach is to use hygiene indicator organisms to identify processing conditions that may have an increased chance of leading to contamination with pathogens. Hygiene indicators are microorganisms that typically occur at substantially higher concentrations in the food than the pathogen. The testing methods quantify these levels, thereby allowing decisions to be based on the concentration of the microorganism rather than its mere presence. However, sometime the presence/absence of a hygiene indicator could be used to determine the effectiveness of a specific process. For example, effective thermal processing should render a food free of any hygiene indicators and their presence could indicate process failure. Quantification of microorganisms may also be important for specific pathogens that can occur more frequently at higher concentrations in some foods without causing public health risks, e.g. *Listeria monocytogenes* (in foods that do not support its growth), *Campylobacter* spp., *Staphylococcus aureus*, or pathogenic *Vibrio parahaemolyticus*.

A key point to remember is that the microbiological methods and the underlying statistics are different for quantifying high levels of microorganisms versus detecting low levels by presence/absence tests. These approaches generate different amounts of information and consequently the associated statistical considerations for their application to acceptance sampling and statistical process control will also differ.

1.2.1 How does microbiological testing differ from chemical testing?

An important aspect of microbiological test methods is the concept of the *lower limit of detection*. This concept originated in the analysis of chemicals but it is not directly applicable to the detection of microorganisms. In this case, when the number of microorganisms is very low, it is not the ability of the method to detect the microorganism that determines the outcome of the test but rather, it is the probability that a cell was actually present in the analytical unit analysed. For example, if a method is analysing 10 µl and the organism is present at a level of 1 cell per ml, then the probability that 10 µl will contain the microorganism is (approximately) 1 in 100 or 1% (Figure 1). This is a consequence of the particulate nature of microorganisms.

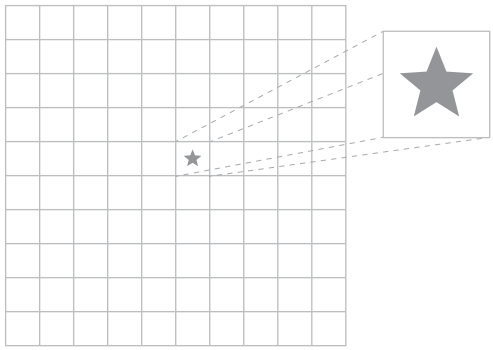


FIGURE 1: Diagram of a 1 ml food sample containing 1 organism, from which a 10 µl aliquot, represented by the small squares, is selected (sub-sampled) for testing, e.g. plating. Only if the aliquot containing the organism is selected (see 'blow-up') will we find the organism during testing.

This phenomenon gives rise to a key difference in the interpretation of the term "homogeneous" between chemical and microbiological testing. In both cases, food samples are thoroughly mixed in the laboratory in such a way as to make the contamination as uniform as possible throughout the sample matrix. This process is referred to as homogenization and may involve a blender, stomacher or similar instruments. This generally works very well for chemical contaminants and the interpretation in this setting is that the contamination is (more or less) uniform throughout the sample matrix.

In contrast, a homogenized microbiological sample will have the target organism 'floating' throughout the sample matrix. In this case, the process of selecting an aliquot for plating involves an element of randomness and thus a probability of different numbers of organisms being found in each sample or sub-sample. Sometimes we might get 4 organisms on a plate and sometimes it might be 5, or 11, or 1, or none. However, if the sample is homogenous in a microbiological sense, then the counts from replicate plates would follow a random pattern that, in statistical terms, is consistent with the Poisson distribution (for example as shown later in Figure 2, the middle). In fact, it is this property on which the Most Probable Number (MPN) approach to microbiological testing is based (Cochran, 1950).

A key point to remember is that, because of the particulate nature of microorganisms, homogeneity in a microbiological setting does not imply uniformity, i.e. constant levels throughout the analytical unit.

1.2.2 How are microorganisms distributed in food?

Microorganisms can be distributed (in a spatial or temporal sense) throughout a food lot or process in a variety of patterns. These are influenced by the microbial growth/death and the type of food processing that occurs during production, e.g. joining and mixing of ingredients or filling and packaging of final product. A good description of the various mechanisms can be found in the document produced by the International Life Sciences Institute Europe Risk Analysis in Food Microbiological Task Force (Bassett *et al*, 2010).

Three examples of the two dimensional spatial distribution of microorganisms in food – regular, random and clustered – are presented in Figure 2. These patterns occur because of the particulate nature of microorganism and how microbiological tests work (see "1.2.1 How does microbiological testing differ from chemical testing?"). The following list provides idealized examples of how these three types of contamination patterns may occur in food, although microbial patterns in actual food samples may be combinations of these three idealized examples.

- A *regular contamination pattern* can be the result of contamination events in the production process, for example, a contaminated filler head. Regular contamination could relate to the interval between contaminated units as well as the number of organisms in a contaminated product. As a consequence, regular contamination is characterized by high prevalence of contaminated units in the lot and a low variability in concentration.
- A *random contamination pattern* is usually the result of thorough mixing or when contamination events occur randomly. There is no specific pattern to the contamination, which may make it harder to find a source. In a microbiological setting this would result in homogeneous contamination (see "1.2.1 How does microbiological testing differ from chemical testing?")
- A *clustered contamination pattern* can be the result of a contamination event that is followed by some growth of the organism and limited mixing of the product. As a result, the individual cells are not widely spread through the food. As a consequence, the prevalence of contaminated units in the lot may be low and the variability in concentration may be high.

These different spatial patterns lead to different statistical distributions that describe the contamination patterns mathematically and detailed information about these distributions can also be found in the ILSI document (Bassett *et al.*, 2010).

A key point to remember is that it is not possible to generalize how microorganisms are distributed throughout a food product, but knowledge about the process and mechanisms for contamination will be important to determine a suitable statistical distribution and appropriate sampling plan.

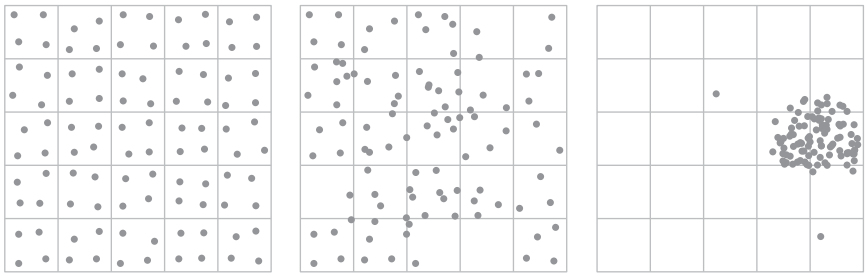


FIGURE 2: Examples of the spatial distribution of 100 microorganisms over 25 portions of food. Source: Jongenburger *et al*. 2012a – reproduced with permission from Food Control. This work was commissioned by the Risk Analysis in Food Microbiology Task Force of the European branch of the International Life Sciences Institute (ILSI Europe).

1.2.3 Why do we use $\log_{10}$ numbers and why do we need to be careful when interpreting them?

Microorganisms grow by dividing in two, resulting in doubling of the number of organisms during each replication cycle. For example, 1 organism will grow to 2, then 4, 8, 16, 32, etc. This is known as exponential growth and can result in rapid increases in microbial populations under suitable conditions. Because microbial populations can be large, e.g. a billion organisms per gram, or 10^9 cfu/g, food microbiologists commonly convert the arithmetic numbers to log numbers to simplify the data analysis and interpretation. While any log-transformation will accomplish this, the log base 10 ($\log_{10}$) has a simple interpretation and is commonly used in food microbiology, as compared to other science areas where log base *e* (ln) or log base 2 ($\log_2$) are commonly used. This conversion can be performed in Excel using the function $\log_{10}$ or simply log, and the reverse transformation from log numbers to the arithmetic numbers can be achieved using 10^, as shown in Example 1.

However, this log-transformation can cause confusion if you do not understand its consequences. For example, it is common to add or subtract log numbers to calculate the effect of increases or reductions in the microbial population, but we must not forget that doing so is equivalent to multiplying or dividing the original arithmetic values, as shown in Example 2, respectively.

In fact, a 1-unit increase in the $\log_{10}$ number is equivalent to a 10-fold increase on the arithmetic scale, so an increase from 2 $\log_{10}$ cfu/g to 3 $\log_{10}$ cfu/g is equivalent to an increase from 100 cfu/g to 1 000 cfu/g. Similarly, a 1 unit decrease in the $\log_{10}$ number is equivalent to a 10-fold decrease (Example 3), or a 90% reduction, on the

Example 1
Converting to and from $\log_{10}$

The following table provides a quick guide to some $\log_{10}$ numbers.

Arithmetic number	$\log_{10}$ Number
$0.01 = 10^{-2}$	-2
$0.1 = 10^{-1}$	-1
$1.0 = 10^{0}$	0
$10.0 = 10^{1}$	1
$100.0 = 10^{2}$	2
$1\,000.0 = 10^{3}$	3

Converting the value 150 cfu/g to the $\log_{10}$ scale results in 2.176 $\log_{10}$ cfu/g. Converting back and forth between arithmetic and log numbers in Excel or LibreOffice can be done by typing the following formulae in quotes into a cell in the spreadsheet:

- Convert to $\log_{10}$: "=$\log_{10}$(150)"
- Convert from $\log_{10}$: "=10^2.18" (not exactly 150, more decimals will give a closer result)
- Finally, given the accuracy of microbiological testing, one digit after the decimal point is sufficient to report a final result, but at least two (or more) digits should be included for all intermediate calculation steps.

A video showing how to perform this calculation can be found at http://youtu.be/mGNRmGDgNOU.

Example 2
The effect of combining analytical units on the $\log_{10}$ concentration

A 10 g unit of food has a concentration of a microbe of 10 000 cfu/g (4 $\log_{10}$ cfu/g) and another 10 g unit has a concentration of the same microbe of 1 000 cfu/g (3 $\log_{10}$ cfu/g). The two analytical units contain a total of 100 000 cfu and 10 000 cfu of the microbe, respectively.

Mixing these two units results in a total of 110 000 cfu (1.1×10^5 cfu) in the 20 g of food. This is obtained by adding the two arithmetic numbers (100 000 and 10 000), and hence the concentration (dividing by 20 g) equals 110 000/20 = 5 500 = 5.5×10^3 cfu/g (or 3.74 $\log_{10}$ cfu/g).

It would be incorrect to add the two $\log_{10}$ numbers (4 $\log_{10}$ cfu/g and 3 $\log_{10}$ cfu/g) and conclude that the concentration of the mixed sample is 7 $\log_{10}$ cfu/g (i.e. 10 000 000 cfu/g). It would also be incorrect to assume that the final concentration of the mixture of two equal parts could be estimated by taking the average of the two $\log_{10}$ concentrations (3 $\log_{10}$ cfu/g and 4 $\log_{10}$ cfu/g) and assuming that the final concentration will be 3.5 $\log_{10}$ cfu/g. This average taken on the log-scale, which equals 3 162 cfu/g, erroneously underestimates the true average of 5 500 cfu/g. The magnitude of these errors can be quite significant and must be avoided.

Example 3
Interpretation of $\log_{10}$ increases and reductions

An antimicrobial intervention is shown to achieve a 2-log reduction of the microbial concentration in the food product. This means that 99% of the microorganisms are removed or inactivated as a result of the intervention. Notice that this reduction is independent of whether the starting concentration is 4 $\log_{10}$ cfu/g or 2 $\log_{10}$ cfu/g.

Similarly, a 3 log reduction is equivalent to a 99.9% reduction, 4 log is equivalent to 99.99% reduction, and so on. For example, a 12 log reduction, which is a commonly used process criterion in the canning industry to inactivate *Clostridium botulinum* spores, is a 99.9999999999% reduction.

arithmetic scale, e.g. a decrease from 3 $\log_{10}$ cfu/g to 2 $\log_{10}$ cfu/g is equivalent to a decrease from 1 000 cfu/g to 100 cfu/g. Importantly, this effect is independent of the starting concentration.

Of particular importance is the effect of such a log-transformation on the statistical distribution that is used to characterize the microbial contamination in the food product, which must be taken into account when developing an MC. Microbiological populations in foods are often described using a $\log_{10}$-normal distribution[2] which is a right-skewed distribution, shown in Figure 3a. When you take a $\log_{10}$ transformation (see also Example 1) of numbers that conform to a $\log_{10}$-normal distribution (Figure 3a), the $\log_{10}$ values take on a normal distribution (Figure 3b). Consequently, the $\log_{10}$ transformation is useful for analysing microbiological data, as a range of classical statistical methods can be applied to the normally distributed data, e.g. confidence intervals, regression analysis, t-test and analysis of variance. Therefore, microbial counts should first be $\log_{10}$-transformed before calculating the mean and SD of these transformed counts – otherwise the right skewed nature of the distribution will "inflate" these two statistics, which can result in incorrect results.

However, care must be taken when the calculated statistics – mean and SD – are transformed back to the arithmetic scale and how the corresponding values are interpreted. It is tempting to transform the average $\log_{10}$ count, via direct exponentiation, to the arithmetic scale. However, this would lead to an underestimation of the "average" number of microorganisms in the food and subsequent misinterpre-

[2] In some microbiological text this distribution is sometimes simply referred to as a lognormal or log-normal distribution and should not be confused with the lognormal distribution used in statistical texts which uses natural logarithm (ln) as the basis of the transformation.

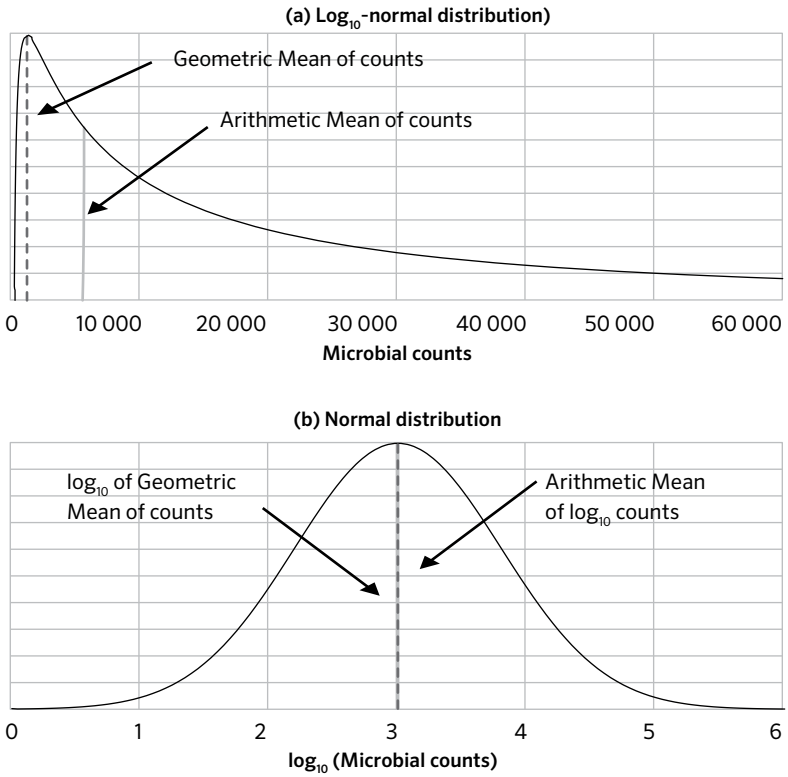


FIGURE 3: Plots of a $\log_{10}$-normal distribution (left) and normal distribution (right). See also Example 1 for an explanation of the $\log_{10}$ transformation. The geometric mean and arithmetic means are shown in these plots because of their special relationship, as illustrated in Example 4.

tation of the risk, which is based on arithmetic mean (see Figure 3 and Example 4). The appropriate calculation of the arithmetic mean from the mean of the $\log_{10}$ counts can be performed using the companion spreadsheet.

A key point to remember is that microbial concentrations are usually $\log_{10}$ transformed for data analysis or for other conveniences such as graphing. However, care needs to be taken in interpreting $\log_{10}$ transformed numbers, including any statistics and mathematical manipulations of them, especially when converting back to the arithmetic scale to assess risk, as shown in Example 4.

Example 4
Relationship between means on the arithmetic and log[10] scales

Consider a microbial concentration in a food which follows a $\log_{10}$-normal distribution with mean 3 and standard deviation 0.8 $\log_{10}$ cfu/g. That is, on the $\log_{10}$-scale the distribution is normal with mean 3 $\log_{10}$ cfu/g and standard deviation of 0.8 $\log_{10}$ cfu/g (see b).

Taking the exponential of the mean on the log10 scale value gives the *geometric mean* on the arithmetic scale, i.e. $10^3 = 1\,000$ cfu/g.

However, the conversions to find the arithmetic mean and standard deviation on the arithmetic scale are more complicated. In this case, the arithmetic average concentration is over 5 455 cfu/g and assuming that the arithmetic average concentration is adequately described by 1 000 cfu/g (using the exponentiation of the $\log_{10}$ mean shown above) would underestimate the arithmetic average by a factor of more than 5!

The companion spreadsheet contains a calculator on the Mean and SD calculator Tab and the mathematical detail can be found in Annex A1.1 Converting the Mean and Standard Deviation from the log10 scale to the arithmetic scale".

A video showing you how to use the calculator can be found at http://youtu.be/iQWCnykNKWQ.

1.2.4 What are the important aspects that characterize the statistical distribution of microorganisms in food?

There are many statistical distributions that can be used to describe and model microbial contamination of food. The applicability of a range of distributions in relation to the patterns shown in Figure 2 were discussed by Bassett *et al.*, (2010) in relation to modelling real food systems and this is also an area of ongoing research (Busschaert *et al.*, 2010; Commeau *et al.*, 2012; Gonzales-Barron *et al.*, 2010; Gonzales-Barron and Butler 2011; Jongenburger, 2012; Jongenburger *et al.*, 2012a, 2012b).

However, all these distributions require the estimation of the following important characteristics (or at least some of them):

- **Prevalence** The percentage of food units in the lot that are contaminated. The importance of this parameter depends on the organism of interest, and is usually used in conjunction with pathogens that may occur at very low concentrations and a presence-absence based microbiological test. It is estimated by the analytical unit detection probability.
- **Mean/Average (μ)** The 'typical' ($\log_{10}$) microbial count that can be expected

over the long term. The mean may be calculated from contaminated units only or from all units, depending on the distribution being modelled.
- **Standard deviation (SD)** The variability between the ($\log_{10}$) microbial counts between sample / analytical units of the same food. The SD may also be calculated from contaminated units only or from all units, depending on the distribution being modelled.
- **Shape** The shape of the distribution of counts (e.g. Figure 3) influences the mathematical distribution that may be used to model them. The histogram is useful to visualize the shape.

There often is no simple answer as to what distribution might be most appropriate in any particular circumstance. In fact, the only way this question can be answered is through data collection and analysis. That is because data allows us to estimate the parameters identified above and this allows us to make realistic assumptions when developing MC and choosing a sampling plan. More information on this topic is provided under "1.2.5 How do we get a better description of the microbial levels in our food product?".

1.2.4.1 What if we have no data to determine the distribution of microorganisms in our food?

Sometimes there is a need to develop a sampling plan without the benefit of data needed to evaluate the applicability and validity of all the parameters described above. While caution needs to be applied in such circumstances, the following approach may help.

1. Is there existing literature for similar food products/production using similar microbiological test methods? If so, decide whether mean, SD, analytical unit detection probability and shape estimates from these sources might provide a sensible starting point.
2. If no published information exists assume that the distribution of microbial counts between different food unit can be described by a $\log_{10}$-normal distribution (see Figure 3).
3. If no published information exists choose a value for the variability in $\log_{10}$ microbial counts (SD) between food units within a lot. Experience and research show that sensible starting values exist for the variability in microbial counts for different food products and the ICMSF provided some guidance in relation to this. For example, van Schothorst *et al*., (2009) discuss scenarios where they use a SD of
    - 0.2 $\log_{10}$ cfu/g for well mixed foods, such as liquids;
    - 0.4 $\log_{10}$ cfu/g in reasonably mixed foods, such as ground meat; and
    - 0.8 $\log_{10}$ cfu/g for less well mixed foods.

While larger SDs may be appropriate when clumping occurs, an SD of 0.8 $\log_{10}$ cfu/g is generally a reasonable starting value.

Consequently, if there is little variability in the microbial concentration in the food, i.e. in processes that are better controlled, then any given sampling plan will be more discriminating (see "2.8.4 What is meant by discrimination and stringency of sampling plans?"), that is, the probability of acceptance drops quickly from 100% to 0%, as we will show in Part 2 in relation to the various sampling plans.

Once you start applying the sampling plan it is important that you capture the data that are being generated, e.g. in a spreadsheet or database. These data should be assessed on a regular basis to obtain better estimates of the contamination levels in the food and the pattern of contamination in the food. In that way, the sampling plans can be refined over time.

A key point to remember is that it is better to start with informed and sensible guesses of the detection probability, mean, SD or shape and to refine the initial assumptions once better data become available than it is to do nothing.

1.2.4.2 Does it matter what statistical distribution we use to describe the microbial contamination?

The reason we use a distribution to describe the microbial contamination in the food lot is that this describes the pattern of contamination levels (high, low, detect, not detect, etc.) to be expected among the randomly chosen sample units (from the same lot). Knowing the pattern allows us to calculate the probability with which such lots would be rejected when a particular sampling plans is used (see "2.8.3 What is the Operating Characteristic Curve?" later).

Consequently, it is important that an appropriate statistical distribution, or combination of statistical distributions, is used[3] to calculate how frequently lots with a particular microbiological concentration can be expected to be accepted and rejected. As noted above, a common default is to assume a $\log_{10}$-normal distribution if no better or other information is available. However other distributions, and their combinations, are possible (e.g. Bassett *et al.*, 2010; Habraken, Mossel and van der Reek, 1986; Jongenburger *et al.*, 2012a) and these will influence the shape of the OC Curve.

However, the statistical principles are unchanged, irrespective of which statistical distribution is used. Consequently, we illustrate these principles here by making

[3] Based on theoretical or empirical justification.

some simplifying assumptions, which allow for easy probability calculations using the companion spreadsheet. However, different statistical distributions can be evaluated using more advanced tools, such as the FAO/WHO web-based tools (http://www.fstools.org/sampling) or those from the ICMSF (http://www.icmsf.org/).

1.2.5 How do we get a better description of the microbial levels in our food product?

The only way to get a good description of the microbial levels in *your* food product is to sample the food, test it for the target organism and collect and analyse the data. You simply cannot expect to make informed decisions without relevant data.

Historical data from routine monitoring may be a good starting point. Alternatively, new data need to be collected from one or more stages of the production process (raw ingredients, different stages of production, final product) depending on what information you seek to obtain and where the MC is to be applied. But keep in mind that concentration data will be of greatest value and hence you may need to enumerate the most suitable organism for your circumstances. These data will allow you to estimate the mean, SD and prevalence and to determine the shape of the microbial count distribution and thus assess whether the $\log_{10}$-normal distribution is appropriate in your situation.

Collecting small sets of sample units, e.g. twice a day over a few weeks, will give you some information to get started. In developing a sampling plan for a new food product or manufacturing line, a "process control study" is often performed to establish an initial baseline estimate of performance when the process is "under control." However, you cannot expect to do this just once and do nothing further. You should think about and assess the various effects that other factors can have on the microbial contamination levels and variability, e.g. suppliers, staff on different shifts, seasonality, etc. The important point is to get started and to refine your sampling programme as better information becomes available.

A key point to remember is that you need data to make informed decisions. Use historical data where appropriate, but keep in mind how those data have been collected and what the limitations are for interpreting them.

1.2.6 What is random sampling and what are the alternatives?

We are often interested in drawing a conclusion about the microbial quality or safety of a lot of food or the control of the underlying production process. However, it would be impractical to test each food unit in the lot or from the pro-

duction process, because there would be nothing left to sell. Hence we must make a decision about the lot or process based on a *sample*, i.e. a set of units obtained in some predetermined way. The units of food that constitute the *sample* are referred to as *sample units*, each of a particular weight, volume or area. While some people refer to this weight, volume or area as the sample size, it is better to refer to this as the *sample unit amount* since the term *sample size* is more commonly used to denote the number of sample units (n) in the sample. A portion of each sample unit, referred to here as the analytical unit amount (w), is subsequently used as the analytical unit in the microbiological test. A graphical representation of the sampling and microbiological testing process is shown in Figure 4, where a sample of size $n = 5$ is selected from the lot, i.e. the sample consists of 5 sample units. Each sample unit consists of a single unit of the food product, so that for a 100 g packet the sample unit amount is therefore 100 g. An analytical unit weighing 25 g, equal to the analytical unit amount w, is then sub-sampled from each sample unit (potentially after homogenizing the sample unit). The analytical unit is then tested for the target microorganism using either a quantitative (enumeration) or qualitative (presence / absence) test. If the 100 g sample unit is first mixed before the 25 g analytical unit is sub-sampled, then this will have implications on the enumeration result or the probability of detection. However, such situations require an additional probability calculation (to account for homogenizing and sub-sampling of the analytical unit) and are beyond the scope of this document. We would encourage readers to keep the sample unit and analytical units the same, as far as practicable.

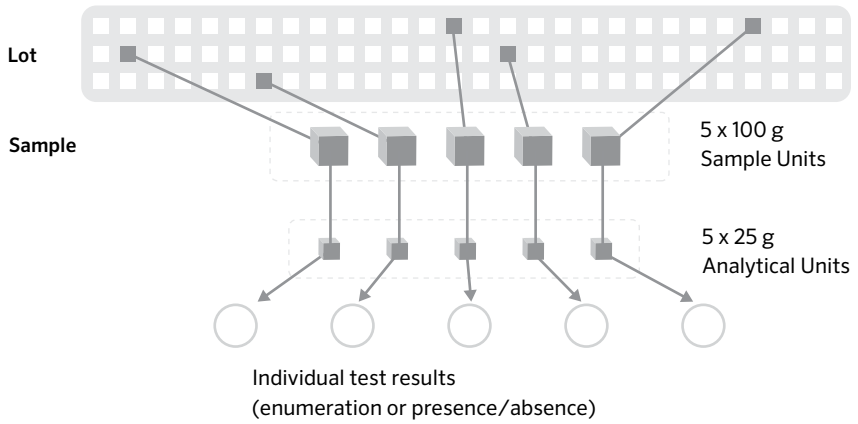


FIGURE 4: Graphical representation of a lot of food units, from which a sample of $n = 5$ sample units is selected. Each sample unit consists of a 100 g (the sample unit amount) from which an analytical unit weighing 25 g (the analytical unit amount) is sub-sampled for enumeration or presence / absence testing.

The most common way to select a sample is *random sampling*, where each possible sample has the same probability of being selected. This is because the assumption of random sampling underpins any statistical data analysis. In practical terms, when the population from which we sample is large, this is equivalent to saying that each sampling unit has the same chance of being selected. To ensure that conscious or

Example 5
Random Sampling

To truly use random sampling, you need to use a random number generator to create a set of suitable random numbers that can be used to select the necessary sample units. Here we use a widely available option – the website www.random.org – although random number tables, which can be found in most undergraduate statistics books, or spreadsheet programs like Microsoft Excel or LibreOffice, can also be used.

Sampling discrete food units: Assume that we have 1 000 food units from which we want to select a random sample of 20 units. The following steps explain how to do this.

- Number the food units sequentially from 1 to 1 000.
- To generate 20 random numbers between 1 and 1 000 visit http://www.random.org/integer-sets/, which will allow you to generate non-repeating random numbers.
- Enter the following details into the form:
  - Generate 1 set with 20 unique random integers in each.
  - Each integer should have a value between 1 and 1 000.
  - Leave the remaining options as they are.

For us this created the following output:
Set 1: 9, 25, 49, 72, 119, 156, 172, 257, 325, 338, 496, 595, 603, 607, 639, 798, 846, 862, 914, 966.

Consequently, we would sample food units 9, 25, 49, and so on. A video showing how to generate a random sample can be found at http://youtu.be/AVnQdTqBqDA.

Sampling bulk product or from a production line: If the product we are interested in sampling is in bulk form then the best way is to sample the product is before it is stored or as it is removed from bulk storage. This is then similar to sampling from a production line.

In this case, we do not have discrete units to sample, so instead we randomly select discrete time intervals (after a given starting time). These intervals can be as large as desired, e.g. 1 second, 1 minute, 5 minutes and so on. We then number the time intervals sequentially and use the same process described above for discrete food units to select a random sample.

subconscious biases do not play a role in the selection of sample units, the random selection should be performed using a random number table or random number generator, as illustrated in Example 5.

One alternative to random sampling is *systematic sampling*. In this method, sample units are collected at fixed, equal intervals throughout a lot, where the interval is defined by time or number of units. This type of sampling is usually applied during processing or production to ensure coverage of the whole lot and is illustrated in Example 6.

However, there is a potential downside to the systematic approach. If there is an underlying periodic phenomenon in the process, then this could give us a distorted view of the process if the periodicity of the process and the sampling interval are related (or multiples of each other). For example, what would happen if the process described in Example 6 involved a 10-head rotary filler? If we were to take every 50th unit from the production line, then these would always come from the same filler head. Hence we would not obtain a sample that is representative of the process, but is only representative of that particular filler head.

Another alternative to random sampling is *stratified random sampling*. This takes into account potential additional sources of variation in the process, e.g. raw materials, multiple filling lines/heads, shifts, time intervals, etc. These sources of variation are usually referred to as strata, and the aim of stratified sampling is to obtain a representative sample throughout the lot or production process, taking into account the various strata.

Example 6
Systematic Sampling

Consider a lot with 1 000 units from which a sample of 20 units is to be selected systematically. This could occur during the production run, as one unit after another comes off the production line, e.g. prior to packing, or after packing provided the order of food units is identifiable.

Sample units need to be taken every 1 000/20 = 50 units (the interval size) to cover the whole lot. The first sample unit is selected randomly from the first interval, i.e. the first 50 units. After that every 50th unit is sampled.

So, if a randomly generated number between 1 and 50 (the first interval) is 26, for example, then the first food unit sampled is unit 26. The second sample unit is the 26+50 = 76th unit, the third is the 76+50 = 126th unit, and so on.

A video showing how to generate a systematic sample can be found at http://youtu.be/6VudQ3g9oyw.

Example 7
Stratified Random Sampling

Consider a milk powder process, where a lot is produced over an eight-hour shift. If we were to choose a random sample, then we might end up with all sample units having to be collected in the first two or three hours. However, we would like to ensure that we obtain coverage of the whole 8-hour shift. So we could divide the production into eight 1-hour intervals (or strata). Assuming that a sample of n = 24 sample units is to be selected we would need to select 24/8 = 3 units from each stratum. Within each stratum, the three sample units could then be selected randomly. This could be done by generating three different random numbers between 1 and 60 (representing all the minutes in one hour) for each stratum. Alternatively, we could divide the shift into 24 strata of 20 minutes each and randomly, or systematically, select a single sample unit from each stratum.

A video showing how to generate a stratified random sample can be found at http://youtu.be/EE8-rwLGyI0.

Note that strata do not have to be equal in size, in which case sample units are allocated proportionally to the size of the stratum. For example, consider the production process from Example 7. Assume that two sources of milk are used, one after the other, and that source 1 results in 75% of powdered milk and source 2 in the remaining 25%. To represent each source proportionally in the sample, we would collect 75% of sample units from the period covered by source 1 and the remaining 25% of sample units from the period covered by source 2.

The difference in the three types of sampling plans is depicted in Figure 5 using the context of the milk powder production used in Example 7. For random sampling, some one hour periods require more than 3 sample units while the last interval has no sample unit collection allocated to it. For systematic sampling a random starting time is selected in the first 20 minutes and then a sample unit is collected every 20 minutes thereafter (3 per hour). For stratified random sampling, the three sample units required within each one-hour stratum are selected randomly.

A key point to remember is that all three sampling approaches – random, systematic and stratified random – all have a random element to them which underpins and supports any data analysis that is subsequently performed. In addition, all three have their advantages and disadvantages and these need to be assessed carefully in conjunction with any additional information that is available.

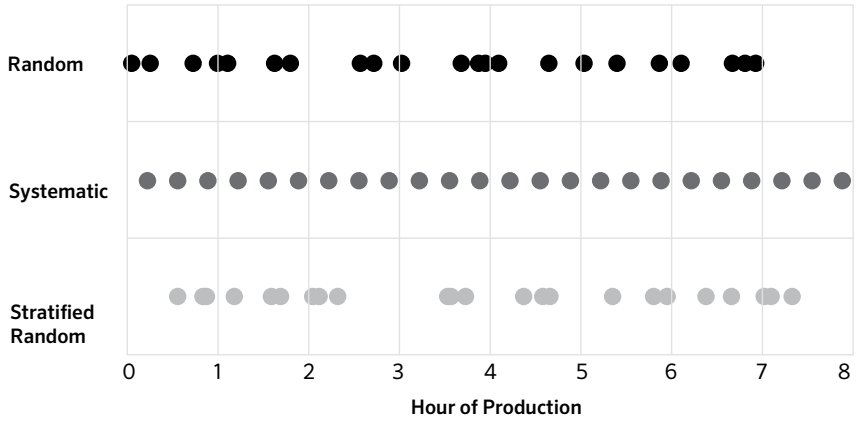


FIGURE 5: Illustration of random, systematic and stratified random sampling for an 8-hour (480 minute) production process of milk powder. For each sampling scheme a point indicates a time at which one of the 24 sample units is to be collected. This example was adapted from Jongenburger (2012).

1.3 WHAT ARE THE KEY TYPES OF SAMPLING PLANS?

There are many different sampling plans that have been developed over time, often to take account of the specific circumstances in which a particular plan is applied. The most common types of sampling plans used in a food context are two-class attributes plans, three-class attributes plans and variables plans. While the origins of these sampling plans are in manufacturing of parts and equipment, such as automotive parts and computer components, they can also be used when assessing the microbial aspects of food products.

As indicated above, the aim of all these sampling plans is to make decisions about the acceptability of the lot or production process. A brief look at these sampling plans and their characteristics is provided below, and we provide more detailed information on the statistical aspects in Part 2 (see "2.8 What are the important types of sampling plans?"). The sampling plans are presented in order of the amount of information that is required and level of complexity of their application.

1.3.1 Attributes sampling plans

Attributes plans are used when each sampling unit can be categorized according to some type of attribute or characteristic of interest. In the simplest case there are only two categories, such as acceptable/not acceptable (e.g. when dealing with the levels of a hygiene indicator) or present/absent (e.g. when dealing with a pathogen), which gives rise to *two-class attributes sampling plans*. When there are three categories, such as acceptable/marginal/not acceptable, then we are dealing with *three-class attributes sampling plans*.

1.3.1.1 Attributes sampling plans for tests that detect the presence of at least one organism per sampling unit (two-class presence-absence sampling plans)

Presence-absence testing is based on detecting the organism of interest with no attempt to quantify the levels of contamination. This is frequently used when testing for pathogens and usually involves enrichment of the sample to improve the sensitivity of the method.

When the organism is detected, the analytical unit (and hence sample unit) is often described as 'positive' and when absent, it is labelled 'negative.' **However, it is important to remember that a 'negative' test result means "the organism of interest was *not detected* in the analytical unit by the method used", rather than "there is no contamination (in the lot)."**

Because each analytical unit is classified according to one of two possible outcomes – present or absent – a two-class attributes sampling plan is appropriate. This type of sampling plan is defined by

- the analytical unit amount (mass, volume, area) of each analytical unit (w),
- the sample size, i.e. number of sample units, to be collected (n), and
- the number of analytical units (c) that are allowed to contain the target organism while still considering the lot to be acceptable.[4]

For some pathogens, e.g. *Salmonella*, for the lot to be deemed acceptable and hence $c = 0$, often none of the sample units are allowed to contain the organism of interest. A sampling plan with $c = 0$ is also commonly referred to as a *zero tolerance sampling plan*, which however does not mean that there is no *Salmonella* in the lot. In fact, we prefer the term *zero acceptance number sampling plan*, which is also used by Codex (CAC 2004), to avoid this incorrect inference. In fact, sampling plans can be readily designed that allow for some sample units to be contaminated

[4] While presence-absence sampling plans are treated separately to concentration-based sampling plans, they are in fact very similar, i.e. the microbiological limit (m) is equivalent to 1 organism in the analytical unit amount, w.

($c > 0$), yet can be more stringent (i.e. it would reject more lots) than zero acceptance number sampling plans.

1.3.1.2 Attributes sampling plans for tests that measure the level of contamination in each sample (concentration-based attributes sampling plans)

Concentration-based sampling plans require the determination of the concentration of microorganisms in each analytical unit. These levels are then compared on a unit-by-unit basis to one or more quantitative limits and this process determines how the sample is categorized. When only one microbiological limit (m) is specified then analytical units are classified as acceptable or not acceptable and a *two-class concentration-based attributes sampling plan* is used. In contrast, a three-class attributes sampling plan[5] is used when two microbiological limits – a marginal limit (m) and an unacceptable limit (M) – are of interest. In this case each analytical unit can be categorized as acceptable, marginally acceptable, or unacceptable. It should be noted that if the microbiological limits are specified on the arithmetic scale, then they can be simply log-transformed without complicated adjustment. That is, the same percentage of analytical units would exceed M if assessed on the arithmetic scale as would exceed $\log_{10} M$ when using the $\log_{10}$ scale, as shown in Example 8 (see also "1.2.3 Why do we use $\log_{10}$ numbers and why do we need to be careful when interpreting them?").

Example 8
Transforming microbiological limits

Codex specifies the limit for *Listeria monocytogenes* in Ready-To-Eat foods in which growth of *L. monocytogenes* will not occur as $m = 100 = 10^2$ cfu/g (CAC, 2007). Equivalently, we could specify the limit on the $\log_{10}$ scale as $m = 2 \log_{10}$ cfu/g.

So, if 5% of analytical units have a concentration of *L. monocytogenes* that exceeds m on the arithmetic scale (100 cfu/g), then 5% of analytical units would also have a $\log_{10}$ concentration of *L. monocytogenes* that exceeds the limit on the $\log_{10}$ scale, i.e. $2 \log_{10}$ cfu/g.

[5] We leave out the "concentration-based" part from the names as these sampling plans only occur when concentrations are measured.

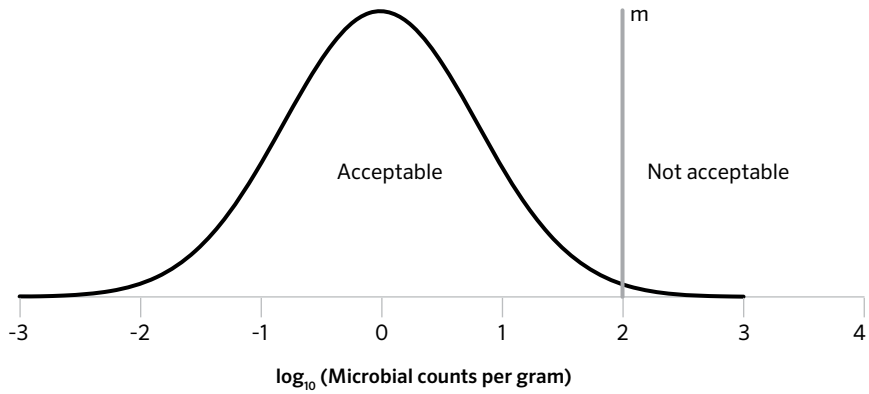


FIGURE 6: An example of a distribution of the $\log_{10}$ concentration of a microorganism in a food lot with the microbiological limit $m = 2$ $\log_{10}$ cfu/g. Analytical units with a $\log_{10}$ concentration less than or equal to $m = 2$ are acceptable and those with a $\log_{10}$ concentration greater than $m = 2$ are unacceptable.

1.3.1.2.1 Two-Class Concentration-Based Plan

If only one microbiological limit (m) is specified, the analytical unit can take only one of two possible outcomes; the concentration of the organism is either less than or equal to the limit (analytical unit is acceptable), or the concentration is greater than the limit *m* (analytical unit is unacceptable). One possible example of such a distribution is graphically depicted in Figure 6. Therefore, the sampling plan is also a two-class attributes plan or more specifically a two-class concentration-based attributes plan to highlight that a concentration based microbiological test, i.e. enumeration test, is used.

This type of sampling plan is defined by

- the analytical unit amount (mass, volume, area) of each analytical unit (*w*),
- the sample size, i.e. number of sample units, to be collected (*n*),
- the microbiological limit (*m*) that determines whether an analytical unit is acceptable or not acceptable, and
- the number of analytical units (*c*) that are allowed to exceed the limit *m* while still considering the lot to be acceptable.

1.3.1.2.2 Three-Class (Concentration-Based) Plan

In some cases, a sampling plan will be developed that uses two quantitative limits, thereby creating three possible categories for each analytical unit, and such a plan is called a three-class sampling plan. Figure 7 is a graphical presentation of the

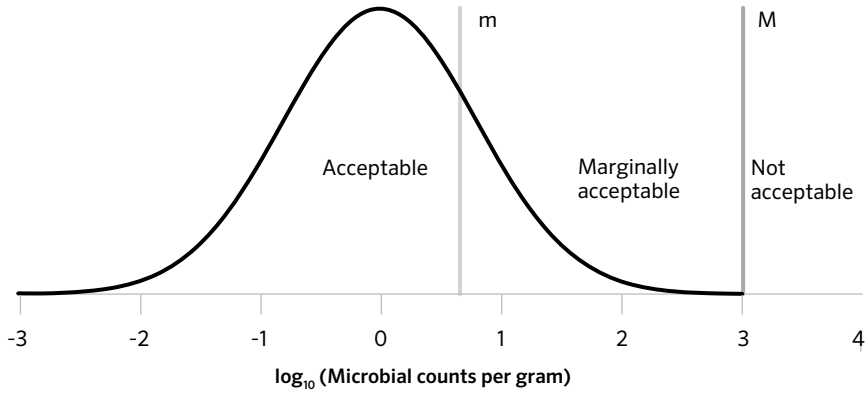


FIGURE 7: An example of a distribution of the $\log_{10}$ concentration of a microorganism in a food lot with the microbiological limits $m = 1.5$ $\log_{10}$ cfu/g and $M = 3$ $\log_{10}$ cfu/g. Analytical units with a $\log_{10}$ concentration less than or equal to m are acceptable and those with a $\log_{10}$ concentration greater than M are unacceptable, while those that have a $\log_{10}$ concentration greater than m and less than or equal to M are marginally acceptable.

three possible outcomes (for one possible underlying distribution). An analytical unit is deemed

- *acceptable* if the ($\log_{10}$) concentration is less than or equal to the marginal limit (m),
- *marginally acceptable* if the ($\log_{10}$) concentration is greater than the marginal limit (m) but less than or equal to the unacceptable limit (M), and
- *unacceptable* if the ($\log_{10}$) concentration exceeds the unacceptable limit (M).

This type of sampling plan is defined by

- the analytical unit amount (mass, volume or area) of each analytical unit (w),
- the sample size, i.e. number of sample units, to be collected (n),
- the marginal and unacceptable microbiological limits (m and M) that determine whether an analytical unit is acceptable, marginal or unacceptable, and
- the number of analytical units (c) that are allowed to exceed the limit m, but not M, while still considering the lot to be acceptable.

Note that it is possible to create sampling plans that allow some analytical units to exceed M without rejecting the lot, but these are not commonly used in relation to MC in foods.

Lots can be rejected because more than *c* analytical units exceed *m* (but are less than M) or because at least one analytical unit exceeds *M*. This distinction is not important for the application of the sampling plans as each lot will either be accepted or rejected – but this information may be useful in terms of trend analysis (discussed later, in Part 3) and when deciding whether there is a need to find the root cause for any unacceptable analytical units.

1.3.2 Variables Sampling Plans

Variables sampling plans provide a natural extension to the concentration-based attributes sampling plans. In attributes sampling plans, the specific level of the microorganism is used only to determine into which category each analytical unit is assigned and the actual concentrations are not further considered. This results in a loss of information (Example 9) that the variables plan tries to overcome.

This type of sampling plan is defined by

- the analytical unit amount (mass, volume, area) of each analytical unit (*w*),
- the sample size, i.e. number of sample units, to be collected (*n*),
- the microbiological limit (*m*) that determines whether an analytical unit is acceptable or not acceptable, and
- a critical value *k* that is calculated from the sample size (*n*) and a *Consumer's Risk Point*, that is, a combination of a tolerable probability of acceptance and percentage of concentrations exceeding the limit *m* (see later "2.8.5 What are the Producer's and Consumer's Risk Points?").

In variables sampling plans, the actual concentrations of the analytical units are used to generate summary statistics, which describe all of the sample results. These

Example 9
Classifying analytical units under two-class attributes sampling plans

Assume that we are sampling a food product and that the microbiological limit (*m*) is 5 000 cfu/g. Under a two-class attributes plan we would treat an analytical unit with 5 010 cfu/g and one with 100 000 cfu/g the same – both are unacceptable.

Similarly, an analytical unit with 10 cfu/g and one with 4 990 cfu/g would both be considered acceptable.

However, notice that the two analytical units that are closest in concentration – 4 990 and 5 010 cfu/g – are treated differently. The small difference in counts could possibly be a result of the uncertainty in the microbiological method.

statistics (sample mean and SD) are then compared to a pre-determined limit to decide whether the lot is acceptable. This way, the actual concentrations have a direct influence on whether a lot is acceptable or not.

A key point to remember is that the variables sampling plan makes better use of the available information than a two- or three-class attributes sampling plan and hence requires a smaller sample size to a comparable attributes sampling plan.

1.3.3 What factors should we consider when choosing a sampling plan?

There are many factors that affect the selection of a sampling plan (and hence MC) and these factors may differ from one application to another. For example, the ICMSF developed 15 sampling plan cases in relation to microbial hazards (ICMSF, 2002). However, there are no definitive rules that dictate what type of plan should be used, but some of the factors that need to be considered include the following:

- **Concentration of the target organism:** If the concentration of the target organism is expected to be very low then a presence-absence test, using enrichment and a large analytical unit amount, may be preferable, e.g. for a highly infective microorganism we are concerned about a low concentration. However, a suitable enumeration test method with a low limit of detection may also be suitable.
- **Microbiological test capability:** The availability of a suitable (validated) microbiological test that can enumerate the target organism is a requirement for any concentration-based plan.
- **Ability to discriminate:** Given the same sample size (n), two-class attributes sampling plans are less able to discriminate between acceptable and unacceptable lots than a three-class or variables sampling plan.
- **Additional information:** Is the information generated from microbiological testing to be utilized for trend analysis or statistical process control purposes? If so, then information about the actual microbial concentrations is preferable to a simple indication of whether the organism is detected.
- **Cost of sample collection and microbiological testing:** If the cost of sample collection or microbiological testing is high, then it may be preferable to utilize a plan where fewer sample units are needed to achieve a desired performance, e.g. three-class or variables sampling plans.

In addition, the following information may assist you to choose an appropriate plan in a particular situation.

1.3.3.1 Two-Class Presence-Absence Sampling Plans

This type of plan is used when the level of the target organism is so low that a large proportion of individual sample or analytical units will not contain the organism. Pathogens that are more likely to cause illness at low levels (e.g. *E. coli* O157:H7) or for which there is a very low level of tolerance (e.g. *Salmonella*) are often addressed using this type of plan.

1.3.3.2 Two-Class Concentration-Based Sampling Plans

This type of plan is often used when low levels of contamination of the target organism are acceptable. This plan may be applied to hygiene indicator organisms that occur at higher numbers or pathogens that are unlikely to cause illness at low levels.

1.3.3.3 Three-Class Sampling Plans

This type of plan is also used when low levels of contamination with the target organism are acceptable. This may be applied to hygiene indicator organisms, and pathogens that are unlikely to cause illness at low levels. In contrast to the two-class concentration-based plan, this type of plan is employed where a clear upper limit exists that defines unacceptable concentrations that should not be exceeded.

1.3.3.4 Variables Sampling Plans

This type of plan is also used when low levels of contamination are tolerable and the target organism is expected to be present most of the time (so it can be enumerated). This may be applied to hygiene indicator organisms and pathogens that are less likely to cause illness at low levels. In particular, the additional information that is derived by using the actual concentrations, i.e. determination of the mean and SD, allows this plan to yield similar discrimination between acceptable and unacceptable lots with smaller sample size (n) than an attributes plan. As such it may be preferable when sample collection is difficult or costly or when the microbiological testing is expensive. In addition, trends can be more readily assessed when the mean and SD of the microbial contamination are available.

A key point to remember is that there are a range of factors that will affect the decision as to which sampling plan is most appropriate in any particular circumstance. These factors need to be considered together to ensure that the sampling plan will be able to meet your requirements.

PART 2

Making decisions about an individual lot

In this part we cover the aspects of MC as they relate to acceptance sampling, that is, to determine whether to accept or reject a lot. We cover in more detail the various acceptance sampling plans introduced in the previous part. In particular, we show how aspects such as sample size, analytical unit amount, and variability in microorganisms affect the likelihood of accepting a lot.

2.1 WHAT IS A LOT?

In the broadest sense, a lot is a predefined quantity of food product. Commonly this is achieved by defining the lot based on production timeframes, production conditions, raw materials, application of cleaning regimes or even geospatial information, such as a field or a paddock. It is assumed that each lot is produced under similar conditions and hence that the units in the lot have experienced similar conditions. However, it is important that the lot is defined *before* a sampling plan is applied to the lot.

In addition, it is also assumed that each individual lot is independent of any other lot, which allows us to make decisions about the lot without having to be concerned about the effect on other lots (see "2.4 What makes lots independent?" for more details).

We have provided some examples of how a lot might and might not be defined in Table 1. However, these examples should not be taken as absolute, but rather as

informative and guiding in nature – each situation will be unique and you need to determine whether a particular lot definition is appropriate.

TABLE 1: Examples of what is and what is not a lot

What is a lot	What is not a lot
All foods formulated with the same ingredients and raw materials (e.g. a batch of food)	When foods are formulated with various ingredients and raw materials
All foods produced between two cleaning breaks	Foods produced in different cleaning periods, e.g. during different days
For continuous production processes, all foods produced over a predefined time frame	All foods produced in different timeframes
All foods produced on the same production line.	Any food produced in different production lines
For infant formula, all foods produced on the same line, holding tank, manufacturing conditions without a cleaning break	For infant formula, foods produced with different holding tanks, lines, manufacturing conditions or cleaning breaks
In the case of fresh produce, a field or part of a field.	Fresh produce from different geographical locations

2.2 CAN WE REDEFINE THE LOT AFTER DETECTING A PROBLEM?

Sometimes a problem is found with a lot of food as a result of sampling and the lot cannot be accepted. It may be tempting to redefine the lot by breaking it into sub-lots and retesting each of these, especially when the lot is large. The aim in this case is to identify one or more contaminated sub-lots (and simultaneously one or more uncontaminated sub-lots) and thereby reduce the amount of product that may be subject to control action, e.g. withdrawn from commerce.

However, irrespective of the type of sampling plan that is used, it is not appropriate to repeatedly test a lot (CAC, 2013a; ICMSF 2002), nor to redefine a lot in this way, i.e. after a lot has been defined, sampled and subsequently rejected. This is because of the uncertainty in the sampling outcome. Sampling can never guarantee identification of the problem in one or more of the sub-lots, which could only be achieved by testing all food product in each sub-lot. To illustrate this point, consider a lot where the prevalence of a pathogen is 1%. As we will see later, a 2-class attributes

sampling plan with $n = 15$ and $c = 0$ has a probability of only 14% of detecting the pathogen in the lot. Consequently, redefining the lot into sub-lots and retesting each of these will generally not change the probability because each lot/sub-lot is very large compared to the total amount sampled. Hence the pathogen is unlikely to be detected again in any of the sub-lots, despite the knowledge that the pathogen was present in one, or all, of the sub-lots.

Consequently, the decision as to how big a lot should be needs to be considered in an economic or public health context and prior to the application of a sampling plan. That is, the cost of testing and the cost of rejecting lots should be evaluated prior to sampling.

Nevertheless, it is possible to re-sample and re-test a lot for investigational purpose, i.e. to better understand the extent and cause of the microbial contamination. However, this can only provide better information about the contamination and cannot provide evidence for releasing part(s) of the original lot into the market place.

2.3 CAN MICROBIOLOGICAL TESTING BE USED TO DEFINE A LOT (E.G. FOR CONTINUOUS PRODUCTION)?

No. As indicated in "2.2 Can we redefine the lot after detecting a problem?" the lot is defined prior to sampling, based on production knowledge, and not as a result of sampling. Even in continuous production you need to determine in advance how much product constitutes a lot, e.g. 1 hour of production or maybe all product produced during a shift (see also Table 1 for some examples). Once the extent of the lot has been established, appropriate sampling and testing can then be applied.

2.4 WHAT MAKES LOTS INDEPENDENT?

An important aspect related to MC is the independence of lots, which is often not well understood. Similar to the fact that lots are not defined as a result of sampling, independence is not achieved as a result of sampling. Instead whether two lots are independent is assessed by knowing how those two lots were produced.

Independence has a particular statistical definition in terms of probabilities. However, it basically means that two lots are not related in time and space and hence knowing that one lot is contaminated does not change the likelihood of the other lot being contaminated. This is illustrated in Example 10.

Example 10
Lot independence

Consider a production process where a lot is normally defined as 'the food produced between two cleaning breaks.' Also assume that a cleaning break results in complete sanitization of the processing environment.

Now, imagine that during one production period, there is a breakdown in processing hygiene and hence the affected lot fails to be accepted when it is sampled. Because of the effectiveness of the cleaning as specified above, this should not have any bearing on lots processed in any other production period, including the one following the loss of process control.

In contrast, a different company producing a similar product might define lots differently – not on production period with cleaning breaks, but based on other quality characteristics of the product. By combining food units from various production periods into lots, they believe that they can better meet customer specifications. However, such lots are not independent! A breakdown during one production period will potentially affect all lots that contain product from this period. **No amount of testing can overcome this problem and make these related lots independent.**

2.5 CAN WE DEFINE A LOT GEOGRAPHICALLY?

Yes. For example, in field harvest situations people have defined lots as 'all product produced in a field or part of a field' (Table 1). In such a [field harvest] situation it may be important to also consider post-harvest processing. For example, parts of a field that are processed (cleaned and packed) in a single shift or day, may also be defined as a lot. However, as for other production processes the lot needs to be defined in advance and be independent of other lots.

2.6 WHAT IS MEANT BY BETWEEN-LOT TESTING AND LOT-BY-LOT TESTING?

Lot-by-lot testing means that *every* lot is tested using a specified sampling plan for the purpose of acceptance/rejection decision-making. It may apply to every lot produced or every lot sent to, or received by, a customer under a commercial supply agreement. In industry terms this testing of every lot is commonly referred to as "test and hold", which tends to be more common when testing for foodborne pathogens.

"Between-lot testing" is a colloquial term which denotes any form of testing multiple lots, either for acceptance/rejection or process verification. It is not a

standard term and lacks precise definition, though people often use it incorrectly to denote acceptance sampling or lot-by-lot testing.

Every lot may also be tested as part of process verification, but not using a sampling plan designed for lot acceptance/rejection. This is not *lot-by-lot testing*, but is referred to as process verification testing (see also "3.2.2 What is the difference between lot-by-lot testing and verification testing?").

It is also worthwhile to note that sampling in a microbiological context is no different to sampling for other quality or safety characteristics of food, or of non-food products. Codex has developed guidelines on sampling (CAC, 2004), which includes references to various related International Organization for Standardization (ISO) standards. These international standards include information on various "sampling systems" that allow the lot-by-lot sampling requirement to be relaxed when a history of good compliance has been demonstrated, e.g. skip-lot sampling. However, such systems are outside the scope of this document and interested readers are referred to Montgomery (2012), for example. Nevertheless, the more generic term *acceptance sampling* (and testing) encompasses these alternatives systems, which are also based on same underlying statistical methodology.

2.7 WHAT IS THE PURPOSE OF LOT-BY-LOT TESTING AND WHO DOES THIS?

The purpose of acceptance sampling testing is to determine the acceptability (quality and/or safety) of a lot in relation to predetermined acceptance criteria, e.g. as specified in an MC. In the traditional sense of acceptance sampling, this approach often emphasizes protection of the producer. That is, lots with acceptable quality, or better, should be accepted most of the time and rejected only infrequently, i.e. such lots should have a high probability of acceptance and hence a low *producer's risk*. However, in a microbiological context, the focus is generally on food lots with unacceptable quality or safety having a low probability of acceptance. Such lots should be rejected most of the time and only accepted infrequently, i.e. to ensure low *consumer's risk* (see also "2.8.5 What are the Producer's and Consumer's Risk Points?").

Acceptance sampling is primarily undertaken by FBOs for high-risk foods, but may also be utilized to verify compliance with hygiene limits, for example, under commercial supply agreements.

While regulatory agencies may also use MC to test for compliance with food safety standards, they generally do not test each lot, but instead sample subset of lots. This

is sometimes known as "isolated lot testing" (CAC 2004) where no history of good compliance and controlled production exists, e.g. at port-of-entry or for export certification.

The calculations of the probability of acceptance are identical for both situations, although traditionally, in acceptance sampling, the sample size (n) calculations are based on either protecting the producer or protecting the consumer. However, no such distinction (between lot-by-lot and isolated lot testing) needs to be made in a microbiological context of pathogens in food because the focus should foremost be on protecting the consumer or customer, while also considering the implications for the producer. Hence, lot-by-lot testing and isolated lot testing are not discussed separately here.

It is noteworthy that there are various reasons for requiring a testing scheme to be implemented, including the following:

- **Meeting supply agreement requirements:** Sampling and testing are useful to demonstrate an agreed level of microbiological control.
- **Demonstrating due diligence:** FBOs are, within reason, expected to do everything under their control to assure food safety. A history of good compliance and suitable corrective actions when problems are identified provides significant evidence.[6]
- **Providing incentive for process control and improvement:** No FBO wants to recall their product because of noncompliance with a food safety standard. Sampling and testing are important tools for customers to encourage suppliers to improve process control.
- **Cost:** The cost of non-compliance and product recalls can far outweigh the costs of demonstrating compliance through food safety programmes and associated product testing.

2.8 WHAT ARE THE IMPORTANT TYPES OF SAMPLING PLANS?

In Part 1 we briefly introduced the four important types of sampling plans used in the food industry: two-class presence-absence sampling plans; two-class concentration-based sampling plans; three-class concentration-based sampling plans and variables sampling plans. Before we look at these sampling plans in more detail, we first discuss the importance of probability in acceptance sampling. We also explain the meaning of sampling plan performance and how performance can be visualized using the OC curve.

[6] Utility depends on the legal situation / interpretation used in a country.

2.8.1 Why do we need to worry about the probability of accepting and rejecting lots?

We have already seen that sampling can never provide a guarantee of safety. Instead, for a particular level of microbial contamination, the outcome from sampling may be that the lot is either accepted or rejected. Due to the uneven distribution of microorganisms in foods and their particulate nature, if we were to sample two lots with the same microbial contamination using the same sampling plan, then it is possible that one lot might be accepted while the other might be rejected.

However, irrespective of the type of sampling plan used, lots with low levels of microbial contamination will be accepted most of the time. Such lots have a high probability of acceptance, P(accept). Similarly, highly contaminated lots will be rejected most of the time and hence they have a small P(accept). Alternatively, we can say that highly contaminated lots are rejected most of the time and therefore they have a high probability of rejection, P(reject).

Because there are only two possible sampling outcomes – we either accept or reject the lot – the two corresponding probabilities always add up to 100% (Example 11) and we can write

$$\mathbf{P(accept)+P(reject)=100\%}$$

Once the type of sampling plan has been selected the probability of acceptance can be calculated for a range of levels of microbial contamination using the appropriate

Example 11
Interpretation of the probability of accepting/rejecting a lot

A zero acceptance number sampling plan ($c = 0$) with $n = 15$ results in a P(accept) of 73.86% when 2% of the food units in a lot are contaminated with a pathogen (see "2.8.6 Two-class presence-absence sampling plans").

This means that 73.86% of all lots with 2% contaminated units would be accepted in the long run. Consequently, the remaining 100 - 73.86 = 26.14% of lots with 2% contamination would be rejected.

Alternatively, it can be said that if we were sampling an isolated lot (with 2% of contaminated units) using the same sampling plan, then we would have a 73.86% chance of failing to detect the contamination (accept the lot) and a 26.14% chance of detecting the contamination (and rejecting the lot).

statistical distributions. A common tool that is used to better understand sampling plans and their associated probabilities of acceptance is the OC curve, which we discuss next.

2.8.2 What is meant by the performance of a sampling plan?

It is important to know the performance of a sampling plan to understand its effect on the control of the process or the safety or quality of the product. The performance of the sampling plan is determined by the probability of accepting or rejecting a lot as a function of the level of microbial contamination. The performance can be visualized and assessed using OC curves (see "2.8.3 What is the Operating Characteristic Curve?").

The performance of a sampling plan is affected by the parameters of the sampling plan, namely analytical unit amount, sample size, acceptance number, and microbiological limit(s). Changing the value of any one of these parameters implies changing the sampling plan.

The performances of the different types of sampling plans are described below and the mathematical details for performing the calculations can be found in Annex 1: Mathematical ."

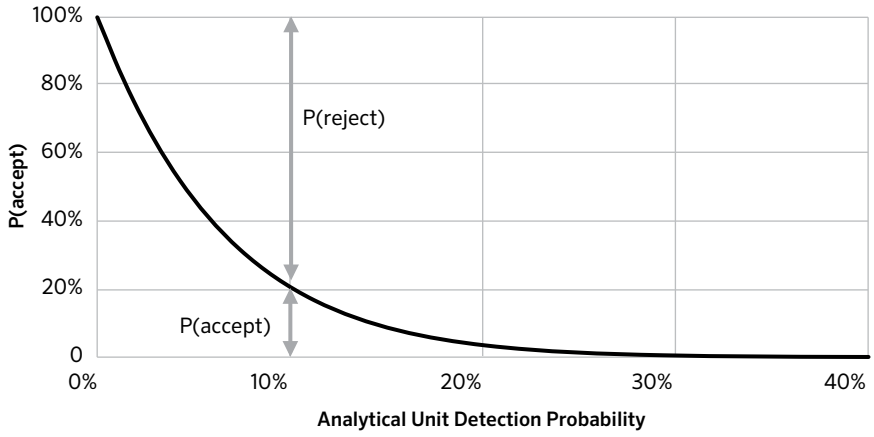


FIGURE 8: Example OC curve for a two-class presence–absence sampling plan with $n = 15$ and $c = 0$. The plot shows how the probability of acceptance reduces as the prevalence of the contamination in the lot increases.

2.8.3 What is the Operating Characteristic Curve?

The *Operating Characteristic (OC) Curve* is a visual representation of how the probability of acceptance reduces with increasing microbial contamination as measured by analytical unit detection probability or average concentration of a microorganism. An example of an OC curve is provided in Figure 8.

While in practice we are often more interested in the probability of rejection, i.e. the probability of detecting unacceptable levels of contamination in a lot, it is standard practice to plot the OC curve with P(accept) on the Y-axis. However, you may prefer to plot P(reject) on the Y-axis for particular applications; this is valid, provided you make it clear what you are plotting.

2.8.3.1 Should we use the $\log_{10}$ geometric mean or the arithmetic mean for the X-axis of the OC curve?

Due to the common use of the $\log_{10}$ transformation in microbiology it is also common to use the mean $\log_{10}$, or equivalently the $\log_{10}$ geometric mean, as the X-axis of the OC curve, see for example ICMSF (2002). However, this practice can be misleading as the variability in microbial counts is not taken into account (see "1.2.3 Why do we use $\log_{10}$ numbers and why do we need to be careful when interpreting them?") and therefore we believe that **the use of the $\log_{10}$ geometric mean as the X-axis should be avoided**. To illustrate this point, consider the following two-class concentration-based sampling plans for three food products with different standard deviations:

- Plan 1 is applied to a food product with SD = 0.3 $\log_{10}$ cfu/g and consists of $n = 5$, $c = 0$, $m = 1.5$ $\log_{10}$ cfu/g and;
- Plan 2 is applied to a food product with SD = 0.6 $\log_{10}$ cfu/g and consists of $n = 20$, $c = 5$, $m = 1.5$ $\log_{10}$ cfu/g; and
- Plan 3 is applied to a food product with SD = 0.9 $\log_{10}$ cfu/g and consists of $n = 40$, $c = 13$, $m = 1.5$ $\log_{10}$ cfu/g.

The OC curves for these three sampling plans are shown below using the $\log_{10}$ geometric mean as the X-axis (Figure 9a) and arithmetic mean as the X-axis (Figure 9b). From these OC curves it is evident that if we were to use the $\log_{10}$ geometric mean as the X-axis we would think that the three sampling plans applied to the different food products are approximately equivalent as the OC curves are almost identical (Figure 9a). However, because of the different SDs of the three food products, and the effect the SD has on the arithmetic mean ("1.2.3 Why do we use $\log_{10}$ numbers and why do we need to be careful when interpreting them?"), the OC curves look quite different when viewed using the arithmetic mean as the X-axis (Figure 9b) as the OC curve is shifted further to the right the larger the SD. These OC curves demonstrate that the overall level of contamination, as best rep-

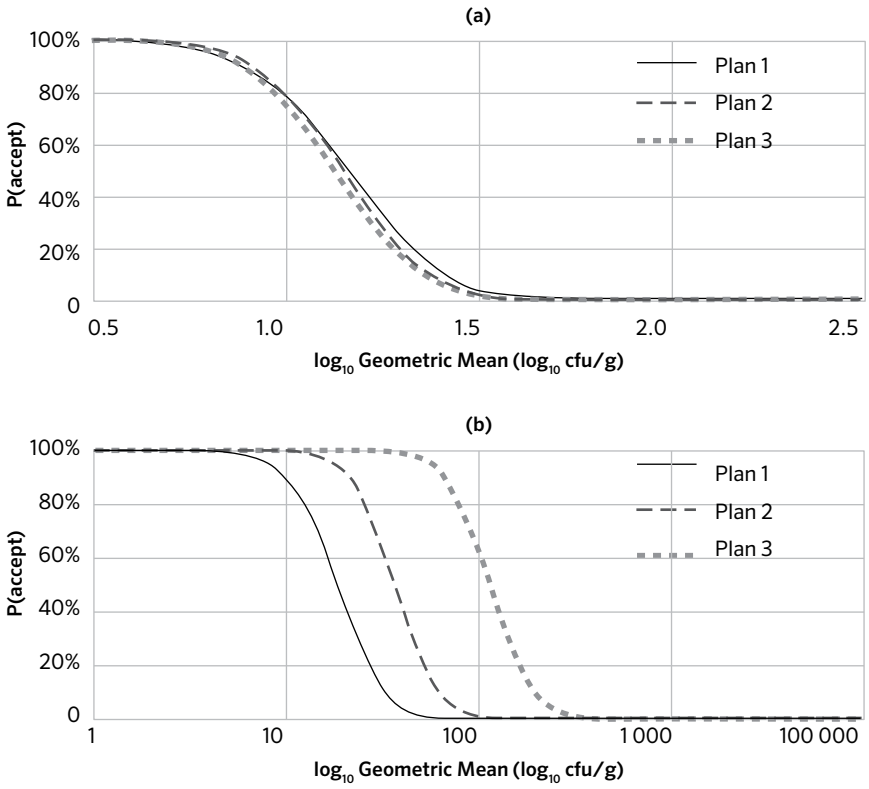


FIGURE 9: Two-class concentration-based sampling plan OC curves with n = 5, c = 0, m = 1.5 $\log_{10}$ cfu/g (Plan 1); n = 20, c = 5, m = 1.5 $\log_{10}$ cfu/g (Plan 2); and n = 40, c = 13, m = 1.5 $\log_{10}$ cfu/g (Plan 3) applied to foods with SDs equal to 0.3, 0.6 and 0.9 $\log_{10}$ cfu/g, respectively

resented by the arithmetic mean concentration – which is directly proportional to the total number of organisms – is quite different at any probability of acceptance between 1 and 99%. The use of the $\log_{10}$ geometric mean on the X-axis suppresses this very different level of control and hence its use can contribute to misinformed decision-making.

In addition, the typical use of an OC curve is to derive a single value to characterize the sampling plan, often referred to as the *level of control*. Based on the three sampling plans and products, the level of control associated with a probability of acceptance of 5% is approximately 1.44 $\log_{10}$ cfu/g if we are using the $\log_{10}$ geometric mean (or mean $\log_{10}$) concentration. However, the arithmetic means

associated with a 5% probability of acceptance are approximately 37, 70 and 227 cfu/g when taking into account the SDs of 0.3, 0.6 and 0.9 $\log_{10}$ cfu/g, respectively, which indicates an up to 6-fold difference in the level of control achieved for these three food products.

2.8.4 What is meant by discrimination and stringency of sampling plans?

Ideally, we would be able to differentiate perfectly between lots that are considered 'acceptable' and those that are considered 'not acceptable', i.e. acceptable lots are accepted 100% of the time and unacceptable lots are accepted 0% of the time. An example OC curve for such an idealized situation is shown in Figure 10 (solid line) where the 'cut-off' prevalence[7] between 'acceptable' and 'unacceptable' lots is 2%. However, in practice no sampling plan will result in such an OC curve and instead the transition between 100% and 0% probability of acceptance is more gradual, as shown by the dashed line in Figure 10. Consequently, the steeper the OC curve (more similar to the idealized situation) the more *discriminating* we say a sampling plan is, implying that we can discriminate (or differentiate) between lots that are considered acceptable (prevalence is less than 2%) and lots that are considered not acceptable (prevalence is 2% or more).

Above we considered a prevalence of 2% to be the limit between acceptable and unacceptable lots. If we wanted to be more *stringent*, then we might consider a smaller prevalence, e.g. 1%, while a larger prevalence, e.g. 3%, would be less stringent.

It should be noted that this idea of "idealized" also has its origins in the traditional theory of acceptance sampling. However, in a microbiological setting, this idea needs to be interpreted in the context of the particular situation and target organisms and hence what "ideal" means in that situation. For example, if risk increases steadily and continuously with increasing prevalence or mean concentration, then a more gradual change, commensurate with risk, may be more ideal in a practical sense. In addition, it is rare to have a situation where a clear, single limit exists between acceptable and unacceptable lots – in practice the transition is more gradual.

2.8.5 What are the Producer's and Consumer's Risk Points?

As we have seen in the above, sampling plans provide a gradual transition between accepting lots of acceptable quality or safety and rejecting lots with unacceptable quality or safety. At any particular level of microbial contamination there is a chance of accepting lots and a chance of rejecting lots – the size of these probabilities depends on the MC (or sampling plan) that has been chosen.

[7] This could equally refer to a proportion of analytical units exceeding a microbiological limit, *m*.

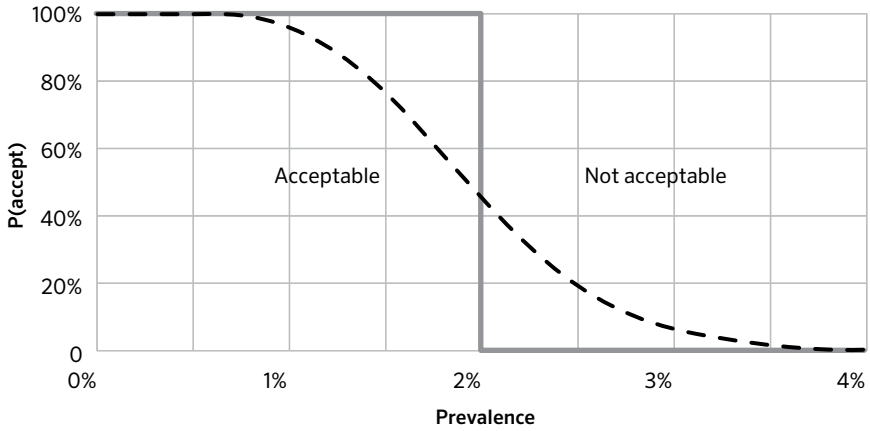


FIGURE 10: Idealized and actual OC curves for a food product that is assumed to be unacceptable when the prevalence of contamination is greater than 2%. The idealized OC curve is represented by the line and the actual OC is represented by the dashed line.

It should be noted that the producer's and consumer's risks are traditional terms defined in the acceptance sampling literature, including the Codex Guidelines on Sampling (CAC, 2004). However, these terms do not refer strictly to risks, but probabilities (as pointed out above), as they do not take into account the resulting severity.

Lots with acceptable levels of microbial contamination, or better, should be accepted most of the time and rejected only infrequently. Such lots should have large probability of acceptance and hence a small probability of rejection, i.e. the *producer's risk*. Consequently, the *Producer's Risk Point* is chosen to have a large probability of accepting lots at an acceptable level of contamination (or better), as shown in Figure 11. Throughout this document, we use P_0(accept) to denote the probability of acceptance and either p_0 or μ_0 to denote the analytical unit detection probability or mean concentration, respectively, when referring to the Producer's Risk Point.

Similarly, lots with unacceptable levels of microbial contamination, or worse, should be rejected most of the time and accepted infrequently, i.e. the *consumer's risk*. Consequently, the *Consumer's Risk Point* is chosen to have a small probability of accepting lots at an unacceptable level of contamination (or worse), as shown in Figure 11. Throughout this document, we use P_1(accept) to denote

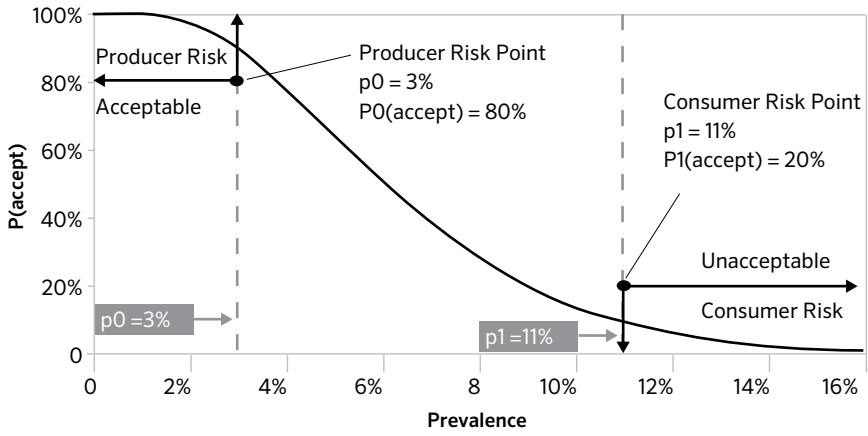


FIGURE 11: OC curve showing the Producer's and Consumer's Risk Points. The Producer's Risk Point is chosen to have a large probability of accepting lots at an 'acceptable' level of contamination (or better), while the Consumer's Risk Point is chosen to have a small probability of accepting lots at an 'unacceptable' level of contamination (or worse).

the probability of acceptance and either p_1 or μ_1 to denote the analytical unit detection probability or mean concentration, respectively, when referring to the Consumer's Risk Point.

From a statistical perspective, a suitable sampling plan will have an OC curve that lies above the producer's risk point and below the consumer's risk point (depending on whether one or both are specified), and such a plan is found by changing the parameters until these criteria are met.

2.8.6 Two-class presence-absence sampling plans

This type of sampling plan is applicable when each test has only one of two possible outcomes, namely, the microorganism is either detected in the sample unit (presence) or it is not (absence). The two-class presence-absence sampling plan is usually used when you are interested in detecting the presence of a pathogen, possibly at very low concentrations, and hence the microbiological test usually involves an enrichment step to improve the sensitivity of the detection method. It is also possible to relate the detection of an organism to a concentration (under certain assumptions), as we will see later (see "2.8.6.1 How does the analytical unit amount (weight, volume or area) affect the probability of acceptance?").

This sampling plan is defined by

- the analytical unit amount (w), i.e. the amount of each analytical unit (mass, volume, area);
- the sample size (n), i.e. number of sample units that are collected;[8] and
- the acceptance number (c), i.e. the number of analytical units that are allowed to contain the target organism while still considering the lot to be acceptable.

For most pathogens, especially those that are highly pathogenic and require relatively few organisms to result in a high likelihood of illness, the acceptance number usually equals zero, i.e. $c = 0$. This is because it is generally not appropriate to detect them among only a few analytical units and still declare the lot as acceptable. Consequently, under this type of plan, acceptance is only possible when there are no organisms detected in any of the analytical units, while rejection will occur as long as there is at least one organism detected in any one of the analytical units, so we can write

P(accept) = P(no organisms detected in the analytical units)

P(reject) = P(at least 1 organism detected in the analytical units) = 100% - P(accept)

Now, provided that the following criteria and assumptions are met we can calculate the probability of acceptance.

1. **The sampling process is random** This is needed for probability calculations to be valid. This is primarily achieved by random sampling, although this assumption will be reasonable under systematic and stratified random sampling[9] (see "1.2.6 What is random sampling and what are the alternatives?").
2. **Sample units are independent of each other** This is something that we generally do not know, but can assume provided care was taken to ensure that the sampling was random.
3. **Each sample unit has the same probability of yielding a detection** (a 'positive' test result) This can generally be assumed provided there are no known strata that might affect the level of contamination.

The probability of acceptance can then be calculated using the binomial distribution – the mathematical details are provided in the Annex A1.3 Two-class presence-absence sampling plans" for those who are interested. However, you do not need to understand the mathematical details to use this sampling plan as the calculations

[8] For some sampling plans, e.g. ISO 2859, the sample size (n) is obtained from the size of the lot. However, this is done for non-statistical reason.

[9] Additional care needs to be taken when strata are of different size as this will affect the probability calculations.

and associated OC curve for this sampling plan can be found in the companion spreadsheet on the Two-class Presence-Absence (1) Tab. This spreadsheet can be used to determine whether a particular sampling plan under consideration will result in the desired lot acceptance probability, P(accept), as shown in Example 12.

It should be noted that not detecting the target organism does not imply that "there is no contamination (in the lot)" but simply that "the organism of interest was not detected by the microbiological method used in the analytical units that were sampled from the lot."

It is often assumed that the microbiological test is 'perfect' and does not result in false positives (specificity = 100%) or false negatives (sensitivity = 100%). That is, the test can correctly differentiate analytical units that contain at least one target microorganism and those that are truly without the target microbe. However, this is generally not true and in this case the Rogan-Gladen estimator can be used to adjust the analytical unit detection probability for a lack of sensitivity and/or specificity[10] (Rogan and Gladen 1978), although estimates of sensitivity and specificity are rarely available. The web-based sampling tool (http://www.fstools.org/sampling, in the advanced mode) offers the opportunity to consider different levels of sensitivity, where false negatives are possible, but does not consider issues of specificity (i.e.it is assumed to be 100%).

Let us now have a look at how the analytical unit amount, average rate of contamination, sample size and acceptance number affect the probability of acceptance when this sampling plan is used.

2.8.6.1 How does the analytical unit amount (weight, volume or area) affect the probability of acceptance?

In many circumstances there exists a standardized test for a microorganism that has been properly validated, e.g. ISO or Association of Analytical Communities (AOAC) International. These methods stipulate the analytical unit amount (*w*), i.e. how much of a food should be tested. The analytical unit amount could be specified as an actual weight, such as for sprout seeds, or alternatively as a volume, such as for milk or water, or even as a sample area, such as when beef carcasses are swabbed or when surface slices of beef trim are collected. In addition, the ICMSF (2011) provides guidance on the analytical unit amount (and sample size) for a range of food products.

[10] As the sensitivity reduces from 100% the chances of accepting a lot, given a particular sampling plan, will increase and as specificity reduces from 100% the chances of accepting a lot, given a particular sampling plan, will decrease.

Example 12
A two-class presence-absence sampling plan with desired performance

Consider that we are interested in sampling a food product for *Salmonella* using a two-class presence-absence sampling plan. Assume that we can tolerate accepting lots at most 10% of the time when the contamination is such that the analytical unit detection probability is 5%. That is, P(accept) should not be greater than 10% when the detection probability is 5%.

We use the Two-class Presence-Absence (1) Tab in the companion spreadsheet. Because we cannot tolerate detections of *Salmonella* in the sample we set $c = 0$. To find a plan that meets our criteria we change the value of n until P(accept) is less than 10% at a 5% detection probability. This happens when n equals 45 (or more).

A video showing you these calculations can be found at http://youtu.be/e3SRSnQ7s4g.

Example 13
Calculate the analytical unit detection probability

Using the calculator on the Analytical Unit Detect. Prob. Tab in the companion spreadsheet we can calculate what analytical unit detection probability we are likely to see for different analytical unit amounts and mean concentrations. For example, a 5 g analytical unit would give a detection probability of 9.52% when the concentration of the microorganism in the food is on average 0.02 cfu/g. In contrast, at the same concentration, a 10 g analytical unit would have a detection probability of 18.13%, while a 25 g analytical unit has a detection probability of 39.35%.

A video showing you these calculations can be found at http://youtu.be/4wxpNFarikU.

But what if there is no standard method or you want to better understand how the analytical unit amount affects the sampling plan? To explore this, we need to first assume that the food from which the sample is collected is well mixed. That is, microbiological contamination is homogeneous (recall from earlier that this is not the same as uniform contamination) with a given contamination rate of organisms per unit of food (g, ml or cm2). We can then calculate the probability of detecting at least one organism in a sample of amount w for any given contamination rate by using a standard statistical distribution, the Poisson distribution (Example 13). The mathematical details are provided in the Annex A1.2 Calculating the Analyti-

cal Unit Detection Probability given the Analytical Unit Amount" and a calculator is included in the accompanying spreadsheet on the Analytical Unit Detect. Prob. Tab.

As we can see from Example 13, when using enrichment methods the larger the analytical unit amount, the more likely the detection of the target organism in the food, even with just a single sample unit. This is also demonstrated in Figure 12 for three different analytical unit weights for a range of concentrations. From this plot it can be seen that, irrespective of the concentration, the larger analytical unit amount makes it more likely to capture the organism in the analytical unit and hence detect the contamination. This effect is more noticeable at low concentrations than it is at higher concentrations, because at high concentrations a large analytical unit amount will contain more than one cell of the target microbe which, however, adds no additional information to the detection (assuming that the method will be able to detect a single organism from the analytical unit). Note that the same applies, irrespective of the unit of measurement – g, ml or cm^2 – provided the homogeneity assumption holds.

Now that we have seen how the analytical unit amount affects the probability of detecting the target organism in a single analytical unit, what happens when we take more than one sample unit, that is, when we apply a sampling plan to the food? The companion spreadsheet includes a tool to do that on the Two-class Pres-

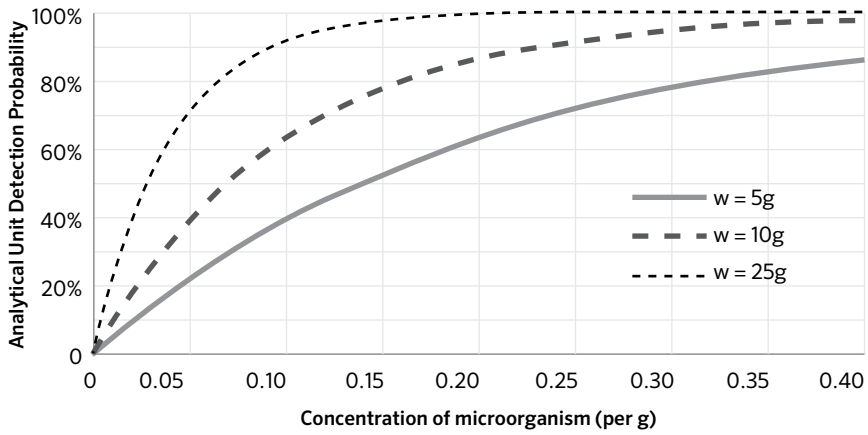


FIGURE 12: Effect of analytical unit amount (*w*) on the probability of detecting the organism (analytical unit detection probability) at various concentrations from a single sample, assuming homogeneous distribution of contamination in the food.

ence-Absence (2) Tab, where we use the analytical unit amount to relate an underlying concentration to the analytical unit detection probability and subsequently the sample size (n) and acceptance number (c) to the probability of accepting the lot. Holding sample size and acceptance number constant at $n = 15$ and $c = 0$, we see from Figure 13 that the larger analytical unit amount makes it less likely that lots are accepted, i.e. more likely that contamination is detected.

However, you may have noticed that while the number of analytical units remains the same, the total amount of material that is tested differs between the three sampling plans: a total of 75g, 150g and 375g are sampled and tested using analytical unit amounts of 5g, 10g, and 25g, respectively. Provided the analytical method has been appropriately validated (for 75, 150 and 375 g), as outlined earlier, it is entirely reasonable to enrich and test a single composite of the total amount, instead of testing 15 individual analytical units, which would be a lot more expensive (ICMSF, 2011; Jarvis, 2007). An example of this approach is "N-60 testing" for beef trimmings in the USA, where 60 small surface slices of beef trimmings are composited and tested for *E. coli* O157 (USDA FSIS, 2012).

But the question now arises as to whether it is better to collect fewer large sample units or more small sample units? From Example 14 we see that, theoretically for perfectly homogeneous contamination levels, it makes no difference how we sample provided the total amount of food remains constant. However, the as-

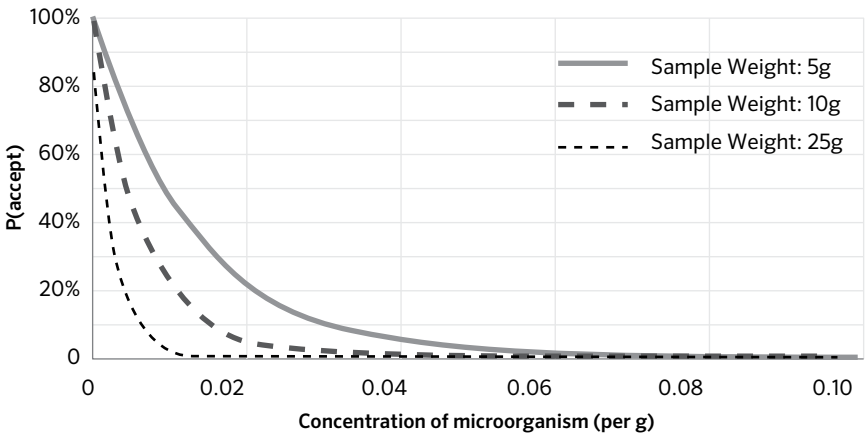


FIGURE 13: Effect of analytical unit amount on the probability of accepting lots , P(accept), when a sampling plan with $n = 15$ and $c = 0$ is used and a homogeneous distribution of contamination in the food is assumed.

Example 14
Fewer large sample units or more small sample units?

Let us assume that we have a validated enrichment method for testing 375g of food. Is it better to collect and test 15 sample units of 25g each or 75 sample units of 5g each?

Using the Two-class Presence-Absence (2) calculator in the companion spreadsheet we find that both give the same P(accept) for any given concentration. So provided we are dealing with *very well mixed* food it makes no difference whether we take fewer large sample units or more frequent small sample units provided the total amount of food tested remains constant.

But this result, which is due to the Poisson distribution, relies on the homogeneity assumption. In practice this assumption may not hold unless you are dealing with very well mixed foods, e.g. liquids. **Consequently, it is generally preferable to test a larger number of small sample units than an equivalent amount of the lot as fewer large sample units, because the chances of detecting contamination 'pockets' is increased by sampling this way.**

sumption of homogeneity used for the calculations is only valid in some very well mixed foods. In most foods there is greater variation in the microbiological contamination levels. As such, it is becoming more common to calculate the OC curve using a Poisson-$\log_{10}$-normal mixture distribution. This approach is provided in the ICMSF tool (http://www.icmsf.org/) and the Basic Mode of the web-based sampling tool (http://www.fstools.org/sampling), and other distribution choices are provided in the Advanced Mode.

2.8.6.2 How does the concentration of microorganisms in food affect the probability of acceptance?

We have already seen that two-class presence-absence sampling plans are usually used when dealing with pathogens. We use the ability to detect the organism as a surrogate for the variable of interest, namely its overall, and presumably low, concentration in the lot. It is the presumption that the overall concentration is low, and the need to enrich the sample to allow us to detect the one or very few organisms in an analytical unit, that requires the use of presence-absence based sampling.

From Figure 12 we know that the higher the level of contamination the more likely a single analytical unit will contain the target organism. Consequently, when multiple sample units are collected as part of a sampling plan, the resulting OC curve will look similar to those in Figure 13, where we see that as the underlying

Example 15
Effect of concentration on the probability of lot acceptance

Consider again a sampling plan with $n = 15$, $c = 0$ and w = 10 g. From the companion spreadsheet (Two-class Presence-Absence (2)) we can see that at a concentration of 0.001 cfu/g, the analytical unit detection probability is 1% and P(accept) = 86.07%. As the concentration increases 10-fold (1 $\log_{10}$ cfu/g) to 0.01 cfu/g the detection probability increases to 9.52% and the probability of acceptance reduces to 22.31%.

concentration increases the probability of accepting the lot decreases (Example 15). That is, 'highly' contaminated lots are infrequently accepted (frequently rejected), although we must remember that accepting a lot is not the same as the lot being free from contamination.

2.8.6.3 How does the number of sample units affect the probability of acceptance?

The sample size (n) indicates how many sample units are collected from a lot to make a decision about its acceptability. For a given detection probability the probability of acceptance decreases as the sample size (n) increases, as is shown in Figure 14 for a sampling plan with $c = 0$. From this figure it can also be seen that the probability of acceptance drops very slowly for small n, e.g. $n = 10$. In contrast,

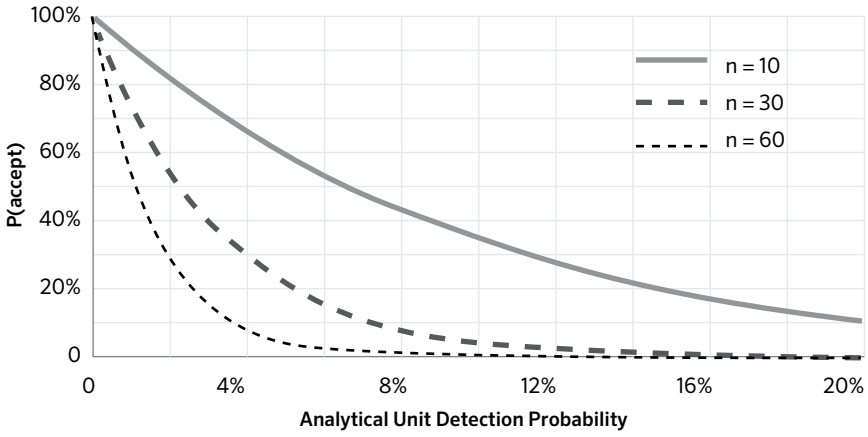


FIGURE 14: Effect of sample size (n) on the probability of accepting lots, P(accept), when a zero acceptance number sampling plan ($c = 0$) is used.

Example 16
A discriminating sampling plan

Consider the sampling plan from Figure 14 with $n = 60$ and $c = 0$. At a detection probability of 0.4% we have P(accept) = 78.62%, which drops to 29.76% for a detection probability of 2%. In contrast, the difference in P(accept) is much smaller for the sampling plan with $n = 10$: 96.07% versus 81.71%. This makes the sampling plan with $n = 60$ more discriminating.

for a large sample size, e.g. $n = 60$, the probability of acceptance drops very quickly. Consequently, sampling plans with a large sample size are more discriminating, that is, the probability of acceptance drops quickly from 100% to 0%, as shown in Example 16.

We also reiterate the comment we made above, that it is usually preferable to collect more small sample units than to rely on fewer larger sample units, because more smaller sample units increase the chances of 'hitting' contaminated spots, especially since the microbial contamination in food products is generally not truly homogeneous.

2.8.6.4 How does the acceptance number affect the probability of acceptance?

The acceptance number (c) determines how many analytical units in the sample are allowed to contain the target organism while still accepting the lot. For highly pathogenic organisms it is common to have an acceptance number $c = 0$ such that the lot is only acceptable when the target organism is not detected in any of the analytical units. For such zero acceptance number sampling plans it is typical that the probability of acceptance drops rapidly from 100% (Figure 15). From Figure 15 it can be observed that as c increases, the probability of acceptance remains close to 100% for longer before dropping, thus shifting the OC curve toward the right. For example, when $c = 2$, it is more likely that a lot can be accepted at higher analytical unit detection probability, because a lot will be acceptable when a sample contains either zero, one or two analytical units in which the target organisms is detected.

The acceptance number c also relates to the stringency of the sampling plan ("2.8.4 What is meant by discrimination and stringency of sampling plans?"). The smaller the value of c, the more stringent the sampling plan.

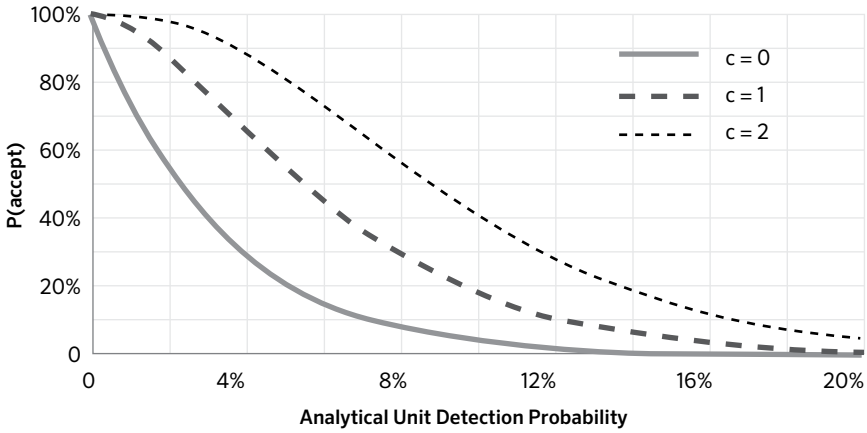


FIGURE 15: Effect of acceptance number (*c*) on the probability of accepting lots, P(accept), when the sample size is $n = 30$.

2.8.6.5 Putting it all together: two-class presence-absence sampling plans

The two-class presence-absence sampling plan is the most commonly used sampling plan when dealing with highly pathogenic microorganisms. Consequently, we are concerned with (very) low levels of microbial contamination, which can cause illness. However, we have also seen that it is never possible to guarantee that a lot is free from microbiological contamination through sampling. Consequently, when an MC is used for a food product in such circumstances it is common to not accept the lot if the target organism is detected in any of the analytical units, that is, a zero acceptance number sampling plan ($c = 0$) is used.

While a larger analytical unit amount will give a better chance of capturing the target organism in the analytical unit, it is generally preferable to collect more small sample units (and hence use smaller analytical unit amounts) than it is to collect fewer but larger sample units (and use larger analytical unit amounts) as this approach provides better coverage of the lot and protects you when the food lot is not homogeneous (i.e. not well mixed). An assumption of at least some level of heterogeneity is now more common than an assumption of homogeneity.

With the analytical unit amount set – usually specified in standardized microbiological tests – the basic statistical approach, using the Two-class Presence-Absence (1) Tab in the companion spreadsheet, to create a sampling plan that meets the

consumer (and producer) risk point is as follows. It should be noted though that this approach should not be applied blindly, but in the context for which the MC is to be developed.

1. Decide whether you cannot accept detection of the target organism in any of the analytical units or whether it would be acceptable to have some analytical units contain the organism. Hence *c* either equals zero or takes on a value that is greater than zero.
2. Decide on an 'unacceptable detection probability' (p_1) in a lot so that lots with this analytical unit detection probability (or greater) should be rejected most of the time.
3. Decide how frequently such lots (with detection probability p_1 or greater) could be accepted, i.e. what is the maximum value of P(accept), call it P_1(accept), that you can tolerate when the detection probability is p_1.[11]
4. Let *c* equal zero (as a starting point).
5. Pick a practical value for *n*.
6. Enter the values of *n* and *c* into the companion spreadsheet.
    a. If P(accept) for your choice p_1 is greater than the P_1(accept) you selected in Step 3 then *increase* the value of *n* by 1 (or more) and repeat Step 6 until the value of the P(accept), shown in cell D12 of the worksheet, is less than or equal to P_1(accept).[12]
    b. If P(accept) for your choice p_1 is less than P_1(accept) you selected in Step 3 then *decrease* the value of *n* by 1 (or more) and repeat Step 6 until the value of P(accept) just exceeds P_1(accept). Now increase *n* by 1 to ensure that P(accept) is less than or equal to P_1(accept).
7. If *c* is allowed to be greater than 0, then check that the OC Curve is acceptable. In particular, check if at a low analytical unit detection probability (p_0) the probability of acceptance P_0(accept) is tolerable and not too small.[13] If the probability of acceptance is tolerable then you are finished. If however P(accept) is too small, i.e. less than P_0(accept), then you will need to increase *c* by 1 and repeat the process from Step 5, including further increases in *c*, until the plan meets your requirements, i.e. the probability of acceptance is less than or equal to P_1(accept) at analytical unit detection probability p_1 and greater than or equal P_0(accept) at detection probability p_0.

A video demonstrating this process of determining a sampling plan can be found at http://youtu.be/YnncxY7imyw.

---

[11] The combination of p_1 and P_1(accept) is referred to as the *Consumer's Risk Point* (see "2.8.5 What are the Producer's and Consumer's Risk Points?"). The corresponding values can be entered into cells B12 and C12 of the worksheet.

[12] The actual P(accept) shown in cell D12 of the worksheet will change from red font to green.

[13] The combination of p_0 and P_0(accept) is also known as the *Producer's Risk Point* (see "2.8.5 What are the Producer's and Consumer's Risk Points?"). The corresponding values can be entered into cells B11 and C11, respectively, of the worksheet.

It should be noted that if there is flexibility in the analytical unit amount that is used, then this step-by-step approach can still be followed. In this case however the effect of different analytical unit amounts can be assessed using the Two-class Presence-Absence (2) Tab in the companion spreadsheet under the assumption that the contamination is homogeneously distributed. Alternate approaches, such as the Poisson-lognormal distribution can be investigated using the web-based (http://www.fstools.org/sampling/) or ICMSF tools (http://www.icmsf.org/).

2.8.7 Two-class concentration-based sampling plans

Two-class concentration based sampling plans are quite similar to two-class presence-absence-based sampling plans as we utilize the binomial distribution to calculate the probability of lot acceptance for both types. However, as we mentioned in Part 1, the concentration-based sampling plans are generally used when we are dealing with hygiene indicators, e.g. *Enterobacteriaceae*, pathogens that are not highly pathogenic and can therefore be tolerated to some degree in the food, e.g. *Listeria monocytogenes* in some ready-to-eat foods (CAC 2007), or where a product will receive a process like cooking that will reduce the numbers of the pathogens to a safe level.

The key difference to a presence-absence based plan is that we are now specifically interested in determining the concentration of the microbes in the food, and not simply their presence, i.e. whether they exceed the detection threshold concentration.

The two-class concentration based sampling plan is defined by

- the analytical unit amount (w), i.e. the amount of each analytical unit (mass, volume, area);
- the sample size (n), i.e. number of sample units that are collected;[14]
- the microbiological limit (m) that determines whether an analytical unit is acceptable or not acceptable; and
- the acceptance number (c), i.e. the number of analytical units that are allowed to contain a concentration that exceeds m while still considering the lot acceptable.

In addition, to assess the performance of the two-class concentration based sampling plan an estimate of the SD of the $\log_{10}$ concentration between analytical units is needed.

The common assumption that is made is that the $\log_{10}$ concentration of the target organism in the food is normally distributed (see Figure 6). However, if evidence

[14] For some sampling plans, e.g. ISO 2859, the sample size (n) is obtained from the size of the lot. However, this is done for non-statistical reason.

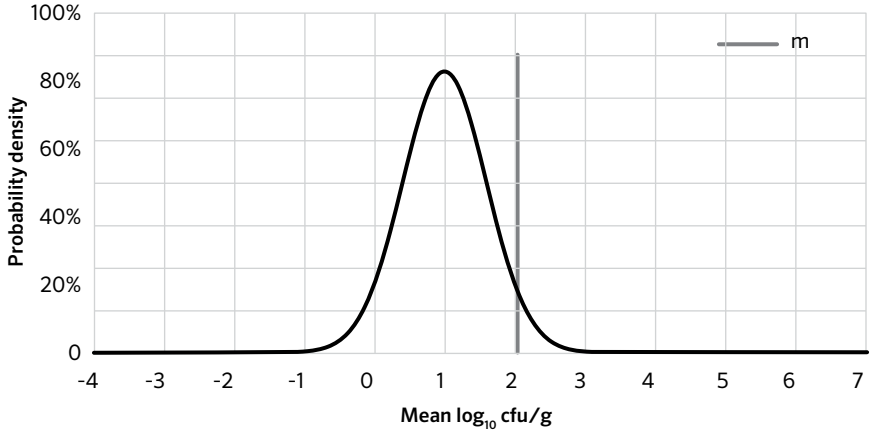


FIGURE 16: Plot of a normal distribution with mean = 1 $\log_{10}$ cfu/g, SD = 0.6 $\log_{10}$ cfu/g and m = 2 $\log_{10}$ cfu/g. The area under the curve to the right of m is the probability that the concentration in an analytical unit of the food exceeds m (see Figure 17).

against this assumption exists, then it is possible to use alternative distributions, e.g. Gamma, though these are less commonly used and beyond the scope of this document. However, the web-based tools (http://www.fstools.org/sampling) do provide flexibility for examining alternative distributions.

So, assuming that the $\log_{10}$-normal distribution holds, we can calculate the probability of a sample exceeding the limit m for any mean $\log_{10}$ concentration. All that is needed is a reasonable estimate of the variability (SD) in the $\log_{10}$ concentration (Figure 16). The resulting probability gives an indication of how frequently sample units from the lot will be found unacceptable (Figure 17). The probability of lot acceptance is then calculated in an analogous way to the two-class presence-absence plan, but using the probability of exceeding m instead of the analytical unit detection probability. The OC curve can then be plotted with either the mean $\log_{10}$ concentration on the X-axis (Figure 18) or with the arithmetic mean concentration on the X-axis (Figure 19). As noted above ("2.8.3.1 Should we use the $\log_{10}$ geometric mean or the arithmetic mean for the X-axis of the OC curve?") we suggest you use the latter approach to highlight the level of control, using actual concentrations in the food, that is being achieved (see also Example 4). The mathematical details are provided in the Annex A1.4 Two-class concentration-based sampling plans" and the calculations of P(accept) and the associated graphs are provided in the companion spreadsheet in the Two-class Concentration Tab. The calculations that were used to generate Figure 16 to Figure 19 are shown in Example 17.

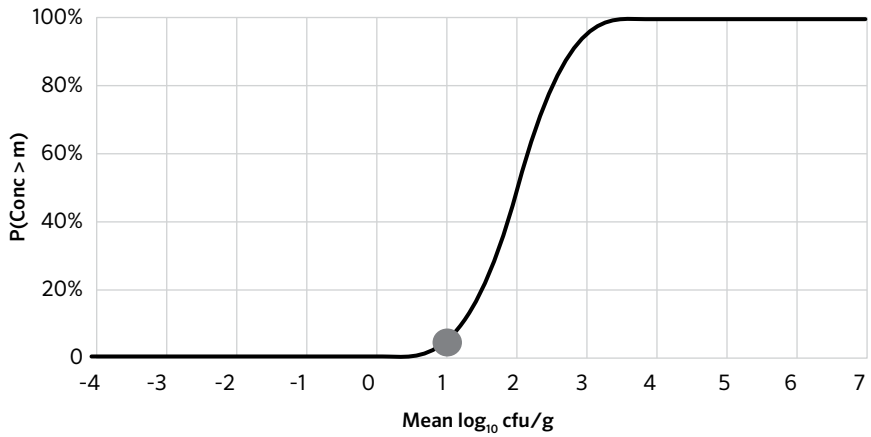


FIGURE 17: Plot of the probability that the concentration in the food exceeds $m = 2$ $\log_{10}$ cfu/g when SD = 0.6 $\log_{10}$ cfu/g. The dot point indicates the probability when the mean $\log_{10}$ concentration ($\log_{10}$ geometric mean) in the lot equals 1.0 $\log_{10}$ cfu/g (Example 17).

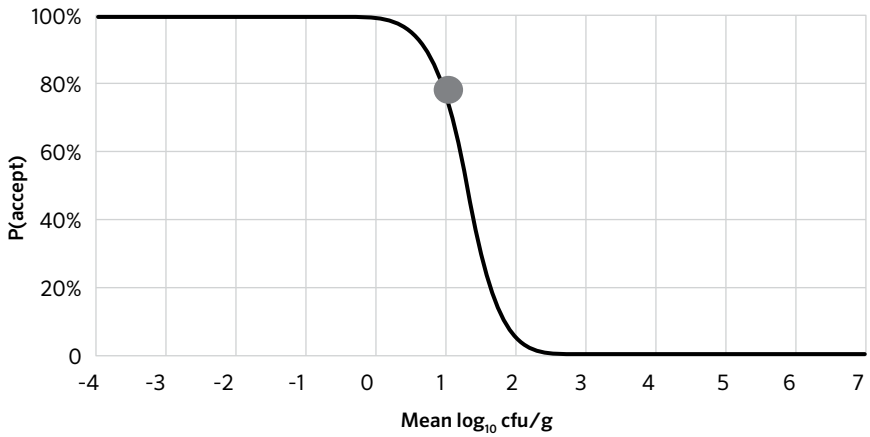


FIGURE 18: Two-class concentration-based sampling plan OC curve using the mean $\log_{10}$ concentration ($\log_{10}$ geometric mean) on the X-axis for a sampling plan with $n = 5$, $c = 0$, $m = 2$ $\log_{10}$ cfu/g and SD = 0.6 $\log_{10}$ cfu/g. The dot indicates P(accept) when the mean $\log_{10}$ concentration ($\log_{10}$ geometric mean) in the lot equals 1.0 $\log_{10}$ cfu/g (Example 17).

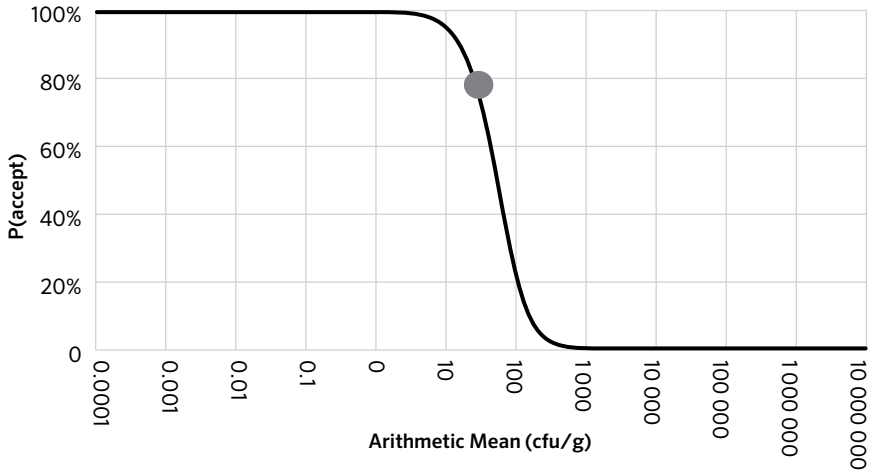


FIGURE 19: Two-class concentration-based sampling plan OC curve using the arithmetic mean concentration on the X-axis for a sampling plan with $n = 5$, $c = 0$, $m = 2$ $\log_{10}$ cfu/g and SD = 0.6 $\log_{10}$ cfu/g. The dot indicates P(accept) when the arithmetic mean = 26 cfu/g (Example 17; equivalent to a $\log_{10}$ geometric mean of 1.0 $\log_{10}$ cfu/g).

Example 17
Concentration-based P(accept)

Consider the sampling plan for *Listeria monocytogenes* provided in CAC (2007) with $n = 5$, $c = 0$ and $m = 100$ cfu/g ($m = 2$ on the $\log_{10}$ scale).

Assuming that *L. monocytogenes* is normally distributed on the $\log_{10}$ scale with a standard deviation 0.6 $\log_{10}$ cfu/g we can use the calculations in the Two-class Concentration Tab in the companion spreadsheet to create the associated OC curve.

From the spreadsheet we can see that if the ($\log_{10}$ geometric) mean equals 1.0 $\log_{10}$ cfu/g (arithmetic mean equals 26 cfu/g) then we can expect that 4.78% of sample units would exceed the limit m = 2.0 $\log_{10}$ cfu/g and hence the probability of acceptance equals P(accept) = 78.28%.

We can also use the tool to change the mean $\log_{10}$ concentration until the value of P(accept) drops below a desired value, e.g. 5%. This is achieved when the mean $\log_{10}$ concentration equals 1.93 (to two decimals) and P(accept) = 4.87%. The arithmetic mean equals 221 cfu/g.

A video showing you these calculations can be found at http://youtu.be/vpbXB3kqYNM.

2.8.7.1 How does the analytical unit amount (weight, volume or area) affect the probability of acceptance?

Unlike for two-class presence-absence sampling plans the analytical unit amount does not have as much of an effect here. However, there are advantages and disadvantages irrespective of how big an analytical unit amount (w) is tested.

If a large analytical unit amount is used, then it may be more difficult to enumerate the organism successfully, especially when the concentration is low. We have already seen this effect in the two-class presence-absence plan where the probability of detecting a microorganism is low when the concentration in the food product is low. Similar to the example in Figure 1, imagine that we have 100 g of a food product which contains exactly 1 organism of interest. If we take a 1 g analytical unit amount, then we have a 1 in 100 (= 1%) chance of capturing the organism. In contrast, if we select a 10 g analytical unit amount the chances are now 1 in 10 (= 10%) of capturing the organism, while testing the whole 100 g will give us a 100% chance. The same can of course happen for higher concentrations when only a small analytical unit amount is used. However, testing a large analytical unit (and hence a large sample unit taken from the lot) also results in an 'averaging effect' and it is then difficult to determine whether the contamination was consistent across the sample unit or due to a highly contaminated clump or cluster. Knowing about the pattern of contamination may provide you with information about the mechanism of contamination and allow you to better estimate the statistical distribution.

As for two-class presence-absence sampling plans we would encourage users to utilize validated analytical unit amounts, i.e. using standardized and validated methods. However, it is important to keep the reason for sampling in mind, whether for lot acceptance or statistical process control, and to select a suitable analytical unit amount for that purpose. In addition, it is worthwhile to point out that enumeration methods often are thought of as 'accurate' because they result in a count. However, the colony count on a plate or the estimate from a MPN method is subject to analytical variability, which is method dependent, and lack of sensitivity and specificity may also affect the result. Consequently, it is important to understand the overall performance (sensitivity, specificity, etc.) of an analytical test, so that the results can be interpreted appropriately. A more detailed discussion of this topic is outside the scope of this document and we refer interested readers to the work by AOAC International (2006) and also Cowell and Morisetti (1969).

2.8.7.2 How does the level of concentration affect the probability of acceptance?

Generally speaking, the higher the mean concentration in the food, the less likely it is for the lot to be accepted, as can be seen from Figure 18. However, P(accept)

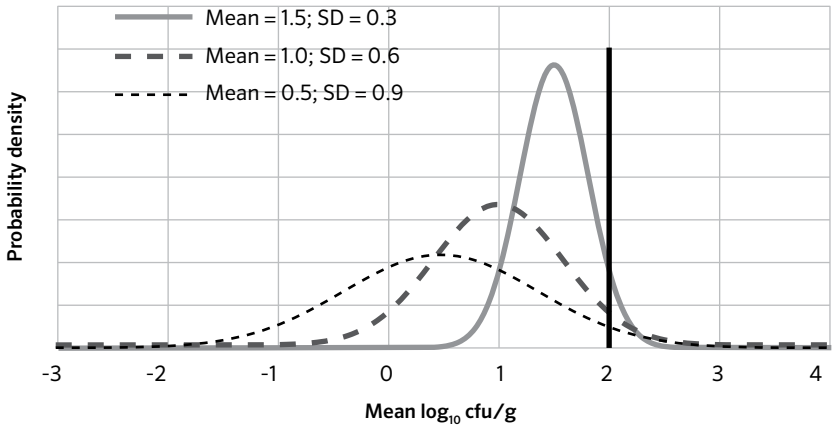


FIGURE 20: Three normal distributions with different means and SDs. All three distribution have a probability of exceeding the limit $m = 2$ $\log_{10}$ cfu/g that equals 4.78%.

depends not only on the mean concentration, but also on the unacceptable limit (*m*) and the SD. The relationship between the mean, SD and unacceptable limit is illustrated in Figure 20, where all three normal distributions result in the same probability of exceeding *m*.

However, keep in mind that when you want to make inferences about the actual number of organisms in the food lot, e.g. to assess the level of control that is being achieved and hence the risk, then you will need to do this on the arithmetic scale and not the $\log_{10}$ scale (see also "1.2.3 Why do we use $\log_{10}$ numbers and why do we need to be careful when interpreting them?").

2.8.7.3 How does the unacceptable limit (*m*) affect the probability of acceptance?

As we have seen in the previous section the value of *m* is integrally connected with the mean and SD. If we have a look at the OC curves in Figure 21, then we can see that increasing the value of *m*, while holding the SD and other parameters (*n*, *c*) constant, has the effect of shifting the OC curve to the right. However, the shape of the OC curve remains unchanged.

2.8.7.4 How does the variability in concentration affect the probability of acceptance?

We have already seen that the SD, mean and unacceptable limit (*m*) are interconnected. The effect of changing the SD on the OC curve, while keeping *m* constant is

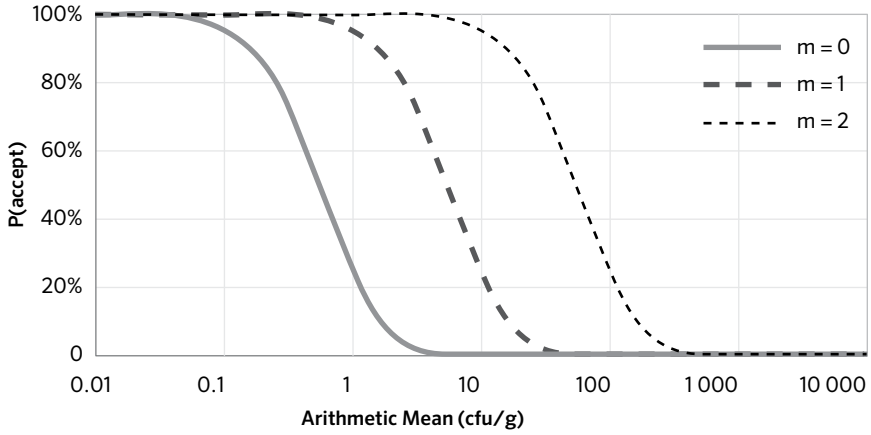


FIGURE 21: Two-class concentration-based sampling plan OC curves with $n = 5$, $c = 0$, SD = 0.6 $\log_{10}$ cfu/g for three different unacceptable limits (m).

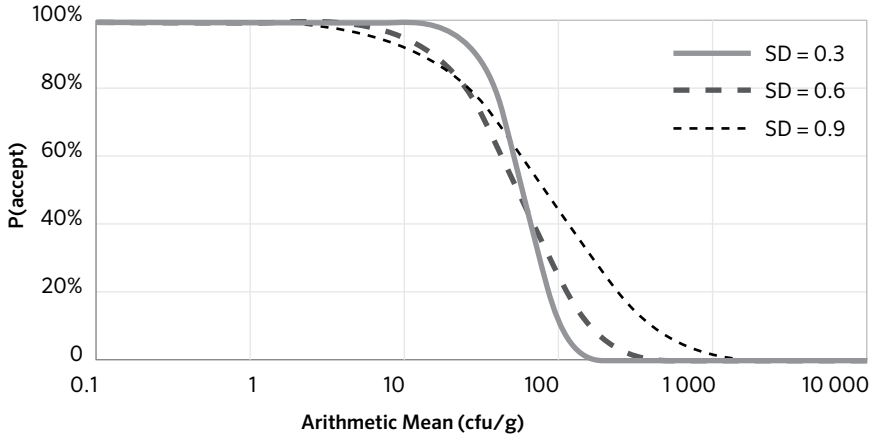


FIGURE 22: Two-class concentration-based sampling plan OC curves with $n = 5$, $c = 0$, $m = 2$ $\log_{10}$ cfu/g for three different SDs.

shown in Figure 22. From this plot it can be seen that the effect on the OC curve is more complicated, because the SD also influences the value of the arithmetic mean ("1.2.3 Why do we use $\log_{10}$ numbers and why do we need to be careful when interpreting them?"). However, when the SD is large, the OC curve starts to drop at a much lower arithmetic mean concentration (here approximately 1 cfu/g) and takes longer to reach a probability of acceptance that is close to zero. This is because a

small percentage of analytical units would have a concentration greater than the limit m when the arithmetic mean is small, and this percentage is larger when the SD is large than when it is small. Therefore, the same sampling plan, here $n = 5$, $c = 0$ and $m = 100$ cfu/g (equivalent to 2 $\log_{10}$ cfu/g), is better able to discriminate between two arithmetic mean concentrations that are close together when the SD is small, compared with when the SD is large.

Consequently, if there is little variability in the microbial concentration in the food, then any given sampling plan will be more discriminating, i.e. drop quickly from 100% to 0%.

2.8.7.5 How does the number of sample units affect the probability of acceptance?

The effect of the sample size (n) is similar to that described for two-class presence-absence sampling plans, that is, a sampling plan with a larger n is more discriminating than one with a small n. This is illustrated in Figure 23, where we can see that a large sample size results in an OC curve that drops more quickly.

Notice that we now have two parameters that both influence how discriminating a sampling plan is: the sample size (n) and the SD. So, if we have a food product with a highly variable microbial concentration (large SD), then we can make the

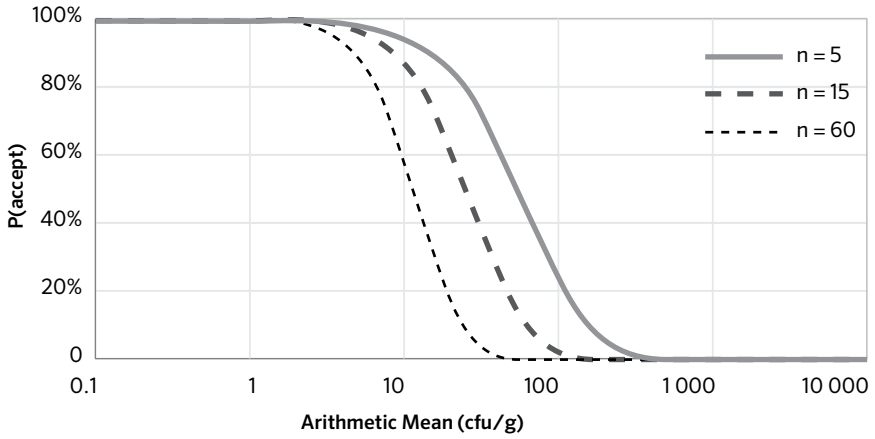


FIGURE 23: Two-class concentration-based sampling plan OC curves with $c = 0$, $m = 2$ $\log_{10}$ cfu/g, SD = 0.6 $\log_{10}$ cfu/g for three different sample sizes.

OC curve more discriminating by increasing the sample size (n). That way it will be possible to achieve the required degree of discrimination despite considerable variability in the food. However, we note that increasing the sample size n comes at an obvious cost, namely the cost for sample collection and microbiological testing. In contrast, attempting to reduce the variability through improved process control also comes at a cost, but is preferable in the long-term over the alternative of increasing the sample size.

2.8.7.6 How does the 'acceptance number' affect the probability of acceptance?

The final parameter that we can change in the two-class concentration based sampling plans is the acceptance number (c). It has a similar effect as the acceptance number in two-class presence-absence based sampling plans, namely the OC curve moves to the right as c increases and the shape also changes slightly (Figure 24).

In the previous section we have alluded to the fact that it is possible to generate sampling plans with similar performance for food products that differ in the variability of microbial concentration. To illustrate this point, consider a food product which is manufactured by three different processors, with different SDs, namely:

- **Processor 1** produces the food with an SD of 0.3 $\log_{10}$ cfu/g;
- **Processor 2** produces the food with an SD of 0.6 $\log_{10}$ cfu/g; and
- **Processor 3** produces the food with an SD of 0.9 $\log_{10}$ cfu/g.

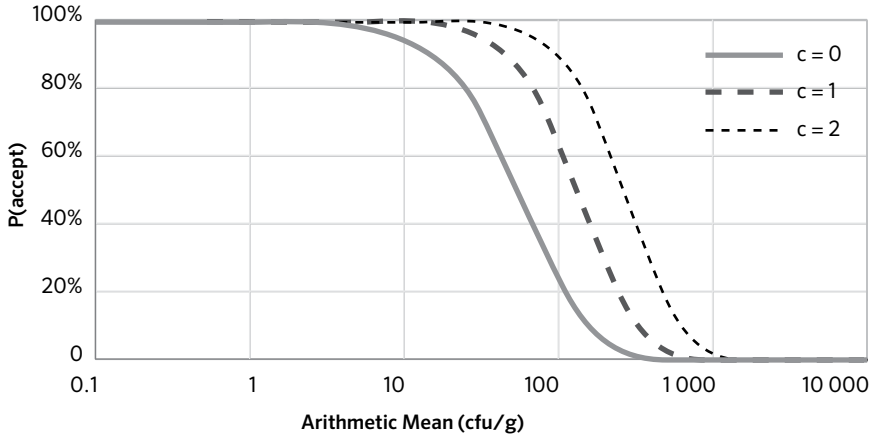


FIGURE 24: Two-class concentration-based sampling plan OC curves with $n = 5$, $m = 2$ $\log_{10}$ cfu/g, SD = 0.6 $\log_{10}$ cfu/g for three different acceptance numbers.

Also assume that the desired MC has a microbiological limit of $m = 2 \log_{10}$ cfu/g. With some trial and error we can find sampling plans that are similar in their performance as shown in Figure 25. These curves are only meant to be indicative and closer agreement between them can be achieved by changing the values of the sample size *n* and acceptance number *c*. While this example is quite extreme it does illustrate that you have considerable flexibility in specifying sampling plans, through *n*, *c*, and even *m*, such that they result in the desired performance for any particular situation. However, the desired performance ultimately needs to be balanced against the costs that such a sampling plan will incur.

2.8.7.7 Putting it all together: two-class concentration-based sampling plans

The two-class concentration-based plan is often used when low levels of contamination are acceptable. Consequently, this sampling plan may be applied to hygiene indicator organisms, pathogens that are less likely to cause illness at low levels and for products that are expected to receive further processing to reduce the microbial contamination to acceptable levels.

The basic statistical approach, using the Two-class Concentration Tab in the companion spreadsheet, to create a sampling plan that meets the consumer (and producer) risk point is largely analogous to that illustrated for two-class presence-

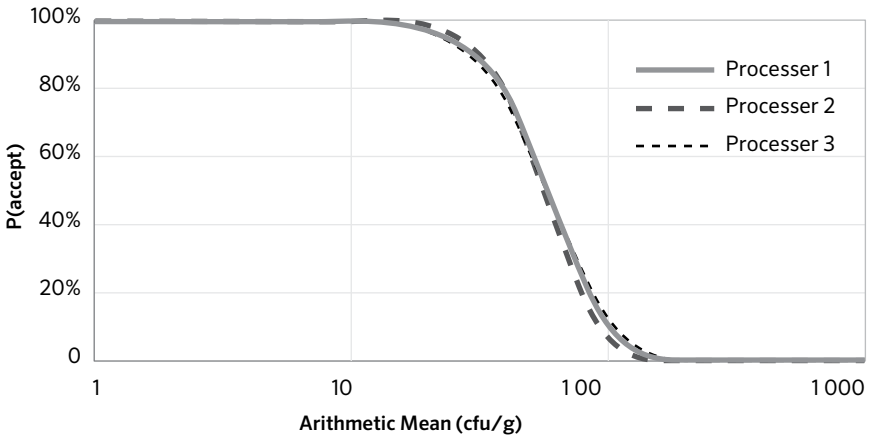


FIGURE 25: Two-class concentration-based sampling plan OC curves for three products with SD = 0.3, 0.6 and 0.9 $\log_{10}$ cfu/g and a microbiological limit of m = 2 $\log_{10}$ cfu/g. The sampling plans that achieve similar performance require *n* = 5 & c = 0 (Processor 1); *n* = 27 & c = 3 (Processor 2); and *n* = 48 & c = 4 (Processor 3).

absence sampling plans ("2.8.6.5 Putting it all together: two-class presence-absence sampling plans"). As before, this approach should be applied in the context for which the MC is to be developed.

1. Decide on an 'unacceptable average concentration' (μ_1) in a lot so that lots with this average concentration (or greater) should be rejected most of the time.
2. Decide how frequently such lots (with average concentration μ_1 or greater) could be accepted, i.e. what is the maximum value of P(accept), call it P_1(accept), that you can tolerate when the mean concentration is μ_1.[15]
3. Enter a suitable value for the SD into the companion spreadsheet.
4. Let *c* equal zero (as a starting point).
5. Pick a practical value for *n*.
6. Enter the values of *n* and *c* into the companion spreadsheet.
    a. If P(accept) for your choice μ_1 is greater than the P_1(accept) you selected (Step 2) then increase the value of *n* by 1 (or more) and repeat Step 5 until the value of P(accept) is less than or equal to P_1(accept).[16]
    b. If P(accept) for your choice μ_1 is less than P_1(accept) you selected (Step 2) then decrease the value of *n* by 1 (or more) and repeat Step 6 until the value of P(accept) just exceeds the P_1(accept). Now increase *n* by 1 to ensure that P(accept) is less than or equal to P_1(accept).
7. If *c* is allowed to be greater than 0, then check that the OC Curve is acceptable. In particular, check if at a low mean concentration (μ_0) the probability of acceptance P_0(accept) is tolerable and not too small.[17] If the probability of acceptance is tolerable then you are finished. If however P(accept) is too small, i.e. less than P_0(accept), then you will need to increase *c* by 1 and repeat the process from Step 5, including further increases in *c*, until the plan meets your requirements, i.e. the probability of acceptance is less than or equal to P_1(accept) at a mean concentration μ_1 and greater than or equal to P_0(accept) at mean concentration μ_0.

A video demonstrating this process can be found at http://youtu.be/w0SMpG-hokXo.

2.8.8 Three-class (concentration-based) sampling plans

Three-class concentration-based sampling plans (Bray, Lyon and Burr, 1973) are

[15] The combination of μ_1 and P_1(accept) is referred to as the *Consumer's Risk Point* (see "2.8.5 What are the Producer's and Consumer's Risk Points?"). The corresponding values can be entered into cells B16 and C16 in the worksheet.

[16] The actual P(accept) shown in cell D16 will change from red to green font.

[17] The combination of μ_0 and P_0(accept) is referred to as the *Producer's Risk Point* (see "2.8.5 What are the Producer's and Consumer's Risk Points?"). The corresponding values can be entered into cells B15 and C15, respectively, in the worksheet.

similar to two-class concentration-based sampling plans as we assume the normal distribution of the $\log_{10}$ concentration of microorganisms. However, instead of a single limit (*m*) that is used to differentiate between acceptable and unacceptable concentrations, we now have two limits: the marginal limit (*m*), which differentiates acceptable from marginally acceptable concentrations, and the unacceptable limit (*M*), which differentiates marginally acceptable from unacceptable concentrations (Figure 7). Consequently, three-class sampling plans are also used when low levels of contamination are acceptable, such as for hygiene indicator organisms or pathogens that are less likely to cause illness at low levels. In contrast to the two-class concentration-based plan, this type of plan is employed where a scientific or operational basis exists for establishing a clear upper limit (*M*) that defines unacceptable concentrations. In this situation it may be possible to set the value of *m* as the maximum allowable concentration of the target organisms under Good Manufacturing Practice (GMP) (Dahms and Hildebrandt, 1998).

This type of sampling plan is defined by

- the analytical unit amount (w), i.e. the amount of each analytical unit (mass, volume or area);
- the sample size (*n*), i.e. number of sample units that are collected;
- the marginal and unacceptable microbiological limits (*m* and *M*) that are used to determine whether an analytical unit is acceptable, marginally acceptable or unacceptable; and
- the acceptance number (*c*), i.e. the number of analytical units that are allowed to exceed the limit *m* but not *M*, while still considering the lot to be acceptable.

In addition, to assess the performance of the three-class sampling plan an estimate of the SD of the $\log_{10}$ concentration between analytical units is needed.

While it is possible to create sampling plans that allow some sample units to exceed *M* without rejecting the lot, these are not commonly used in relation to MC in foods. [18]

It should be noted that, similar to two-class concentration-based sampling plans, we are assuming that the $\log_{10}$ concentrations are normally distributed. However, this assumption is important only insofar as to allow us to calculate the proportion of concentrations in the food that are expected to be acceptable, marginally acceptable and unacceptable, given some value for the mean and SD. However, if you have evidence against this assumption, then it is possible to use a different

[18] Note that in some publications the definition of *c* includes the numbers of unacceptable (or bad) sample units and hence differs from the definition used here. However, in the microbiological setting it is common to allow no unacceptable units in the sample and hence the two definitions are equivalent in this case.

distribution, although this situation is less common and beyond the scope of this document. The web-based tools (http://www.fstools.org/sampling) do provide flexibility for alternative distributions.

Because there are now three possible outcomes for each analytical unit (acceptable, marginally acceptable or unacceptable) the binomial distribution is no longer applicable and the trinomial distribution is used instead. As long as the mean and SD have been provided (Figure 26) the probability with which each of the three outcomes is expected to occur can be calculated from the normal distribution (Figure 27). These probabilities are then used for the calculation of the probability of acceptance using the trinomial distribution.

The OC curve can be plotted with the mean $\log_{10}$ concentration on the X-axis (Figure 28) or the arithmetic mean concentration on the X-axis (Figure 29). As pointed out previously ("2.8.3.1 Should we use the $\log_{10}$ geometric mean or the arithmetic mean for the X-axis of the OC curve?") we suggest you use the latter approach to highlight the level of control that is being achieved, using actual concentrations in the food, (see also Example 4). The mathematical details are provided in Annex A1.5 Three-class sampling plans" and the calculations of P(accept) and the associated OC curves are provided in the companion spreadsheet in the Three-class Concentration Tab. The calculations that were used to generate Figure 26 to Figure 29 are shown in Example 18.

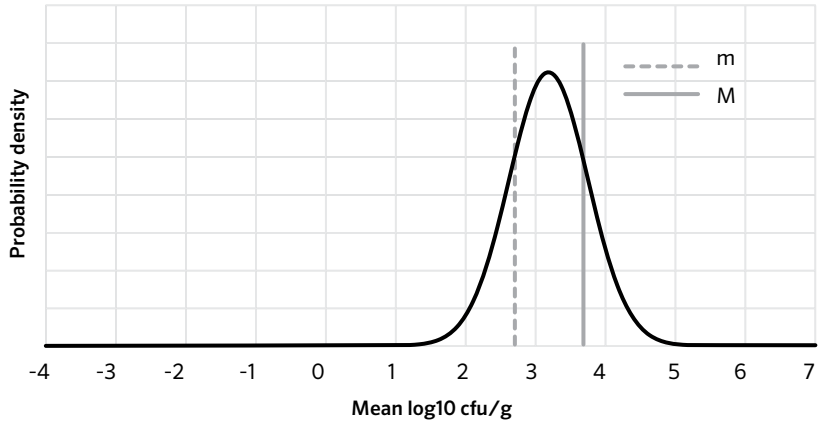


FIGURE 26: Plot of a normal distribution with mean = 3.2 $\log_{10}$ cfu/g, SD = 0.55 $\log_{10}$ cfu/g and microbiological limits m = 2.7 $\log_{10}$ cfu/g and M = 3.7 $\log_{10}$ cfu/g. The probabilities of an analytical unit being marginally acceptable (m < Conc ≤ M) and unacceptable (Conc > M) are shown in Figure 27.

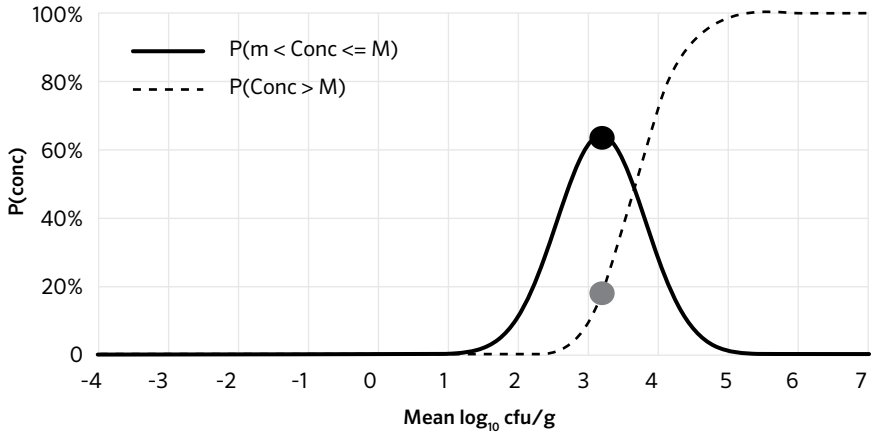


FIGURE 27: Probabilities that the concentration in the food is marginal, P(m < Conc ≤ M), and unacceptable, P(Conc > M), with SD = 0.55 $\log_{10}$ cfu/g, m = 2.7 $\log_{10}$ cfu/g and M = 3.7 $\log_{10}$ cfu/g. The dots indicate the probabilities when the mean $\log_{10}$ concentration ($\log_{10}$ geometric mean) in the lot equals 3.2 $\log_{10}$ cfu/g (Example 18).

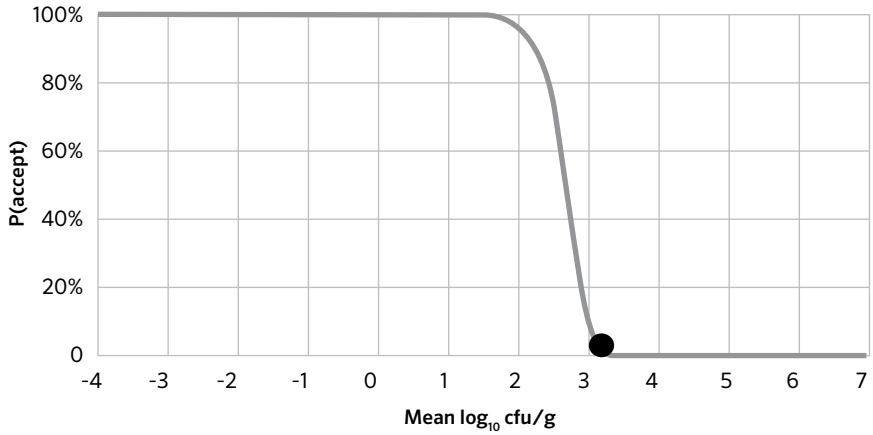


FIGURE 28: Three-class concentration-based sampling plan OC curve using the mean $\log_{10}$ concentration ($\log_{10}$ geometric mean) on the X-axis for a sampling plan with n = 5, c = 2, m = 2.7 $\log_{10}$ cfu/g, M = 3.7 $\log_{10}$ cfu/g and SD = 0.55 $\log_{10}$ cfu/g. The dot indicates P(accept) when the mean $\log_{10}$ concentration ($\log_{10}$ geometric mean) in the lot equals 3.2 $\log_{10}$ cfu/g (Example 18).

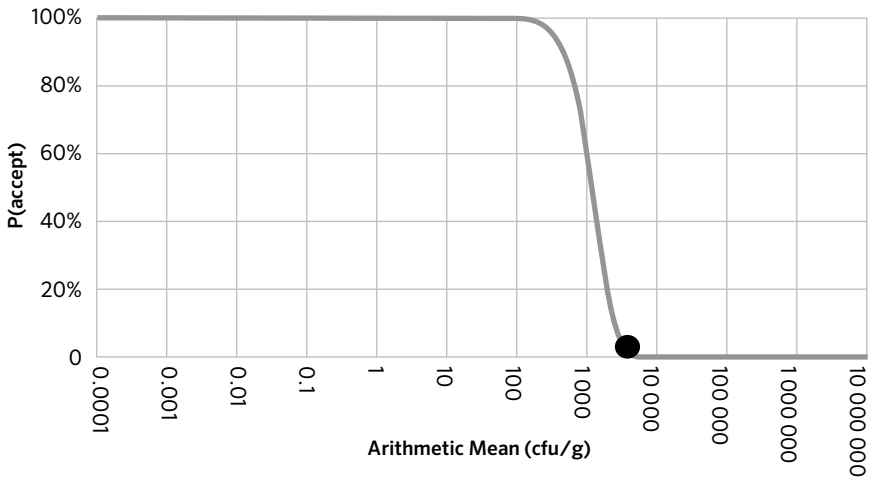


FIGURE 29: Three-class concentration-based sampling plan OC curve using the arithmetic mean concentration on the X-axis for a sampling plan with $n = 5$, $c = 2$, $m = 2.7$ $\log_{10}$ cfu/g, $M = 3.7$ $\log_{10}$ cfu/g and SD = 0.55 $\log_{10}$ cfu/g. The dot indicates P(accept) when the arithmetic mean = 3534 cfu/g (Example 18; equivalent to a $\log_{10}$ geometric mean of 3.2 $\log_{10}$ cfu/g).

2.8.8.1 How does the analytical unit amount (weight, volume or area) affect the probability of acceptance?

Similar to two-class concentration-based sampling plans, the analytical unit amount is of less importance than for two-class presence-absence sampling plans. In addition, the results from enumeration methods are subject to analytical variability and hence the same advice applies here as it did for two-class concentration-based plans ("2.8.7.1 How does the analytical unit amount (weight, volume or area) affect the probability of acceptance?"). That is, the sample unit amount and analytical unit amount need to be relevant to the reason for sampling so that the microbiological enumeration results are relevant and the performance of the analytical method needs to be understood so that the results can be interpreted appropriately.

2.8.8.2 How does the level of concentration affect the probability of acceptance?

As for two-class concentration-based sampling plans, the higher the mean concentration in the food, the less likely it is for the lot to be accepted (Figure 29).

Example 18
Three-class attributes sampling plan P(accept)

Consider the sampling plan for the Aerobic Plate Count (APC) for powdered infant formula provided in ICMSF (2011) with $n = 5$, c = 2, m = 500 cfu/g and M = 5 000 cfu/g (m and M are 2.7 and 3.7 on the $\log_{10}$ scale, respectively).

Assuming that the APCs are normally distributed on the $\log_{10}$ scale with a standard deviation 0.55 $\log_{10}$ cfu/g we can use the calculations on the Three-class Concentration Tab in the companion spreadsheet to create the associated OC curve.

From the spreadsheet tool we can see that if the mean $\log_{10}$ concentration equals 1.0 $\log_{10}$ cfu/g (the arithmetic mean equals 22.3 cfu/g) then we expect that 99.90% of sample units have a concentration less than m, 0.10% have a concentration between m and M, and 0.00% of sample units exceed M. Hence the probability of acceptance equals P(accept) = 100.0%.

Again, if we are looking for the mean $\log_{10}$ concentration that yields a small P(accept), e.g. 5%, then we can change the $\log_{10}$ mean until P(accept) becomes just less than 5%. Here the probability of acceptance drops to P(accept) = 4.82% when the mean $\log_{10}$ concentration equals 3.13 $\log_{10}$ cfu/g (to two decimals); the arithmetic mean equals 3008 cfu/g. For this scenario we expect that only 21.66% of sample units have a concentration less than m, 63.29% have a concentration between m and M, and 15.05% exceed M.

A video showing you these calculations can be found at http://youtu.be/tU-RbLu_sBw.

However, the situation is now a bit more complicated as we have two limits (m and M) instead of just one (m).

So, as the mean $\log_{10}$ concentration, and hence arithmetic mean, increases, the percentage of marginally acceptable concentrations will also increase and eventually so will the percentage of unacceptable concentrations (Figure 27). When the mean $\log_{10}$ concentration is half way between the m and M, the percentage of marginally acceptable concentrations (those greater than m and less than or equal to M) will reach a maximum, while the percentage of unacceptable concentrations (those greater than M) will continue to increase as the mean $\log_{10}$ concentration increases.

It is again worthwhile to point out that when you want to make inferences about the actual number of organisms in the food, e.g. to assess the level of control and/or risk that is being achieved, then you will need to do this on the arithmetic scale and not the $\log_{10}$ scale (see also "1.2.3 Why do we use $\log_{10}$ numbers and why do we need to be careful when interpreting them?").

2.8.8.3 How do the marginal and unacceptable limits (*m* and *M*) affect the probability of acceptance?

As pointed out above, it is sensible to select values of *m* and *M* that reflect maximum tolerable concentrations under GMP and an unacceptable limit of contamination for the food, respectively. In some circumstances there may not be such natural limits, when the application of the three-class sampling plan is still desired. Dahms and Hildebrandt (1998) provide some guidance on how to determine *M* once *m* and the SD are provided, or how to calculate *m* from *M* and SD.

Similar to two-class concentration-based sampling plans the effect of *m* and *M* is relative to the mean concentration and in particular the size of the SD. The effects of different choices of *m* and *M*, while keeping the standard deviation (SD = 0.55), sample size ($n = 5$) and acceptance number ($c = 2$) fixed, can be seen from Figure 30 and Figure 31. In Figure 30 we have three OC curves for increasing values of *m* and *M*, but the distance between them remains constant at 1 $\log_{10}$. The effect is that the OC curve simply shifts to the right with increasing *m* and *M*, because higher levels of contamination are acceptable. The shape of the curve however is the same for all three curves.

In contrast, for Figure 31 we have kept the value of *M* constant and only changed the value of *m*. From this plot we can see that for the smallest value of *m*, namely $m = 2.1$, the OC curve is furthest to the left. This is due to the probability of acceptance being primarily determined by the number of analytical units that are

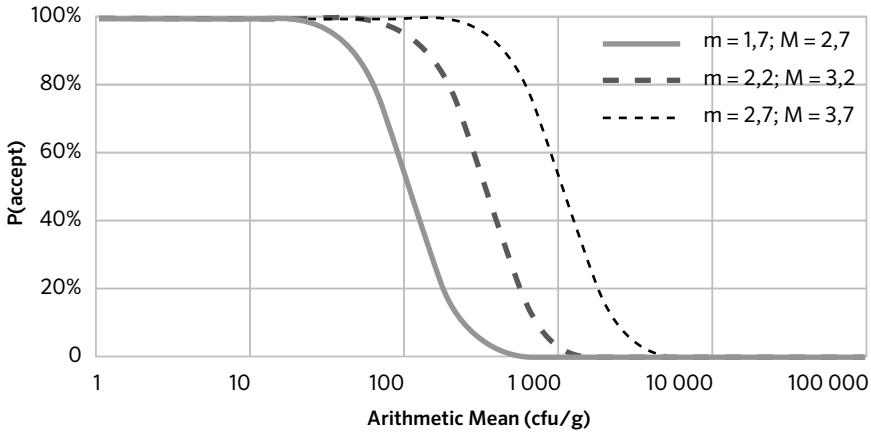


FIGURE 30: Three-class concentration-based sampling plan OC curves with $n = 5$, $c = 2$, SD = 0.55 $\log_{10}$ cfu/g for three different, equally spaced combinations of *m* and *M*.

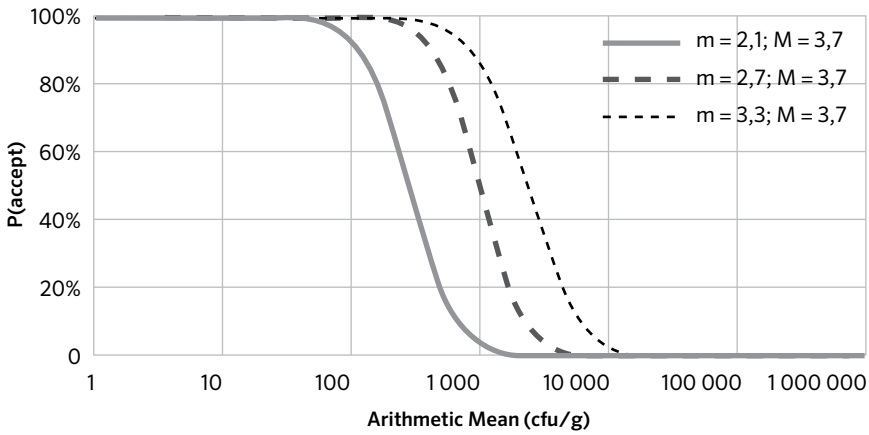


FIGURE 31: Three-class concentration-based sampling plan OC curves with $n = 5$, $c = 2$, SD = 0.55 $\log_{10}$ cfu/g for three different, unequally spaced combinations of m and M.

expected to exceed m. Because M is far away, lots are rejected by too many marginally acceptable sample units (more than $c = 2$) rather than observing unacceptable concentrations (those greater than M). This situation changes as m is chosen closer to M, which also slightly affects the shape of the OC curve. If m is chosen very close to M, then rejection happens mainly because of concentrations in the food exceeding M. In this case the sampling plan becomes similar to a two-class concentration-based plan with only a single limit.

As noted above, the effect of the distance between m and M is relative to the SD of the distribution of microorganisms in the food. From experience a difference between m and M that equals approximately 1.5 to 2.5 multiples of the SD works well in practice, e.g. 0.5 $\log_{10}$ units for well mixed foods such as milk and 1.0 $\log_{10}$ units for more heterogeneous foods such as ground meat (Dahms and Hildebrandt, 1998).

2.8.8.4 How does the variability in concentration affect the probability of acceptance?

As we have seen above, the shape of the OC curve is influenced by the distance between m and M, relative to the variability (SD) in the $\log_{10}$ concentrations found in the food. However, to explore the effect of different SDs on the OC curve we let $m = 2$ $\log_{10}$ cfu/g and $M = 3$ $\log_{10}$ cfu/g (Figure 32).

The effect of the SD is similar to the effect we saw for two-class concentration-based sampling plans, namely, the OC curve becomes steeper and hence more

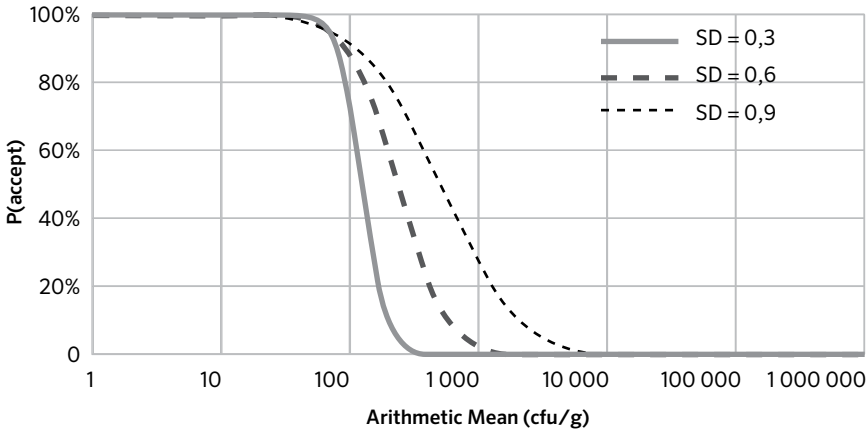


FIGURE 32: Three-class concentration-based sampling plan OC curves with $n = 5$, $c = 2$, $m = 2 \log_{10}$ cfu/g, and $M = 3 \log_{10}$ cfu/g, for three different SDs.

discriminating (between two mean concentrations) as the SD gets smaller. This is because the smaller SD allows small changes in the mean concentration to 'be more obvious'. However, unlike for the two-class concentration-based sampling plan (Figure 22) smaller SDs also result in the OC curve being shifted to the left (i.e. the sampling plan is more stringent). This effect is due to the presences of *two* microbiological limits, *m* and *M*, and in particular the lower limit *m*, as for small SDs lots will be rejected due to too many marginally acceptable analytical units (those with a concentration exceeding *m* but not *M*), rather than due to unacceptable analytical units (those with a concentration exceeding *M*).

2.8.8.5 How does the number of sample units affect the probability of acceptance?

The effect of the sample size (*n*) is again similar to that described for two-class sampling plans, namely a sampling plan with larger sample size *n* is more discriminating than one with small *n*. This is illustrated in Figure 33, where we can see that a large sample size results in an OC curve that drops quicker, i.e. is more discriminating.

2.8.8.6 How does the 'acceptance and marginal number' affect the probability of acceptance?

The last parameter that we can change in the three-class sampling plan is the acceptance number (*c*). It has a similar effect as the acceptance number in two-class presence-absence based sampling plans, namely the OC curve moves to the right (is less stringent) as *c* increases and the shape of the OC curve also changes slightly.

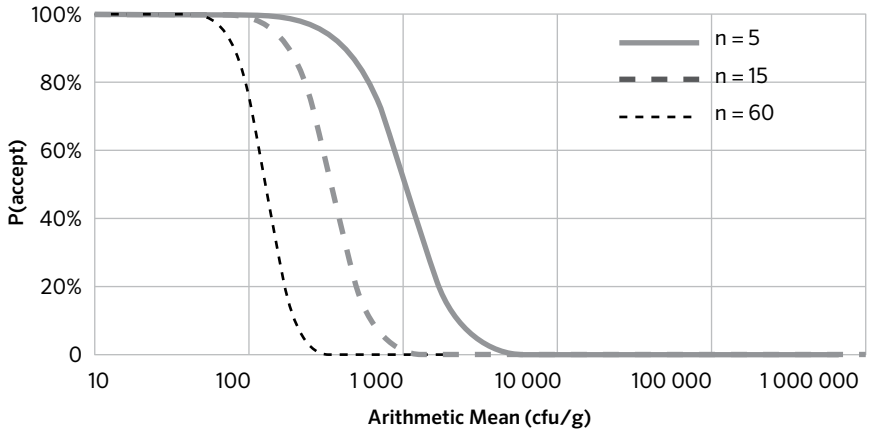


FIGURE 33: Three-class concentration-based sampling plan OC curves with $c = 2$, $m = 2.7$ $\log_{10}$ cfu/g, $M = 3.7$ $\log_{10}$ cfu/g, SD = 0.55 $\log_{10}$ cfu/g for three different sample sizes.

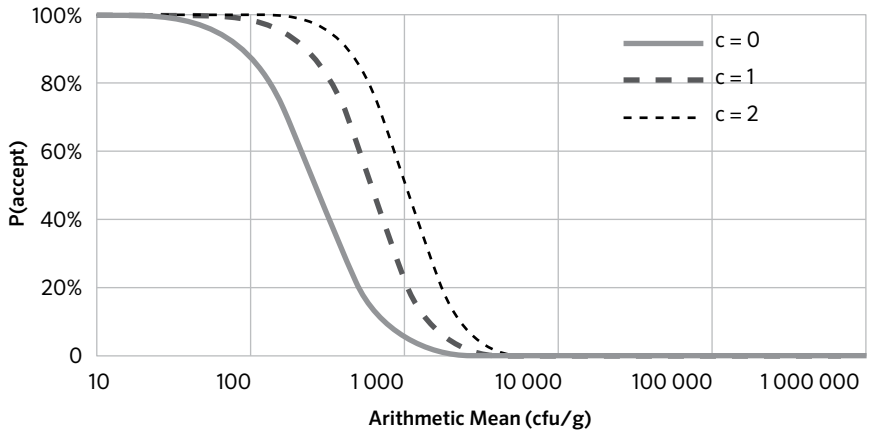


FIGURE 34: Three-class concentration-based sampling plan OC curves with $n = 5$, $m = 2.7$, $M = 3.7$, SD = 0.55 $\log_{10}$ cfu/g for three different number of marginally acceptable units (c).

Again, we have multiple parameters that influence how discriminating a sampling plan is: the sample size (n), acceptance number (c) and SD, as well as the distance between m and M. Consequently, if we have a food product with a large variability in microbial concentration, then we can compensate for this by increasing the sample size (and acceptance number). However, because of the influence of m and M on the probabilities of marginal and unacceptable concentrations it is not as simple to find equivalent sampling plans for different SDs.

2.8.8.7 Putting it all together: three-class sampling plans

Similar to the two-class concentration-based sampling plan, the three-class sampling plan can also be used when low levels of contamination are acceptable. Consequently, these sampling plans are suitable for hygiene indicator organisms and pathogens that are less likely to cause illness at low levels. However, this type of plan is employed where a scientific or operational basis exists for establishing a clear upper limit M that defines unacceptable concentrations and a marginally acceptable limit m.

The basic statistical approach, using the Three-class Concentration Tab in the companion spreadsheet, to create a suitable sampling plan that meets the consumer (and producer) risk point is largely similar to the approach illustrated for two-class concentration-based sampling plans ("2.8.7.7 Putting it all together: two-class concentration-based sampling plans"). However, the approach is more complex and may involve more trial-and-error than for the simpler two-class concentration-based sampling plans. Consequently, the steps outlined below are less 'prescriptive' than for two-class sampling plans. As noted earlier, ideally the marginally acceptable limit m and the unacceptable limit M should, as always, be determined on scientific or operational grounds. If only one of these limits is naturally available, then the other may be determined from it. Experience indicates that a difference between m and M of approximately 1.5 to 2.5 SDs works well in practice, i.e. a difference of 0.5 $\log_{10}$ units for well mixed foods such as milk and 1.0 $\log_{10}$ units for more heterogeneous foods such as ground meat (Dahms and Hildebrandt, 1998).

1. Enter a suitable value for the SD into the companion spreadsheet.
2. Decide on suitable values for the marginally acceptable limit (m) and the unacceptable limit (M) and enter these into the spreadsheet.
3. Decide on an 'unacceptable average concentration' (μ_1) in a lot so that lots with this average concentration (or greater) should be rejected most of the time.
4. Decide how frequently such lots (with average concentration μ_1 or greater) could be accepted, i.e. what is the maximum value of P(accept), call it P_1(accept), that you can tolerate when the mean concentration is μ_1.[19]

[19] The combination of μ_1 and P_1(accept) is referred to as the *Consumer's Risk Point* (see "2.8.5 What are the Producer's and Consumer's Risk Points?"). The corresponding values can be entered into cells B19 and C19 of the worksheet.

5. Let $c = 1$ (as a starting point) because we already have an unacceptable limit (M) that no sample units are allowed to exceed, so there is no point in setting c to zero.
6. Pick a practical value for n.
7. Enter the values of n and c into the companion spreadsheet.
    a. If P(accept) for your choice μ_1 is greater than the P_1(accept) you selected in Step 4, then *increase* the value of n by 1 (or more) and repeat Step 7 until the value of P(accept) is less than or equal P_1(accept).[20]
    b. If P(accept) for your choice μ_1 is less than P_1(accept) you selected in Step 4, then *decrease* the value of n by 1 (or more) and repeat Step 7 until the value of P(accept) just exceeds the P_1(accept). Now increase n by 1 to ensure that P(accept) is less than or equal to P_1(accept).
8. Check that the OC Curve is acceptable. In particular, check if at a low mean concentration (μ_0) the probability of acceptance P_0(accept) is tolerable and not too small.[21] If the probability of acceptance at μ_0 is tolerable then you are finished. If however P(accept) is too small, i.e. less than P_0(accept), then you will need to increase c by 1 and repeat the process from Step 6, including further increases in c, until the plan meets your requirements, i.e. the probability is acceptance is less than or equal to P_1(accept) at a mean concentration $\mu 1$ and greater than or equal P_0(accept) at mean concentration μ_0.

A video demonstrating this process can be found at http://youtu.be/A1Gf7X-BUsXU.

2.8.9 Variables sampling plans

Variables sampling plan have a long history although they are not as frequently used as two and three-class concentration-based attributes sampling plans (Smelt and Quadt, 1990; ICMSF, 1986; Kilsby, 1982; Lieberman and Resnikoff, 1955; Owen, 1967). The variables sampling plan is also suitable when low levels of contamination are acceptable and the target organism is expected to be present most of the time (so they can be enumerated). This type of sampling plan may be applied to hygiene indicator organisms and pathogens that are less likely to cause illness at low levels. In particular, the additional information that is considered by using the actual concentration data allows this plan to yield similar discrimination between acceptable and unacceptable lots with fewer sample units (smaller n) than a similarly performing attributes plan. As such this type of sampling plan may be

[20] The actual P(accept) shown in cell D19 will change from red to green font.

[21] The combination of μ_0 and P_0(accept) is referred to as the *Producer's Risk Point* (see "2.8.5 What are the Producer's and Consumer's Risk Points?"). The corresponding values can be entered into cells B18 and C18, respectively, in the worksheet.

preferable when sample collection is difficult or costly or when the microbiological testing is expensive. In addition, trends can be assessed better because the mean and SD of the microbial contamination in the sample units can be calculated.

The variables sampling plan is defined by

- the analytical unit amount (w), i.e. the amount of each analytical unit (mass, volume or area);
- the sample size (n), i.e. number of sample units that are collected;[22]
- the microbiological limit (m) that determines whether an analytical unit is acceptable or not acceptable; and
- the critical value k that is calculated from the sample size (n) and consumer's risk point, i.e. the combination of a tolerable probability of acceptance and associated percentage of concentrations exceeding the limit m.

As for the other concentration-based sampling plans, an estimate of the SD of the $\log_{10}$ concentration between analytical units is needed to assess the performance of the variables sampling plan. In addition, the $\log_{10}$ concentration of the target organism in the food is again assumed to be normally distributed (see Figure 4). However, if there exists evidence against this assumption, then it is possible to use alternate distributions (Takagi, 1972) although these are even less commonly used than the normal distribution in the context of variables sampling plans and are beyond the scope of this document.

The decision about whether a lot is accepted is based on how close the mean $\log_{10}$ concentration from the sample that we selected is to the microbiological limit m. If the sample mean is too close to m (given the value of the SD of the distribution of microorganisms in the food, SD), then the lot is rejected; otherwise the lot is accepted.

The construction of the variables sampling plan is not as straight forward as the creation of two or three-class concentration based attributes sampling plans, and this probably explains why they have been much less commonly used. For example, the critical value k does not have an intuitive meaning and has to be obtained from published tables, such as those by the ICMSF (2002), or calculated using the companion spreadsheet (see also Annex A1.6 Variables sampling plans"). An additional complication arises because these sampling plans can be created by assuming either that the SD in the microbial concentrations is known or that it is estimated from the sampling sample. The assumption of whether the SD is known

[22] For some sampling plans, e.g. ISO 3951, the sample size (n) is obtained from the size of the lot. However, this is done for non-statistical reason.

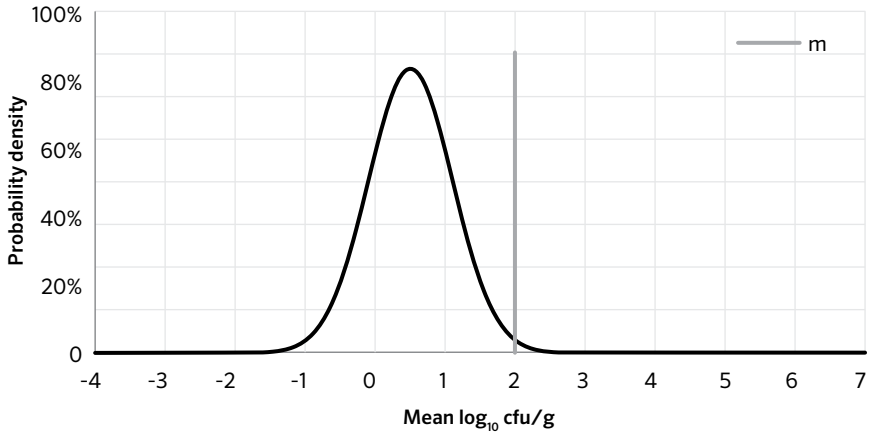


FIGURE 35: Plot of a normal distribution with mean = 0.5 $\log_{10}$ cfu/g, SD = 0.6 $\log_{10}$ cfu/g and microbiological limit m = 2 $\log_{10}$ cfu/g. The area under the curve to the right of m is the probability that the concentration in an analytical unit of the food exceeds m (see Figure 36).

or estimated affects the calculation of the critical value of k. This calculation can be performed in most spreadsheets when SD is assumed known. However, when the SD is estimated from the sample, specialized statistical functions are needed and these are generally not available in most spreadsheet software, including Microsoft Excel and LibreOffice. For these reasons we focus on the situation where SD is assumed known (not estimated from the sample), which is what we have implicitly done for two and three-class concentration-based sampling plans. With more specialized software, variables sampling plans where the SD is estimated from the sample are analogous to use and interpret as those discussed here.[23]

Similar to the two-class concentration-based sampling plan, the variables sampling plan has a single microbiological limit (m) that differentiates acceptable from unacceptable microbial concentrations in the food (Figure 6). Assuming that normality of the $\log_{10}$ concentrations holds, the critical value k can then be computed once you specify three pieces of information, namely:

- the sample size n;
- a maximum tolerable percentage (p_1) of concentrations in the food that exceed the limit m; and
- how frequently such lots (with percentage p_1 or greater of concentrations exceeding m) should be rejected, i.e. what is the maximum value of P(accept),

[23] When the standard deviation is estimated from the sample, the value of k will differ from those used here.

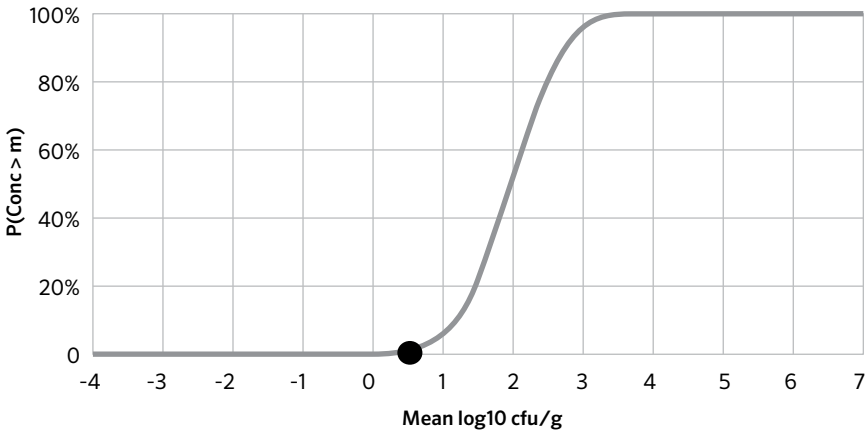


FIGURE 36: Plot of the probability that the concentration in the food exceeds $m = 2$ $\log_{10}$ cfu/g when SD = 0.6 $\log_{10}$ cfu/g. The dot point indicates the probability when the mean $\log_{10}$ concentration ($\log_{10}$ geometric mean) in the lot equals 0.5 $\log_{10}$ cfu/g.

Example 19
Variables sampling plan P(accept)

Consider a variables sampling plan for the APC with $n = 5$, $m = 100$ cfu/g (= 2 $\log_{10}$ cfu/g) that is to achieve probability of acceptance of P_1(accept) = 5% when the percentage of unacceptable concentrations equals $p_1 = 10\%$.

Assuming that the APCs are normally distributed on the $\log_{10}$ scale with a standard deviation of 0.6 $\log_{10}$ cfu/g we can use the calculations on the Variables Tab in the companion spreadsheet to create the associated OC curve.

Entering the above values in the spreadsheet we find that the critical value k = 2.017 for the specified consumer's risk point. Subsequently, when we take a sample of n = 5 units and calculate the sample mean $\log_{10}$ concentration we would reject the lot if the sample mean is greater than 0.79 (= m - k×SD = 2 - 2.017×0.6) $\log_{10}$ cfu/g, and accept the lot otherwise.

From the spreadsheet we can also see that if the mean $\log_{10}$ concentration in the lot equals 0.5 $\log_{10}$ cfu/g (arithmetic mean equals 8.2 cfu/g) then we expect to accept such lots 86.0% of the time (Point of Interest). Note that at this mean $\log_{10}$ concentration the percentage of unacceptable concentrations is only 0.62%.

A video showing you these calculations can be found at http://youtu.be/8jiab43VB94.

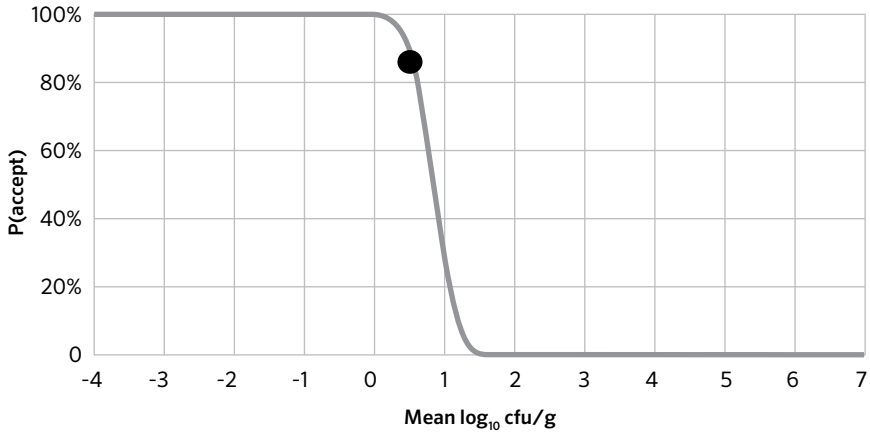


FIGURE 37: OC curve using the mean $\log_{10}$ concentration ($\log_{10}$ geometric mean) on the X-axis for a variables sampling plan with $n = 5$, SD = 0.6 $\log_{10}$ cfu/g, $m = 2$ $\log_{10}$ cfu/g and $k = 2.017$, which corresponds to a consumer's risk point with $p_1 = 10\%$ and P_1(accept) = 5%. The dot indicates P(accept) when the mean $\log_{10}$ concentration ($\log_{10}$ geometric mean) in the lot equals 0.5 $\log_{10}$ cfu/g (Example 19).

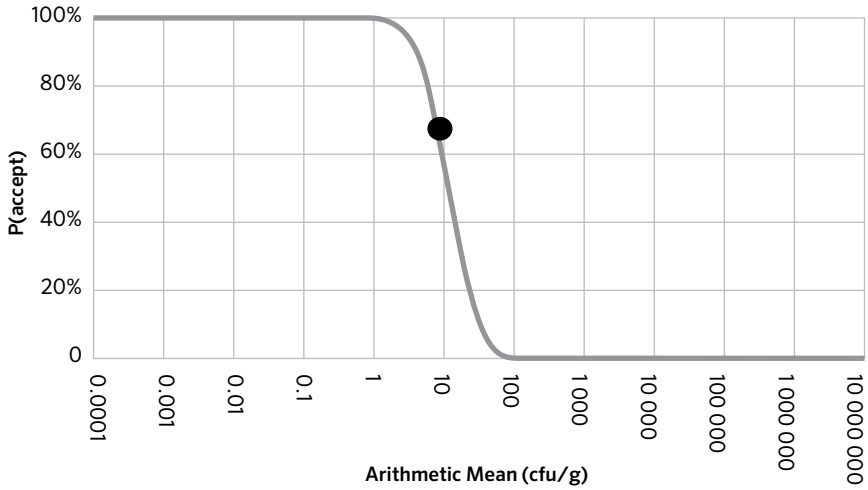


FIGURE 38: OC curve using the arithmetic mean concentration on the X-axis for a variables sampling plan with $n = 5$, SD = 0.6 $\log_{10}$ cfu/g, $m = 2$ $\log_{10}$ cfu/g and $k = 2.017$, which corresponds to a consumer's risk point with $p_1 = 10\%$ and P_1(accept) = 5%. The dot indicates P(accept) when the arithmetic mean = 8.2 cfu/g (Example 19; equivalent to a $\log_{10}$ geometric mean of 0.5 $\log_{10}$ cfu/g. See "1.2.3 Why do we use $\log_{10}$ numbers and why do we need to be careful when interpreting them?").

call it P_1(accept), that you can tolerate when the percentage of concentrations in the food exceeding m is p_1.[24]

We can now calculate the probability of how frequently sample units from the lot are unacceptable for any mean $\log_{10}$ concentration (Figure 35 and Figure 36). The sample size n and critical value k are then used to calculate the probability of acceptance, which is plotted on the Y-axis of the OC curve and either the mean $\log_{10}$ concentration or the arithmetic mean concentration are plotted on the X-axis, as shown in Figure 37 and Figure 38, respectively. We strongly suggest that you use the latter to indicate the level of control, using actual concentrations in the food, that is being achieved (see also Example 4 in "Why do we use $\log_{10}$ numbers and why do we need to be careful when interpreting them?"). The mathematical details are provided in Appendix A1.6 Variables sampling plans" and the calculations of P(accept) and associated graphs are provided in the Variables Tab of the companion spreadsheet. The spreadsheet also contains the calculation of the critical value k for a given consumer's risk point and sample size, as well as the maximum sample mean that would allow a lot to be accepted. The calculations that were used to generate Figure 35 to Figure 38 are shown in Example 19.

2.8.9.1 How does the analytical unit amount (weight, volume or area) affect the probability of acceptance?

Similar to two and three-class concentration-based sampling plans, the analytical unit amount is of less importance than for two-class presence-absence sampling plans. In addition, the results from enumeration methods are subject to analytical variability and hence the same advice applies here as it did for two-class concentration-based sampling plans ("2.8.7.1 How does the analytical unit amount (weight, volume or area) affect the probability of acceptance?"). That is, the sample unit amount and analytical unit amount need to be relevant to the reason for sampling so that the microbiological enumeration results are relevant and the performance of the analytical method needs to be understood so that the results can be interpreted appropriately.

2.8.9.2 How does the level of concentration affect the probability of acceptance?

The effect of changes in the mean $\log_{10}$ concentration, and hence arithmetic mean, on the probability of acceptance depends on the various parameters of the sampling plan, including the sample size, consumer's risk point (these determine the critical value k) and the SD in the $\log_{10}$ microbial concentrations. Nevertheless, just as was the case for two and three-class concentration-based sampling plans, the higher the mean concentration in the food, the smaller the probability of accepting the lot (Figure 37).

[24] The combination of p_1 and P_1(accept) is referred to as the *Consumer Risk Point*.

However, **it is worthwhile to pointing out that you can reject a lot even if none of the individual concentrations in the sample exceed the microbial limit *m***. This can happen if the sample mean concentration is too close to the limit, i.e. exceeds the maximum allowable sample mean for the specified SD, sample size and consumer's risk point. The fact that a lot can be rejected without any of the sample concentrations exceeding *m* is likely to be another source for confusion and indicates why variables sampling plans have not been readily adopted.

It is again important to note that when you want to make inferences about the actual number of organisms in the food, e.g. to assess the level of control that is being achieved, then you will need to do this on the arithmetic scale and not the $\log_{10}$ scale (see also "1.2.3 Why do we use $\log_{10}$ numbers and why do we need to be careful when interpreting them?").

2.8.9.3 How does the unacceptable limit (*m*) affect the probability of acceptance?

As we have seen for the various concentration-based sampling plans, the effect of the value of *m* is integrally connected with the mean and SD. However, if we have a look at the OC curves in Figure 39, then we can see that increasing the value of *m*, while holding the SD and other parameters (*n* and *k*) constant, has the effect of shifting the OC curve to the right without affecting the shape of the OC curve.

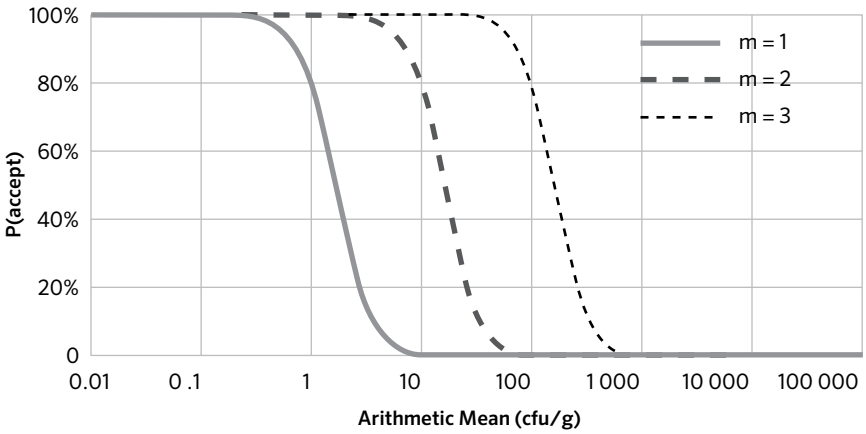


FIGURE 39: Variables plan OC curves for three different unacceptable limits (*m*) with *n* = 5, SD = 0.6 $\log_{10}$ cfu/g and *k* = 2.017, which corresponds to a consumer's risk point with p_1 = 10% and P_1(accept) = 5%.

2.8.9.4 How does the variability in concentration affect the probability of acceptance?

As for other sampling plans, a smaller SD results in a steeper OC curve and hence makes the sampling plan more discriminating. The effect of changing the SD on the OC curve, while keeping m and k constant is shown in Figure 40. However, unlike for the two-class concentration-based sampling plan (Figure 22)smaller SDs result in the OC curve also being shifted to the right. This effect is because the SD affects the proportion of unacceptable concentrations as well as the variability of the sample mean, known as the standard error, which is used to determine lot acceptance and rejection. Consequently, as the SD increases, the maximum tolerable sample mean reduces. Thus, for the SDs of 0.3, 0.6 and 0.9 $\log_{10}$ cfu/g used for Figure 40 these maximum tolerable sample means are 1.39, 0.79 and 0.18 $\log_{10}$ cfu/g, respectively.

2.8.9.5 How does the number of sample units affect the probability of acceptance?

The effect of the sample size (n) is again similar to the other sampling plans, that is, a sampling plan with a larger n is more discriminating than one with a small n. This is illustrated in Figure 41, where we can see that a large sample size results in an OC curve that drops quicker. However, it is worthwhile to point out that greater discrimination can be achieved with much smaller sample sizes than for two or

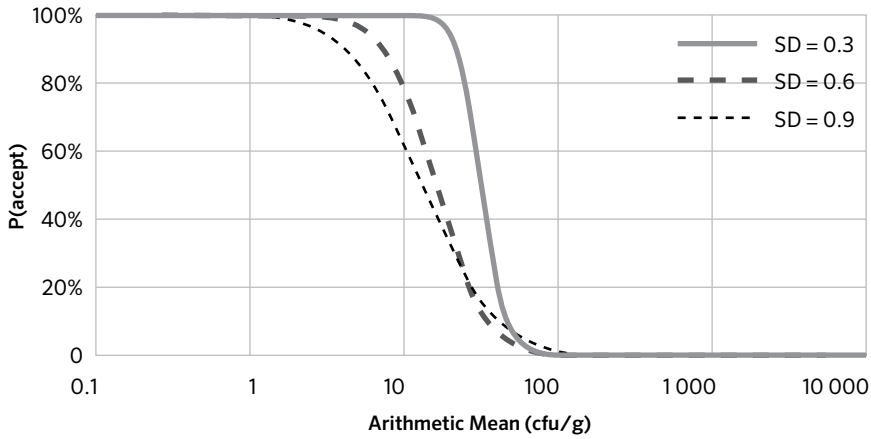


FIGURE 40: Variables plan OC curves for three different SDs with $n = 5$, $m = 2$ $\log_{10}$ cfu/g and $k = 2.017$, which corresponds to a consumer's risk point with $p_1 = 10\%$ and $P_1(\text{accept}) = 5\%$.

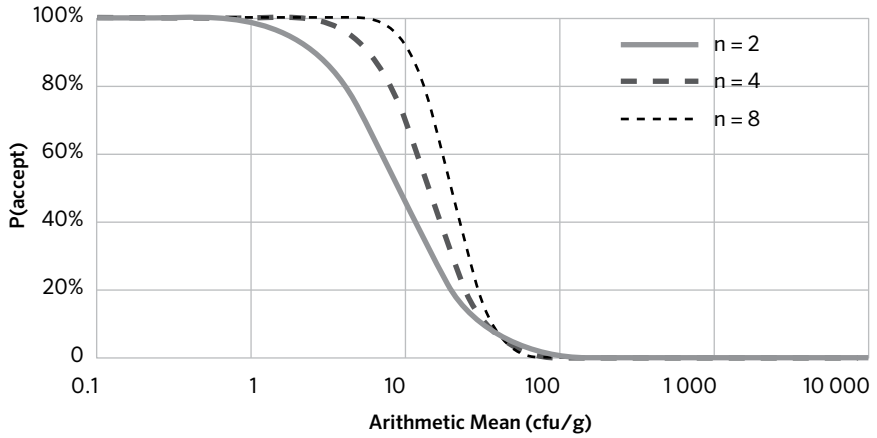


FIGURE 41: Variables plan OC curves for three different sample sizes (n) with SD = 0.6 $\log_{10}$ cfu/g, m = 2 $\log_{10}$ cfu/g and k = 2.017, which corresponds to a consumer's risk point with p_1 = 10% and P_1(accept) = 5%.

three-class attributes sampling plans. This is because for variables sampling plans we are using all the information available to us in the form of $\log_{10}$ concentrations, instead of losing information by categorizing the $\log_{10}$ concentration into acceptable, marginal or unacceptable. Using all available information is preferable and more cost effective over using less information.

It is for this reason that variables sampling plans can achieve similar performance to attributes sampling plans with much smaller sample sizes, but at increased complexity. For example, consider the two OC curves in Figure 42 which are practically identical. Both are based on $\log_{10}$ normal distributions with SD = 0.6 $\log_{10}$ cfu/g and a microbiological limit of m = 2.0 $\log_{10}$ cfu/g. The two-class concentration-based OC curve results from a sampling plan with n = 60 and c = 2. At an arithmetic mean concentration of 26 ($\log_{10}$ geometric mean of 1.0 $\log_{10}$ cfu/g) we have a proportion of concentrations exceeding m of p_1 = 4.78%. The resulting probability of acceptance is P_1(accept) = 44.86%. There is nothing special about these values other than p_1 is very close to 5% and they suit this discussion. To generate a variables sampling plan that is equivalent to the two-class sampling plan we let n = 5 (as a starting point) and use the above calculated values of p_1 and P_1(accept) to specify the consumer's risk point. This yields a critical value k = 1.724, but the OC curves do not quite agree. After a few changes of n we arrive at n = 12, k = 1.704 and the OC curves shown in Figure 42. For all intents and purposes the two sampling plans are equivalent though

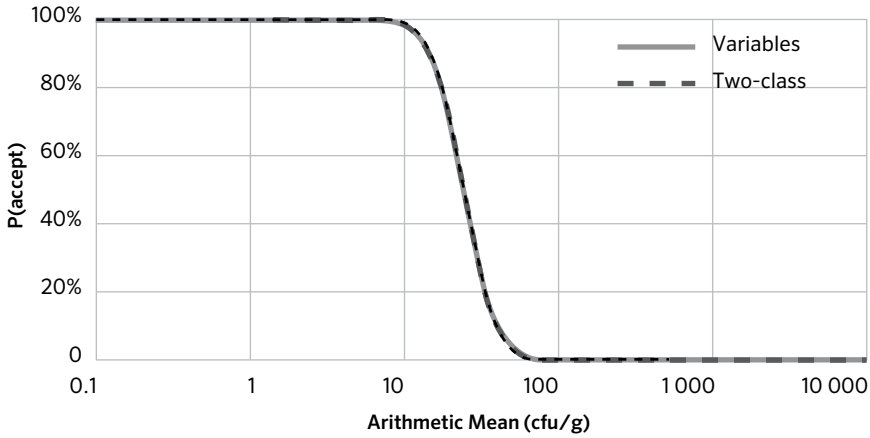


FIGURE 42: OC curves for two different sampling plans for a food with SD = 0.5 $\log_{10}$ cfu/g and unacceptable limit m = 2.0 $\log_{10}$ cfu/g. The two-class concentration-based sampling plan is for a sample size n = 60 and c = 2 and yields a consumer's risk point p_1 = 4.78% and P_1(accept) = 44.86%, while the corresponding variables sampling plan has n = 12 and k = 1.704.

the variables sampling plan requires only a fifth (n = 12) of the number of sample units compared to the two-class concentration-based plan (n = 60). As pointed out above, the variables sampling plan is therefore more cost efficient than the two-class sampling plan because it uses all the available information.

2.8.9.6 How does the consumer's risk point affect the probability of acceptance?

To explore the effect of changing the Consumer's Risk Point (see "2.8.5 What are the Producer's and Consumer's Risk Points?") on the OC curve, we look at two separate cases – the effect of changing the percentage of unacceptable concentrations p_1 while keeping P_1(accept) fixed and vice versa. The resulting OC curves are shown in Figure 43 and Figure 44. Because the consumer's risk point influences the value of k, it follows that the shape of the OC curve changes. However, as can be seen from the two figures, this change is barely perceptible. Instead the main effect of increasing either p_1 or P_1(accept) is to shift the OC curve to the right, i.e. make the sampling plans less stringent.

2.8.9.7 Putting it all together: variables sampling plans

The variables sampling plan can be used when low levels of contamination are ac-

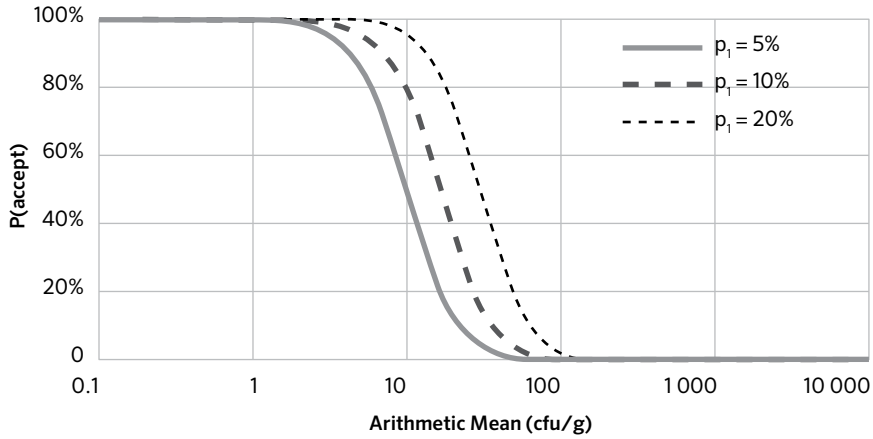


FIGURE 43: Variables plan OC curves for three different p_1 values of the consumer's risk point specifications with $n = 5$, SD = 0.6 $\log_{10}$ cfu/g, $m = 2$ $\log_{10}$ cfu/g and P_1(accept) = 5%.

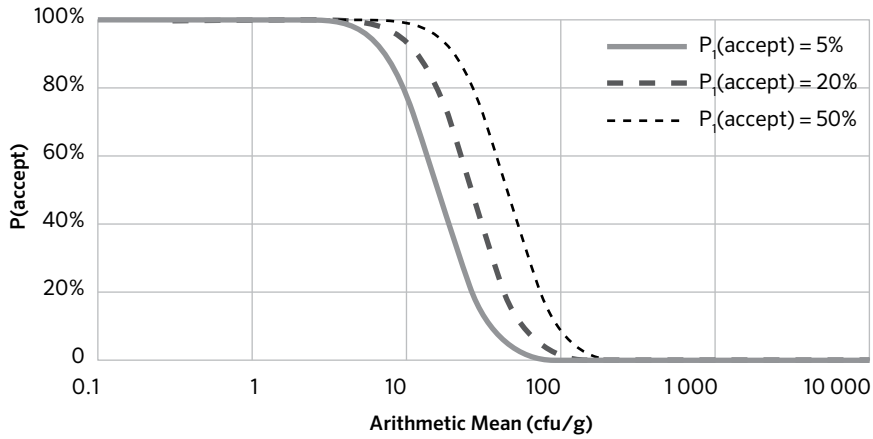


FIGURE 44: Variables plan OC curves for three different P_1(accept) values of the consumer's risk point specifications with $n = 5$, SD = 0.6 $\log_{10}$ cfu/g, $m = 2$ $\log_{10}$ cfu/g, and $p_1 = 10\%$.

ceptable and the target organism is expected to be present most of the time (so it can be enumerated). Consequently, this type of sampling plan is applicable for use with hygiene indicator organisms and pathogens that are less likely to cause illness at low levels.

As we have seen above, the additional information, i.e. mean and SD, that is utilized with this type of sampling plan, allows similar discrimination between acceptable and unacceptable lots with much fewer sample units (smaller *n*) than a comparable attributes plan. As such the variables sampling plan may be preferable especially when sample collection is difficult or costly or when the microbiological testing is expensive. In addition, trends can be assessed better (see Part 3) because the mean and SD of the microbial contamination are available.

The basic statistical approach, using the Variables Tab in the companion spreadsheet, to determine a suitable sampling plan is as follows. As before, this approach should be applied in the context for which the MC is to be developed.

1. Decide on an 'unacceptable percentage' (p_1) of concentrations exceeding *m*, so that lots with this percentage (or greater) should be rejected most of the time.
2. Decide how frequently such lots (with percentage p_1 or greater of concentrations exceeding *m*) could be accepted, i.e. what is the maximum value of P(accept), call it P_1(accept) that you can tolerate when the percentage is p_1.
3. The combination of p_1 and P_1(accept) is referred to as the *Consumer's Risk Point* (see "2.8.5 What are the Producer's and Consumer's Risk Points?"). Enter these into the cells for P(Conc > *m*) and P(accept) in the Consumer's Risk Point section of the companion spreadsheet.
4. Enter a suitable value for the SD and the unacceptable limit (*m*) into the companion spreadsheet.
5. Let $n = 2$ be the starting point.
6. Enter the value of *n* into the companion spreadsheet. The spreadsheet automatically calculates the appropriate value of *k* and the mean concentration (μ_1) that is associated with the percentage p_1. The plan will meet your specified consumer's risk point.
7. Check if at a lower mean concentration (μ_0) the probability of acceptance P_0(accept), the *Producer's Risk Point*, is tolerable and not too small (these values can be entered as the 'Point of Interest'). If the probability of acceptance is tolerable then you are finished. If however P(accept) is too small, i.e. less than P_0(accept), then you will need to increase *n* by 1 and repeat the process from Step 6 until the plan meets your requirements, i.e. the probability of acceptance is greater than or equal to P_0(accept) at mean concentration μ_0.

A video demonstrating this process can be found at http://youtu.be/8jiab43VB94 (see also Example 19).

The spreadsheet also provides the necessary information for you to decide whether a lot should be accepted or rejected under "*Reject lot if sample mean (log_{10}) exceeds.*" So, once you have collected a sample of size *n,* you calculate the mean $\log_{10}$ concentration. If the sample mean exceeds the calculated value, then the lot is rejected; otherwise the lot is accepted. The process is illustrated in Example 20.

Example 20
Variables sampling plan: accepting or rejecting a lot

Consider a variables sampling plan with $m = 2$ $\log_{10}$ cfu/g, SD = 0.6 $\log_{10}$ cfu/g, $n = 3$ and a consumer's risk point with $p_1 = 10\%$ and $P_1(\text{accept}) = 5\%$.

Using the Variables Tab in the companion spreadsheet we calculate that $k = 2.231$ and that the maximum tolerable sample mean is 0.66 $\log_{10}$ cfu/g.

We now collect a sample and obtain the following $\log_{10}$ concentrations: 0.2, 0.8 and 1.4 $\log_{10}$ cfu/g. The sample mean equals 0.8 $\log_{10}$ cfu/g and hence we would reject the lot because this value exceeds the maximum tolerable sample mean of 0.66 ($m - k \times SD = 2 - 2.231 \times 0.6$) $\log_{10}$ cfu/g.

Notice however that none of the individual sample units exceed the microbiological limit $m = 2$, and it may seem incorrect to reject the lot based on these results. **However, the large sample mean provides evidence to indicate that a much greater percentage than p_1 (= 10%) of unacceptable concentrations are present in this lot (given the SD specified), even if we did not capture any of them in our sample**.

PART 3

Making decisions related to process verification

In this part we explore how you can use MC for process verification and statistical process control. We first briefly look at food safety control systems and what we mean by process verification. We then consider some approaches to statistical process control and in particular *moving windows* (given the specific mandate from the CCFH) and how they can be applied together with the information in Parts 1 and 2 as part of process control activities and process verification.

3.1 WHAT IS MEANT BY FOOD SAFETY CONTROL SYSTEM?

Codex defines the term *Food Safety Control System* as

> *"The combination of control measures that, when taken as whole, ensures that food is safe for its intended use."* (CAC, 2008b)

Consequently, all processing aspects that relate to ensuring that the final food product is safe to consume are included in the control system. Therefore, a food safety control system may be applied on a food business basis, e.g. all the business processes and procedures that ensure the safety of the food produced by that business.

When designing a food safety control system, the following nine points outlined by the ICMSF (2002), should be considered.

1. Knowledge of the significant hazards
2. Knowledge of the factors that are necessary for control
3. Knowledge of the extent of variability and factors that influence variability
4. Establishing criteria for the factors that must be controlled
5. Establishing monitoring procedures
6. Organizing and interpreting data
7. Using the data to improve control and measure change
8. Responding to the data
9. Investigating and learning from previously unrecognized factors or unforeseen events

3.1.1 What is meant by performance of a food safety control system?

Performance of a food safety control system is related to how capable the system is to consistently control the hazard(s) to a specified and validated outcome. A well performing food safety control system will be able to consistently achieve a Performance Objective or Food Safety Objective (CAC, 2013a).

3.1.2 When is our process under control?

In essence, a process that is under control will consistently operate the way it was designed to. Consequently, monitoring and verification activities demonstrate that the predefined control measures are and have been operating as intended. In addition, any corrective actions that are applied as a result of detecting significant process deviations, i.e. when the process has gone out-of-control, are achieving the desired effect.

A process that is under control is predictable to a certain degree, depending on the variability that is inherent in the process. Note that control does not imply small variability, as variability will depend on the process and the food product being manufactured. Nevertheless, for a process to be predictable, it is desirable that the variability in microbial concentrations between production lots (called between-lot variability) is small compared to the variability within a lot (within-lot variability). An important part of process control is to understand the sources that contribute to the within and between-lot variability and to quantify their effects.

As part of continuous process improvement you should aim to reduce the variability in the food. This will make your food product more consistent and microbial concentrations more predictable. Reducing variability should be the focus only once control has been achieved. There is no point in trying to reduce variability when you have no understanding of the factors that influence the variability and

when your process is fluctuating widely and inexplicably. You must first understand those production factors that affect variability and understand how they do this before trying to make changes to a process That is, you really need to "know your process."

3.2 WHAT IS MEANT BY VERIFICATION?

Codex defines *verification* as:

> *"The application of methods, procedures, tests and other evaluations, in addition to monitoring, to determine whether a control measure is or has been operating as intended."* (CAC, 2008b)

There are different ways to verify that a process is operating as intended, in the current context, verification means using an MC as a tool to demonstrate that the upstream process is under control. This includes using the MC to identify breakdowns in process control that compromise product safety. It is the responsibility of the FBO to investigate, identify and rectify the sources or factors associated with the loss of control.

3.2.1 What verification is not

Given the above Codex definition of verification we can also identify instances that do not constitute verification, such as

- monitoring of a Process Criterion or Critical Control Point (CCP);
- validation of a food safety control system;
- an unplanned, ad hoc activity involving microbiological sampling and testing; or
- lot-by-lot testing for product release.

However, with respect to the last bullet point, the accumulated data can be used for trend analysis and to verify that the food safety control system is operating as intended. That is, verification is an on-going activity incorporating relevant data, rather than making decisions on individual lots.

3.2.2 What is the difference between lot-by-lot testing and verification testing?

Acceptance sampling has two main effects on food control:

1. It is used directly to determine the disposition of a lot, e.g. testing beef trim for *E. coli* O157. This direct use has an effect on tested lots (that are rejected), but no effect on lots that are not tested.

2. It is used indirectly by regulatory bodies and customers to provide an incentive to producers to control the food production process to avoid rejection or recall of lots and any associated costs or consequences of lot rejection, e.g. increased sampling.

Verification testing is more commonly used than acceptance sampling and applications of verification testing for food control include:

- Testing prescribed or carried out by regulatory bodies or buyers as part of commercial agreements. This type of verification testing is similar to acceptance sampling where usually not every lot is sampled and tested.
- Testing conducted by the FBO to verify that the food production system is operating as expected. This form of verification testing is generally an important component of demonstrating process control (see above).

The remainder of Part 3 is dedicated to the application of verification testing as part of process control activities.

An important distinction between acceptance sampling and process control verification testing is the extent to which the action is considered reactive or proactive risk management. Modern food safety control is, where possible, directed at proactive risk management rather than reactive risk management. The indirect application of acceptance sampling, via the economic incentive to avoid rejection, is proactive in that the effect is on every lot of production, through improved processes and practices. Industrial or statistical process control is the most proactive risk management tool. This type of process control relies heavily on the collection, analysis and interpretation of data and is aimed at reducing variability by systematically detecting and addressing significant process deviations. Ideally this allows the process to be controlled appropriately and to prevent major food safety failures.

3.2.3 Can Microbiological Criteria be used for both safety and quality?

While the sampling plans described in Part 2 can be used for verification of both food safety and quality, it is recommended that the two verification aims are kept separate and that MC are used primarily for food safety, as seen from the Codex definition (CAC, 2013a).

Because of the practical implications, such as costs, MC are primarily used to verify the food safety expectations of a specific food process.

3.2.4 What are some of the benefits of being able to verify process control?

By using and demonstrating effective process control, regulators, customers and the general public may gain confidence in the safety of the food supply. This increased confidence can have a number of tangible benefits for the FBO, such as;.

- access to markets that demand validated and verified food safety control systems;
- demonstration of an appropriate duty-of-care, i.e. due diligence, through proactive food safety risk management;[25]
- potential for reduced sampling and testing costs, because of increased customer confidence; and
- reduced effects during crisis situations, i.e. in a recall situation, because better information is available to make informed decisions.

3.2.5 Where along the food chain can process control verification be applied?

Process control verification can be applied at any point in the food chain. At a chosen point, it can be used to measure the performance of *upstream components* of the food chain, i.e. what happened before the testing point. However, it does not indicate the level of performance of any *downstream components* of the food chain, i.e. what happens to the food after the testing point.

3.2.6 What is being measured when applying a Microbiological Criterion at a certain point in a food system?

When we apply an MC at a point in the supply chain to establish the microbiological status of a component in the food production system then we obtain information on the safety of the process upstream from the point of sampling. That is, we assess the cumulative history of every step that happened before the point of sampling and not just information on that particular process step.

Consequently, the information generated can be used to verify the functioning of the process upstream. However, for the same reason, when an MC is applied at a point in the process no information about what happens after this point, i.e. downstream, is generated. So, if you want to develop a Performance Objective (CAC 2008a) at a processing step such that it links with a Food Safety Objective (CAC 2013b), then you will require additional information about what is expected to happen to the microbial levels in the food downstream. A useful resource for this is the web-based BASELINE project (http://baselineeurope.eu/), which provides a

[25] Utility depends on the legal situation / interpretation used in a country.

decision-support tool for sampling plan selection and can help you assesses how to link these sampling plans with a Performance Objective or Food Safety Objective.

3.3 WHAT PROCESS CONTROL APPROACHES ARE AVAILABLE?

The term *process control* can have a different meaning for different people. In the context that we are using it here we are primarily referring to statistical process control (SPC), that is, using data and statistical methods to make informed decisions about the ability of the process to produce safe food. Unfortunately, we cannot give this area a comprehensive treatment as it beyond the scope of this document. However, there are many books available for those who are interested in the subject (Montgomery, 2012; Wheeler, 2010), although most tend to focus on general manufacturing processes rather than on food microbiology specifically. Exceptions are books by the ICMSF (ICMSF, 2002; ICMSF, 2011), which cover some of the statistical process control concepts, as well as the work by the AOAC International on best practices in microbiological methods, especially Appendix F (AOAC International, 2006).

Before we get into some statistical process control approaches, we note that if you are already performing lot-by-lot testing, then these data can be used for the purposes of statistical process control without additional cost. For this reason, we suggest, where microbiological testing involves enumeration methods, that you use the actual concentration data generated by microbiological testing, instead of only the category that a test result falls into, e.g. acceptable, marginally acceptable or unacceptable, even if you are using two-or three-class sampling plans. There will be additional costs to cover the time needed for graphing, analysing and drawing conclusions, but we anticipate that these costs are small once the system has been established and when compared with the long-term benefit gained from a process that is better controlled.

3.3.1 What is meant by trend analysis?

In the broadest sense *trend analysis* is an approach to detect a pattern, or changes in a pattern, in microbiological data that have been collected over a period of time. Often this involves relatively long periods, e.g. days or weeks depending on the frequency of data collection and the underlying temporal patterns in the process.

Trend analysis can be applied to many types of data collected and recorded in a process, including the results of microbiological testing and how these results compare against an MC. Consequently, trend analysis may be able to detect loss

of process control due to slow and gradual changes or to larger and more sudden changes in the microbial concentrations in the process.

Trend analysis may show changes or patterns in the data that are a result of unwanted changes or events in the manufacturing process, enabling the FBO to take corrective actions before the process is out of control. The trends, or patterns,

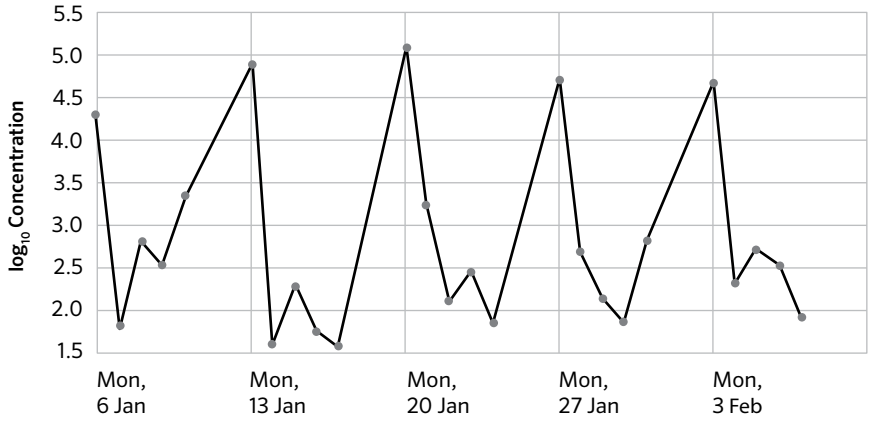


FIGURE 45: Example of a cyclic pattern using date on the X-axis. In this process the $\log_{10}$ concentrations are higher on Mondays than any other day of the week, the reason for which should be investigated.

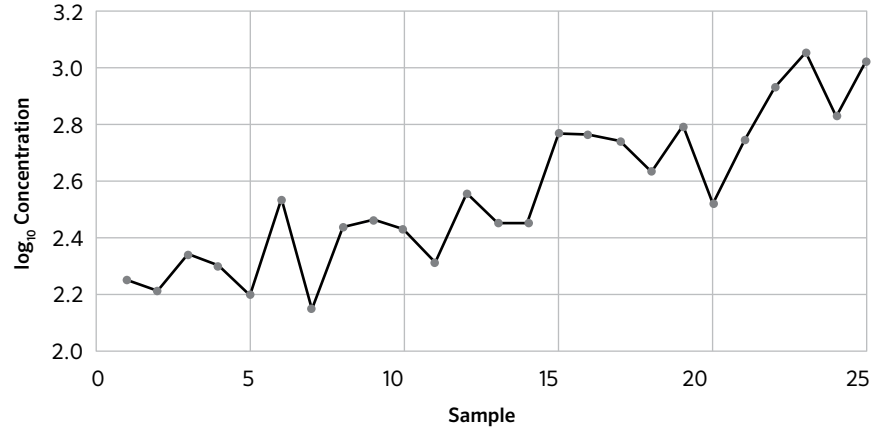


FIGURE 46: Example of increasing microbial concentrations over time, where sample number is plotted on the X-axis.

can be visualized by plotting the test results over time as shown in Figure 45 and Figure 46. As can be seen from these two plots, the time aspect (X-axis) can be displayed differently depending on the circumstances, e.g. date, time or sample number.

3.3.2 What are control charts?

Control charts represent one type of tool to conduct trend analysis in statistical process control. A control chart is a plot of the data over time together with a centreline and lower and upper control limits (LCL and UCL). An example control chart is shown in Figure 47. The centreline is used to identify the process average (mean) and the control limits are used to show the extent of the natural variability in the process. These are usually chosen to lie three SDs (mean ± 3 SD) on either side of the centreline. Assuming the $\log_{10}$-normal distribution holds, only a very small proportion of observations (0.27% or 2.7 in every 1 000 sample units) are expected to fall outside these control limits by chance alone.

Test results within the control limits indicate that the process is in statistical control: they are part of the *common cause variation*. In contrast, points that fall outside the control limits are usually referred to as *special causes,* or due to *special cause variation*, and their reason should be investigated. In addition, patterns

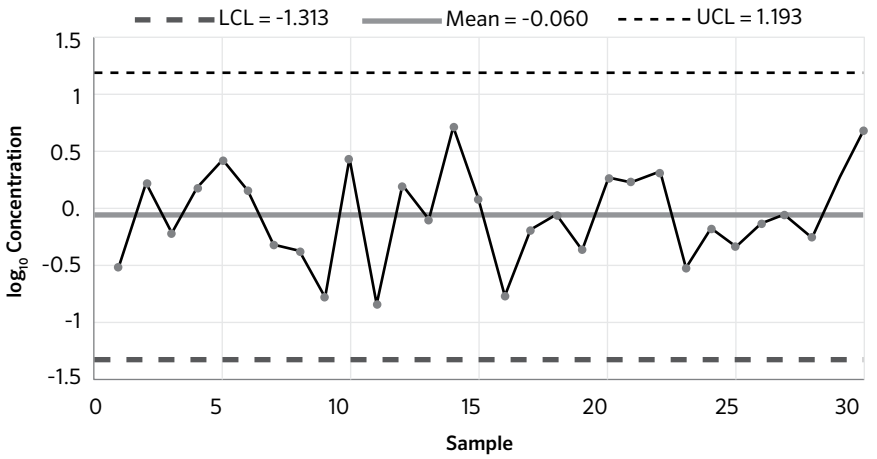


FIGURE 47: Example of a control chart showing the average $\log_{10}$ concentration as the solid centreline (mean). The lower and upper control limits (dashed) indicate the extent of the natural variability (mean ± 3 SD) in $\log_{10}$ concentrations that is inherent in the process.

seen in control charts can also be examples of special cause variation. Decision rules such as the Western Electric Rules (W.E.C., 1956) or Nelson's Rules (Nelson, 1984) can be used to identify significant patterns in control charts, e.g. gradual and sudden changes in the mean or variability, that indicate a loss of statistical control.

It should be noted that the control limits are calculated from the microbiological concentration data obtained directly from the process and these limits have nothing to do with the *specification limit(s)* that might be specified as part of an MC by the FBO or a customer. Therefore, a process may be in statistical control, yet may not be able to meet the specifications that have been set for the food product – the process would be considered to be 'not capable'. The opposite is also possible, that is, a process may not be in *statistical* control, but be 'capable' of meeting product specifications.

Control charts have two essential uses. Firstly, providing an on-going analysis of process performance and secondly gaining, monitoring and maintaining control of a process. Common types of control charts include:

- Individual and Moving Range charts for tracking individual microbial concentrations;
- Mean (X-bar) and Range (R) charts for tracking groups (samples) of measurements – such as those collected as part of variables sampling plan;
- p- and np-charts for tracking proportions or analytical unit detection probabilities, e.g. detect or not detect, such as those collected as part of two- or three-class[26] sampling plans (presence/absence as well as concentration); and
- u- and c-charts for tracking frequency of certain occurrences or events.

In many situations more than one type of control chart may be applicable although we prefer those that make full use of microbial concentrations as pointed out previously. Detailed information about the various control charts and how to establish them can be found in texts on statistical quality control (Montgomery, 2012; Wheeler, 2010).

It should be noted that statistical process control techniques help to provide experience in 'process thinking', which is a central tenet of Hazard Analysis and Critical Control Points (HACCP). These techniques are used to develop a historical record of performance, evaluate the long-term stability of a process and determine process

[26] For three-class sampling plans we could either plot the proportion of marginally acceptable and unacceptable results, i.e. the proportion of all sample units that are not 'acceptable', or consider only the proportion of unacceptable sample units.

capability, i.e. whether the process can actually meet specifications most of the time, as well as judge the effectiveness of process improvement actions.

3.3.3 What actions are typically taken in response to process control deviations?

As part of setting up your food safety control system and control charts, you should develop an *Out-of-Control Action Plan* (OCAP). This plan can be in the form of a flow chart and it provides guidance on what you should do, and in what order, when the process goes out of control. The OCAP is predicated on the assumption that you know your process well, that you can conduct a Root Cause Analysis (see for example Rooney and Vanden Heuvel, 2004), and take corrective actions, e.g. sanitation, to bring the process back into control.

In addition, you should take measures to prevent the recurrence of any unacceptable microbiological contamination. Those measures may include modifications to the HACCP-based procedures, process re-design, equipment alterations or other food hygiene control measures.

In combination, these techniques are used proactively to improve the long-term stability and control of the process.

3.3.4 What is the relationship between trend analysis, control charts and the moving window approach?

As we have pointed out above, trend analysis is an all-encompassing term that is used to describe the process of looking for temporal patterns in data. Both control charts and the moving windows approach are more formal methods of trend analysis that also include decision rules, i.e. how significant patterns are detected and identifying when action is required.

3.4 MOVING WINDOWS

3.4.1 What are moving windows?

The moving window approach is a formalized method, used in process control, that uses a decision rule to determine when the process is considered to be out of control. The moving window approach has been utilized in various meat slaughter regulations, such as the USDA's *Pathogen Reduction/HACCP Systems rule* (USDA FSIS, 1996), the European Union's process hygiene criteria for *Salmonella* on carcasses (EC, 2005), or the Australian *E. coli* and *Salmonella* Monitoring (ESAM) program (AQIS, 2003), as well as for juice production (USFDA, 2001).

In the moving window approach a number of sample units (n) is collected over a period of time, i.e. the w*indow*. Every time a new test result becomes available, it is included in the window and the oldest observation is removed, so that only the most recent n test results are in the window and thereby the window moves; hence the term *moving window*. When a new test result is added to the window, the n observation in the window are compared with the microbiological limit(s) (m,

Example 21
Moving windows

The moving window approach has been utilized in various meat slaughter regulations, such as the USDA's *Pathogen Reduction/HACCP Systems rule* (USDA FSIS, 1996) or the Australian *E. coli* and *Salmonella* Monitoring (ESAM) program (AQIS, 2003).

For example, under the ESAM program, the moving window for *E. coli* on cow/bull carcasses was based on a three-class sampling plan with n = 15, c = 3, m = 0 cfu/cm^2 (i.e. a detection of *E. coli* exceeds m) and M = 20 cfu/cm^2. Sample units of a total of 300 cm^2 are obtained at a frequency of 1 random carcass in every 300 slaughtered.

Consequently, a slaughter establishment that slaughters 900 animals per day would accumulate 3 new test results each day and hence the moving window would cover the previous five days (n = 15 sample units) of production. For example, consider the following sets of three sample results (cfu/cm^2) collected over five days.

Week 1 - Monday:	0, 0, 0
Week 1 - Tuesday:	0, 0, 0
Week 1 - Wednesday:	0, 0, 12
Week 1 - Thursday:	0, 5, 0
Week 1 - Friday:	0, 15, 0

Based on these results the process is considered to be in control – three detections of *E. coli* were recorded in Week 1, all below the limit of M = 20, which is acceptable.

On Monday in Week 2 a new set of three sample units is collected and when these become available they are added to the window and the Week 1 - Monday sample units are dropped. The window now looks as follows:

Week 1 - Tuesday:	0, 0, 0
Week 1 - Wednesday:	0, 0, 12
Week 1 - Thursday:	0, 5, 0
Week 1 - Friday:	0, 15, 0
Week 2 - Monday:	1, 0, 0

The process is now considered to be out of control as we have four detections i.e. all exceed m = 0). Alternatively, the process would also be considered out of control if one result exceeded the unacceptable limit M.

M) using the acceptance number (*c*), similar to the application of two- and three-class concentration-based sampling plans. The moving window approach can also be applied to a set of results, e.g. when multiple tests are carried out in a day, so that more than one, but less than *n*, observations are added at every update of the window, as shown in Example 21.

Spreading the *n* sample units out over time, makes the moving window approach practical and cost effective when verifying process control. In addition, this approach can be used with MC in a similar way to acceptance sampling, though interest now lies in the acceptability of the process, rather than individual lots, so that appropriate interventions can be initiated when unacceptable shifts in the process occur.

The moving windows approach is commonly used in combination with two-class and three-class attributes sampling plans, largely because they frequently require a large sample size (*n*), as seen in Part 2. So, when we look at the most recent *n* sample units – the *window* – we check how many of them are unacceptable, or marginally acceptable in the case of a three-class plan, and if this number exceeds the acceptance number *c*, then the process is signalled to be out of control (Example 21).

3.4.2 What happens when a Moving Windows-based criterion signals "out-of-control" and how do we regain "in-control" status?

When the number of unacceptable sample units exceeds the acceptance number (*c*) for the window or if one sample unit exceeds the value of *M*, the system goes from an acceptable or "in-control" state to an "out-of-control" state, as shown in Example 21. This is an indication that the process is no longer operating as intended, i.e. something happened that has resulted in an increase in microbial levels.

The first task now is to undertake a root cause analysis to identify the cause of the deviation and correct it, which should be a step in your OCAP as discussed earlier. This is often easier said than done, especially since microbiological testing is not real-time and test results become available one or more days after the sample unit was collected. Nevertheless, it is important to attempt to identify the cause of the deviation so that future occurrences can be prevented through appropriate control measures. Consequently, any unusual observations and events during production should be recorded as detailed production records will be helpful in undertaking a root cause analysis.

However, even once the cause is identified and rectified, the process *remains* in the out-of-control state based on the moving window criterion until the window no

longer contains an excess number of unacceptable sample units, i.e. more than *c*. This can lead to an extended time period during which the process is still considered out of control despite the deviation being corrected, as shown in Example 22.

One possibility for overcoming this limitation, and shortening the "out-of-control" time, is to allow an increased rate of sampling and testing, provided the root cause of the process deviation has been found and rectified (Example 23).

Example 22
Moving windows: returning to "in-control"

Following on from Example 21, assume that the reason for the process failure was investigated and rectified. At the end of Week 2 the results (cfu/cm^2) might look like this:

Week 2 - Monday:	1, 0, 0
Week 2 - Tuesday:	0, 0, 0
Week 2 - Wednesday:	0, 1, 0
Week 2 - Thursday:	0, 2, 8
Week 2 - Friday:	0, 0, 0

The process remains out of control at this point and does not return to the in-control state until at least Monday in Week 3, when the Week 2 - Monday results drop out of the moving window. Provided there are no further *E. coli* detections on Monday in Week 3 the process is then considered to be in-control again.

However, if *E. coli* are detected on Monday in Week 3 then there are more than the $c = 3$ tolerable detections in the window, and the process remains out of control.

Example 23
Moving windows : returning to "in-control" more quickly

Consider the process from Example 21, and again assume that the underlying cause for the process deviation has been identified and rectified. Provided the competent authority is agreeable, the establishment increases the sampling rate from 1 in 300 carcasses to 1 in 150 carcasses. Hence, six new sample results become available each day and the results (cfu/cm^2) might now look as follows:

Week 2 - Monday:	1, 0, 0
Week 2 - Tuesday:	0, 0, 0, 2, 0, 0
Week 2 - Wednesday:	0, 1, 0, 0, 12, 0

Using this approach, the process may return to in-control by Thursday in Week 2, when the Week 2 - Monday results drop from the window, provided there are no further *E. coli* detections. At this point, the establishment would also return to three sample units per day, i.e. 1 sample unit per 300 carcasses.

Another possibility for overcoming this limitation is to assume that the corrective action has adequately addressed the problem and hence that the 'new' process is in control. Consequently, the window is reset after the corrective action has been implemented, i.e. contains no observations within the window (Example 24). This approach for example is used in Australia (AQIS, 2003).

3.4.3 What factors are considered in establishing the length or other properties of moving window-based criteria?

For a moving window-based MC the size of the window is equal to the sample size (n) that is determined using the acceptance sampling procedures described in Part 2. Consequently, the same considerations that affect the choice of sampling plan, as described in Part 2, also play a determining role here. In particular, we need to consider the combination of n, m, M and c, for a particular within-lot SD and distribution of microbial concentrations.

In addition, once the window size (n) has been determined, we also need to decide on how fast we want to be able to respond. The desired response time in turn will determine how frequently sample units need to be collected. For example, if we have a window size of $n = 15$, then taking a sample once per shift would stretch the window out over 15 shifts. This would mean that once a change in the process has happened it could take some time for the out-of-control signal to occur, although the bigger the change in the process the sooner the moving window will signal "out-of-control." Therefore, if we wanted the response time to be quicker, then we would have to collect more sample units per shift, e.g. three. In determining the size of the moving window consideration should be given to the combination of the production frequency and sampling frequency necessary to obtain a sufficient number of results that enables appropriate verification of performance of a process or a food safety control system. Once again, the deciding factor of how quick we can respond may be practicality and costs.

Another factor to keep in mind is in relation to the variability in the process. We know from Part 2 that not every lot is identical in terms of its level of contamination. However, we also pointed out earlier that a stable, a process under statistical process control requires predictability, which is achieved, at least partially, when lots can be produced consistently, that is, when the variability in the process – within and between lots – is stable and predictable over time. Clearly, both components of product variability need to be assessed and estimated. **Practically, the moving windows approach is appropriately applied when the between-lot variability is (much) less than the within-lot variability.**

A final factor that needs to be considered in the establishment of a moving windows-based MC is the control action that is to be taken as a result of the processes going out-of-control. As noted above, increased sampling, steps in the OCAP and other such measures are only helpful if you know your process and the factors that contribute to its variability. Here, the emphasis is not on a single lot which is to be accepted or rejected, but instead on the process' ability to produce food of acceptable safety and quality.

3.4.4 How do we characterize the performance of a moving windows-based criterion?

The OC curve, which we used to describe the performance of a sampling plan ("2.8.2 What is meant by the performance of a sampling plan?"), can also be used to describe the performance of the moving windows approach. However, in this case the within and between-lot variabilities need to be taken into account collectively to estimate the appropriate SD and develop the sampling plan. In addition, how quickly an out-of-control status is detected and how long it takes to recover from an out-of-control status are important characteristics of the performance of the moving windows-based criterion.

3.4.5 What are the benefits of moving windows-based criteria?

The moving window approach focuses on the verification of process control and not on proving safety. Consequently, the moving window approach is consistent with the HACCP approach to controlling microbiological hazards. In contrast, acceptance sampling, and especially lot-by-lot testing, is often used (incorrectly) for proving safety, e.g. with a zero acceptance number sampling plan, even though this is not actually possible. **As we have pointed out earlier ("1.3.1.1 Attributes sampling plans for tests that detect the presence of at least one organism per sampling unit"), not *detecting* contamination is not the same as "there is no contamination in the lot" or stated differently "absence of evidence is not evidence of absence."**

Because the moving windows approach spreads microbiological testing over time, it may reduce the cost of sampling and testing without compromising our ability to verify process control. Consequently, there is a cost benefit compared with lot-by-lot testing, especially if individual sample units can be analysed without extra costs for shipping (i.e. an in-house or close laboratory is available for the FBO). However, the aims of the approaches – acceptance sampling/lot-by-lot testing and moving windows – also differ.

Finally, the moving windows approach can provide a measure of when basic assumption used in the design of a food process are no longer valid despite the

monitoring of Process Criteria or CCPs still indicating adequate control, e.g. temperature. For example, if the contamination of incoming raw materials exceeds the levels for which a Process or CCP has been validated, then it is possible that concentrations in the end product are higher than anticipated. Consequently, the moving windows approach can provide a potential indicator of when a food safety control system needs to be re-validated.

3.4.6 What are the limitations of moving windows-based criteria?

While the moving windows approach has some distinct advantages in terms of cost, it may also be more difficult to convince customers and stakeholders that a smaller number of sample units taken at one point in time is adequate for verifying process control. However, for some FBOs, e.g. those without in-house sample analysis capacity, the cost of frequently shipping individual sample units (e.g. $n = 1$, weekly) for laboratory analysis can be bigger than shipping multiple sample units less frequently (e.g. $n = 5$, monthly).

In addition, the moving windows approach is potentially slower to indicate an out-of-control process than statistical process control charts. We say potentially, because how quickly the out-of-process signal is obtained depends how far the process has gone out of control, as well as the frequency of testing and the acceptance number *c*. Another possibility for overcoming this limitation is to assume that the corrective action has adequately addressed the problem and hence that the 'new' process is in control. Consequently, the window is reset after the corrective action has been implemented, i.e. contains no observations within the window (Example 24). This approach for example is used in Australia (AQIS, 2003).

Alternative statistical process control methods, such as X-bar and R-charts introduced in "3.3.2 What are control charts?" may be quicker at detecting small changes in the process.

Another potential limitation of the moving windows approach is that the efficacy of this approach is reduced when between-lot variability is greater than within-lot variability. When this is the case then the microbial safety or quality of lots can vary widely, which implies that there is little control in the process to produce consistent lots. It is quite possible that the microbiological status of the raw material(s) has a strong influence on the microbiological status of the food product.

As we have seen above, when the window size (n) is large then it can take a long time before an out-of-control process can be considered "in-control," even when the cause of the process deviation has been found (see "3.4.2 What happens when

Example 24
Moving windows: returning to "in-control" by resetting the window

Following on from Example 21, assume that the reason for the process failure was investigated and rectified. Consequently, the moving window is reset after Monday of Week 2 and the process is assumed to have returned to 'in control.' At the end of Week 2 the results (cfu/cm^2) might look like this:

Week 2 - Tuesday:	0, 0, 0
Week 2 - Wednesday:	0, 1, 0
Week 2 - Thursday:	0, 2, 8
Week 2 - Friday:	0, 0, 0

Provided there are no further *E. coli* detections on Monday in Week 3 the process is then considered to be in-control. However, if *E. coli* are detected on Monday in Week 3 then there are more than the $c = 3$ tolerable detections in the window, and the process goes out of control again.

a Moving Windows-based criterion signals "out-of-control" and how do we regain "in-control" status?").

Finally, we have made an implicit assumption that you know your process well enough to utilize the moving windows approach. In particular, there is a need for the process to be stable and in control and that there is little variability between lots compared with the variability within lots. However, to establish that this is actually the case you need to conduct a *process control study*, or in the case of national standards, a *national baseline study*. These studies aim to find out enough about the process, or national situation, so that significant sources of variability and factors affecting process performance are well understood and quantified. Unfortunately, these types of studies may be expensive, but the process understanding that you develop can be significant and may be well worth the effort and cost in the long term.

3.4.7 How can these limitations be overcome?

Open and effective communication may be required to communicate the strengths and limitations of the moving window approach with stakeholders.

As we pointed out in the previous section, the first undertaking is a process control study, which will provide the required information on between- and within-lot variability and other factors that will have significant influence on the microbial characteristics of the food produced. The information gained this way can then

be used to design the moving window with respect to the required reaction time, number of sample units (n), microbiological limits (m and M), and suitable acceptance number (c).

Lastly, we noted above that a process can remain in the out-of-control state for a long time after a moving window has been failed. A way to overcome this limitation is to increase the frequency at which sample units are collected, as shown in Example 23.

Concluding remarks

As pointed out before, while the food-specific aspects of MC are well understood, the mathematical and statistical aspects of MC are less well understood, and this hinders the consistent and appropriate application of MC in the food industry. However, when assessing the microbiological quality or safety of food products through sampling it is important to recognize that the microbiological and statistical aspects are integrally linked. Consequently, it is imperative that the performance and the limitations of the microbiological test methods are understood so that the results can be interpreted correctly.

Considerable information is required on the microbiological contamination of a food product to properly assess the performance of an MC, such as the statistical distribution of microorganisms in the food, including the average concentration, variability between food units and the shape of the distribution. In particular, enumeration data provide much more information than simple compliance with a microbiological limit, and hence the raw data should be recorded whenever possible, so they can be used to estimate the statistical distributions, and their parameters, and so that trends can be analysed. In many instances only limited data may be available at the early stages of developing an MC and hence it may be tempting to want to collect more data before constructing a sampling plan. However, we recommend that a sensible guess about the statistical distribution and variability between food units may still provide a useful starting point for developing a sampling plan. This plan should later be refined once more data become available as part of the routine data collection and application of the MC.

Microbiological concentration data are usually $\log_{10}$ transformed for the purpose of data analysis and graphing. However, special care needs to be taken when transforming the summary statistics, and especially the mean, back to the arithmetic scale. While traditionally the OC curve is presented with the mean $\log_{10}$ concentration on the X-axis (which is equivalent to the $\log_{10}$ geometric mean concentration), we recommend that OC curves are presented using the arithmetic mean on the X-axis. Using the arithmetic mean allows the MC to be properly interpreted with respect to the level of control that is being achieved.

When the concentration of the target organism is expected to be very low it is common to enrich the sample units to allow the target organism to grow so that it can be detected. Because of the heterogeneity of microbes in food, i.e. the uneven spatial distribution of microorganisms throughout the food, it is preferable to collect more small sample units than fewer large sample units, as this approach increases the chances of detecting contamination in the food lot.

However, irrespective of which approach is used it is important to remember that not detecting contamination in the few sample units does not imply that there is no contamination in the lot – sampling can never *guarantee* safety. Consequently, MC should be considered to form part of a food safety control system, one which should also include ongoing 'monitoring' of the system through trend analysis, such as moving windows or control charts or both.

References

Almiron, M.G., Lopes, B., Oliveira, A.L.C., Medeiros, A.C. & Frery, A.C. 2010. On the numerical accuracy of spreadsheets. *Journal of Statistical Software,* 34(4). Available at: http://www.jstatsoft.org/v34/i04

AOAC [Association of Analytical Communities] International. 2006. *Final Report and Executive Summaries from the AOAC International Presidential Task Force on Best Practices in Microbiological Methodology*, 481 North Frederick Avenue, Suite 500, Gaithersburg, MD 20877-2417, USA: AOAC International. Available at: http://www.fda.gov/food/foodscienceresearch/laboratorymethods/ucm124900.htm.

AQIS [Australian Quarantine and Inspection Service].2003. AQIS Notice Number MEAT 2003/6: Revised ESAM Program. Available at: http://www.daff.gov.au/__data/assets/pdf_file/0003/113970/2003_06.pdf.

Bassett, J., Jackson, T., Jewell, K., Jongenburger, I. & Zwietering, M.H. 2010. *Impact of Microbial Distributions on Food Safety*, International Life Sciences Institute (ILSI) Europe. Available at: http://www.ilsi.org/Europe/Publications/Microbial Distribution 2010.pdf.

Bower, J.A. 2013. Statistical Methods for Food Science: Introductory Procedures for the Food Practitioner 2nd ed., Wiley Blackwell.

Bray, D.F., Lyon, D.A. & Burr, I.W. 1973. Three class attributes plans in acceptance sampling. *Technometrics*, 15(3), pp. 575-585.

Busschaert, P., Geeraerd, A.H., Uyttendaele, M. & van Impe, J.F. 2010. Estimating distributions out of qualitative and (semi)quantitative microbiological contamination data for use in risk assessment. *International Journal of Food Microbiology*, 138(3), pp. 260-269.

CAC [Codex Alimentarius Commission]. 2004. CAC/GL 50-2004: *General Guidelines on Sampling*, Codex Alimentarius Commission. Available at: http://www.codexalimentarius.org/download/standards/10141/CXG_050e.pdf

CAC. 2006. *CAC/GL 47-2003: Guidelines for Food Import Control Systems (Revision 1-2006)*, Codex Alimentarius Commission. Available at: http://www.codexalimentarius.org/download/standards/10075/CXG_047e.pdf

CAC. 2007 CAC/GL 61-2007: *Guidelines on the Application of General Principles of Food Hygiene to the Control of Listeria monocytogenes in Foods*, Codex Alimentarius Commission. Available at: http://www.codexalimentarius.org/download/standards/10740/CXG_061e.pdf.

CAC. 2008a. CAC/GL 63-2007: *Principles and Guidelines for the Conduct of Microbiological Risk Management*. Codex Alimentarius Commission. Available at: http://www.codexalimentarius.org/download/standards/10741/CXG_063e.pdf

CAC. 2008b. CAC/GL 69-2008: *Guidelines for the Validation of Food Safety Control Measures*, Codex Alimentarius Commission. Available at: http://www.codexalimentarius.org/download/standards/11022/CXG_069e.pdf

CAC. 2013a. CAC/GL 21-1997: *Principles and Guidelines for the Establishment and Application of Microbiological Criteria Related to Foods (Revision 1-2013)*, Codex Alimentarius Commission. Available at: http://www.codexalimentarius.org/download/standards/394/CXG_021e.pdf

CAC. 2013b. *Codex Alimentarius Commission Procedural Manual* 22nd ed., Codex Alimentarius Commission. Available at: ftp://ftp.fao.org/codex/Publications/ProcManuals/Manual_22e.pdf

Cochran, W.G. 1950. Estimation of bacterial densities by means of the "most probable number. *Biometrics*, 6(2): 105-116.

Commeau, N., Parent, E., Delignette-Muller, M.L. & Cornu, M. 2012. Fitting a lognormal distribution to enumeration and absence/presence data. *International Journal of Food Microbiology*, 155: 146-152.

Cowell, N.D. & Morisetti, M.D. 1969. Microbiological techniques: Some statistical aspects. *Journal of the Science of Food and Agriculture*, 20(10): 573-579.

Dahms, S. & Hildebrandt, G. 1998. Some remarks on the design of Three-Class sampling plans. *Journal of Food Protection*, 61(6): 757-761.

EC [European Commission]. 2005. *Commission Regulation (EC) No 2073/2005 on microbiological criteria for foodstuffs*, Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2005R2073:20130701:EN:PDF.

Gonzales-Barron, U., Kerr, M., Sheridan, J.J. & Butler, F. 2010. Count data distributions and their zero-modified equivalents as a framework for modelling microbial data with a relatively high occurrence of zero counts. *International Journal of Food Microbiology*, 136(3): 268-277.

Gonzales-Barron, U. & Butler, F. 2011. A comparison between the discrete Poisson-gamma and Poisson-lognormal distributions to characterise microbial counts in foods. *Food Control*, 22(8): 1279-1286.

Habraken, C.J.M., Mossel, D.A.A. & van der Reek, S. 1986. Management of *Salmonella* risks in the production of powdered milk products. *Netherlands Milk and Dairy Journal*, 40(2-3): 99-116.

ICMSF [International commission for Microbiological Specifications for Foods]. 2002. *Microorganisms in Food 7: Microbiological testing in food safety management*, Kluwer Academic / Plenum Publishers.

ICMSF. 2011. *Microorganisms in Food 8: Use of Data for Assessing Process Control and Product Acceptance*, Springer.

ICMSF. 1986. *Microorganisms in Foods 2: Sampling for Microbiological Analysis: Principles and Specific Applications*, University of Toronto Press.

Jarvis, B. 2007. On the compositing of samples for qualitative microbiological testing. *Letters in Applied Microbiology*, 45(6): 592-598

Jongenburger, I. 2012. *Distribution of microorganisms in foods and their impact on food safety.* PhD Thesis, Wageningen University. Available at: http://edepot.wur.nl/196895

Jongenburger, I., Bassett, J., Jackson, T., Zwietering, M.H. & Jewell, K. 2012a. Impact of microbial distributions on food safety I. Factors influencing microbial distributions and modelling aspects. *Food Control*, 26(2): 601-609.

Jongenburger, I., Bassett, J., Jackson, T., Gorris, L.G.M., Jewell, K. & Zwietering, M.H. 2012b. Impact of microbial distributions on food safety II. Quantifying impacts on public health and sampling. *Food Control*, 26(2): 546-554.

Kilsby, D.C. 1982. Meat Microbiology. *In* M. H. Brown, (ed.). Applied Science Publishers.

Lieberman, G.J. & Resnikoff, G.J. 1955. Sampling plans for inspection by variables. *Journal of the American Statistical Association*, 50(270): 457-516.

McCullough, B.D. & Yalta, A.T. 2013. Spreadsheets in the Cloud — Not Ready Yet. *Journal of Statistical Software*, 52(7): 1-14. Available at: http://www.jstatsoft.org/v52/i07

Montgomery, D.C. 2012. *Introduction to Statistical Quality Control* 7th ed. Wiley, New York, USA.

Moore, D., McCabe, G.P. & Craig, B. 2012. *Introduction to the Practice of Statistics* 7th ed., W. H. Freeman.

Nelson, L.S. 1984. Technical Aids. *Journal of Quality Technology*, 16(4): 238-239.

Owen, D.B. 1967. Variables sampling plans based on the normal distribution. *Technometrics*, 9(3): 417-423.

Rogan, W.J. & Gladen, B. 1978. Estimating Prevalence from the results of a screening test. *American Journal of Epidemiology*, 107(1): 71-76.

Rooney, L.L. & Vanden Heuvel, L.N. 2004. Root Cause Analysis for Beginners. Quality Progress, 2004: 45-53, Available at: https://www.env.nm.gov/aqb/Proposed_Regs/Part_7_Excess_Emissions/NMED_Exhibit_18-Root_Cause_Analysis_for_Beginners.pdf

Schilling, E.G. & Neubauer, D.V. 2009. *Acceptance Sampling in Quality Control* 2nd ed., CRC Press.

van Schothorst, M., Zwietering, M.H., Ross, T., Buchanan, R.L. & Cole, M.B. 2009. Relating microbiological criteria to food safety objectives and performance objectives. *Food Control*, 20(11): 967-979.

Smelt, J.P.P.M. & Quadt, J.F.A. 1990. A proposal for using previous experience in designing microbiological sampling plans based on variables. *Journal of Applied Bacteriology*, 69(4): 504-511.

Takagi, K. 1972 On Designing Unknown-Sigma Sampling Plans Based on a Wide Class of Non-Normal Distributions. *Technometrics*, 14(3): 669-678.

USFDA [United States Food and Drug Administration]. 2001. Hazard analysis and critical control point (HACCP); procedures for the safe and sanitary processing and importing of juice; final rule. *Federal Register*, 66(182):6138-6202. Available at: http://www.gpo.gov/fdsys/pkg/FR-2001-01-19/pdf/01-1291.pdf.

USDA FSIS [United States Department of Agriculture Food Safety Inspection Service]. 1996. Pathogen Reduction: Hazard Analysis and Critical Control Point (HACCP) Systems. *Federal Register*, 61(144): 38806-38989. Available at: http://www.gpo.gov/fdsys/pkg/FR-1996-07-25/pdf/96-17837.pdf.

USDA FSIS. 2012. *Compliance Guideline for Establishments Sampling Beef Trimmings for Shiga Toxin-Producing Escherichia coli (STEC) Organisms or Virulence Markers,* Food Safety and Inspection Service. Available at: http://www.fsis.usda.gov/wps/wcm/connect/e0f06d97-9026-4e1e-a0c2-1ac60b836fa6/Compliance_Guide_Est_Sampling_STEC_0512.pdf?MOD=AJPERES.

W.E.C. [Western Electric Rules]. 1956. *Statistical Quality Control Handbook*. Western Electric Co., Indianapolis, Indiana, USA.

Wheeler, D.J. 2010. *Understanding Statistical Process Control* 3rd ed. SPC Press.

Annexes

Annex 1

Mathematical details

The information provided in this mathematical annex is not essential to the understanding of the concepts discussed in this document. They have been included here for those who are interested and to be transparent about the calculations that have been performed. In addition, the companion spreadsheet tools utilize these calculations extensively.

A1.1 CONVERTING THE MEAN AND STANDARD DEVIATION FROM THE LOG_{10} SCALE TO THE ARITHMETIC SCALE

Denote the concentration of microorganisms in the food by Y and assume that it is $\log_{10}$-normally distributed with mean μ and standard deviation σ. That is, μ and σ are the mean and standard deviation on the $\log_{10}$ scale. We can then write

$$Y \sim \log_{10} N(\mu, \sigma^2),$$

or alternatively on the $\log_{10}$ scale

$$X = \log_{10} Y \sim N(\mu, \sigma^2).$$

Consequently, the mean and standard deviation on the arithmetic scale, which we denote by μ' and σ' can be calculated as

$$\mu' = 10^{\mu + 0.5 \ln(10)\sigma^2}$$

and

$$\sigma' = \mu' \sqrt{\exp\{(\sigma \ln(10))^2 - 1\}}$$

where 'ln' denotes the natural logarithm and 'exp' denotes the exponential function.

A1.2 CALCULATING THE ANALYTICAL UNIT DETECTION PROBABILITY GIVEN THE ANALYTICAL UNIT AMOUNT

Denote the concentration of a microorganism in the food product by λ cfu/g. Provided the food is well-mixed we can use the Poisson distribution to describe the

probability distribution of the number of organisms (X) that might be contained in a 1 g sample of the food. That is, we can write

$$X \sim \text{Po}(\lambda),$$

where λ is referred to as the rate parameter.

If instead of sampling 1 g we test an analytical unit amount of w grams, then the number of organisms (X_w) in this larger analytical sample is also Poisson distributed, but with a rate parameter $w\lambda$, i.e.

$$X_w \sim \text{Po}(w\lambda).$$

The probability that x microbes are contained in the analytical unit is then given by

$$P(X_w = x) = \frac{(w\lambda)^k}{k!} e^{-w\lambda}.$$

Consequently, the detection probability equals the probability of an analytical unit containing one or more target microbes, i.e. $P(X_w \geq 1)$. This probability is given by

$$P(X_w \geq 1) = 1 - P(X_w = 0) = 1 - e^{-w\lambda}.$$

This equation forms the basis of the calculations in the Analytical Unit Detect. Prob. Tab in the companion spreadsheet.

These calculations can be extended to the situation where homogeneity cannot be assumed (van Schothorst *et al.*, 2009) and the rate parameter λ is itself a random variable with a distribution, such as $\log_{10}$-normal or Gamma. The web-based tool and the ICMSF tool provide this option.

Note: These calculations work equally if the concentration is on a volume (cfu/ml) or area (cfu/cm^2) basis.

A1.3 TWO-CLASS PRESENCE-ABSENCE SAMPLING PLANS

The two-class presence-absences sampling plan is based on the binomial distribution, because each sample unit can only result in one of two possible outcomes – the presence or absence of the target microbe. The binomial distribution applies provided the following assumptions are met (see also "Two-class presence-absence sampling plans").

1. The sampling process is random. This is needed for probability calculations to be valid. This is primarily achieved by random sampling, although this assumption will be reasonable under systematic and stratified random sampling[27] (see "What is random sampling and what are the alternatives?").
2. Sample units are independent of each other. This is something that we generally do not know, but can assume provided care was taken to ensure that the sampling was random.
3. Each sample unit has the same probability of yielding a detection (a 'positive' test result). This can generally be assumed provided there are no known strata that might affect the level of contamination.

Let X equal the number of detections in the sample of size n, and denote the analytical unit detection probability by p. The lot is accepted when the number of detections is less than or equal to the acceptance number c. The probability of acceptance is obtained from the probability mass function and is given by

$$P(\text{accept}) = P(X \le c) = \sum_{i=0}^{c} \binom{n}{i} p^i (1-p)^{n-i}$$

where $\binom{n}{i} = \frac{n!}{i!(n-i)!}$. If $c = 0$, then this simplifies to $P(\text{accept}) = (1-p)^n$.

A1.4 TWO-CLASS CONCENTRATION-BASED SAMPLING PLANS

The calculation of the probability of acceptance is analogous to the calculation for the two-class presence-absence sampling plan – the only difference is that the value of p is calculated from the normal distribution.

Denote the $\log_{10}$ concentration by Y and assume that it is normally distributed with mean μ and standard deviation σ. The probability of exceeding the microbiological limit m (on the $\log_{10}$ scale) is given by

$$P(Y > m) = 1 - P(Y \le m) = 1 - \Phi\left(\frac{m-\mu}{\sigma}\right)$$

where Φ() denotes the cumulative distribution function of the standard normal distribution. The probability $P(Y > m)$ is then used in place of the analytical unit detection probability p in the calculation of P(accept) using the binomial distribution shown above ("A1.3 Two-class presence-absence sampling plans").

[27] Additional care needs to be taken when strata are of different size as this will affect the probability calculations.

A1.5 THREE-CLASS SAMPLING PLANS

The three-class sampling plan is generated in a similar way to the two-class concentration-based sampling plan, with the main difference being that there are now three outcomes instead of just two, i.e. acceptable, marginally acceptable and unacceptable.

$$p_a = P(Y \text{ is acceptable}) = P(Y \le m) = \Phi\left(\frac{m-\mu}{\sigma}\right)$$

$$p_m = P(Y \text{ is marginally acceptable}) = P(Y \le M) - P(Y \le m) = \Phi\left(\frac{M-\mu}{\sigma}\right) - \Phi\left(\frac{m-\mu}{\sigma}\right)$$

$$p_u = P(Y \text{ is unacceptable}) = P(Y > M) = 1 - P(Y \le M) = 1 - \Phi\left(\frac{M-\mu}{\sigma}\right)$$

where m and M are the microbiological limits (on the $\log_{10}$ scale) that differentiate acceptable from marginally acceptable and marginally acceptable from unacceptable, respectively.

The probability of acceptance is subsequently calculated for the three-class sampling plan using the trinomial distribution. In the special case where no sample units are allowed to exceed the unacceptable limit M, the probability of acceptance

$$P(\text{accept}) = P(X \le c) = \sum_{i=0}^{c} \binom{n}{i} p_a^{n-i} p_m^i$$

where X is the number of marginally acceptable sample units and $\binom{n}{i} = \frac{n!}{i!(n-i)!}$ (Bray *et al*. 1973).

A1.6 VARIABLES SAMPLING PLANS

We denote the $\log_{10}$ concentration by Y and assume that Y is normally distributed with mean μ and standard deviation σ, that is $Y \sim N(\mu, \sigma^2)$.

The probability of $\log_{10}$ concentrations exceeding the unacceptable limit m is set equal to θ under the consumer's risk point, that is

$$P(Y > m) = 1 - P(Y \le m) = 1 - \Phi\left(\frac{m-\mu}{\sigma}\right) = \theta$$

where Φ denotes the standard normal cumulative distribution function. Consequently, the mean can be written as

$$\mu = m - z_{1-\theta}\,\sigma$$

where $z_{1-\theta}$ is the $1-\theta$ quantile of the standard normal distribution.

Because *Y* is normally distributed it follows that the sample mean, calculated from a sample of size *n*, is also normally distributed with the same mean μ and standard deviation $\sigma/\sqrt{n}$, also known as the standard error, that is, $\bar{Y} \sim N(\mu, \sigma^2/n)$.

We reject the lot if the sample mean is too close to the unacceptable limit *m*, and hence accept the lot if the sample mean is less than a critical limit, which we denote by *L*. We want this probability of acceptance to be α under the consumer's risk point. That is,

$$P(\text{accept}) = P(\bar{Y} \leq L) = \alpha.$$

Consequently, we can solve this equation for L, which is given by

$$L = \mu + z_{\alpha}\frac{\sigma}{\sqrt{n}}$$

And substituting for μ (from above) we get

$$L = m - \left\{ z_{1-\beta} - \frac{z_{\alpha}}{\sqrt{n}} \right\} \sigma$$

The term in braces is the critical value k.

Annex 2

Resources

The ***FAO/WHO Microbiological Sampling Plan Analysis Tool*** is a web-based sampling tool which we have referred to several times in the document. There is a basic version and an advanced version of the tool. The latter provides additional functionality including:

- the effect of between-lot variability on the probability of acceptance, and
- determining the appropriate sample size to achieve a specified consumer's risk point.

These tools can be found at http://www.fstools.org/sampling/. User accounts are required for the more advanced version but these can be obtained by following simple registration steps. Having an account allows a user to save the sampling plans and scenarios under which such sampling plans are assessed.

The ICMSF has a wealth of resources available on its website at http://www.icmsf.org/. In particular, there are two free Excel spreadsheets available for download and a number of presentations and articles related to MC and sampling.

The BASELINE project (http://baselineeurope.eu/) is an EU funded project for "Selection and Improving of Fit-For-Purpose Sampling Procedures for Specific Foods and Risks" and the software developed in this project can be accessed at http://www.baselineapp.com. User accounts are freely available and upon login the user can evaluate a range of sampling plans adapted to different food matrices and can take into account a heterogeneous contamination for sampling plan developments. In addition, the web application includes a decision-support tool for sampling plan selection and assesses how these sampling plans can be linked to a Performance Objective or Food Safety Objective (CAC 2013a). Finally, the application incorporates the examples of Codex Alimentarius on sampling plans.

The R package AcceptanceSampling allows the calculation of the probability of acceptance under two-class presence-absence, two-class concentration-based and variables sampling plans. It also has the capability to calculate the sample size to achieve a specified consumer and/or producer's risk point. Because of the command line interface and R's 'steep' learning curve, this resource is recommend-

ed for the more advanced users. Both R and the AcceptanceSampling package can be obtained from http://www.r-project.org/.

The Engineering Statistics Handbook is an online resource covering a range of statistical methods, including acceptance sampling and statistical process control (http://www.itl.nist.gov/div898/handbook/). It is produced jointly by the US National Institute of Standards and Technology (NIST) and SEMATECH.

Some basic information in statistics can be found in several FAO manuals. One such example is available at http://www.fao.org/docrep/w7295e/w7295e08.htm.

Annex 3

Links to companion tools for this document

Microbiological Criteria and Sampling Plan analysis tool

A Microsoft Excel version can be downloaded from: http://fao.org/2/jABhk and http://www.who.int /foodsafety/publications/mc-tool.xlsx

A LibreOffice Calc version can be downloaded from: http://fao.org/2/xBhEF

Training Videos on the use of the Microbiological criteria and sampling plan analysis tool

Full playlist available at:

https://www.youtube.com/playlist?list=PLzp5NgJ2-dK5pegu_aA0r0lTUfjM8Z7El

Index of videos:

Converting to and from log10. Available at: https://www.youtube.com/watch?v=mGNRmGDgNOU&index=1&list=PLzp5NgJ2-dK5pegu_aA0r0lTUfjM8Z7El

Converting summary statistics. Available at: https://www.youtube.com/watch?v=iQWCnykNKWQ&index=2&list=PLzp5NgJ2-dK5pegu_aA0r0lTUfjM8Z7El

Simple random sampling. Available at: https://www.youtube.com/watch?v=AVnQdTqBqDA&index=3&list=PLzp5NgJ2-dK5pegu_aA0r0lTUfjM8Z7El

Stratified Random Sampling - Statistical Aspects of Microbiological Criteria Related to Foods. Available at: https://www.youtube.com/watch?v=EE8-rwLGyl0&index=4&list=PLzp5NgJ2-dK5pegu_aA0r0lTUfjM8Z7El

A two-class presence-absence sampling plan with desired performance. Available at: https://www.youtube.com/watch?v=e3SRSnQ7s4g&list=PLzp5NgJ2-dK5pegu_aA0r0lTUfjM8Z7El&index=5

Calculate the analytical unit detection probability. Available at: https://www.youtube.com/watch?v=4wxpNFarikU&index=6&list=PLzp5NgJ2-dK5pegu_aA0r0lTUfjM8Z7El

Putting it all together: Two-class presence-absence sampling plan. Available at: https://www.youtube.com/watch?v=YnncxY7imyw&list=PLzp5NgJ2-dK5pegu_aA0r0lTUfjM8Z7El&index=7

A two-class concentration-based sampling plan. Available at: https://www.youtube.com/watch?v=vpbXB3kqYNM&index=8&list=PLzp5NgJ2-dK5pegu_aA0r0lTUfjM8Z7El

Putting it all together: Two-class concentration-based sampling plan. Available at: https://www.youtube.com/watch?v=w0SMpGhokXo&index=9&list=PLzp5NgJ2-dK5pegu_aA0r0lTUfjM8Z7El

A three-class concentration-based sampling plan. Available at: https://www.youtube.com/watch?v=tU-RbLu_sBw&index=10&list=PLzp5NgJ2-dK5pegu_aA0r0lTUfjM8Z7El

A variables sampling plan. Available at: https://www.youtube.com/watch?v=8jiab43VB94&list=PLzp5NgJ2-dK5pegu_aA0r0lTUfjM8Z7El&index=12

Systematic Sampling. Available at: https://www.youtube.com/watch?v=6VudQ3g9oyw&list=PLzp5NgJ2-dK5pegu_aA0r0lTUfjM8Z7El&index=13

FAO/WHO Microbiological Risk Assessment Series

1 Risk assessments of *Salmonella* in eggs and broiler chickens: Interpretative Summary, 2002

2 Risk assessments of *Salmonella* in eggs and broiler chickens, 2002

3 Hazard characterization for pathogens in food and water: Guidelines, 2003

4 Risk assessment of *Listeria monocytogenes* in ready-to-eat foods: Interpretative Summary, 2004

5 Risk assessment of *Listeria monocytogenes* in ready-to-eat foods: Technical Report, 2004

6 *Enterobacter sakazakii* and microorganisms in powdered infant formula: Meeting Report, 2004

7 Exposure assessment of microbiological hazards in food: Guidelines, 2008

8 Risk assessment of *Vibrio vulnificus* in raw oysters: Interpretative Summary and Technical Report, 2005

9 Risk assessment of choleragenic *Vibrio cholerae* 01 and 0139 in warm-water shrimp in international trade: Interpretative Summary and Technical Report, 2005

10 *Enterobacter sakazakii* and *Salmonella* in powdered infant formula: Meeting Report, 2006

11 Risk assessment of *Campylobacter* spp. in broiler chickens: Interpretative Summary, 2008

12 Risk assessment of *Campylobacter* spp. in broiler chickens: Technical Report, 2008

13 Viruses in food: Scientific Advice to Support Risk Management Activities: Meeting Report, 2008

14 Microbiological hazards in fresh leafy vegetables and herbs: Meeting Report, 2008

15 *Enterobacter sakazakii* (*Cronobacter* spp.) in powdered follow-up formula: Meeting Report, 2008

16 Risk assessment of *Vibrio parahaemolyticus* in seafood: Interpretative Summary and Technical Report, 2009

17 Risk characterization of microbiological hazards in food: Guidelines, 2009.

18 Enterohaemorragic *Escherichia coli* in meat and meat products: Meeting Report, 2010

19 *Salmonella* and *Campylobacter* in chicken meat: Meeting Report, 2009

20 Risk assessment tools for *Vibrio parahaemolyticus* and *Vibrio vulnificus* associated with seafood: Meeting Report and Follow-up, In press

21 *Salmonella* spp. In bivalve molluscs: Risk Assessment and Meeting Report, In press

22 Selection and application of methods for the detection and enumeration of human pathogenic *Vibrio* spp. in seafood: Guidance, 2016

23 Multicriteria-based ranking for risk management of food-borne parasites, 2014

24 Statistical aspects of microbiolgical criteria for foods: A risk managers guide, 2016

25 A risk based approach for the control of *Trichinella* in pigs and *Taenia saginata* in beef: Meeting Report, In press

26 Ranking of low moisture foods in support of microbiological risk management: Meeting Report and Systematic Review, In press

27 Microbiological hazards associated with spices and dried aromatic herbs: Meeting Report, In press

28 Microbial Safety of lipid based ready-to-use foods for the management of moderate acute and severe acute malnutrition: First meeting report, 2016

29 Microbial Safety of lipid based ready-to-use foods for the management of moderate acute and severe acute malnutrition: Second meeting report, In press

30 Interventions for the Control of Non-typhoidal *Salmonella* spp. in Beef and Pork: Meeting Report and Systematic Review, 2016

Microbiological Risk Assessment Series

Microbiological Criteria have been used in food production and the food regulatory context for many years. While the food-specific aspects of microbiological criteria are well understood, the mathematical and statistical aspects are often less well appreciated, which hinders the consistent and appropriate application of microbiological criteria in the food industry. This document has been developed to begin redressing this situation.

A particular aim of this document is to illustrate the important mathematical and statistical aspects of microbiological criteria, but with minimal statistical jargon, equations and mathematical details. It is hoped that the resulting document and support materials make this subject more accessible to a broad audience.

This volume and others in this *Microbiological Risk Assessment Series* contain information that is useful to both food safety risk assessors and risk managers, the Codex Alimentarius Commission, governments and regulatory agencies, food producers and processers and other institutions and individuals with an interest in Microbiological Criteria. This volume in particular aims to support food business operators, quality assurance managers, food safety-policy makers and risk managers.

9 789241 565318