2

Sampling techniques

D. Legan and M. H. Vandeven, Kraft Foods North America, USA

2.1 Introduction: common definitions

There are many good texts describing the statistics of sampling including some designed to address the specific needs of microbiologists (ICMSF, 1974, 1986, 2002; Jarvis, 1989, 2000). They will supply additional background information and explanation. We intend to review the art and science of sampling for pathogens in and around foods indicating when to use statistical methods for selecting the sampling plan and when to prefer other methods. We will use examples from some common sampling circumstances for illustration. We will describe the 'attributes' sampling plans advocated by The International Commission on Microbiological Specifications for Foods (ICMSF) since 1974 to show why they are the most practical solution to the problem for which they were developed, namely port-of-entry sampling of foods in international trade. We will demonstrate how these plans can be rescaled to perform in terms of concentration of organisms when the log counts from sampling follow a normal distribution, and how this can provide a quantitative approach to food safety management. We will indicate how variables sampling plans may allow more value to be extracted from the data derived from sampling when the result of testing gives countable numbers of organisms and the distribution of organisms in the foods is known. We will consider environmental sampling as part of a program of food safety management within a manufacturing facility. Finally we will make some suggestions on how to extract more value from the historical records of test results produced in quality control laboratories at great expense.

Some terms that will appear frequently in this chapter that may not be familiar to all readers are explained below.

Accept. Allow the food to be used for its intended purpose. This decision is made about the lot, not about an individual test result.

Acceptable. A lot that is suitable for its intended purpose. A sample that gives a 'passing' test result – at or below the maximum acceptable level specified.

Acceptance-sampling plan. A pre-determined procedure to determine whether a lot is acceptable. Such a plan should include statements about the number, size and type of samples to take, the laboratory methods to be used for analysis, and the criteria for acceptability.

Attribute. A property of the sample used to assign it to one of the classes in the sampling plan.

c. See two-class plan and three-class plan, below.

Class. One of the conditions, such as 'acceptable', 'marginally acceptable' or 'defective' defined by quantitative limits.

Colony forming unit (cfu). A bacterial or yeast cell, or small group of joined cells, or a fungal spore, cell or fragment of mycelium that gives rise to a single colony when grown on a solid medium.

Consumer's risk. The probability of accepting in error a lot that is defective.

Defective. A lot that is not suitable for its intended purpose. A sample that gives a 'failing' test result – above the maximum acceptable level specified.

Food Safety Objective (FSO). The frequency or maximum concentration of a microbiological hazard in a food considered acceptable for consumer protection.

Good Manufacturing Practice. A formal approach to managing food quality.

HACCP. Hazard Analysis Critical Control Point. A formal tool for managing food safety by identifying hazards and taking steps to ensure that they are controlled.

ICMSF. International Commission on Microbiological Specifications for Foods: an international expert body with the primary role of giving guidance on appraising and controlling the microbiological safety and quality of foods.

Lot. A defined quantity of a food or ingredient. A lot can be large and with a limited history, e.g. a shipload of grain, or small and with a clear history, e.g. a quantity of product made by batch mixing ingredients and further processing, e.g. 500 loaves of bread, with traceable records of all the ingredients. Or it could be somewhere in between, e.g. a truck load of flour from a mill that we deal with frequently or product made on a single production line during a recorded period of time (e.g. one hour, one shift).

m. See two-class plan and three-class plan, below.

M. See three-class plan, below.

Marginally acceptable. A sample that gives a test result greater than the limit defined by good manufacturing practice but less than the limit defined by safety.

n. See two-class plan and three-class plan, below.

Operating characteristic (OC) curve. A graph relating the probability of accepting the lot to a measure of the underlying condition of the lot such as

proportion defective or mean concentration of organisms.

Producer's risk. The probability that an acceptable lot is rejected in error.

Probability of acceptance (*Paccept*). The likelihood of accepting the lot (on the basis of the test results).

Probability of rejection (*Preject*). The likelihood of rejecting the lot (on the basis of the test results).

Reject. Do not allow the food to be used for its intended purpose. This does not imply that the food may not be used for some other purpose – for example diverted to animal feed or processed to destroy the organisms of concern and then used in a different product.

Sample. One dictionary definition of a sample is 'a portion, piece or segment that is representative of a whole' (Anon, 1982). The critical concept is that the portion represents the whole – if this is not true, no useful decision can be made based on properties of the sample. When sampling for pathogens we can think of a hierarchy of samples. At the highest level, there is the 'lot sample'. This is a collection of individual units taken from the lot. Statisticians tend to refer to this collection of units as the sample, whereas microbiologists tend to use sample to indicate the individual units. Throughout this chapter we will defer to the microbiologists and use 'sample' to indicate the individual unit and the ICMSF term 'population sample' to represent the collection of units. However, it is important to be aware of this potential for confusion in the dialogue between microbiologists and statisticians, a dialogue that we fully recommend. Sometimes we may need to take a 'subsample' from each individual sample. For example, when testing for a very non-uniformly distributed contaminant we may initially take a sample of several kilograms and mix it thoroughly to distribute the contaminant as uniformly as possible in the sample. We can then take a subsample for ease of handling and storage. This type of procedure is used in aflatoxin testing (FAO, 1993; Whitaker *et al*., 1994). There is another important term 'analytical sample' which may be the whole sample, or subsample or a small portion of either, but which represents the actual material processed through the laboratory testing procedure.

Two-class plan. An acceptance-sampling plan in which results for individual samples are assigned to one of two classes, 'acceptable' or 'defective', depending on whether they exceed the acceptable limit (ICMSF term m). A number of samples, n, are tested and the lot is rejected if more than c samples exceed m (usually $c = 0$).

Three-class plan. An acceptance-sampling plan in which results for individual samples are assigned to one of three classes, 'acceptable', 'marginally acceptable' or 'defective', depending on whether they exceed the acceptable (ICMSF term m) or marginally acceptable (ICMSF term M) limit. A number of samples, n, are tested and the lot is rejected if more than c samples exceed m or if any exceed M. (c decreases as the required plan stringency increases.)

2.2 The purpose of sampling

The point of sampling is usually to allow us to make a decision about the lot based on properties of the sample(s). The decision will be sound only to the extent that the sample *truly* represents those properties of the lot that concern us. So before we begin to sample anything, we need to understand clearly what our purpose is. When detecting pathogens in food, we usually intend to ensure that the food is safe to eat. Historically we would randomly select a food product at the end of a production line, or from a retail store. We would test it for the presence and/or concentration of pathogens and, if the laboratory results were 'acceptably low', we would conclude that the food was safe. This approach leaves us unsure about how truly representative the sample was. Hence, we are uncertain about whether we made a good decision. In this situation a formal, statistically-based, acceptance-sampling plan can tell us a great deal about how likely we are to make the 'right' or the 'wrong' decision. We will discuss this in more detail later in the chapter.

We may have the fundamentally different aim of ensuring that a food manufacturing process can consistently make safe products. In this case, our emphasis should be on managing the process to eliminate pathogens or reduce their incidence and/or concentration in the product to an acceptable level. Management focus is on controlling those points in the process that play a critical role in ensuring product safety. A HACCP plan is the tool that helps to identify those points, sets appropriate control limits, and ensures that they are under control by appropriate monitoring and verification. We may use sampling of incoming raw materials as part of this control scheme. A statistically-based sampling plan here may help us to understand how likely we are to make the 'right' or the 'wrong' decision about whether to accept the incoming material. We may also use environmental sampling within the plant to help us minimize the possibility of cross-contamination from the environment (everything in the plant that is not food) to the food. The environmental sampling protocol will be based on experience of which parts of the plant environment have most contact with the product and which are most likely to harbor pathogens. Hence, environmental sampling protocols are not statistical sampling plans, though they may have a statistical element. For example, the protocol may stipulate increased levels of sampling once a test result shows that a sampling site is 'out of control'. Within the HACCP plan we may be sampling finished product, but this now serves simply to verify that the process is under control. By statistical analysis of historical data from the process, we can establish whether the process is, in fact, capable of the performance that we require (Harris and Richardson, 1996; Hubbard, 1996).

Occasionally we may already know, or suspect, that a foodstuff has an unacceptable level of pathogen contamination. In this instance, product sampling is a waste of available resources since, at best, it confirms what we already know or suspect whilst at worst it may lead us to conclude, incorrectly, that the food is safe. Rather we are interested in investigating the source of

contamination so that we may subsequently eliminate it. Another circumstance concerns the decision of whether or not to accept a delivery of a material when we have no knowledge of its history. For an individual manufacturer this may be a consignment of a commodity raw material. For a regulatory agency it may be one of the many consignments of internationally-traded foodstuffs unloaded daily at ports and airports throughout the world.

2.3 Sampling and the problem of pathogen distribution

Consider that we are concerned with a toxin-forming pathogen, e.g. *Staphylococcus aureus*, where we can tolerate a measurable population of, say, 10^4 colony-forming units (cfu) per gram and still have a substantial safety margin before any toxin will be formed in the food. What issues do we face in deciding whether a lot is acceptable?

Example 1. Consider a hypothetical continuously stirred tank containing 100 l of milk contaminated with *S. aureus* at a mean concentration of 10^6 cfu/ml. The mixing disperses the cells so that the concentration of cells in any particular 1 ml of milk will be roughly the same as in any other (they are uniformly distributed). There is no particular difficulty with sampling since any 1 ml sample would yield a count of about 10^6 cfu/g, which is substantially higher than our acceptable level of 10^4 cfu/g (Fig. 2.1a) and we would reject the milk.

Example 2. Imagine that we refill the same tank with fresh milk. It is still well stirred so that the *S. aureus* is (approximately) uniformly distributed, but now it contains a mean concentration of about 1 cfu/ml. Again, there is no particular difficulty because any sample will contain a concentration of *S. aureus* well below our acceptable level of 10^4 cfu/ml (Fig. 2.1b) and we would accept the milk.

Example 3. Now imagine that we have replaced the milk with a heavy cream, turned off the stirrer and allowed the temperature to rise until bacterial growth can occur. Just 5000 isolated *S. aureus* cells contaminated the tank but cell growth and division occur until the mean concentration is 2000 cfu/g, still well below our acceptable level of 10^4 cfu/g. However, because the cream is viscous and not stirred those cells are contained in 5000 micro colonies, each of 4×10^4 cells and each much smaller than 1 ml (Fig. 2.1c). Now we have a considerable difficulty, because most of the cream (95,000 1 ml portions) would give a result well below our acceptable level of 10^4 cfu/g, but 5000 portions of 1 ml would give a count above the acceptable level. There is a significant chance that we will make the wrong decision based on the result from a single 1 ml sample.

Example 4. If we were to turn on the stirrer to mix the cream from example 3, the *S. aureus* cells would begin to disperse throughout the cream. If we took samples shortly after beginning to stir we would find that the number of cells in

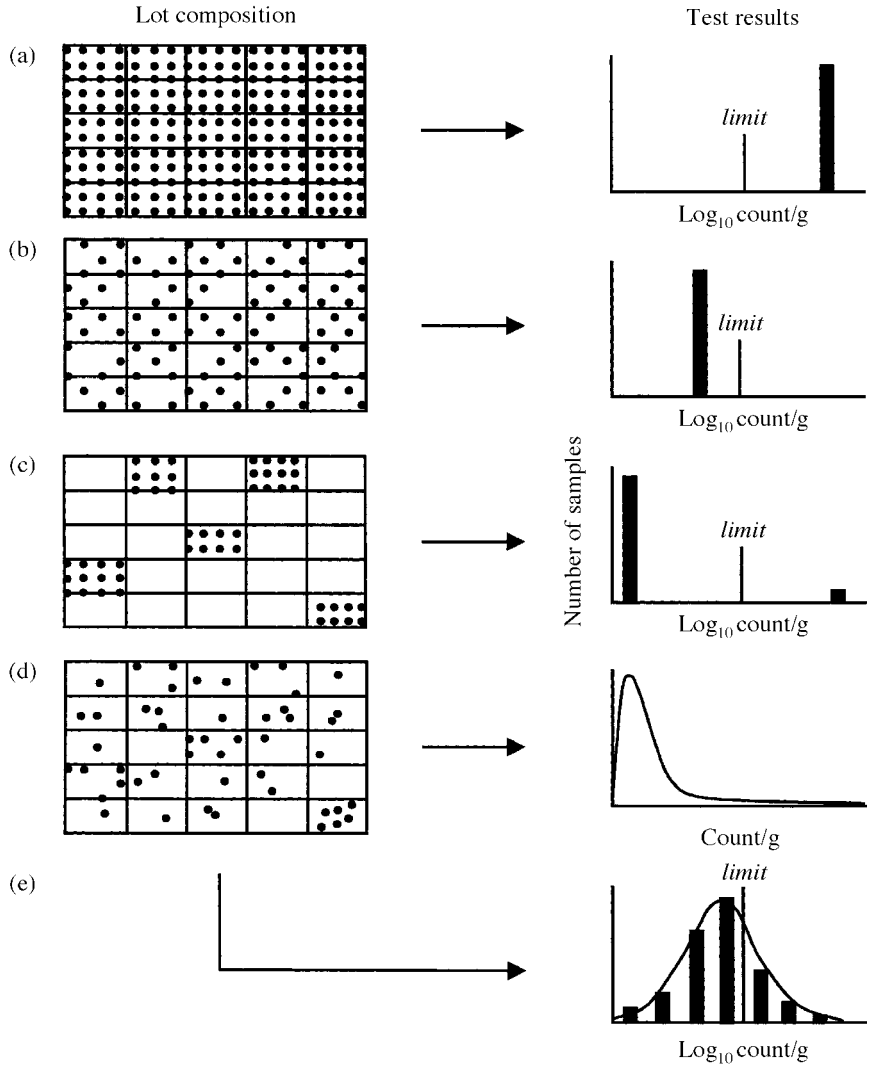

Fig. 2.1 Schematic illustration (not to scale) of the relationship between hypothetical distributions of microorganisms in foods, the pattern of test results on sampling them and a hypothetical acceptable limit.

individual 1 ml aliquots of cream varies, with a distribution skewed to the right (Fig. 2.1d). In this example, taking logs of the counts we see that the log cfu/ml approximately follows a normal distribution (Fig. 2.1e). We could express this by indicating that the results were 'log-normally distributed. If we stirred for long enough the cells would eventually become uniformly distributed. But in a solid food the opportunity for continuous mixing does not exist and the lognormal distribution of microbial load is most commonly seen. These

examples are all idealized and ignore the element of randomness seen in any real system that would mean some distribution of results would be seen in all cases – albeit that the distribution would be very narrow in a well-mixed system.

Now assume that we are not concerned about *S. aureus*, but instead about an infectious pathogen such as *Salmonella* where our acceptable level is very low. In fact, we would like to be able to say that *Salmonella* is not present at all.

Example 5. For the situation in example 1, any sample from the tank would yield *Salmonella* and we would correctly reject the milk.

Example 6. For the situation in example 2, although the mean concentration is 1 cfu/ml, there is no assurance that any particular 1 ml sample will contain any *Salmonella* and some samples may contain more than one. Now, if we take a single 1 ml sample we are at risk of making the 'wrong' decision and accepting the milk even though it contains a significant number of *Salmonella*.

Example 7. For the situation in example 3, the risk of making the wrong decision to accept the cream is even greater than that in example 6 because most of the potential samples would test negative even though there is a significant level of *Salmonella* in the lot.

Example 8. For the situation in example 4 we are now at less risk than in example 7 of making the wrong decision because the mixing means that most of the individual portions contain (small numbers of) *Salmonella* but the risk is still significant.

So how do we improve our chances of making the 'right' decision? And can we understand the risk of making the 'wrong' decision? A statistically-based acceptance sampling plan is the answer – though even the best plan, short of total destructive testing of the entire quantity of food, which is never a realistic option, cannot guarantee the 'right' decision.

In 1974, the ICMSF published guidance on the use of sampling plans and microbiological criteria for foods in international trade (ICMSF, 1974) that was later updated (ICMSF, 1986). The plans were designed for sampling at ports of entry, where the sampler knows little or nothing about the history of the food before sampling and may know nothing about the distribution of pathogens in the food. The ICMSF scheme for port of entry sampling uses 'attributes' sampling plans, so called because the microbiological test results are used to attribute the samples to one of a number of classes: 'acceptable' or 'defective' (in a two-class plan) or 'acceptable', 'marginally acceptable' or 'defective' (in a three-class plan) by comparison with predetermined criteria. The decision to accept or reject the product is indicated by the number of test sample results in each class. Over the years, the attributes plans advocated by the ICMSF have become widely accepted to the point that they now form the basis for microbiological specifications in many trading contracts, and are written into food law in some countries. The ICMSF advice has been further updated

(ICMSF, 2002) to include a more risk-based approach to food safety management, including situations where the history of the food is known and/or the distribution of pathogens in the food is understood.

2.3.1 Sampling and the degree of risk

Given that we are always constrained by the availability of resources, how do we put our sampling effort where it will do the most good? The ICMSF (1974, 1986, 2002) presented a comprehensive scheme for use of acceptance sampling plans based on degree of health risk associated with a food and the change in health risk that is expected to occur between sampling and consumption. There are five levels of risk and three condition levels, leading to 15 cases (Table 2.1). Plans are selected depending on the case. Three-class plans are not very stringent and are appropriate where health risk is relatively low (cases 1 to 9). The plans' stringencies are adjusted by changing the number of samples, n, the

Table 2.1 ICMSF cases based on factors influencing risk and recommended attributes plans (ICMSF, 1986; table 10, p. 74). Reproduced with permission from The University of Toronto Press

	Conditions in which food is expected to be handled and consumed after sampling, in the usual course of events		
Degree of concern relative to utility and health hazard	Conditions reduce degree of concern	Conditions cause no change in concern	Conditions may increase concern
No direct health hazard	Increase shelf-life	No change	Reduce shelf-life
Utility, e.g. shelf-life and spoilage	Case 1	Case 2	Case 3
	3-class, $n = 5$, $c = 3$	3-class, $n = 5$, $c = 2$	3-class, $n = 5$, $c = 1$
Health hazard	Reduce hazard	No change	Increase hazard
Low, indirect (indicator)	Case 4	Case 5	Case 6
	3-class, $n = 5$, $c = 3$	3-class, $n = 5$, $c = 2$	3-class, $n = 5$, $c = 1$
Moderate, direct, limited spread	Case 7	Case 8	Case 9
	3-class, $n = 5$, $c = 2$	3-class, $n = 5$, $c = 1$	3-class, $n = 10$, $c = 1$
Moderate, direct, potentially extensive spread	Case 10	Case 11	Case 12
	2-class, $n = 5$, $c = 0$	2-class, $n = 10$, $c = 0$	2-class, $n = 20$, $c = 0$
Severe direct	Case 13	Case 14	Case 15
	2-class, $n = 15$, $c = 0$	2-class, $n = 30$, $c = 0$	2-class, $n = 60$, $c = 0$

upper limit on the 'acceptable level', m, and the upper limit on the 'marginally acceptable level', M, and the maximum tolerable number of results, c, that exceed m (Table 2.1). Two-class plans with $c = 0$ are usually used for situations where the health risk is significant and the decision-maker is simply interested in whether or not a pathogen is present in a lot of foodstuffs (cases 10 to 15). The result is that the greatest sampling effort can be placed against the greatest hazard.

Imagine that we are concerned about *Staphylococcus aureus* in milk powder. *S. aureus* would be considered a moderate hazard: as a foodborne illness it is not usually life threatening and symptoms are usually of short duration and self-limiting. Hence it would be assigned to cases 7–9. If we were receiving the milk powder for use as an ingredient in a baking operation we can see that subsequent processing (baking) reduces the risk (because the baking step would kill the pathogen). Accordingly this would be assigned to case 7 where a three-class plan having $n = 5$ and $c = 2$ is called for. What if, instead, we were concerned about *Salmonella* in pâté? *Salmonella* is a serious hazard of moderate duration and would be assigned to cases 10–12. In the case of pâté there is no process to kill the *Salmonella* but growth could occur during storage, increasing the risk. Thus we would assign this to case 12 which calls for a two-class plan with $n = 20$, $c = 2$.

2.3.2 Effect of lot size

For acceptance sampling where there is no requirement for larger lots to meet a more stringent standard – which is usually the case in food inspection – the number of samples taken is independent of lot size. This is because the probability of acceptance, based on the proportion of defective units in the lot, barely changes with lot size unless the number of samples taken represents at least one-quarter to one-half of the lot – a circumstance exceedingly rare in food examination (ICMSF, 2002, p. 125). If the *number* of defective units in a lot is more important than the *proportion* defective then the sampling plans described in this chapter are not appropriate. If the aim of the sampling exercise is to estimate the proportion of defective units in the lot, rather than to make an accept/reject decision remember that the confidence limits on that estimate will become narrower as the number of sample units tested increases but, again, the plans described here were not designed for this purpose.

2.4 Acceptance sampling when the history of the material is not known

The ICMSF guidance, which has stood the test of more than 30 years of practical experience, can help us to select the sampling plan to use. But how can we assess the risk of making a 'wrong' decision?

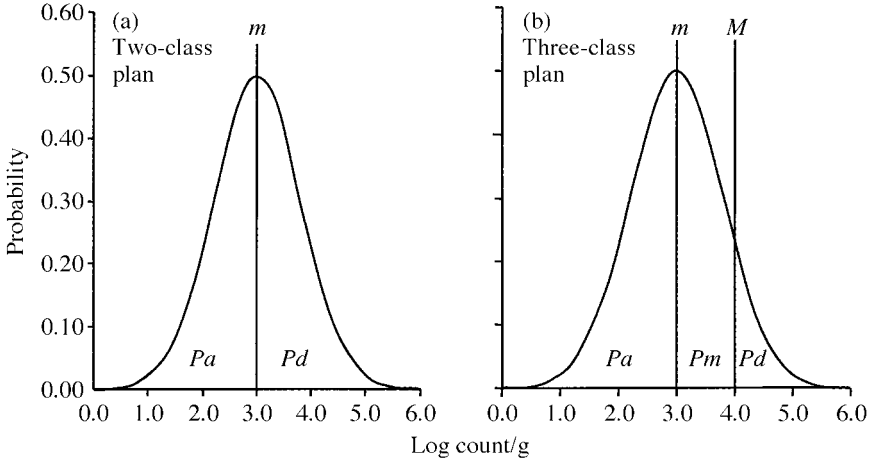

Fig. 2.2 The relationships between acceptable and defective log concentrations for a two-class plan and acceptable, marginally acceptable and defective concentrations for a three-class plan, when $m = 3.0, M = 4.0$ and the distribution of organisms has mean 3.0 and $\sigma = 0.8$. P_a is the proportion of acceptable material, P_m is the proportion marginally acceptable and P_d is the proportion defective. (Reprinted from Legan *et al*., 2001, with permission from Elsevier Science.)

2.4.1 Plan operating characteristics

We select the sampling plan according to the risk-based case, and choose the limits of acceptable, *m,* (and marginally acceptable, *M*) concentration of the pathogen. Figure 2.2 illustrates how the proportion defective relates to the limit of acceptability for a lot where the microorganisms are lognormally distributed. We can then calculate the probability of acceptance (or conversely of rejection). Legan *et al*. (2001) described the details of this calculation using the multinomial distribution, which can be generalized to either the binomial or trinomial as appropriate. We could use the binomial distribution for a two-class plan and the trinomial distribution for a three-class plan. When we plot the probability of acceptance against the proportion of defective samples in the lot (for a two-class plan) we create the operating characteristic (OC) curve, which shows us how well the plan discriminates between acceptable and defective material and what the decision risk is with different proportions of defective units in the lot (Fig. 2.3).

Two-class plans

The operating characteristic curve allows us to recognize two categories of decision risk. The 'producer's risk' is the probability that the plan will indicate rejection of a lot of food that is truly acceptable, and the 'consumer's risk' is the probability that the plan will indicate acceptance of a lot which is truly unacceptable. The ideal OC curve would give a *Paccept* of 1 at all defect rates up to the acceptable level and a *Paccept* of 0 at all defect rates above the acceptable level. All realistic sampling plans fail to meet this ideal. The

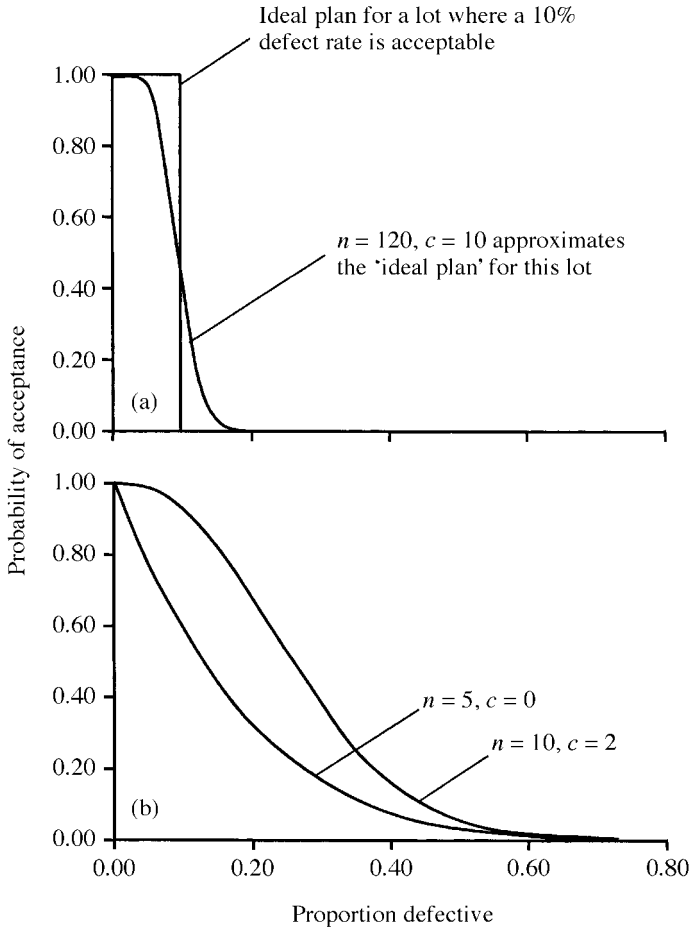

Fig. 2.3 Operating characteristic curves for two-class sample plans: (a) OC curves for an ideal plan for an acceptable defect rate of 10% and an approximation to the ideal defined by $n = 120, c = 10$ and (b) OC curves for practical plans with $n = 5, c = 0$ and $n = 10, c = 2$.

steepness of the OC curve and plan stringency both increase as the number of samples increases. Plan stringency decreases as c increases, so it is possible, in principle, to approach the 'ideal' curve for an acceptable defect rate by using a large number of samples to increase the steepness of the curve and setting c accordingly (Fig. 2.3a). This approach would be indicated where sampling and testing are fast and inexpensive, the costs to both parties of the wrong decision are very high *and* there is a tolerable defect rate. This situation is unlikely ever to occur when sampling foods for pathogens. Hence we are forced to consider the compromises involved in more realistic sampling plans (Fig. 2.3b). The operating characteristic curves of these plans differ dramatically from the ideal. They are much less steep, and hence less discriminating, and when $c = 0$ there is

no 'shoulder', implying that the plan will occasionally indicate rejection of a lot with a defect rate so low that we may, in fact, consider it acceptable. Nevertheless, these plans offer a realistic compromise between the degree of protection required and the resources needed for sampling and testing.

Two-class plans are used when sampling for pathogens that we want to be absent from the food. For example if concerned about *Salmonella*, we would consider the two-class plans associated with cases 10–15 depending on the conditions of use and target consumer. These plans would give $n = 5$, $n = 10$, $n = 20$, $n = 15$, $n = 30$ and $n = 60$ respectively, all with $c = 0$.

Three-class plans

Three-class plans are useful when we have some tolerance for 'marginally defective' material, which implies material that is outside the limits achieved by good manufacturing practice, so long as it is not hazardous. Just as we created an operating characteristic curve for a two-class plan (Fig. 2.3) so we can represent the operating characteristics of a three-class plan in a diagram; but now we have three variables to consider. These are the proportion defective (P_d), the proportion marginally defective (P_m) and the proportion acceptable (P_a) (see Fig. 2.2b). By plotting the *Paccept* against P_d and P_m we create the operating characteristic *surface*. This gives an overview of the plan operating characteristics but it is difficult to estimate *Paccept* accurately at any particular point on the surface (Fig. 2.4a). We could plot slices through the surface as two-dimensional OC curves appearing similar to Fig. 2.3b by fixing P_m and plotting *Paccept* against P_d or by fixing P_d and plotting *Paccept* against P_m. Alternatively, we can use a contour plot (Fig. 2.4b) that allows us to estimate accurately all the values of P_d and P_m that give a desired *Paccept*.

Three-class plans can be used when sampling for a pathogen that we can tolerate in food at low levels. For example if we are concerned about *S. aureus*

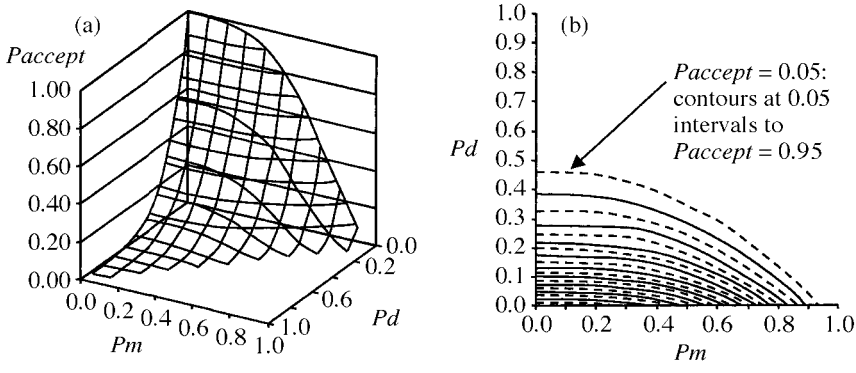

Fig. 2.4 (a) Operating characteristic surface showing an overview of plan performance for a three-class plan with $n = 5, c = 3$, (b) Contour plot for the same sampling plan linking combinations of proportion marginally acceptable and proportion defective that give equal probabilities of acceptance. This plan would be recommended for case 1 or case 4 in the ICMSF scheme.

we would consider the three-class plans associated with cases 7–9. If the food may support the growth of *S. aureus*, thereby increasing the risk, we would be in case 9 (Table 2.1) and we could use a sampling plan where $n = 10$ and $c = 1$. If we could be sure that no growth would occur and hence there would be no increase in risk we could use the plan for case 8 where $n = 5$ and $c = 1$.

2.4.2 Setting *m* and *M* (acceptable and marginally acceptable levels)

Setting the values of *m* (limit of acceptable count in two- and three-class plans) and *M* (limit of marginally acceptable count in three-class plans) plays a critical role in determining the final performance of the sampling plan. This judgment is normally somewhat subjective and is based on our best knowledge about the pathogen of concern, the nature of the food, the status of the intended consumer and the demands of any relevant laws or regulations. Fortunately there are many prior examples of expert judgment that we can use to guide us (ICMSF, 1974, 1986; Shapton and Shapton, 1991).

In a two-class plan we have only to set *m.* We typically use two-class plans when we have no tolerable level for the pathogen of concern. Accordingly we set *m* at zero. We must realize, however, that the true value of *m* defaults to the sensitivity of the analytical method used. An enrichment culture method applied to a 25 g sample, as in *Salmonella* testing, has a maximum theoretical sensitivity of one cell in the entire 25 g, which is equivalent to a concentration of 0.04 cells/g and *m* is then equivalent to a concentration < 0.04 cells/g.

For other pathogens such as *Staphylococcus aureus* where we do have some tolerance we typically use a three-class plan. Now we set *M* to the maximum tolerable level based on the hazard. We set *m* to a value that shows the food was produced under conditions of good hygienic practice (GHP). Where we have access to plant records these can be used to establish directly the level of *m*. The value of *m* is always lower than the value of *M* (or we have created a two-class plan) but otherwise there need be no quantitative relationship between them when used for pathogens. When a three-class plan is used for spoilage organisms a procedure developed by Dahms and Hildebrandt (1998) to set the difference between *m* and *M* can improve the overall level of control if the sample test results follow a lognormal distribution of known standard deviation.

2.4.3 Beware of 'retest' results

When testing detects a pathogen and a two-class plan indicates rejection of a lot, we may be tempted to repeat the test using a retained portion of the sample or using new samples taken from the lot. We wish only to 'confirm' that the original result was correct before causing the disruption that may be associated with rejection of the lot. However, there are many reasons why a retest may fail to detect the pathogen that was correctly detected initially (ICMSF, 2002, pp. 191–7; Flowers and Curiale, 1993). Some of these are connected with the distribution of organisms in foods that we have previously discussed. Others

may be connected with a low incidence of contaminated units within a batch or a change in the concentration of the pathogen in the food making it harder to detect, even though it is still present. Hence an initial positive result should not simply be discarded in favor of a negative retest result unless there is a clear indication of a laboratory or sampling error associated with the first result.

Even when retest results are considered in addition to the original results it is important to understand the nature of the resulting sampling plan. For example, at a defect rate of 5% the two plans $n = 60$, $c = 0$ and $n = 95$, $c = 1$ give the same probability of acceptance (in this instance 0.05). However, if one positive result is obtained from 60 samples tested and another 35 samples are then examined and all test negative we have not, in fact, used the plan $n = 95$, $c = 1$ which would justify accepting the lot. Rather we have used the two-stage plan $n_1 = 60$, $c_1 = 1$ plus $n_2 = 35$, $c_2 = 0$ which gives a higher probability of acceptance (actually 0.07; ICMSF, 2002). Montgomery (1996) shows how to calculate probabilities for two-stage plans. Even if we selected a two-stage plan giving the same *Paccept* as the plan with $n = 60$, $c = 0$ we should not forget that the implications of the plans can still be different. The decision to accept when no pathogens have been detected allows the possibility that the lot truly contains no pathogens. The decision to accept when one sample out of a larger number tested contained a detectable pathogen does not.

2.5 Acceptance sampling when the history of the material is known

The ICMSF continues to recommend attributes sampling plans for circumstances where we have no knowledge of product history and/or have no knowledge of the distribution of pathogens in the food. However, it is increasingly realized that preventative measures such as Good Hygienic Practice (GHP) and Hazard Analysis and Critical Control Point (HACCP) systems are much more effective tools for managing food safety than end product testing. ICMSF advocated the use of HACCP in 1988 (ICMSF, 1988). Later, ICMSF published its scheme for managing microbiological risks (ICMSF, 1997). In this scheme a sampling plan is used to assure that a food is safe only when this assurance cannot be provided more effectively, for example, when there is no knowledge that GHP and HACCP have been properly applied. This advice has been considerably expanded in the latest ICMSF guidance (ICMSF, 2002), which advocates a risk-based approach wherever feasible. The heart of this risk-based approach is the Food Safety Objective (FSO: van Schothorst, 1998), which provides a functional link between risk assessment and risk management. Developments in quantitative risk assessments in microbiology (Buchanan *et al*., 1997; Whiting and Buchanan, 1997) have made it possible to link the exposure assessment of a pathogen to likely public health outcomes. It is then for society, through public officials charged with policy, to decide how much risk is tolerable and hence how much protection is required.

The question of how to join the 'traditional' and familiar use of attributes plans with the risk-based approach then arises. Legan *et al*. (2001) proposed a rescaling of the operating characteristic curves of attributes plans to show probability of acceptance at different concentrations of organisms based on the assumption that the logarithm of test results from microbiological sampling follow a normal distribution. In manufacturing, or other ongoing circumstances such as a relationship with a particular supplier, historical data are often available. These data can allow us to understand the typical distribution of test results, including testing the assumption that test results are lognormally distributed, and to use this knowledge when designing sampling plans and setting criteria.

2.5.1 Determining the concentration of microorganisms controlled by attributes plans

Expressing the performance of attributes plans in terms of concentration of organisms has been suggested several times. Foster (1971) wrote, 'at very low levels of contamination it may be more meaningful to talk in terms of concentration per unit'. Later, Hildebrandt *et al*. (1995) and Dahms and Hildebrandt (1998) discussed the performance of three-class attribute sampling plans in relation to the variability of the lot and the chosen values for m and M, assuming a lognormal distribution of counts. Their approach clearly makes the connection between the distribution of organisms in the food and the probability of finding defective units by sampling.

Legan *et al*. (2001) built on the approach of Hildebrandt *et al*. (1995) using a statistical distribution of microorganisms to establish the proportion of defective samples in a lot, then using the calculated proportion of defective samples to determine the probability of acceptance in the usual way (i.e. cumulative binomial or multinomial distribution). Assuming a normal distribution for the log concentration of microbes, the area under the normal density function below m is used to define the value for proportion acceptable (P_a; Fig. 2.2). The area between m and M defines the value for proportion marginally acceptable (P_m; Fig. 2.2b), and the area above M (or m for a two-class plan) defines the value for proportion defective (P_d). The statistical methods for calculating the values of P_a, P_m and P_d are detailed in Legan *et al*. (2001).

Operating characteristic curves relating probability of acceptance (*Paccept*) to lot mean concentration were developed by fixing the standard deviation σ, and then 'sliding' the mean of a normal distribution through a range of values. Figure 2.5 illustrates this for a two-class plan with $n = 5$, $c = 0$, $m = 1.0$ log cfu/g. Figure 2.5a shows a normal distribution with $\sigma = 0.8$ and three different means. Material with mean levels of contamination above m in each position is defective. When m is fixed, each different mean level of contamination gives a different P_d. Figure 2.5b plots P_d for all mean levels of contamination in the range -2.0 to $5.0\log_{10}$ cfu/g to show how it increases with mean log count. Finally the *Paccept* at each P_d was computed using the binomial distribution and

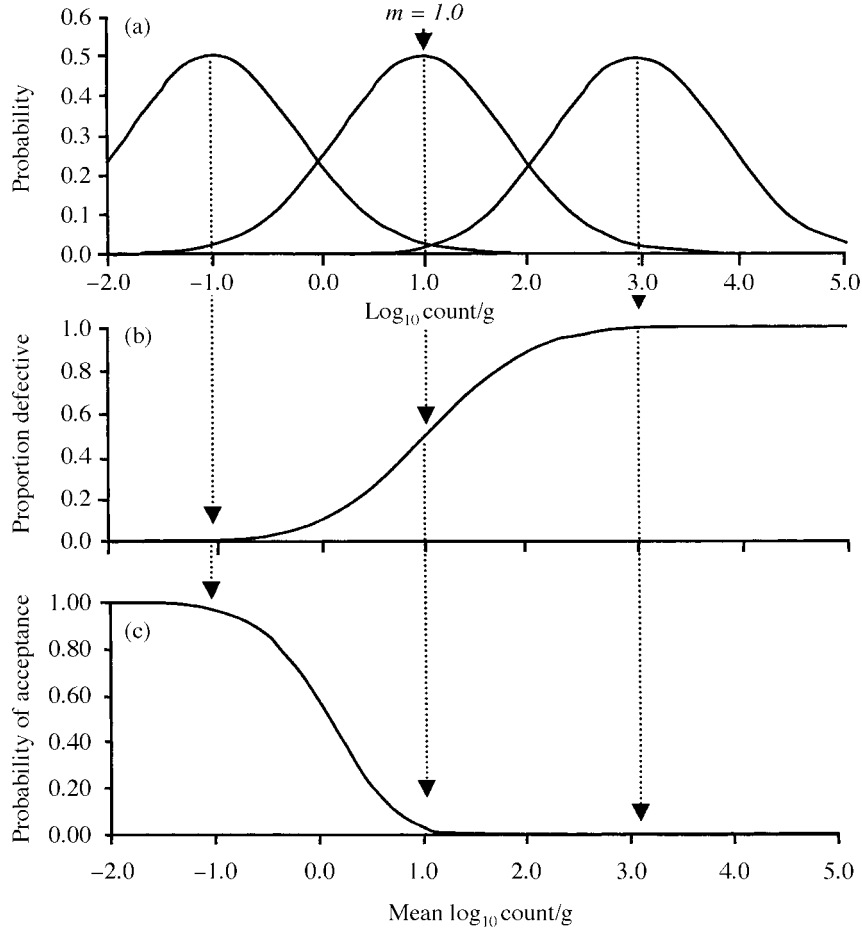

Fig. 2.5 Illustration of the rescaling from concentration to probability of acceptance for a two-class plan with $n = 5$, $c = 0$, m= log count/g of 1.0. a) The distributions have $\sigma = 0.8$ but different means. All parts of each distribution above m are defective. b) The proportion defective in each distribution illustrates the relationship between mean log count and proportion defective. c) The probability of acceptance determined for the relevant proportion defective using the operating characteristic curve (illustrated for $n = 5, c = 0$ in Fig. 2.3) gives the probability of acceptance for each mean log count/g. (Reprinted from Legan *et al.*, 2001, with permission from Elsevier Science.)

plotted in Fig. 2.5c to give the operating characteristic curve that relates probability of acceptance to mean concentration of microorganisms. The last step is equivalent to reading the proportion defective associated with each mean log count from Fig. 2.5b and finding the associated probability of acceptance from the $n = 5$, $c = 0$ curve in Fig. 2.3b.

This approach allows microbiologists to assess easily the performance of the sampling plans in terms of the familiar units of concentration of organisms.

Even the most stringent practical sampling plan (case 15: $n = 60$, $c = 0$) will accept a defect rate of 2% on 30% of sampling occasions. In this case defective means containing an unacceptably high concentration of the pathogen of concern, and a 2% defect rate implies 200 potentially hazardous units in a batch of 10,000, yet the sampling plan would indicate acceptance of the batch nearly one time in three! In the context of an ongoing relationship, a rejection rate of two in three would create pressure that rapidly leads to an improvement in future performance, but in the short term the acceptance rate of one in three may not provide the protection that we require. Rescaling this plan in terms of concentration of organisms, when $m = \log_{10}$–1.4 (a concentration of 0.04 cfu/g, equivalent to one cell in 25 g) and using σ as 0.8 $\log_{10}$ cfu/g there is a *Paccept* of 0.04 when the concentration is $\log_{10}$–2.7 (equivalent to one cell in 500 g). In other words, testing 60 samples and getting no positives only leaves us 96% sure that the concentration is below 1 cfu in 500 g. To put this into a risk-based perspective, one of the performance criteria recommended for managing *Listeria monocytogenes* in frankfurters is that 'the concentration of *L. monocytogenes* after cooking (within the manufacturing plant) shall be $\leq$ 1 cfu/kg' (ICMSF, 2002, p. 298). In other words, the most stringent sampling plan (case 15: $n = 60$, $c = 0$) would detect a process failure allowing this performance criterion to be exceeded by a factor of two on 96% of sampling occasions.

Rescaling OC curves into units of concentration shows vividly the relationship between all of the elements of the sampling plan and the mean population of microorganisms that can be controlled by sampling. It makes very clear how high these concentrations are with plans involving practicably small numbers of samples when the detection limit of the analytical method is not a limiting factor. When sampling for pathogens and the test sensitivity is an issue, the concentrations of organisms that can be controlled are far from zero.

For control of quality, spoilage or safety where the concern relates to relatively high numbers of microorganisms, the plans can be effective. Examples would include high numbers of food spoilage organisms, or *S. aureus* where it is generally regarded that 10^4–10^5 cfu/g are needed before toxin is formed. The plans are less effective in circumstances where concern relates to low numbers of microorganisms unless stringent cases, involving large numbers of samples, are used. The emergence of foodborne pathogens with a very low infective dose, e.g. pathogenic *Escherichia coli*, makes the application of even stringent microbiological criteria unlikely to ensure the safety of food with respect to the presence of these organisms.

When historical data are available we can easily estimate whether or not the microbial population is described by a normal distribution of log counts. If it is not, many other distributions could be used. For example, the negative binomial may be useful if the distribution is heavily skewed to the right and has been used to describe the distribution of aflatoxin in peanut lots (Whitaker *et al.*, 1994). It follows that examination of baseline data for different product and/or raw material streams to confirm the assumption of lognormality and to obtain estimates of σ would facilitate selection of appropriate values of n, c, m (and M).

2.5.2 Variables sampling plans

When we are concerned about the mere presence of a pathogen, as in *Salmonella* testing, we have no choice but to use a two-class attributes plan with $c = 0$ and n selected to give the discrimination that we require. However, if we are concerned about spoilage organisms, or pathogens where we have some tolerance for countable numbers, such as *S. aureus*, we may have the option of using a variables plan. The key advantage of a variables plan is that it captures all of the information in the test results (recorded counts and variance between multiple samples), whereas in an attributes plan, results are judged to be above or below the critical limit(s) m and M and the information on by how much is lost, as is the between sample variance. For example, in a three-class plan with $m = 10^2$ and $M = 10^3$, all results between 101 and 999 will carry the same weight. This ability to capture more of the information in the test results may make variables plans more cost-effective to operate than attributes plans. Variables plans may also remain more discriminating than attributes plans as the standard deviation of the data increases (ICMSF, 2002). Variables plans are, however, fundamentally different from attributes plans and may not, therefore, give generally equivalent indications. Variables plans are constructed on the assumption that the distribution of results is known. If this is not true then a variables plan cannot be used. Also, the decision rule in a (lognormal) variables plan is sensitive to the magnitude of the standard deviation. If the standard deviation in the test lot is greater than that used to set up the plan then the plan will not perform as expected.

Microbiological test results can often be transformed to follow a normal distribution by taking logs of the count: we can say that the results follow a lognormal distribution. In a manufacturing environment, or when there is an ongoing relationship with a particular supplier, we may have enough accumulated data to show that in the long run the results follow a lognormal distribution. It is still possible that the log counts within a single batch may not be normally distributed but it is hard to test the lot distribution unless the number of samples is large. The lognormal distribution is used as the basis of published variables plans for microorganisms in foods (Kilsby *et al*., 1979; Kilsby, 1982; Malcolm, 1984; ICMSF, 1986, 2002; Jarvis, 2000).

Once the underlying distribution (of log counts) is known to be normal, and the standard deviation has been established, the variables plan can be defined. The critical limit (C), that defines the boundary between acceptable and unacceptable material, is set (some authors use V to indicate this quantity). The value of C is the safety (or quality) limit in log concentration. The proportion of values that exceed the critical limit then depends on the population distribution such that:

$$(C - \mu)/\sigma = k$$

where μ is the population mean, σ is the population standard deviation and k is a quantity to be used with normal probability tables found in most introductory statistical methods books (e.g. Montgomery, 1996) to find the proportion of the population that exceeds the critical limit C (Table 2.2).

Table 2.2 Values of k for different proportions of a lot exceeding the critical limit, C

% > C	Approximate k	% > C	Approximate k
1	2.32	50	0
2	2.05	55	−0.12
5	1.65	60	−0.25
10	1.30	65	−0.38
15	1.05	70	−0.52
20	0.85	75	−0.68
25	0.67	80	−0.85
30	0.52	85	−1.05
35	0.38	90	−1.30
40	0.25	95	−1.65
45	0.13	99	−2.32

Since k becomes smaller as the proportion > C increases, small values of k imply high proportions > C and we would want to reject any lot where the value of k was smaller than the value corresponding to the acceptable proportion > C. However, we cannot directly establish μ and σ. We can, instead, estimate them using the sample mean ($\bar{x}$) and standard deviation (s). This allows us to use $\bar{x}$ and s to estimate the proportion > C through the quantity k_0 where:

$$(C - \bar{x})/s = k_0$$

The estimates $\bar{x}$ and s carry a degree of uncertainty and k_0 is about equally likely to fall above as below k. To compensate for this uncertainty the tolerance intervals k_1 (for defined probability of rejection: Table 2.3) and k_2 (for defined probability of acceptance: Table 2.4) have been calculated (Malcolm, 1984). Plans can then be designed to accept or reject with suitable probabilities based on the tails of the distribution of test results.

Variables plans for safety and quality limits

To use the variables plan for managing a safety or quality limit we must decide on the P_d that we are willing to tolerate and the probability (*Preject*) that we require of rejecting lots with a higher P_d. It is likely that in this case P_d is low and *Preject* is high. We refer to appropriate tables (Table 2.3) to find the value of k_1 that connects *Preject*, P_d, and the number of samples tested, n. Using the sample log counts we calculate the sample mean ($\bar{x}$) and standard deviation (s). We then calculate $\bar{x} + k_1 s$ and if this is greater than C we reject the lot.

Variables plans for good manufacturing practice limits

We can also use the variables plan to control good manufacturing practice limits. In this case we will select some limit C_m that we know is achievable when the operation is running well. This limit is somewhat analogous to the ICMSF limit m in a three-class plan. Now we want a high probability that we will accept the lot (*Paccept)* so long as the proportion > C_m is acceptable. Now both *Paccept*

Table 2.3 Values of k_1 calculated using the non-central t-distribution for safety/quality limits. Reject if $\bar{x} + k_1 s > C$. (Adapted from Malcolm, 1984, with permission. © The Society for Applied Microbiology)

Probability of rejection (*Preject*)	Proportion exceeding C (P_d)	Number of replicates (n)							
		3	4	5	6	7	8	9	10
0.99	0.10			5.4	4.4	3.9	3.5	3.2	3.0
	0.20			4.0	3.3	2.9	2.6	2.4	2.2
	0.30			3.0	2.5	2.2	2.0	1.8	1.7
	0.40			2.3	1.9	1.6	1.5	1.4	1.3
	0.50		2.3	1.7	1.4	1.2	1.1	1.0	0.9
0.95	0.05	7.7	5.1	4.2	3.7	3.4	3.2	3.0	2.9
	0.10	6.2	4.2	3.4	3.0	2.8	2.6	2.4	2.4
	0.30	3.3	2.3	1.9	1.6	1.5	1.4	1.3	1.3
	0.50	1.7	1.2	0.95	0.82	0.73	0.67	0.62	0.58
0.90	0.10	4.3	3.2	2.7	2.5	2.3	2.2	2.1	2.1
	0.25	2.6	2.0	1.7	1.5	1.4	1.4	1.3	1.3

Table 2.4 Values of k_2 calculated using the non-central t-distribution for good manufacturing practice limits. Accept if $\bar{x} + k_2 s < C_m$. (Adapted from Malcolm, 1984, with permission. © The Society for Applied Microbiology)

Probability of acceptance (*Paccept*)	Proportion exceeding C_m (P_m)	Number of replicates (n)							
		3	4	5	6	7	8	9	10
0.95	0.10	0.33	0.44	0.52	0.57	0.62	0.66	0.69	0.71
	0.20	−0.13	0.02	0.11	0.17	0.22	0.26	0.29	0.32
	0.30	−0.58	−0.36	−0.24	−0.16	−0.10	−0.06	−0.02	0.00
0.90	0.05	0.84	0.92	0.98	1.03	1.07	1.10	1.12	1.14
	0.10	0.53	0.62	0.68	0.72	0.75	0.78	0.81	0.83
	0.20	0.11	0.21	0.27	0.32	0.35	0.38	0.41	0.43
	0.30	−0.26	−0.13	−0.05	0.01	0.04	0.07	0.10	0.12
	0.40	−0.65	−0.46	−0.36	−0.30	−0.25	−0.21	−0.17	−0.16
	0.50	−1.09	−0.82	−0.69	−0.60	−0.54	−0.50	−0.47	−0.44
0.75	0.01	1.87	1.90	1.92	1.94	1.96	1.98	2.00	2.01
	0.05	1.25	1.28	1.31	1.33	1.34	1.36	1.37	1.38
	0.10	0.91	0.94	0.97	0.99	1.01	1.02	1.03	1.04
	0.25	0.31	0.35	0.38	0.41	0.42	0.44	0.45	0.46
	0.50	−0.47	−0.38	−0.33	−0.30	−0.27	−0.25	−0.24	0.22

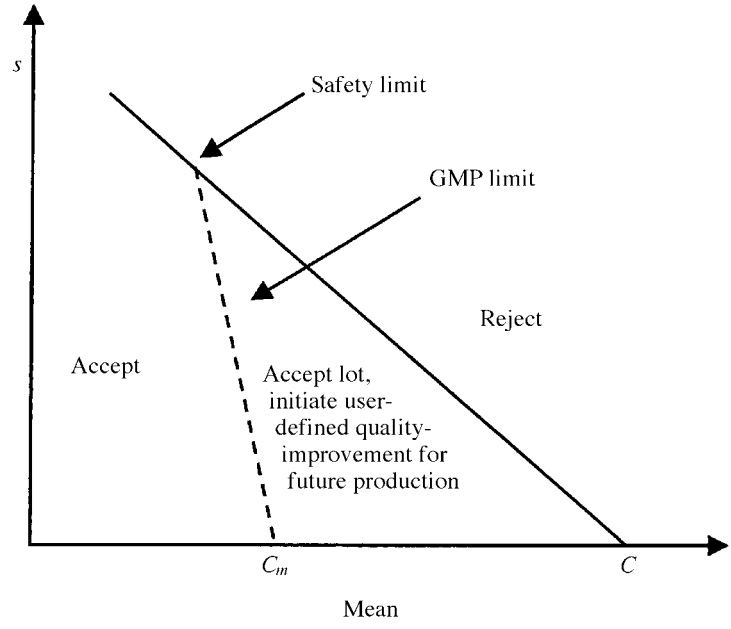

Fig 2.6 Schematic standards diagram showing GMP and safety limits for a variables sampling plan. The safety/quality limit is defined such that a lot has a probability of rejection of at least *Preject* if more than a given proportion p_d exceeds the limit C (see Table 2.3), i.e. the mean and standard deviation (s) of samples from the lot fall to the right of the safety/quality line. If the mean and standard deviation together fall to the left of the safety/quality line the lot will be accepted. But some quality improvement action will be taken unless the standard deviation and mean together fall to the left of the GMP line. The GMP line is defined so that the batch has a probability of acceptance of at least *Paccept* if less than the proportion p_m exceeds the limit C_m (Table 2.4). The GMP line is not operative to the right of the quality/safety line. (Adapted from Kilsby *et al*. (1979) with permission. © The Society for Applied Microbiology.)

and the proportion > C_m could be high. From Table 2.4, we find the value of k_2 that connects *Paccept*, P_d and the number of samples tested, *n*. Now we use the sample mean ($\bar{x}$) and standard deviation (s) to calculate $\bar{x} + k_2 s$ and if this is equal to C_m we accept the lot.

Figure 2.6 shows how the safety and GMP limits can be used together. Lots falling to the left of both the safety and GMP limits are completely satisfactory and lots falling to the right of the safety limit must be rejected. When a lot falls between the two limits this triggers an action to return the process to full control and may indicate additional actions with respect to that particular lot.

2.5.3 The connection between sampling and risk management

We could characterize the 'old-fashioned' approach to sampling as an attempt to 'test in' safety. The development of good manufacturing practice, good hygienic practice and HACCP methodology has moved us to the modern, proactive

approach to food safety management where we design and operate production systems to build in safety. Sampling then became a means of monitoring that the system was performing as expected and detecting any substantial failures. Looking forward, a movement towards risk-based food safety management is emerging. For foods in international trade, the ICMSF scheme for managing microbiological risks proposed the Food Safety Objective (FSO) as a functional link between risk assessment and risk management. The FSO is defined as 'a statement of the frequency or maximum concentration of a microbiological hazard in a food considered acceptable for consumer protection' (van Schothorst, 1998) and allows the equivalence of different control measures to be established. Control measures include killing microorganisms during processing, preventing them from growing, e.g. by formulation, or temperature control during storage, and excluding them using sampling together with appropriate microbiological criteria. In order to compare the equivalence of different control measures we must be able to relate their performance in terms of achieving an FSO. In other words, can they maintain the frequency or maximum concentration of a pathogen at, or below, the level required for consumer protection? Variables sampling plans, and attributes plans re-scaled in terms of concentration of organisms, enable the performance achieved by different control measures to be compared directly in terms of concentration.

2.6 Environmental sampling and tightened inspection/skip lot sampling

Environmental sampling may be thought to be strictly outside the scope of detecting pathogens in foods. But environmental pathogens in a food manufacturing facility can potentially become foodborne through cross-contamination so we will consider briefly this aspect of pathogen sampling. For the purpose of this discussion we assume that our hypothetical food processing facility is following the principles of good hygienic practice (GHP) as outlined by Codex (1997). This implies that there are procedures in place to manage the risk of environmental pathogens by, for example, minimizing cross-contamination from raw materials to finished product, having equipment of suitable sanitary design, appropriately maintaining and sanitizing equipment, removing waste, and training personnel. Environmental sampling then becomes a tool to ensure that those procedures are working. More detailed discussion of environmental sampling can be found in Holah (1999), ICMSF (2002) and Tompkin (2002).

2.6.1 Routine inspection

In routine environmental sampling we want to ensure that the operation remains under control by detecting any increase in the risk of cross-contamination. First we need to know something about the 'normal' risk of cross-contamination when

the operation is under control. We may, initially, be interested in a wide range of pathogens, indicators or even spoilage organisms but will likely select just one for ongoing monitoring. We would probably start by taking process flow samples at various stages of the operation to determine how the microbiological condition of the food changes within the process. Food residues (crumbs, shavings and the like) may be sampled as a kind of composite sample of the process flow at a particular location and a guide to the potential for recontamination from held-up product. Product contact surfaces may be sampled at points in the process flow using swabs or sponges to assess a large area. When appropriate, equipment should be shut down so that covers can be removed and samples taken from inside. Samples should be taken at different times (e.g. beginning, middle and end of each shift; before and after cleaning, etc.). Samples may also be taken from drains, cleaning equipment, joins between floors and process equipment, etc. These samples can indicate sources of ongoing contamination.

Results from all the samples should be arranged so that the 'normal' state of different locations when the process is under control can be seen and the most sensitive locations (and marker organisms) identified. The normal state may show a degree of seasonal variability, which will only become apparent over a number of years and is impossible to detect during an initial sampling exercise

Environmental sampling protocols are not statistically-based; rather they are designed to pay most attention to those areas known to pose the highest risk of product contamination. However, some statistical concepts are applicable. Trend analysis can indicate unfavorable changes over time, even if individual changes are small and differential inspection rates allow the frequency of inspection to alter in response to the history of particular sampling locations.

It is common practice to designate a number of zones within the processing plant, where different zones have different levels of risk of contaminating the product. ICMSF (2002) considers four zones:

- Zone 1 includes all product contact surfaces that can potentially directly contaminate the food. Even in zone 1, we are most concerned with surfaces that can recontaminate the food after any microbial kill step and before packaging so not all surfaces are created equal.
- Zone 2 includes equipment that is close to the product flow and could indirectly contaminate the food.
- Zone 3 includes items within the processing area that could harbor contamination and/or transmit contamination to items in Zones 1 or 2.
- Zone 4 includes items outside the processing area that, if not appropriately maintained, could serve as a reservoir of infection to the other three zones.

The exact balance of samples from the four zones is based on experience of the sites most likely to indicate that the operation is 'out of control' in terms of good hygienic practice. The selection of sites can shift over time as operation of the sampling protocol improves knowledge of the operation.

Knowing when to sample can be as important as knowing where to sample. In some operations, the most critical time may be immediately after startup. In

others, control may be lost slowly over time. The frequency of sampling may be increased if evidence indicates an increased risk and special sampling can be implemented in response to known risk factors such as construction.

2.6.2 Investigating a problem

We may start an investigation in response to information suggesting that a problem already exists. Here we probably want to know the source of the problem so that we can take steps to correct it. Random sampling is not an efficient approach to an investigation. Rather the investigator must apply knowledge of microbiology, process operations and equipment design, along with information gained from visual inspection of the operation and the trigger for the investigation to determine those sampling sites most likely to harbor the organism(s) of concern. Examples include, but are not limited to, those sites suggested to establish a baseline for a routine monitoring program. Even whilst applying this pre-existing knowledge, it is important not to jump to conclusions, particularly when resources are limited and time is constrained (as is usually the case). To maintain a degree of flexibility, investigational sampling is very likely to be iterative. For example, the pathogen of concern may seem to enter the process at a particular point. We can then go back to examine in more detail the influences between the last 'point of absence' and the first 'point of detection'. Over time the investigation works towards identifying the source and this process can sometimes be very lengthy.

2.6.3 Tightened inspection/skip lot sampling

When we have an ongoing relationship with a supplier, or other ongoing sampling situation, we can, over time, develop a level of confidence in the performance of that supplier. That may lead us to relax the rate of sampling, including skipping sampling of some lots altogether, using the freed sampling resources where they can be more beneficial. If, however, the reduced rate of sampling finds a 'defective' lot we will revert to the initial sampling rate, or even to a more stringent level of sampling, until we develop confidence that the supplier's performance has returned to an acceptable level. Formal methods for establishing such plans are described by Schilling (1982) and Montgomery (1996).

2.7 Taking samples

The process of physically selecting and taking samples is outside the scope of this chapter. Nevertheless, it may be helpful to keep some principles in mind:

- For acceptance sampling, where a key assumption is that samples are representative of the lot, sample units should be selected randomly using a pre-determined scheme. It is easiest to sample when material is in motion and therefore accessible.

- For investigations it is more efficient to use targeted sampling.
- Samples should be taken aseptically, and anyone charged with sampling for microbiological analysis needs appropriate training. Sampling equipment must be sterile, and must not violate management policies for other aspects of the operation. For example, it is common for glass to be banned from food production facilities and, where this is so, glass sample jars, etc., should not be used.
- All results should be traceable back to the original sample. Appropriate labeling and documentation should be used.
- Where samples must be stored and/or transported for analysis, they must be protected from contamination and physical damage, and maintained to prevent or minimize changes in the microbial population present. Wet or perishable samples may need to be maintained at chill temperatures in a refrigerator and shipped in an insulated container with ice or freezer packs. Frozen samples may need to be packed with dry ice to prevent thawing in transit.
- The analyzing laboratory, whether in-house or a contractor or consultant should be capable of performing the required analyses reliably. Laboratories should have some formal procedures in place to assure this reliability. Formal quality management systems include those based on the ISO 9000 series, ISO Guide 25 or the principles of Good Laboratory Practice (Wood *et al*., 1998).
- Samples should be as large as reasonably practical to minimize the impact of the distribution of microorganisms within the sample. For packaged food products the sample unit will be one or multiple product units (more product units for smaller products). For unpackaged food products, food within process, or bulk raw materials such as flour, sugar, etc., sample units in the region of 200–500 g are reasonably practicable. For environmental samples taken from a processing plant it may only be possible to recover 1–2 g, for example if scraping material from a gasket in a piece of process equipment. If it seems important to sample the site based on pre-existing knowledge or observations, the small sample should still be taken.
- The sample unit is usually larger than the amount actually analyzed (the analytical sample). Where the nature of the sample permits, it should be aseptically blended before removing the analytical unit. For example shaking or squeezing may blend liquids or powders in a container with some headspace. This mixing is best done immediately before removing the analytical sample.
- In qualitative analysis for the presence of a pathogen, combining multiple analytical samples into the enrichment phase of an analysis can significantly reduce the total amount of work when it has been shown to give equivalent recovery to not compositing. This does not sacrifice the discriminating power of the sampling plan since, with a two-class sampling plan where $c = 0$, a single positive result is enough to indicate rejection of the lot. The critical condition is whether the total mass of analytical samples contains at least one detectable pathogen. Compositing is generally not sound in quantitative

testing, whether by a two- or three-class plan, because it has the effect of averaging the count across all the samples.

For further detailed discussion on this topic, see Chapter 3.

2.8 Maximizing the value of test results

It is common practice to use test results once only to support a decision based on the plan criteria, e.g. to accept or reject a lot. The historical value of individual data points is often overlooked. Consequently quality laboratories may have files containing microbiological test results generated at great expense. Some simple techniques for plotting and statistically analyzing those data may greatly improve the return on the initial investment in testing. These techniques come from the field of statistical process control (Montgomery, 1996), which is used to great effect in many manufacturing operations but is still rarely applied to microbiological data.

2.8.1 Control charts

Control charts are sequential plots of the data or some derivative such as the mean, range, standard deviation, etc., to allow trends to be observed. Depending on the exact nature of the chart, various kinds of warning or action limits or tests can be established.

Variables data

The simplest kind of control chart is known as an 'X', or 'individual measurement', chart. It is just a sequential plot of numerical results (e.g. log counts of an indicator or spoilage organism) against order of testing. This chart can show trends that may otherwise be missed, potentially allowing appropriate intervention before the count becomes unacceptably high. Equally it may help to identify times when the process results are better than normal and there is the opportunity to investigate to determine the cause and permit an ongoing improvement (Fig. 2.7a). A 'control limit' can be determined using 'expert judgment' and be drawn on the chart to give an immediate visual indication of an 'out of control' event. This control limit is typically the count at which the result is considered marginally acceptable. The criterion defining 'out of control' is then based on the number of marginally acceptable results observed within a pre-determined period of time or number of samples analyzed. Figure 2.7a shows hypothetical test results for a spoilage organism in samples from a process that is showing a trend towards loss of control through January. The pre-determined criterion of two sequential marginally acceptable results shows that control was first lost in late January, but no action is taken until the acceptable limit is breached by a large margin in mid February. That triggered an investigation of the cause and a corrective intervention in late February resulting

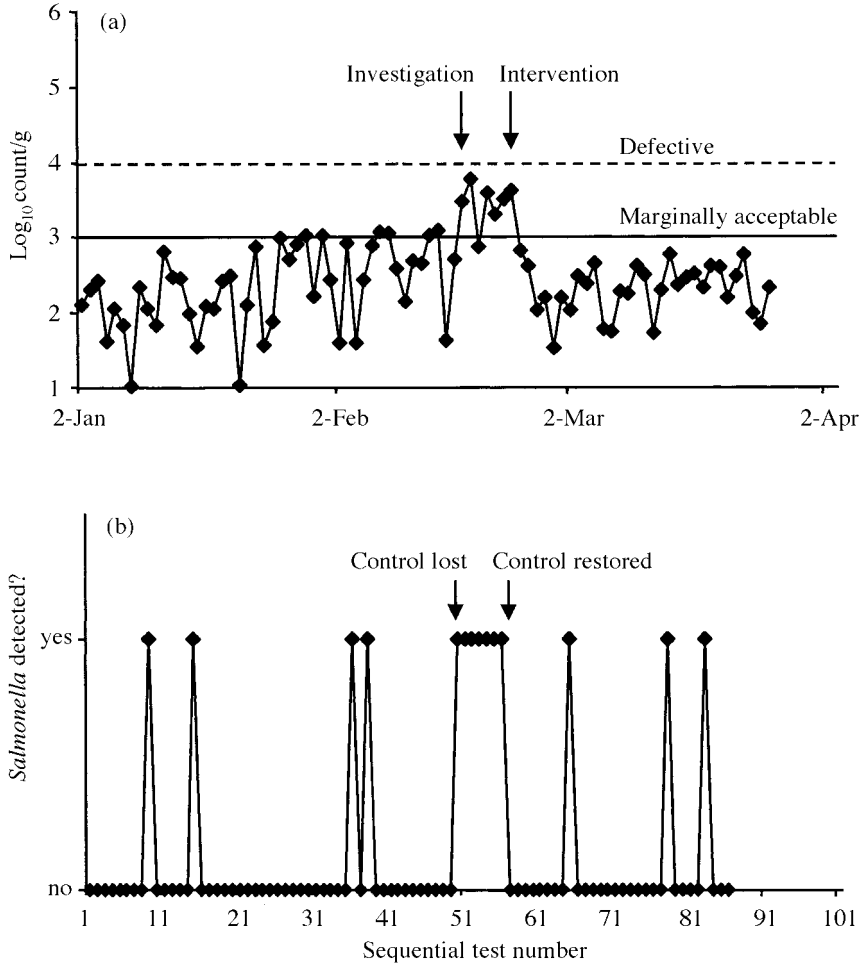

Fig. 2.7 Hypothetical examples of control charts for (a) individual measurements ('X') and (b) attributes.

in control being regained. Note that the intervention reduced the variability as well as the mean of the test results. Also note that a response to the first observation of two sequential marginally acceptable results, even though individually they only just exceeded the limit, would have prevented the more serious later breach of the limit.

Attributes data

When the test gives only a positive or negative result, as when testing for the presence of *Salmonella*, it is still possible to plot the results in sequence. Now the 'out of control' condition is indicated when more than a pre-determined number of positive results are observed in sequence (Fig. 2.7b). Since this type

of chart can only show the direction of a result (positive or negative) it responds only to the sequence. Hence an 'out of control' condition will only be identified after a delay. Figure 2.7b shows hypothetical results of daily *Salmonella* tests for a process with a 10% incidence of positive results when in control. On the basis of past experience, positive results from three sequential tests are considered to indicate loss of control. The chart shows that control was lost on day 50 but, because the test method introduces a delay, this was not detected until day 56 when corrective action was immediately implemented and control was regained.

More sophisticated control charts use formal statistical rules to derive the control limits, although the limits can be modified subsequently in the light of experience. Their use for microbiological data has been described by Jarvis (2000) and ICMSF (2002) and their statistical principles discussed in detail by Montgomery (1996). Using these charts we can:

- determine whether the process is capable of consistently delivering the required performance (e.g. food safety objective, trading specification)
- estimate the probability of severely out of control conditions based merely on the process variability (Peleg, 2002)
- access a number of additional decision rules to identify out of control conditions on the basis of trends in the data, even when the control limits are not breached.

Of course, within the context of food safety management, the use of control charts need not be restricted to microbiological results. They can be used to show data for any type of measurement used to manage a critical control point within a HACCP program and Khongsak and Hourigan (2002) gave an example for critical control points in shell egg washing. A full discussion of these approaches is outside the scope of this chapter, but the reader is encouraged to investigate further.

2.8.2 Process variability reduction

The variability in a sequence of data is, itself, an important source of information. After statistical analysis to determine the distribution that best describes the observed variability it is possible to estimate the future frequency of extreme events, though not to predict when they will occur (Peleg and Horowitz, 2000; Nussinovitch and Peleg, 2000; Engel *et al*., 2001; Peleg, 2002). It is apparent that results are likely to reach extreme values more frequently when processes have large variability than when the variability is small. Hence, management intervention that reduces the variability of a process also reduces the likelihood of an extreme event such as an outbreak of food poisoning. The information on variability is already present in our records of test results over time. Such approaches are widely used in engineering but they are still quite rare for microbiological control. For example, Reneau (2001) described the use of an individual measurement control chart to plot bulk tank somatic cell count (BTSCC) in milk from a Wisconsin dairy herd. The BTSCC correlates with milk

microbiological quality and other factors. The chart allowed the reduced variability in somatic cell count caused by a management intervention to be identified rapidly so that the improvement in milking practice could be captured.

2.9 Future trends

Consumers in all developed countries are increasingly aware of food safety. Particularly since the emergence of foodborne listeriosis in the 1980s and of pathogenic *Escherichia coli* in the 1990s this has increased the pressure for pathogen testing and seems likely to increase the total number of samples tested. However, industry and regulators have known for many years that it is impossible to 'test in' safety. Rather, food safety is 'designed in' to products and assured by appropriate control procedures. Thus it seems likely that we will continue to see sampling resources increasingly used to monitor or verify process controls and assure good hygienic practice. As this shift progresses we would expect that 'foods', certainly finished products, will represent a smaller proportion of all samples tested for pathogens during food safety management. The trend towards Food Safety Objectives as risk-based targets will maintain the tendency towards control of the process and distribution system. It will also, in time, drive a more quantitative understanding of food safety and encourage the use of quantitative sampling plans that help to demonstrate the equivalence of different control measures. There is clearly value in the re-use of data through charting techniques and we expect that this will become commonplace in time. Indeed this is an area that seems poised for growth, initially through the advocacy of a small number of influential groups and later gathering momentum as the power of these approaches becomes more widely recognized. In short, we expect sampling to remain a critical element of food safety management for the foreseeable future, but with a more quantitative and proactive role than in the past.

2.10 Sources of further information and advice

Advice on the risk-based approach to food safety management is given in ICMSF (2002). This book gives detailed guidance on most of the topics discussed in this chapter. Jarvis (1989, 2000) gives more detail on the underlying statistics of sampling from the perspective of an experienced microbiologist. Information on ICMSF members, programs and publications can be found at http://www.dfst.csiro.au/icmsf.htm. For more background on attributes plans we recommend ICMSF (1986). Now out of print, part two of this book, giving sampling plans and criteria for a wide range of foodstuffs, is available as a free download from the ICMSF website at http://dfst.csiro.au/icmsf/book2.htm. A Microsoft Excel workbook for exploring performance of ICMSF recommended plans in terms of concentration of microorganisms based on the approach of

Legan *et al*. (2001) can also be downloaded free at http://www.dfst.csiro.au/icmsf/samplingplans.htm. Another Excel workbook, based on the same principles but taking a different approach to designing and improving sampling plans, is available with a review from Campden & Chorleywood Food Research Association (Goddard *et al*., 2001). Many statistical packages can be used to create control charts. An easier approach for some people might be to use the Microsoft Excel workbook and accompanying guide available from Campden & Chorleywood Food Research Association (Jewell *et al*., 2002).

2.11 Acknowledgements

We would like to thank Joe Meyer for helpful comments on the manuscript, Martin Cole for prompting us to work on acceptance sampling in the first place, Keith Jewell for thoughts on use of control charts for microbiological data, and Lynne Hare for support and encouragement.

2.12 References

ANON (1982) *The American Heritage Dictionary New College Edition*. Houghton Mifflin Publishers, Boston.

BUCHANAN, R.L., DAMERT, W.G., WHITING, R.C. and VAN SCHOTHORST, M. (1997) Use of epidemiologic and food survey data to estimate a purposefully conservative dose-response relationship for *Listeria monocytogenes* levels and incidence of listeriosis. *Journal of Food Protection*, **60**(8) 918–22.

CODEX (1997) Codex Alimentarius Commission, J*oint FAO/WHO food standards programme Codex committee on food hygiene, supplement to volume 1B- 1997. Recommended international code of practice, general principles of food hygiene.* CAC/RCP-1-1969, Rev. 3.

DAHMS, S. and HILDEBRANDT, G. (1998) Some remarks on the design of three-class sampling plans. *Journal of Food Protection*, **61**(6) 757–61.

ENGEL, R., NORMAND, M.D., HOROWITZ, J. and PELEG, M. (2001) A qualitative probabilistic model of microbial outbursts in foods. *Journal of the Science of Food and Agriculture*, **81** 1250–62.

FAO (FOOD AND AGRICULTURE ORGANISATION) (1993) *Sampling plans for aflatoxin analysis in peanuts and corn. FAO Food and Nutrition Paper no. 55.* FAO, Rome.

FLOWERS, R.S. and CURIALE, M.S. (1993) Qualitative microbiological testing: discrepancies between original and retest results. *Scope*, **8** 1–5. Silliker Laboratories, Homewood, IL.

FOSTER, E.M. (1971) The control of *Salmonellae* in processed foods: a classification system and sampling plan. *Journal of the AOAC,* **54**(2) 259–66.

GODDARD, M., JEWELL, K., MORTON, R.S., PAYNTER, O., RUEGG, J. and VOYSEY, P.A. (eds) (2001) *Designing and improving acceptance sampling plans – a tool*. Review No. 27. Campden & Chorleywood Food Research Association, Chipping Campden, UK.

HARRIS, C.S.M and RICHARDSON, P.S. (1996) *Review of principles, techniques and benefits*

of statistical process control. Review No. 4. Campden & Chorleywood Food Research Association, Chipping Campden, UK.

HILDEBRANDT, G., BÖHMER, L. and DAHMS, S. (1995) Three-class attributes plans in microbiological quality control: a contribution to the discussion. *Journal of Food Protection,* **58** 784–90.

HOLAH, J (ed.) (1999) *Effective microbiological sampling of food processing environments.* Report no. 20. Campden & Chorleywood Food Research Association, Chipping Campden, UK.

HUBBARD, M.R. (1996) *Statistical quality control for the food industry*. Chapman & Hall, London.

ICMSF (INTERNATIONAL COMMISSION ON MICROBIOLOGICAL SPECIFICATIONS FOR FOODS) (1974) *Microorganisms in Foods 2. Sampling for microbiological analysis; principles and specific applications*. University of Toronto Press, Toronto.

ICMSF (INTERNATIONAL COMMISSION ON MICROBIOLOGICAL SPECIFICATIONS FOR FOODS) (1986) *Microorganisms in Foods 2. Sampling for microbiological analysis; principles and specific applications* (2nd edn). Roberts, T.A., Bryan, F.L., Christian, J.H.B., Kilsby, D., Olson, J.C. Jr. and Silliker, J.H. (eds). University of Toronto Press, Toronto.

ICMSF (INTERNATIONAL COMMISSION ON MICROBIOLOGICAL SPECIFICATIONS FOR FOODS) (1988) *Microorganisms in Foods 4. Application of the hazard analysis critical control point (HACCP) system to ensure microbiological safety and quality.* Blackwell Scientific Publications, Oxford.

ICMSF (INTERNATIONAL COMMISSION ON MICROBIOLOGICAL SPECIFICATIONS FOR FOODS) (1997) Establishment of microbiological safety criteria for foods in international trade. *World Health Statistics Quarterly*, **50** 119–23.

ICMSF (INTERNATIONAL COMMISSION ON MICROBIOLOGICAL SPECIFICATIONS FOR FOODS) (2002) *Microorganisms in foods 7. Microbiological testing in food safety management.* Tompkin, R.B., Roberts, T.A., van Schothorst, M., Cole, M.B., Gram, L., Buchanan R.L. and Dahms S. (eds). Kluwer Academic/Plenum Publishers, New York.

JARVIS, B. (1989) Statistical aspects of the microbiological analysis of foods. *Progress in Industrial Microbiology*, vol 21. Elsevier, Amsterdam.

JARVIS, B. (2000) Sampling for microbiological analysis. In *The microbiological safety and quality of Food*, Vol 2. pp. 1691–733. Lund, B.M., Baird-Parker, T.C. and Gould, G.W. (eds). Aspen Publishers, Gaithersburg, MD.

JEWELL, K., VOYSEY, P AND THE STATISTICS AND SAMPLING WORKING PARTY (2002) *Statistical quality assurance: how to use your microbiological data more than once.* Review No. 36. Campden & Chorleywood Food Research Association, Chipping Campden, UK.

KHONGSAK, S. and HOURIGAN, J.A. (2002) The use of statistical process control (SPC) to enhance the validation of critical control points (CCPs) in shell egg washing. *Food Control*, **13** 263–73.

KILSBY, D. (1982) Sampling schemes and limits. In Brown, M.H. (ed.), *Meat Microbiology*, pp. 387–421. Applied Science Publishers, London.

KILSBY, D., ASPINALL, L.J. and BAIRD-PARKER, A.C. (1979) A system for setting numerical microbiological specifications for foods. *Journal of Applied Bacteriology*, **46** 591–9.

LEGAN, J.D., VANDEVEN, M.H., DAHMS, M.H. and COLE, M.B. (2001) Determining the concentration of micro-organisms controlled by attributes sampling plans. *Food Control*, **12** 137–47.

MALCOLM, S. (1984) A note on the use of the non-central *t*-distribution in setting numerical microbiological specifications for foods. *Journal of Applied Bacteriology,* **57** 175–7.

MONTGOMERY, D. (1996) *Introduction to Statistical Quality Control*, 3rd edn, John Wiley & Sons, New York.

NUSSINOVITCH, A. and PELEG, M. (2000) Analysis of the fluctuating patterns of microbial counts in frozen industrial food products. *Food Research International,* **33** 53–62.

PELEG, M. (2002) Interpretation of the irregularly fluctuating microbial counts in commercial dairy products. *International Dairy Journal*, **12** 255–62.

PELEG, M. and HOROWITZ, J. (2000) On estimating the probability of aperiodic outbursts of microbial populations from their fluctuating counts. *Bulletin of Mathematical Biology*, **62** 17–35.

RENEAU, J.K. (2001) Somatic cell counts: measures of farm management and milk quality. *National Mastitis Council Annual Meeting Proceedings (2001)*, 29–37.

SCHILLING, E.G. (1982) *Acceptance Sampling in Quality Control,* Marcel Dekker, New York.

SHAPTON, D.A. and SHAPTON, N.F. (1991) *Principles and Practices for the Safe Processing of Foods*. Butterworth-Heinemann Ltd., Oxford.

TOMPKIN, R.B. (2002) Control of *Listeria monocytogenes* in the food-processing environment. *Journal of Food Protection*, **65** 709–25.

VAN SCHOTHORST, M. (1998) Principles for the establishment of microbiological food safety objectives and related control measures. *Food Control*, **9** 379–84.

WHITAKER, T.B., GIESBRECHT, F.G., WU, J., HAYLER, W.H. and DOWELL, F.E. (1994) Predicting the distribution of aflatoxin test results from farmer's stock peanuts. *Journal of AOAC International*, **77** 659–66.

WHITING, R.C. and BUCHANAN, R.L. (1997) Development of a quantitative risk assessment model for *Salmonella enteritidis* in pasteurized liquid eggs. *International Journal of Food Microbiology*, **36**, 111–25.

WOOD, R., NILSSON, A. and WALLIN, H. (1998) *Quality in the Food Analysis Laboratory,* pp 16–35. Cambridge, RSC.