

Codex Alimentarius Concepts of Sampling

P. G. MARTIN

The Public Analyst's Laboratory, 220 Elgar Road, Reading

The work of the Codex Alimentarius in relation to the sampling of foods is briefly described.

I would like to preface my remarks with a brief description of the way in which the Codex Alimentarius works. The idea of codifying food standards originated with the Codex Alimentarius Europeanus in 1958. The attempt at codification was taken over by the Food and Agriculture Organisation and the World Health Organisation of the United Nations in the early sixties and was thereafter on a world basis, rather than exclusively European. Although the programme of work is run jointly by the two U.N. agencies, the Secretariat is within the Food Policy and Nutrition Division of F.A.O. in Rome. The executive body, the Codex Commission, meets every two years. This makes recommendations on the basis of reports submitted to it by a number of Committees. Some of these are commodity Committees, such as the Oils and Fats Committee, hosted by the U.K. Others refer to particular aspects of food in general such as the Pesticides Residues Committee and the Codex Committee on Methods of Sampling and Analysis (CCMAS), hosted by Hungary. Methods of analysis have often been elaborated by commodity Committee but are then approved by CCMAS.

The work of the Codex is largely carried out by these Committees, made up of members appointed by national governments. The composition of the national delegation, appointment of members to a Committee, consultation and all other Codex work within a country is entirely the responsibility of the national government which is in official communication with the Codex Alimentarius via the National Codex contact point, in our case Miss McElnea at the Ministry of Agriculture, Fisheries and Food.

Generation of a Codex recommended standard is through a long, ten-step process that takes several years and involves two formal requests for government comments, although there are other opportunities along the line for comment. The first and so far only recommended sampling procedure to get, so to speak, past the post is RM42¹ of 1969, Sampling Plans for Prepackaged Foods (AQL6.5). These plans enable one to accept or reject a lot or consignment according to whether it apparently contains more or less than the agreed proportion of defective units. The agreed proportion in this instance is 6.5 per cent., hence the name AQL (Acceptable Quality Level) 6.5. I say "apparently" because the statistics are such that there is only a probability rather than a

certainty of discovering by sampling whether the proportion of defectives in the whole lot is really above or below 6.5 per cent. "Defective" refers to failure to conform to a quality characteristic evaluated by organoleptic or physical means such as colour, flavour, texture, size, appearance etc. These sampling plans are not intended to cover factors which may present a hazard to health, or which are unwholesome or otherwise highly objectionable to the consumer, such as the presence of pesticide residues, contaminants, and foreign matter exemplified by stones or large insects or blown cans. Although intended primarily for quality evaluation, these plans may be suitable for some determinations such as net weight, drained weight and Brix, provided 6.5 per cent. is considered the appropriate AQL.

These plans are only recommended by the Codex Alimentarius in relation to assessing defectives in products covered by certain Codex Standards, there so far being 34 of them, including the recommended International Standards for Canned Fruit Cocktail, for Jams, Jellies and Citrus Marmalade, for Quick-frozen Shrimps or Prawns, and for Canned Crab Meat. For small lots and sampling at the retail level, the ratio of sample size to lot size becomes too high to make these plans the ones of choice.

It is interesting to speculate that concern with the sampling of lots and consignments to the exclusion of problems at the retail level reflects the fact that national delegations are almost invariably derived from central government entities.

TABLE I
SAMPLING PLAN 1 (INSPECTION LEVEL I, AQL = 6.5)

Lot size (<i>N</i>)	Sample size (<i>n</i>)	Acceptance number (<i>c</i>)
4800 or less	6	1
4801 $\approx$ 24,000	13	2
24,001 $\approx$ 48,000	21	3
48,001 $\approx$ 84,000	29	4
84,001 $\approx$ 144,000	48	6
144,001 $\approx$ 240,000	84	9
More than 240,000	126	13
Net weight is greater than 1 kg (2.2 lb) but not more than 4.5 kg (10 lb)		
2400 or less	6	1
2401 $\approx$ 12,000	13	2
12,001 $\approx$ 24,000	21	3
24,001 $\approx$ 42,000	29	4
42,001 $\approx$ 72,000	48	6
72,001 $\approx$ 120,000	84	9
More than 120,000	126	13
Net weight greater than 4.5 kg (10 lb)		
600 or less	6	1
601 $\approx$ 2000	13	2
2001 $\approx$ 7200	21	3
7201 $\approx$ 15,000	29	4
15,001 $\approx$ 24,000	48	6
24,001 $\approx$ 42,000	84	9
More than 42,000	126	13

TABLE II
SAMPLING PLAN 2 (INSPECTION LEVEL II, AQL = 6.5)

Lot size (N)	Sample size (n)	Acceptance number (c)
4800 or less	13	2
4801 $\approx$ 24,000	21	3
24,001 $\approx$ 48,000	29	4
48,001 $\approx$ 84,000	48	6
84,001 $\approx$ 144,000	84	9
144,001 $\approx$ 240,000	126	13
More than 240,000	200	19
Net weight is greater than 1 kg (2.2 lb) but not more than 4.5 kg (10 lb)		
2400 or less	13	2
2401 $\approx$ 12,000	21	3
12,001 $\approx$ 24,000	29	4
24,001 $\approx$ 42,000	48	6
42,001 $\approx$ 72,000	84	9
72,001 $\approx$ 120,000	126	13
More than 120,000	200	19
Net weight greater than 4.5 kg (10 lb)		
600 or less	13	2
601 $\approx$ 2000	21	3
2001 $\approx$ 7200	29	4
7201 $\approx$ 15,000	48	6
15,001 $\approx$ 24,000	84	9
24,001 $\approx$ 42,000	126	13
More than 42,000	200	19

The plans (see Tables I and II) include 2 inspection levels, one for normal use and the second if there is lack of agreement, the lot is suspect and so on. The plan defines, at each of these two levels, for a lot made up of primary units N , the number of samples n , that must be taken at random and the maximum number of these which may be defective (the acceptance number c). The number of samples to be taken and the acceptance number also varies with weight of the primary units. The values of n and c are read off from a table. The sampling plan at an Acceptable Quality Level of 6.5 per cent. will accept a lot or production which has 6.5 per cent. defective units approximately 95 per cent. of the time. Reference to operating characteristic curves shows how reliable results are if more samples are taken or if the actual level of defectives is substantially different from 6.5 per cent. In studying the OC curves (Figure 1) the following conclusions can be drawn.

1. All the curves have the same general slope, although that for when only 6 samples are drawn is flatter.
2. All curves intersect at a point represented by the co-ordinates of "6.5 per cent. defective" and approximately "95 per cent. probability of acceptance".
3. As the sample size increases, the curves become steeper and more discriminating, i.e. lots having defectives in excess of 6.5 per cent. are rejected with greater frequency.
4. The reliability of the larger sample size is not in direct proportion to the increased sample. For example, for a lot that is 20 per cent. defective a

sample size of 6 (curve E) will accept such a lot 65 per cent. of the time: whereas a sample size of 48 (curve L) will accept the same lot 22 per cent. of the time. In this example, increasing the number of samples eight-fold only decreases the likelihood of acceptance by a third.

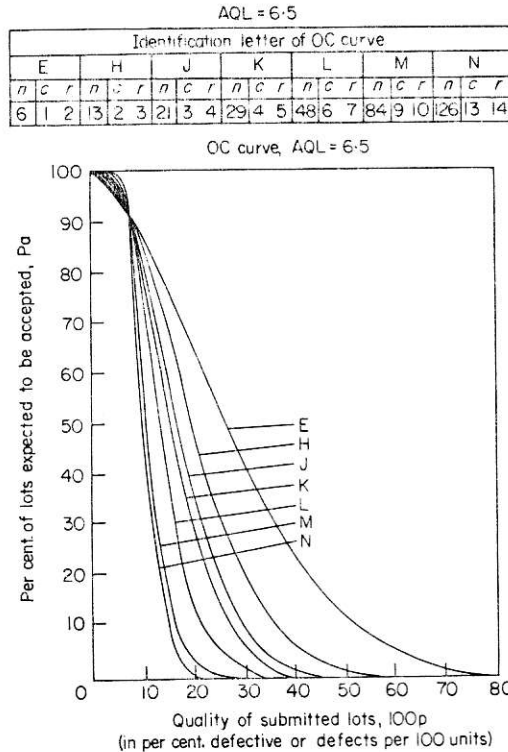


Fig. 1. Operating characteristic curves. AQL = Acceptable Quality Level; OC = operating characteristic.

To illustrate the use of the OC curves (AQL 6.5) suppose a lot is 10 per cent. defective. Looking at curve E, 6 samples being taken, the lot will be accepted 88 per cent. of the time. Taking 84 samples (curve M) the lot will be accepted 65 per cent. of the time. If a lot is 30 per cent. defective, 6 samples will show the lot acceptable 42 per cent. of the time, 21 samples will show the lot acceptable 8 per cent. of the time and 84 samples will always show that the lot fails. Thus although the plan is calculated on 6.5 per cent. as the maximum level of acceptable defectives, it is only really effective in picking out lots about 15–20 per cent. defective and even then only if the higher rates of sampling are used. 6.5 per cent. was picked as the level that could reasonably be maintained by good manufacturing practice without undue hidden cost to the consumer, but the appropriateness of the figure may be questioned. More effective sampling is likely to be very expensive in time and money and therefore a balance has to be sought.

Current discussion on sampling within the CCMAS is on the so-called "Moderate Acceptance Probability" sampling plan. This is an attempt to evolve a scheme incorporating the best of the 50-50 Indifference Quality Level (IQL) plans used in the U.S.A., Canada and Australia and the "High Acceptance Probability" (HAP) plans used by the E.E.C. and some others. The plan was presented to a working group of the International Organisation for Legal Metrology last Spring and brought before the CCMAS meeting in Budapest last October by the Swiss delegation. The plan is for random and infrequent checking of consignments and is not intended for "in-line" quality control. It is intended to establish the acceptability or non-acceptability of lots of pre-packed commodities prior to sale. Among its important features are the following.

1. The standard deviation of the frequency distribution of quality parameters, e.g. net contents must not exceed a specified maximum value, i.e. limits are set on uniformity rather than on setting tolerance limits. This permits reduction in sample size.
2. Inspection can be either by attributes or by variables provided the selected plan has an operating characteristic curve equivalent to the one required or specified by legal requirements, administrative guidelines or other rules.
3. Inspections may be repeated, the level of confidence increasing markedly with each inspection. For example, if sampling is arranged so that the rejection risk on marginal lots (i.e. the risk that a good lot will be failed) is 16 per cent. at the first sampling, a second sampling reduces this to 2.5 per cent. and a third to 0.4 per cent.
4. The plan is such that a decision could be made after the first, or second or third sampling, with increasing levels of confidence at each successive step, so that the number of steps chosen would be appropriate to the importance of the enforcement action contemplated, i.e. warning, severe reprimand or fine, confiscation of goods.
5. The plan differs from many others in that sample size is independent of lot size and is constant once certain choices are made, such as whether sampling is to be by attributes or by variables, and is to be single, double or sequential.

Those two plans, the earlier one complete with assumptions about how to interpret the data (i.e. the implied statement "all consignments with more than 6.5 per cent. defective units are to be rejected") and the second one still at the discussion stage and essentially a sampling plan only without implied value judgements on the results, are attempts at sampling pre-packed goods for quality.

I would now like to turn to the attempts to devise sampling plans for contaminants. A working group of the Codex Committee on pesticide residues have proposed a sampling plan for lots that must be checked for conformity with recommended maximum Pesticide Residue levels. This specifies that between three and fifteen primary samples must be taken depending on the weight of the lot, 3 for less than 50 kilo lots and 15 if the lot weight is over 2000 kilos. If the lot is in small containers, 1-15 containers are taken, the

maximum amount for lots containing over 250 containers. The well-established procedures for materials shipped in bulk, such as grain, may be used instead. These primary samples are mixed to form the bulk sample which is subsequently submitted to the laboratory, or if too large for this, divided by suitable means, but without cutting individual fruits or vegetables. Minimum amounts of material to be submitted to the laboratory are specified and vary from 0.5–2 kilos for different products, minimum numbers also being specified for larger fruits and vegetables, 5 for cabbages, melons, cucumbers, 10 for approximately apple-sized products such as oranges, carrots, potatoes, and 10 for eggs. This procedure is intended to give the average pesticide residue level in the lot and has been praised for its practicability, but I have been unable to trace any discussion of a statistical base for it and I am inclined to view its simplicity with a certain degree of suspicion. I wonder if it would indeed give anything like a reliable estimate of the average pesticide residue, especially when one is required to take only 15 cabbages, say from a lot of several tons, or a total of 1 kilo from 15 different places in a bulk consignment of tons of peas or olives. The Pesticide Residues Committee has decided not to seek endorsement of its methods of analysis and sampling by the Committee on Methods of Analysis and Sampling so we do not at the moment have any benefit that might derive from an interchange of views. I am never quite sure if pesticide residues are additives or contaminants but there is one aspect in which they may differ from some contaminants. An attempt is presumably made to spray pesticide over a crop as evenly as possible (though I wonder about the frequency of the human element such as using up or washing out a knapsack sprayer in a small corner of the field).

Some contaminants, notoriously aflatoxin, are extremely unevenly distributed through the product, for example one or two nuts may contain 1000 mg/kg of aflatoxin and the other million in a lot be uncontaminated. It only needs 5 nuts to be contaminated to that level to make the average for the entire lot equal to the U.S. limit of 5 p.p.b. Since only a small part of the lot would actually be sampled, the chance inclusion of even one of the heavily contaminated nuts would lead to rejection of the lot. F.A.O. have recently published "Mycotoxin Surveillance; A Guideline²". In this it is pointed out that the only statistically-based sampling plan for contaminants of this type is that developed by Whitaker^{3,4} and co-workers for aflatoxin and shelled peanuts and that it is not known whether the plan is applicable to other commodities. On the basis of this the Food and Drug Administration have developed sampling procedures for aflatoxin and these have been accepted by F.A.O. and reproduced in the publication mentioned above. This is an activity entirely separate from the Codex Alimentarius work. The U.S. collect samples for aflatoxin testing as follows:

Commodity	Sample size
Smooth peanut butter	24 subsamples of over 8 oz or 12 × 1 lb subsamples, randomly drawn
Peanuts, roasted in shell	Random sample of 25 lb
Other peanut products	48 units of >1 lb or 48 × 1 lb subsamples from a bulk
Almonds, walnuts, peanuts, figs	20 lb from >20 units or 40 lb from >40 units for lots larger than 200 units or bulk
Fig paste	10 lb total from not less than 10 units

They carry out nearly 1000 analyses yearly on these products so that total weight of samples taken is of the order of 15-20 tons.

They are rather more strict on imported foods, requiring for example that a 50-lb sample is collected at random from not less than 20 per cent. of the available units of consignments of pistachio nuts over 75,000 lbs. In the case of maize, every bag is sampled to give a sample of about 70 kg. This is subdivided to obtain about 5 kg and if examination by U.V. light suggests this is contaminated the entire 5 kg is ground to pass a 20-mesh sieve prior to analysis.

In conclusion, if I may return to the Codex for a moment, the Methods of Analysis and Sampling Committee, which has spent much of its life examining and approving methods of analysis, is now beginning to lay much more stress on sampling. The Codex Alimentarius may prove to be the most satisfactory forum in which to discuss the relative merits of the IQL and HAP sampling plans and develop a viable improvement, as some would consider the "Moderate Acceptance Probability" sampling plan.

The development of sampling methodology within F.A.O. and W.H.O. has been for the solution of specific problems and the approach has therefore been somewhat piecemeal. In general, the results are perhaps best considered not as the views of the U.N. Agencies but as the consensus views of the 115 Member Governments who are able to use the organisations as practical means for discussion and harmonisation.

References

1. "F.A.O./W.H.O. Codex Alimentarius Sampling Plans for Prepackaged Foods", 1969, (AQL6-5), CAC/RM 42-1969.*
2. "Mycotoxin Surveillance, a Guideline". F.A.O. Food Control Series No. 4, F.A.O., 1977.
3. Whitaker, T. B., Dickens, J. W., and Wiser, E. H., *J. Amer. Oil Chemists Soc.*, 1976, **53**, 545.
4. Velasco, J., Whitaker, T. B., and Whitten, M. E., *J. Amer. Oil Chemists Soc.*, 1975, **52**, 191.

*F.A.O. documents may be obtained from H.M.S.O. or Distribution and Sales, F.A.O., Via delle Terme di Caracalla, 00100-Rome. Some Codex documents are available from Miss W. Noble, Codex Secretariat, F.A.O.