WELMEC Publication 6.5

Guidance on Controls by Competent Department's on "e" marked Prepackages

Published by:
WELMEC Secretariat
BEV
Arltgasse 35
A-1160 Vienna
Austria

Tel +43 1 21176 3608
Fax +43 1 49 20 875 8006

Email welmec@metrolgie.at

INDEX

1 Background and Introduction

1.1 WELMEC Working Group 6 was set up to discuss, and propose solutions for, the problems associated with the trading of prepackaged goods between EEA countries. It was decided that a universal manual for inspectors, which could be used by Competent Departments in all EEA countries, should be produced. The intention of the manual is to achieve a uniform level of enforcement.

1.2 The OIML recommendations 79 and 87 on prepackages have been used to produce best practice on some points that are not specified in the Directives, where the OIML requirements do not align with those in the Directives the these are referred to in 'Notes'.

This document looks particularly at the requirements of Directive 76/211/EEC (the Directive), which are similar to those in Directive 75/106/EEC, but where relevant requirements in other Directives are also considered. The guidance in WELMEC document 6.3 is reflected in this document.

The Directive also specifies that checks shall be carried out by Competent Departments to ensure that prepackages meet the above requirements. The methods of working in Member States vary with some only using enforcement agencies to ensure compliance and others using certification bodies for assessing systems. The means of recognizing procedures and the agencies involved in the implementation of the Directive in each Member State are given in WELMEC document 6.0.

1.4 The Directive requires the prepackages between 5 g and 10 kg or 5 ml and 10 L: meet the content requirements, sometimes referred to as the 3 Packer's Rules, meet the labeling requirements with regard to quantity, packer's or importer's identity, and the 'e' mark, have an appropriate quantity control system to ensure that the 3 Packer's Rules are met, and to keep adequate records to show this.

1.5 The 3 Packer's Rules can be summarized as:

a) the average content shall not be less than the stated (nominal) quantity,

b) there shall be no more than 2.5%[1] of prepackages with a content below TU_1, and

c) there shall be no prepackages with a content below TU_2.

Where:

- TU_1 is the nominal quantity less one tolerable negative error (TNE) and
- TU_2 is the nominal quantity less two TNE.

1.6 The TNE in the contents of a prepackage is fixed in accordance with the table below:

Nominal quantity Q_n in grams or milliliters	Tolerable negative error	
	As % of Qn	g or ml
5 to 50	9	
from 50 to 100		4.5
from 100 to 200	4.5	
from 200 to 300		9
from 300 to 500	3	
from 500 to 1 000		15
from 1 000 to 10 000	1.5	

[1] Regarding the second rule, the Directive specifies an acceptable number of prepackages below TU_1 for each of reference test sample size. The proportion of prepackages below TU_1 needs to be sufficiently small, in general it appears that not more than 2.5% below TU_1 is appropriate.

In using the table, the values of the tolerable negative errors shown as percentages in the table, calculated in units of weight or volume, shall be rounded up to the nearest one-tenth of a gram or milliliter.

2 Terms and Definitions

Throughout this document, and other WELMEC series 8 documents the terms used, and their definitions, are those stated in WELMEC documents 6.1 Definitions of Terms and 6.2 Translation of Terms.

3 Duties of Competent Departments

3.1 Checks to ensure that prepackages comply with the requirements of this Directive shall be carried out by the competent departments of the Member States by sampling on the packers' premises, or if this is not practicable, on the premises of the importer or his agent established in the Community.

3.2 The checks should cover the adequacy of the quantity control system, confirm that it was being followed, and that its appropriateness had been regularly reviewed. This will include:

- the labeling of the product,
- the accuracy and suitability of the equipment and whether it was adequately maintained,
- the adequacy of the records, and their accuracy by checking prepackages from the appropriate batch
- the quantity in prepackages.

Checks on 'e' marked products and the quantity control system used for their production should be carried out at packers' and importers' premises generally at least once a year for those importing, exporting or packing prepackages. Member States have various ways of determining the frequency of visits, which include assessing

- the number of prepackages,
- the value of the product packed,
- the quality system in use and complaints received,
- the level of compliance found on visits.

3.3 This shall be done by means of statistical sampling check carried out in accordance with the accepted methods of quality acceptance inspection. Its effectiveness shall be comparable to that of the reference method specified in Annex 1 of Directive 76/211/EEC. The operating characteristic curve of the reference test is in Annex D.8. See also Annex I of WELMEC document 6.3.

3.4 The Directive does not preclude any checks, which may be carried out by the competent departments at any stage in the marketing process, in particular for the purpose of verifying, that prepackages meet the requirements of the Directive.

3.5 The Directives control products packed with a weight or volume quantity declaration between the limits of 5 g or 5 ml and 10 kg or 10 L domestic legislation may control goods outside these limits or sold by reference to length, area and number.

3.6 **Recognition of a Quantity Control System**

The Competent Department shall recognize the quantity control system in the way specified in national legislation. This may also result in an approval to mark prepackages with the 'e' mark. For the methods of recognition used in Member States refer to WELMEC 6.0.

3.7 Where there have been changes in the quality system theses changes need to be recognized by the Competent Department before they are brought into use. Guidance on recognition of the packer's procedure for carrying out production checks is given in WELMEC 6.6

4 Duties of Packers

4.1 The term 'packer' is applied in the Directive with a broad definition as the person responsible for the packing of a prepackage. Domestic legislation may specify whether the company or individual employee is held responsible.

4.2 The packer is generally the last person who has altered the prepackage, its contents or it's labeling in any way. The end of the packing process is when the only intention thereafter is to store or distribute the prepackage to the consumer. At this stage the prepackage is fully labeled and has passed all the checks in the packing process.

4.3 The packer's duties [2] are:

- to ensure the prepackages meet the labeling and contents requirements of the Directive,
- either to measure the quantity of product into each package or to check the actual contents after packing,
- where the actual contents are not measured, to carry out checks so that the quantity of goods is effectively assured.
- to use suitable and legal measuring equipment for these purposes.

4.4 Essentially the duties in 4.4.b) above permit the packer to have one of two systems. The first is to measure the contents of each prepackage, in which case no further checks need to be carried out and no records have to be made, this scheme is most suitable for low volume packers. Records of the maintenance of equipment need to be considered.

4.5 The other option is for a packer to produce prepackages but to have a system of checks in place that will assure the contents of the prepackage. This requirement is fulfilled if the production checks are carried out in accordance with procedures recognized by the competent department in the Member State and the packer holds at their disposal the results of those checks (together with any corrections and adjustments, which they have shown are necessary. The means by which the procedures are recognized may be specified in the domestic legislation.

4.6 The Directive makes it a responsibility of a packer to carry out checks "...So that the quantity of goods is effectively ensured.". As he is also responsible for ensuring that the prepackages meet the Directive (e.g. pass a reference test) there is an inference that his checks should be as effective in detecting non-compliance as the Inspector's reference test. Annexes E, F and G give guidance on this.

4.7 The requirement to use 'suitable' and 'legal' equipment is discussed in Annex B.

[2] Directive 76/211/EEC, annex 1 paragraph 4

4.8 For a system to give the necessary assurance it must:

- cover the setting up, monitoring and review of the quantity control system,
- justify the target quantity and control limits for each product,
- contain procedures to be followed when limits are breached,
- require records to show the system is being followed.

5 Duties of Importers

5.1 For the purposes of the Directive, an importer is someone who brings prepackages into the EEA, therefore movement within the EEA does not involve import / export for the purposes of the Directive. The importer has the same responsibilities as a packer but the Directive recognizes that they may not physically come into contact with the prepackages being imported.

5.2 The Directive states, "In the case of imports from non-EEC countries, the importer may instead of measuring and checking provide evidence that he is in possession of all the necessary guarantees enabling him to assume responsibility." What is considered acceptable is dependent on the national legislation.5.3 Some of the acceptable guarantees include:

a) evidence from a competent department in a Member State,

b) evidence from an EEA accepted competent department in the exporting country,

c) records of checks carried out by a competent sub-contractor at the place of first entry into the EEA,

d) to obtain records from the packer and to carry out checks to verify the data contained in them.

5.4 Evidence referred to in a) and b) above shall state that the quantity control system had been assessed and that the controls and records guarantee compliance with the requirements of the Directive. The certificate needs to specify the type of goods, the nominal quantity and packaging that has been assessed.

5.5 Checks appropriate for c) and d) above are contained in Annex J.

5.6 Any documentation may be subject to inspection by a Competent Department. These guarantees do not replace the responsibility of the Competent Department to carry out reference tests at the importer's, or their agent's, premises.

6 Annex A Labeling of Prepackages

Indication of quantity

A.1 Prepackages shall be marked with an indication of quantity and an 'e' mark to show compliance with the requirements of the Directive. The content shall be equal to or exceed the indicated content of the prepackage taking into consideration the maximum deficiency that is permitted for the indicated quantity.

A.2 Indications of quantity shall be expressed in the legal units of volume, in the case of liquid products, or in the legal units of mass in the case of other products [3]. Community and national legislation may require indications of volume or mass for specified categories of products or types of prepackages [4]. Where required, the quantity shall be of a particular quantity prescribed by legislation (see Directive 80/232/EEC)

If trade practice or national regulations are not the same in all Members States for a category of product or type of prepackage, those prepackages must as a minimum show the metrological information corresponding to the trade practice or national regulations prevailing in the country of destination [5].

OIML recommends that the indication should be expressed in units of

a) volume in the case of liquids or viscous products, and

b) in units of mass for solids, semi-solid or viscous, a mixture of solid and liquid, or the solid part of a mixture of a solid and liquid[6].

A.3 Prepackaged foodstuffs that are normally sold by number are not required to be marked with an indication of volume or mass[7]. The 'e' mark does not apply to such prepackages.

A.4 Prepackaged solid foodstuffs presented in a liquid medium shall be marked with an indication of drained weight in addition to an indication of weight [8]. At present the 'e' mark applies to the weight and not the drained weight.

---

3 Legal units within the meaning of Directive 80/181/EEC.

4 For example, Directive 80/232/EEC specifies that cosmetics in solid or powder form should be marked in grams and cosmetics in liquid and paste form should be marked in milliliters.

5 Directive 76/211/EEC, Article 4.3

6 OIML R79 (1997) *Labeling requirements for prepackaged products* Article 5.3.1

7 Directive 2000/13/EC, Article 8.3

8 Directive 2000/13/EC, Article 8.4

A.5 Products that are prepackaged in aerosol form shall be marked with indications of volume and weight [9]. The 'e' mark applies to both indications. Where the volume of the product and the capacity of the container are equal to one of the range of product volumes and the corresponding container capacity listed in Annex III of Directive 80/232/EEC, an indication of weight is not required if the prepackage is labelled with indication of product volume and container capacity. In this case the 'e' mark applies to the indication of volume. However, for aerosols OIML [10] recommends that an aerosol should be marked with the quantity in mass (weight) of product and expellant.

A.6 Community and national legislation may permit indications other than the content [11] or for indications to be given for quantities above 10 kg or 10 L. The 'e' mark does not apply to such indications.

Manner of marking and presentation

A.7 The markings required shall be affixed on the prepackage in such a manner as to be indelible, easily legible and visible on the prepackage in normal conditions of presentation. Some practices that are not acceptable are:-

- having information hidden in the fold of the prepackaging,
- having markings on a clear container in a similar colour as the product contained in it.
- having markings on the rear of a container.

A.8 For prepackaged foodstuffs, the indication of quantity shall be easy to understand and marked in a conspicuous place in such a way as to be easily visible, clearly legible and indelible. The indication shall not in any way be hidden, obscured or interrupted by other written or pictorial matter. The indication shall appear in the same field of vision as:

- the name under which the product is sold;
- the date of minimum durability or, in the case of foodstuffs which from a microbiological point of view are highly perishable, the 'use by' date; and
- with respect to beverages having an alcohol strength of more than 1.2% vol., the actual alcoholic strength by volume [12].

A.9 'Easily legible' must take into account the size, font, colour and background of the markings. Ideally the mandatory markings shall be discrete from other information on the prepackage, and wherever possible take into account those customers with poor eyesight.

A.10 OIML[13] recommends that the statement of the quantity shall appear on the principal display panel in easily legible boldface type or print that contrasts conspicuously with the background and with other information on a prepackage. When the value of the quantity is blown, embossed or moulded on the surface of the prepackage, then all other required label information shall be provided conspicuously elsewhere on the surface or on a label.

---

9 Directive 75/324/EEC, Article 8.1.e

10 OIML R79 paragraph 5.3.2

11 For example, pre-packed fertilizers may be marked by weight or by gross weight (Directive 76/116/EEC).

12 Directive 2000/13/EC, clause 8

13 OIML R 79 paragraph 5.5.2

Nominal Quantity.

A.11 For collections of items, where the individual items would normally be sold separately, the outer container shall be marked with the content of each item and the number of items present (where this cannot be visually determined). Where the individual items would not be regarded as units of sale then the total quantity shall be marked on the outer container together with the total number of items.[14]

The height of the figures in the numerical value must be at least that stated in the table:

Nominal quantity in g or ml	Minimum height in mm
5 up to and including, 50	2
over 50, up to and including 200	3
over 200, up to and including 1000	4
over 1000	6

For packs destined for the United States of America the minimum height requirements (as stated in OIML R79:1997 Annex B Table 3) are:-

Area of principal display panel (cm^2)	Minimum height of numbers and letters (mm)	Minimum height if blown or molded on surface of container (mm)
$A \leq 32$	1.6	3.2
$32 < A \leq 161$	3.2	4.8
$161 < A \leq 645$	4.8	6.4
$645 < A \leq 2581$	6.4	7.9
$2581 < A$	12.7	14.3

A.12 The unit of measurement shall either be written in full or abbreviated. The only permitted abbreviations are:

	For foodstuffs[11]	For non-foodstuffs	For use only as a supplementary indication which accompanies the metric indication and is not more prominent
volume	ml, mL, cl, cL, l, L	ml, mL, cl, cL, l, L	fl. oz., pt, qt, gal
mass	g and kg	g and kg	oz, lb

A.13 Where a prepackage is marked with more than one indication of quantity [15], all indications shall be in close proximity to each other, shall not be more prominent than the required marking, and the quantity to which each indication refers shall be unambiguous. Where quantity markings are repeated on the packaging they shall all contain the same information.

[14] Directive 2000/13/EC, clause 8

[15] For example, an indication of weight and volume, weight and drained weight, volume and container capacity, a supplementary indication of weight or volume in non-SI units, an indication of a 'portion' or 'dose' of the product expressed in units of weight or volume.

NOTE: OIML recommends that if the prepackaged product is labelled on more than one location of it's packaging, the information on all labels shall be equivalent [16].

A.14 For prepackaged foodstuffs, the labeling and methods used must not be such as could mislead the purchaser to a material degree as to the quantity of foodstuffs [17].

The same prohibition shall apply to the methods by which prepackaged foodstuffs are presented (in particular their shape, appearance or packaging, the packaging material used, the way in which they are arranged and the setting in which they are displayed) and advertised.[18] See Annex I relating to deceptive packaging.

Containers, which contain dangerous preparations offered or sold to the general public, shall not have either a shape and/or graphic decoration likely to attract or arouse the active curiosity of children or to mislead consumers [19].

A.15 If the trade practice or national regulations are not the same in all Member States for a category of products or for a type of prepackage, those prepackages must at least show the metrological information corresponding to the trade practice or national regulations prevailing in the country of destination.

Prepackages bearing the "e" mark.

A.16 The 'e' mark may be used only in respect of prepackages which are intended for sale in a constant nominal quantity, which are equal to values predetermined by the packer. These quantities must be expressed in units of weight or volume and be not less than 5 g or 5 ml and not more than 10 kg or 10 L. Furthermore it must not be possible to alter the quantity of the content without the prepackage either being opened or undergoing a perceptible modification.

Where the "e" mark is used, it must appear in the same field of vision as the quantity indication, be of the form shown in annex II to Directive 71/316/EEC and it must also be at least 3mm in height.

Only one of the quantity declarations is considered to be the 'nominal quantity' and the 'e' mark should be in the same field of vision as this. This combination must be 'easily legible and visible on the prepackage in normal conditions of presentation', this implies that the marking should be visible to the intended purchaser without the prepackage having to be handled.

Where more than one statement is given on a multi-pack, e.g. '4 x 10 g e 40 g' the 'e' applies to the quantity which the packer controls. Where the individual items could be sold separately this would, in this example, be 10 g. Where the individual items are not appropriate for selling singularly it would be the 40 g.

NOTE: OIML recommendation 79 requires that the quantity marking is put on that part of the prepackage, which is most likely to be displayed under normal and customary conditions of display.

Identity Mark or Inscription

A.17 A mark or inscription must be incorporated on the prepackage, which identifies the packer or importer, or the person who arranged for the prepackages to be made up who are established in the EEA.

The minimum requirement is for the name or mark, together with the postcode or a geographical code. This marking shall enable the competent departments to establish who is the packer or importer of the prepackage. .

- Markings must be 'easily legible and visible on the prepackage in normal conditions of presentation'.

---

16 OIML R79 (1997), Article 6.3
17 Directive 79/112/EEC, Article 2
18 As above
19 Directive 88/379/EEC, Article 6.1.b

- Other vertical Directives may require extra information such as the full address, or address of the Registered Office to be supplied. OIML recommends that when the name is not that of the packer or importer the name may be qualified by a phrase that reveals the connection such person has with the product, for example "manufactured for...". [20]
- Other legislation may require the Country of Origin to be stated on the label.

A.18 Where a competent department requires information on a packer or importer resident in another Member State then they should contact the competent department there using the information in WELMEC 6.0.

[20] OIML R79 paragraph 4

7 Annex B Measuring Equipment for the Packer or Importer

B.1 The requirement in the Directive is for the equipment used by packers or importers to be 'legal' and 'suitable'. 'Legal' should be taken to mean verified if controlled by European or national legislation, otherwise to have a recognized certificate of accuracy, the latter should show traceability and indicate the uncertainty of measurement the equipment can operate to. Legal equipment will be subject to re-verification or in-service inspection as determined by national legislation.

B.2 Suitable' ensures that measurements made are do so with an appropriate uncertainty of measurement. This uncertainty will include the uncertainty of calibration together with the error permitted for the equipment. Generally a total uncertainty of measurement (which includes all pieces of equipment involved and the method of test) of less than one-fifth of TNE.

B.3 Where the measurement uncertainty exceeds one-fifth TNE then it should be taken into account in establishing and monitoring the target fill quantity.

B.4 All equipment shall be periodically maintained, the periodicity being set so that records of calibration can show that the equipment remains within permitted tolerances between calibrations. If equipment is adjusted then both the 'before adjustment' and 'after adjustment' figures shall be recorded in order to demonstrate that the calibration period is appropriate.

B.5 Where there is no requirement for equipment to be verified or be maintained within certain limits then the equipment should be maintained to comply with the relevant OIML recommendation, e.g. R51 for checkweighers and R61 for automatic gravimetric filling instruments.

8 Annex C Production Records

C.1 The Directive requires that checks be so organized to effectively ensure the quantity of the contents. That condition is fulfilled if the packer carries out production checks in accordance with procedures recognized by the competent department in the Member State and if he holds at their disposal records of such checks. The records must include any corrections or adjustments made to the process and also show that they have been properly and accurately carried out. The following records should be maintained as appropriate to the process involved.

C.2 Identification and specification

- product identity
- testers identity

C.3 Batch data

batch identity

- batch size
- density, if applicable
- nominal quantity

C.4 Process characteristics

- target value, or set points for checkweighers
- average quantity target and control limits
- process variation limits
- tare variability and other allowances
- for checkweighers- the zone of indecision, the mean error and standard deviation, checks on data collection and calculations

C.5 Production checks

- identification of checkpoint / packing line
- reference to product identity
- batch identity
- time and date of sampling
- number of prepackages in a sample
- tare if applicable
- average and variance of actual contents (sample data)
- average and variance of actual contents (batch data)
- number (%) of prepackages below TU1, and corrective action taken when the batch is likely to have >2.5%
- number (%) of prepackages below TU2, and corrective action taken
- for checkweighers - checks on set points, to ensure no drift.

C.6 Where legal and suitable equipment is being used to ensure each prepackage contains at least the quantity marked on it then no 'production check' records are required.

C.7 Equipment maintenance and calibration records are needed to determine the 'process characteristic' - included in 'process variation limits' - as well as the target fill.

C.8 Any software used for processing of quantity control data shall be:-

- validated prior to use and after every change, records to be kept,
- assigned issue status (e.g. date or version number) which is identified with the records produced from it,
- protected against corruption and unauthorized modification.

Record retention

C.9 All records need to be retained for as long as prepackages remain in the distribution chain, with a minimum of 1 year. Records relating to equipment need to be retained for at least as long as twice the calibration period, the actual period will depend on the national legislation. The calibration period shall be set so that the change in the accuracy of the equipment is less than the error permitted in the quantity control system.

Validation of records

C.10 The Competent Department shall carry out checks to validate the records retained. This shall involve:

- testing the equipment used by the packer or importer for accuracy and suitability,
- using a statistically significant sample, determining the contents of prepackages and comparing the mean and standard deviation of the sample statistically with the same attribute found in the associated check records for those prepackages.
- using the statistics, determining whether the target quantity and control limits used by the packer or importer are appropriate to guarantee compliance with the Directives.

9 Annex D Statistical Principles relating to Control Systems and Reference Tests

D.1 Introduction

D.1.1 This chapter describes the basic statistical concepts and methods involved in reference testing and the setting up and operation of quantity control systems. It does not provide the depth of treatment necessary for a full understanding of the topics, and is intended as a reminder and summary rather than as a text.

D.2 Types of data and levels of measurement

D.2.1 In metrological control, two fundamentally different types of numerical data are encountered. Firstly, the contents of prepackages are *measured* (gravimetrically or volumetrically), to yield a number characteristic of each item (e.g. its weight or volume). These measurements are on an interval or ratio scale, and so differences between them can be used for statistical calculations and, provided that actual contents are determined (i.e. not including any tare weight), their ratios to each other can also be used.

D.2.2 Secondly, in testing for conformity to the Second and Third Rules, the numbers of prepackages in each of three categories are determined, viz:

- Not less than TU_1 (adequate)
- Less than TU_1 (non-standard)
- Less than TU_2 (inadequate)

Such observations of numbers of items in named categories are on a *nominal* scale, though their names and definitions also imply an *order* property.

D.2.3 The importance of this distinction is in the type of summary statistics that may be applied to the data, and in the nature of subsequent statistical analysis to which they may be subjected. Thus nominal data may be summarized by calculating *proportions* (or percentages) of items falling in the various categories, whilst interval or ratio measurements may be summarized by such statistics as means, ranges, standard deviations, etc.

D.3 Distribution

D.3.1 In most natural situations and manufacturing processes, it is found that observations made on similar items or under similar conditions vary to some extent. The pattern of occurrence of these different values forms a distribution. Whilst a great variety of distributions occur in practice, many are found to approximate to a few theoretical distribution models, which thus form a convenient means of processing numerical data, subject to certain assumptions underlying a particular model.

D.3.2 The model, expressed as a mathematical function, permits the calculation of the proportions of values which (in an ideal situation) would occur at various points, or in various regions, along the scale of measurement. In practice, complete recording of all possible values is impracticable, but under the assumption of random sampling (i.e. that all items in a population, sometimes hypothetical, have equal chances of appearing in the sample), these proportions may be treated as probabilities of occurrence. When values, or collections of values, of low probability are observed, this should lead to questioning the basis of the model, its assumptions or its parameters, and hence to some action such as rejection on a reference test or corrective action in quantity control sampling.

D.4 Probability

D.4.1 Because reference tests and quantity control procedures are based on probabilistic arguments and on the use of probability distributions, we state briefly the frequency definition of probability and three rules. In terms of metrological control, the word 'outcome' will refer to the result of making measurements or of counting items, and 'events 'will be the occurrence of measurements within specified intervals, or of counts of a given (integer) value.

D.4.2 The frequency definition states that where an experiment of chance may result in different outcomes, of which some result is an event A, then the probability of occurrence of A is given by

P(A) = (Number of outcomes yielding event A)/(Total number of possible outcomes)

In empirical work P(A) must often be estimated from:

- (Number of observations of A) / (Total number of observations)
- the estimate becoming increasingly reliable as the number of observations increases.

D.4.3 The three rules are:

- *Complement Rule*. For an event $\underline{A}$, defined as the complement of A (i.e. any outcome not classified as A),
$$P(\underline{A}) = 1 - P(A).$$
- *Union Rule.* Where $A \cup B$ indicates the occurrence of A or B or both, and $A \cap B$ indicates the occurrence of both A and B, then
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$
- For events which cannot occur together (mutually exclusive events), $P(A \cap B) = 0$, and then
$$[P(A \cup B) = P(A) + P(B)]$$
- *Intersection Rule*. Where $P(B \mid A)$ is the probability of occurrence of B given that A has occurred (i.e. the probability of B *conditional* on A), then
$$P(A \cap B) = P(A) \times P(B \mid A)$$
- If A and B are *independent* events, then $P(B \mid A) = P(B)$, and in this case
$$P(A \cap B) = P(A) \times P(B).$$

Note that $A \cap B$ may imply either the occurrence of A followed by B as outcomes of successive trials, or of a single outcome simultaneously satisfying two events defined as A and B.

D.4.4 As an example of the interplay of these rules, consider the probability of failure on reference test (using the double sampling plan with 30 + 30 items) of a group, which actually contains just 21 % of non-standard items. We show in section D.5 that for samples of 30 drawn at random from a large bulk containing 1 in 40 non-standard, the probabilities of 0, 1, 2 etc non-standard items appearing in the sample are: -

$$P(O) = 0.46788$$

$$P(I) = 0.35991$$

$$P(2) = 0.13381$$

Now the events 'no non-standard items discovered', 'just one non-standard item found', 'exactly two. . etc,, are mutually exclusive, so that

$$P(\leq 2 \text{ non-standard}) = 0.46788 + 0.35991 + 0.13381 = 0.96160.$$

Here we have used the union rule. In order to find the probability that the group fails the reference test at the first stage (with three or more non-standard items being found in the sample of 30), we use the complement rule,

$$P(\geq 3) = P(> 2) = 1 - P(\leq 2) = 1 - 0.96160 = 0.03840.$$

In order to complete the evaluation, we must consider the case of 'suspended judgment at the first sampling stage and rejection at the second stage'. We require that two non-standard items be found in the first 30, and that at least three *more* occur at the second stage, giving a total of at least five for both stages combined. Let us define the events.

A Occurrence of two non-standard items at first stage.

B Occurrence of three or more non-standard items at second stage.

Now, because the second sample is also of 30 items, we have

P(A) = 0.13381, as already calculated,

P(B) = 0.03840, as obtained by the union and complement rules.

Then $P(A \cap B) = P(A) \times P(B \mid A)$, but in this case A and B are independent (in random sampling, the occurrence of non-standard items in the first stage does not influence their appearance or absence in the second). Thus

$$P(A \cap B) = P(A) \times P(B) = 0.13381 \times 0.03840 = 0.00514.$$

Finally, the overall probability of failure is the probability of the union of A □ B (failure on second stage) with, say, event C, failure on the first stage. We already have P(C)=0.03840, and the events C and $A \cap B$ are mutually exclusive (if failure occurs at the first stage, there will be no second stage). Thus

$$P(C \cup (A \cap B)) = 0.0384 + 0.00514 = 0.04354.$$

In this example, we have in fact evaluated a point on the Operating Characteristic of the sampling plan-a subject covered later in this chapter.

D.5 Probability distributions

D.5.1 Measurements

The most widely used model for measurements made on a conceptually continuous interval or ratio scale is the *Normal distribution,* characterized by its mean μ and standard deviation σ. This distribution is used to describe the relative frequency of occurrence of values within specified intervals, e.g. (x_1, x_2).

The probabilities are obtained from the distribution function, or integral of the density function, appropriate values of this integral being widely tabulated or available as standard functions on most electronic computers and on some calculators.

Any Normal distribution is reduced to the standard form by the transformation

$u = (x - \mu) / \sigma$

yielding the standardized Normal variable.

Well-known features of the Normal distribution are

- About two-thirds (68.7 %) of the distribution lies in the range ($\mu \pm \sigma$),
- Most of the distribution (95.5 %) lies within., $\mu \pm 2\sigma$ (or 95% within ± 1.96σ),
- Almost all of the distribution (99.7%) lies within $\mu \pm 3\sigma$, (or 99.9% within $\mu \pm 3.29\sigma$)

This distribution forms the basis of the reference tests for the average system, and of much quality control practice. Its validity rests on the Central Limit Theorem. The subject is more fully discussed in Section 8.6 below.

D.5.2 A packer filling cans to a nominal 250 g declaration sets his process target quantity to 252 g . If the packed quantities have a Normal distribution with a standard deviation of 5 g, what proportion of cans constitute non-standard prepackages?

Here, we have μ =252 and σ = 5. The value of the variable of interest is the weight corresponding to TU_1 in this case 241 g. Thus the standardized Normal variable is

$$U = (241-252) / 5 = - 2.2$$

The negative sign indicates a value in the lower half of the distribution, and reference to Tables gives a probability of 0.014 for a value at or below U = - 2.2 (because of the symmetry of the Normal distribution, this is the same as the probability for u at or above 2.2). We may interpret this as meaning that, under the conditions stated, about 1.4 % of prepackages will be non-standard.

D.5.3 *Counts of non-standard or inadequate items*

If repeated random samples of the same size, n, are drawn from a batch ('population') of infinite size having a proportion p of 'defectives', the number of defective items in the samples will form a distribution. The theoretical model for the number defective in these circumstances is the *Binomial distribution,* defined by the sample size, n, and population proportion, p. The probability of occurrence of exactly r defectives in a sample of n is given by

$$P(r) = p^r . q^{n-r} . n! / (r! (n-r)!)$$

in practice, samples are drawn from finite batches without replacement, but provided that the sampling fraction is less than 10 % and the batch size at least 100, the Binomial distribution provides a satisfactory model. Indeed, another and simpler model is applicable in most reference testing situations.

D.5.4 Provided that p is at most 0, (i.e. $\leq$ 10 % non-standards), the *Poisson* distribution may be used. With mean non-standards per sample given by m = np, the Poisson distribution gives

$$P(r) = e^{-m} \times m^r / r!$$

with m^0, 0! and 1! all defined as 1.

(Note that the Poisson distribution arises in other contexts, its use as an approximation for the Binomial being one of its applications.)

D.5.5 By way of detailed illustration, we complete the calculations for the reference test example of D.4.4. Under conditions of Binomial sampling, we draw a sample of n = 30 items from a population containing p = 0.025 of non-standard items. We thus have..-

$$P(0) = 0.025^0 \times 0.975^{30}$$

(the remainder of the expression in paragraph D.5.4 equalling 1 when r = 0)

Thus	P(O) = 0.97530 (since 0.0250 = 1)	= 0.46788
Next	$P(I) = 0.025^1 \times 0.975^{29} \times 30/1$	= 0.35991
	$P(2) = 0.025^2 \times 0.975^{28} \times (30 \times 29) / (1 \times 2)$	= 0.13381

and so on.

D .5.7 Alternatively, using the Poisson approximation, we would set m = np = 30 x 0.025 = 0.75 (the average number of non-standard items per sample of 30). Then we have.-

$P(0) = e^{-0.75} \times 0.75^0 / 0!$	$= e^{-0.75}$	= 0.47237
$P(1) = e^{-0.75} \times 0.75^1 / 1$		= 0.35427
$P(2) = e^{-0.75} \times 0.75^2 / (2 \times 1)$		= 0.13285

These values differ only in the third decimal place from those obtained using the Binomial distribution model.

D.5.8 It has to be recognized that even the Binomial model is slightly unrealistic. It assumes that for each selection of an item from the group, the probability of its being non-standard remains constant. In fact, items will be sampled without replacement, so that the probability of any selection being a non-standard item changes continuously, depending on how many nonstandard items have been withdrawn.

Thus for a group of 200 prepackages (which with p = 0.025 would contain 5 non-standard) the probability of a non-standard item at the first selection is 5/200=0.025. However, for the second selection, the probability is either 4/199 or 5/199, depending on the nature of the first item selected.

Distribution	Parameters	P(O)	P(1)	P(2)	P(> 3)
Hypergeometric	N=200, n=30, p=0.025	0.43974	0.39736	0.13800	0.02490
Binomial	n=30, p=0.025	0.46788	0.35991	0.13381	0.03840
Poisson	m=np=0.75	0.47237	0.35427	0.13285	0.04051

The overall effect of this non-independence on the probabilities of 0, 1, 2 etc non-standard items appearing in a sample of 30 is shown in the table above, which compares the Binomial, Poisson and true (Hypergeometric) distribution terms. The assumptions involved are as follows: -

Hypergeometric — Random sampling without replacement of n = 30 items from a group of 200, of which 5 are non-standard.

Binomial — Either sampling with replacement, or from a group of infinite size containing 2.5% non-standard.

Poisson — Approximation valid for large group and small proportion non-standard.

D.6 Sampling and estimation

D.6.1 The object of drawing samples is generally to estimate some features of the population from which they are drawn, in order to make inferences about that population. In reference testing and quantity control, this object can be narrowed down to estimating the average and variation of prepackage contents, or the proportion non-standard, in the batch, group or process.

D.6.2 Repeated sampling from a given batch will inevitably yield differing estimates of the property of interest. Thus sample averages, standard deviations or proportions will vary from sample to sample. The pattern of variation in these estimates is described by *sampling distributions.* The most important of these is the Normal distribution, especially in connection with the distribution of sample means. In this case one form of the Central Limit Theorem states that:

'for independent random samples of size n drawn from a population with mean p and finite variance σ^2, the distribution of sample means tends to the Normal form as n increases, with mean μ and variance σ^2 / n.'

Thus $\underline{x}$ is an unbiased estimate of μ (the mean of its sampling distribution corresponds to the mean of the population), and even for non-Normal parent distributions, sample estimates of the batch (or process) mean tend to the Normal distribution for moderate sample sizes; in many industrial situations, even sample sizes as low as 4 or 5 are sufficient for the reasonable assumption of Normality in the distribution of sample averages.

D.6.3 The precision of an estimate obtained by sampling is measured by its standard error - in its statistical usage, the standard error is treated in a similar manner to the standard deviation, but the term standard error is used in connection with sample estimates of parameters, the standard deviation measuring the variation in the original variable.

The most commonly encountered standard error is that of the sample mean. One would expect the precision of an estimate to improve with increasing sample size, and the standard error of $\underline{x}$ for a sample of size n is given by $\sigma / \sqrt{n}$, i.e. the standard deviation of the underlying variable divided by the square root of the sample size. Standard errors exist for other sample statistics, but the only examples applicable in average quantity control are the following.-

Standard error of the sample median (for samples from a Normal population); = $1.25 \sigma / \sqrt{n}$
and, though rarely used directly,
Standard error of the sample estimate of standard deviation (again for a Normal population) = $\sigma / \sqrt{2n}$

Standard error of a sample proportion, $p = \sqrt{((p (1-p)) / n)}$

D.6.4 Given a sample estimate $\underline{x}$ obtained from a sample of n observations on a population whose standard deviation is □ and invoking the Central Limit Theorem, it is now possible to construct a *confidence interval for* the true (but unknown) population mean μ,

$$\underline{x} - u \sigma / \sqrt{n} \leq \mu \leq \underline{x} + u \sigma / \sqrt{n}$$

It is important to recognize that this statement does not postulate a range over which μ varies, but an interval within which it may be asserted to lie. In making such assertions, there is a risk of error, this risk corresponding to the probability that the Standard Normal deviate lies outside ±u. The particular value of u adopted for constructing the interval is chosen to yield a desired confidence level, and conventional (though not immutable) levels are:

95% yielding	$\underline{x} - 1.96 \sigma / \sqrt{n} \leq \mu \leq \underline{x} + 1.96 \sigma / \sqrt{n}$
99% yielding	$\underline{x} - 2.58 \sigma / \sqrt{n} \leq \mu \leq \underline{x} + 2.58 \sigma / \sqrt{n}$
99.9% yielding	$\underline{x} - 3.29 \sigma / \sqrt{n} \leq \mu \leq \underline{x} + 3.29 \sigma / \sqrt{n}$

The underlying probabilistic principle is that in making such assertions, 100 (1 - α)% will be correct and 100α% incorrect, where α is the *two-tail* probability associated with the Normal deviate u.

D.6.5 In most practical situations the true value of σ is unknown and must (like μ) be estimated from sample data. Under these circumstances, the required sampling distribution is no longer that of

$(\underline{x} - \mu)/ \sqrt{n} / \sigma$, but $(\underline{x} - \mu) \sqrt{n} / s$ where s has a sampling distribution known as Student 't' distribution, (with n-1 degrees of freedom.

The confidence interval then becomes: $\underline{x} - t \times s / \sqrt{n} \leq \mu \leq \underline{x} + t \times s / \sqrt{n}$

where t is the 100 (1 - 0.5 α) % point of the t-distribution with n -1 degrees of freedom (usually symbolized by υ or ϕ, and 100 (1 -α)% is the required confidence level.

(Note that degrees of freedom are of much wider application than considered here, where the definition has been restricted to the particular case of constructing a confidence interval for μ. Also, other sampling distributions exist, such as for standard deviations, ranges, etc, but these cannot be set out in detail here.)

Suppose that a sample of n = 20 has been drawn from a group and that the following sample statistics are obtained:

the mean,	$\underline{x}$ = 248.9 g
and standard deviation,	s = 2.73 g

What range might the true group average be expected to lie ?

It is necessary to choose a confidence level, and we might select 99% so that $\alpha = 0.01$, and we require t for the 0.5 α probability level with 20-1 = 19 degrees of freedom. Table of the t-distribution shows this to be 2.861.
We therefore have $248.9 - (2.861 \times 2.73) / \sqrt{20} \leq \mu \leq 248.9 + (2.861 \times 2.73) / \sqrt{20}$. I.e. with 99% confidence we can assert that μ lies between 247.15 g and 250.65 g. Thus we note that the true group mean compatible with a nominal quantity of 250 g is not ruled out by this confidence interval.

D.7 Hypothesis testing

D.7.1 As well as inference involving point and interval estimates of population or batch parameters, one may wish to test hypotheses about the batch. This is the philosophy underlying reference tests in particular, and is also an aspect of quantity control methods. An initial or working hypothesis is formulated and, using an appropriate sampling distribution, it is tested by calculating the probability (conditional on the truth of the hypothesis) of observing a sample statistic as extreme, or more extreme, than that yielded by the data. If this probability is low, the hypothesis is rejected in favour of an (often vaguely specified) alternative. In terms of the two reference tests, we have the following.

D.7.2 For average contents the initial hypothesis is

$$H_0 : \quad \mu \geq Q_n \qquad \text{where } Q_n \text{ is the nominal quantity.}$$

The limiting form of this hypothesis, where $\mu = Q_n$ is tested using the statistic

$$t = ((\underline{x} - Q_n) \sqrt{n}) / s$$

and rejected if it lies beyond the critical (lower) value corresponding to a one-tail probability of 0.005. This is a one-sided test as the Inspector is not concerned with generous overfill but only with failure to meet the legal requirements. The vague alternative hypothesis (a 'composite' alternative

$$H_1 : \quad \mu < Q_n$$

is adopted and the reference test results in the rejection of the batch and the appropriate action being taken (remedial or disciplinary).

D.7.3 For the test to determine conformity to the second of the 'Three Packer's Rules' in respect of non-standard items, the initial hypothesis is :

$$H_0 : \quad p < 0.025,$$

where p is the batch proportion of prepackages containing less than TU_1. Assuming random sampling from the batch, the acceptance/rejection criteria are based on the number of defective packs in the sample, with probabilities of occurrence generally less than 0.01 under the initial hypothesis, but varying somewhat for the various batch and sample sizes.

D.7.4 Considering again the example of paragraph D.6.6, but using the data to test the hypothesis $\mu \geq 250$, we have:

$$t = (248.9 - 250) \sqrt{20} / 2.73 \qquad = - 1.802$$

Now the reference tests for average quantity use a critical t value at the 0.005 one-tail level, and for 19 degrees of freedom this critical value is 2.861. The absolute magnitude of the observed t-statistic is smaller than the critical value, and the initial hypothesis cannot therefore be rejected at this level of significance-the group would pass the reference test because there is inconclusive evidence of the average contents being below the nominal quantity. In fact, the Directive uses the criterion:

$$\underline{x} \geq Q_n - 0.640s \qquad \text{Accept}$$
$$\text{and} \quad \underline{x} < Q_n - 0.640s \qquad \text{Reject} \qquad \text{for } n = 20.$$

Here, 0.640 is in fact $2.861 / \sqrt{20}$, or $t / \sqrt{n}$.

D.8 Risks, errors, power and the Operating Characteristic

D.8.1 In testing hypotheses, two kinds of risk are involved, firstly, the initial hypothesis may be true, but an unfortunate sample results in its rejection. Secondly, the initial hypothesis may be false, but may be accepted because of inadequate evidence to the contrary, or because a 'lucky' sample yields data favourable to the initial hypothesis.

D.8.2 The first kind of error, known variously as Type I or rejection error or (in terms of Acceptance Sampling, of which the reference test is an example) Producer's Risk. The risk is measured by the probability a (the 'significance level'), used to set up the critical value of the test statistic.

D.8.3 The second kind of error, Type II or acceptance error, the Consumer's Risk in the present context, is measured by probability β. This is the probability that, under some *specific* alternative value of the population parameter, the test will result in acceptance. It is thus conventional to evaluate β over a range of values of the population parameter, yielding the power curve when 1-β (the rejection probability) is plotted against that parameter. The acceptance probability, β, may similarly be plotted and yields the *Operating Characteristic, OC,* which is conventionally adopted as a measure of the effectiveness of a sampling procedure in discriminating between acceptable and unsatisfactory product. A good test yields large β when the initial hypothesis is true (acceptance is then the correct decision) and rapidly diminishing β as the true parameter departs from the value specified under the initial hypothesis.

D.8.4 The simplest presentation of the OC, for the average contents test is obtained by assuming that the individual prepackage contents are Normally distributed about their mean. The OC may then be drawn as the probability of batch acceptance for various proportions of prepackages containing less than the nominal quantity. Where the batch average equals or exceeds the nominal quantity, the proportion below nominal will be less than 50%. As the batch average falls, or as the variability increases (or both together), so the proportion below nominal will increase. This assumes Normality in the underlying distribution of prepackage contents, but it must be stressed that testing for Normality, or the occurrence of non-Normality for any particular product, are not the Inspector's concern, and non-Normality of the distribution does not constitute grounds for disputing the result of a reference test.

8.8.5 For the non-standard prepackages test, the probability of acceptance of the batch is simply plotted against various proportions of non-standard, assuming random sampling.

D.8.6 For both types of test, the larger sample sizes provide better discrimination between satisfactory conformity (to the appropriate rule) and violation. The OC curves for non-standard items apply to single sampling procedures, those for the double sampling procedures being broadly similar. The justification for using double sampling is that, for warehouse testing it may require less sampling and testing effort to achieve the same discrimination as a single sampling procedure (when on-line however, the entire double sample may have to be taken at one time, thus losing the possible saving in effort.)

D.8.7 Operating characteristic curves of the reference tests (directive CEE 76/211, Annex II)

The Operating Characteristic Curve of a statistical test links the probability of lot acceptance (P_A) with the level of the characteristic under control.

D 8.7.1 Average tests

The characteristic under control is the malfunctioning of average expressed by a parameter lambda (λ)

$$\lambda = -\left[\frac{\bar{x} - Qn}{s}\right]$$

$\bar{x}$ = average calculated on the sample;

Q_N = nominal quantity of the pre-prepackage;

s = standard deviation calculated on the sample = $s = \sqrt{\sum_{i=1}^{i=n} \frac{\left(x_i - \bar{x}\right)^2}{n-1}}$;

x_i = quantity measured of the pre-prepackage of rank i, in the sample ;

n is the sample size

The probability of lot acceptance (P_A) is given by the formula

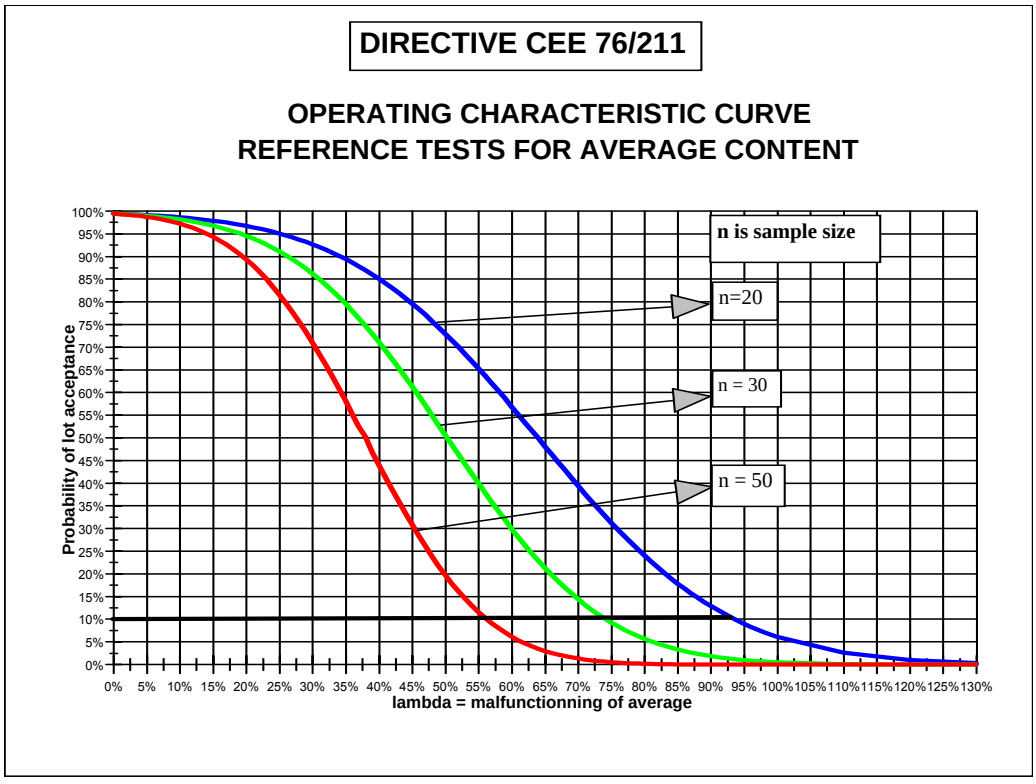
$$P_A = F\left[t_{1-\alpha} - (\lambda \,.\, \sqrt{n})\right]$$

F : distribution function of a Student random variable

(1-α) is the level of confidence of the test = 0,995 according the directive

$t_{1-\alpha}$ = Fractile of order (1-α) of a Student random variable with (n-1) degrees of freedom.

The OC curve is given by the graph below:



D 8.7.2 Tests for the minimal content

The characteristic under control is the rate of non conform items in lots

D.8.7.2.1 Single sampling plans

The probability of lot acceptance (P_A) is given by the formula

$$P_A = \sum_{i=0}^{i=c} C_n^c p^c (1-p)^{n-c}$$

p is the rate of non conform items in the controlled lot

n is sample size ;

c = value of the criteria of rejection = maximum value of non-conforming items admitted in the sample .

D.8.7.2.2 Double sampling plans

The probability of lot acceptance (P_A) is given by the formula

$$P_A = \sum_{i=0}^{i=c1} C_{n1}^i p^i (1-p)^{n1-i} + \left[\left(\sum_{i=c1+1}^{i=r1-1} \left(C_{n1}^i p^i (1-p)^{n1-i} \right) \right) \cdot \left(\sum_{i=0}^{i=c2} \left(C_{n1+n2}^i p^i (1-p)^{(n1+n2)-i} \right) \right) \right]$$

p is the rate of non-conforming items in the controlled lot;

c1 = maximum number of non conform items admitted in the first sample ;

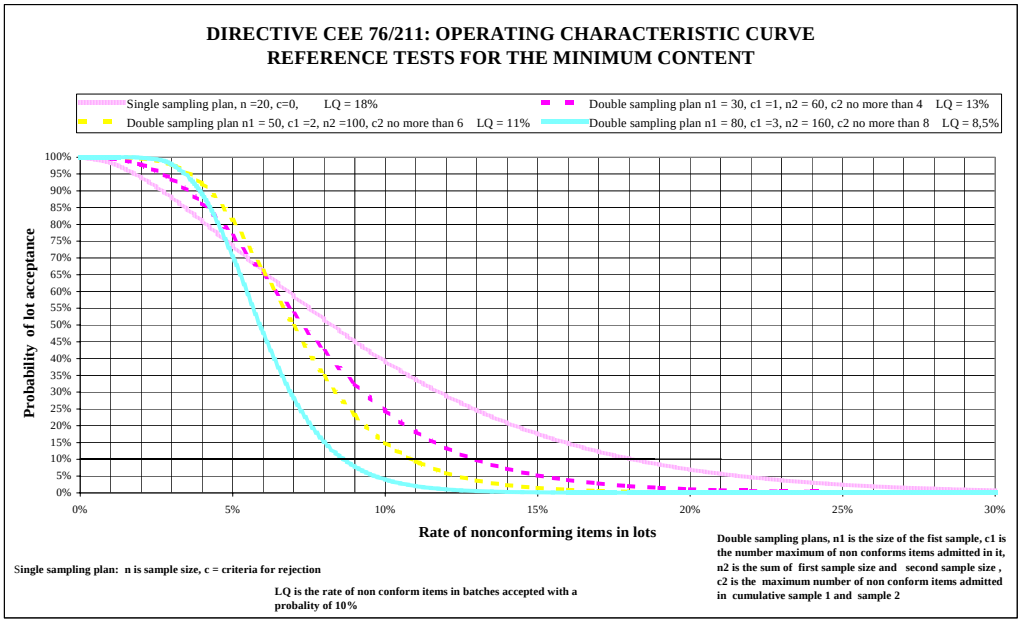
c2 = maximum number of non conform items admitted in the cumulative first and second sample ;

r1 = minimum number of non conform items in the first sample resulting rejection of the lot ;

n1 = size of the fist sample ;

n2 = cumulative size of the first and second sample ;

$c1 \leq r1 \leq c2$.



D.9 Quantity control

D.9.1 The basic principles of hypothesis testing also apply to continuous quantity control methods. The initial hypothesis is that the process is running satisfactorily, and that samples obtained from process observations conform to a target distribution -for example, of individual values, sample means or medians, sample standard deviations, sample ranges, etc. If a sample value (or a sequence of values) provides evidence of departure from the target conditions, corrective action or investigation is initiated. However, if conventional significance levels such as 5% or even 1%, were adopted, Type 1 errors would result in too many false alarms. For this reason, control charts on which sample data is plotted to give a topical display of the state of the process, often carry Action Limits corresponding roughly to 1/1,000 points of the target distribution. They may also have Warning Limits but, whereas a single violation of the Action Limit signals the need for investigation, two or more violations of the Warning Limit within a short sequence of samples are required to trigger the same action. In general terms, one requires k Warning Values in any succession of m values, and often k = m = 2 (i.e. two successive Warning Values constitute a signal).

D.9.2 The most commonly used control charts use Warning Values corresponding to about a 1 in 40 probability of occurrence when the process is at its target level. If one assumes, reasonably, that sample means (samples of generally 3-10 items) are approximately Normally distributed, the lower Action and Warning Values for guarding against any downward shift in average quantity become:

Action limit = $Q_t - 3.09\,\sigma$

Warning limit = $Q_t - 1.96\,\sigma$

In practice, these values are often rounded to $3\sigma / \sqrt{n}$ and $2\sigma / \sqrt{n}$ with little effect on the characteristics of the control chart.

D.9.3 Again taking an earlier example, paragraph D.5.2, suppose that the packer who sets his target quantity at 252 g (with a standard deviation of 5 g, this will avoid producing an excessive proportion of non-standard prepackages) decides to adopt a control chart procedure using samples of size n = 5. What Action and Warning Limits might he adopt?

With Q_t = 252g and $\sigma / \sqrt{n}$ = 2.236, the conventional alternatives are:

Action limit: either	252 - (3x2.236)	= 245.3 g
or	252 - (3.09x2.236)	= 245.1 g
Warning limit: either	252 - (2x2.236)	= 247.5 g
or	252 - (1.96x2.236)	= 247.6 g

These lines would be drawn, along with a further line at the target quantity of 252 g, on a chart and successive sample values entered on the chart. Similarly computer software is available for recording this data.

D.9.4 An alternative, and often more efficient, system is based on Cumulative Sum (Cusum) methods. Here a constant, known as a Target or Reference value, (T) is subtracted from each sample statistic (e.g. mean, range, etc) as it is obtained. The resulting deviations are cumulated, and this cumulative sum is recorded or plotted with $\sum (x - T)$ as ordinate (y-axis) and sample number as abscissa (x-axis). When the process conforms exactly to the target condition, the Cusum hovers around zero, giving a plotted path parallel to the sample number axis. If the average level of the sample statistic exceeds T, the Cusum increases, and its plotted path slopes upward.

Conversely, if the sample estimates lie predominantly below T, the Cusum becomes negative and its path slopes downward. The average level of the sample statistic is thus related to the slope of the Cusum chart, and decision rules, based on steepness of slope and the length of sequence over which it prevails, provide means of detecting the need for corrective action. A chart is not absolutely necessary, and numerical decision rules may be formulated, making Cusum methods suitable for computerized systems, the occurrence of signals being accompanied by estimates (based on the local rate of change, or slope, of the Cusum) of the magnitude of corrective action required.

D.9.5 The above is a description of the use of control chart and Cusum methods for monitoring the average level at which the process operates. Related procedures can also be applied to sample ranges or sample values of standard deviation, in order to monitor the process variability. Any change in variability needs to be taken account of, for two main reasons:

(i) a change in σ affects $\sigma / \sqrt{n}$ proportionately, and thus the control chart limits or Cusum decision rules for sample means need to be amended;

(ii) a change in σ may affect the risk of producing an excessive proportion of non-standard prepackages, and an adjustment to the target quantity may be required.

Further details of control procedure and methods for formulating target quantities to take account of the underlying process variation, appear in Annex E of this Manual.

D.10 Average Run Length

D.10.1 As for the OC curve of acceptance tests, some measure of performance is useful in assessing and comparing quantity control procedures. Probabilities do not offer a useful base because of the continuous nature of quantity monitoring methods.

Thus if a process is exactly at the target level, and a single Action Limit at the 0.001 probability level is drawn, then although there is but a low probability of a signal (which under these conditions would be a false alarm) at any one sampling point, when a sequence of say 100 samples is considered, the probability of a false alarm somewhere in the sequence is given by

$$1 - (1 - 0.001)^{100} = 0.0952$$

and the longer the sequence, the more likely a signal becomes.

An often-preferred criterion is the **A**verage **R**un **L**ength, ARL, until under any specified state of the process a signal is generated. The ARL is measured by the number *of samples* taken, on average, until occurrence of the signal, so that the sampling rate is also relevant to considering the average times to occurrence of the signal.

D.10.2 When the process is at or close to its target level, the ARL should be long, as signals are then false alarms. If the process moves appreciably from the target so that, for example, conformity to one or more of the Three Rules is jeopardized the ARL should be short.

D.10.3 It is impossible to generalize on suitable ARL characteristics as the ease and cost of sampling, and testing, the available knowledge of the behaviour of the process will affect their choice, and methods of using the data obtained from sampling. Often, however, ARLs around 500-1000 samples are preferred when the process is on target, so that unnecessary adjustments are minimized. When the process shifts to an unsatisfactory state, ARLs of 4-10 samples often result, and the condition can be quickly recognized and corrected. This performance will also be governed by the nature and size of the sample taken. Purely by way of example, in monitoring against the average requirement, a procedure based on the sample means in samples of about three to six items is often used. If the standard error of the sample means, which may depend on a medium-term variation as well as variation between items within the samples, is denoted by σ_e, then efficient control procedures may yield the following response to shifts in process average level.

No shift from target:	ARL 500-1000 samples
Shift of 0.5 σ_e	ARL 50-200 samples
Shift of σ_e	ARL 10-50 samples
Shift of 1.5 σ_e	ARL 8-20 samples.
Shift of 2 σ_e	ARL 4-8 samples

D.10.4 The ARL characteristics of typical and widely used control procedures are linked to the operating characteristics curves of the reference test, so as to provide a means of relating target and sampling levels to measures of process variation.

D.11 Components of variation

D.11.1 The measure of variation most commonly adopted to describe the dispersion in a set of numerical values is the standard deviation. Although other measures such as sample ranges are often used in quantity control, they are adopted because of their simplicity in application and in fact provide indirect estimates of the standard deviation. The square of the standard deviation, the variance (denoted by σ^2, or where an estimated value is concerned, by s^2) is of particular importance, especially when variation may arise from a number of contributory sources. Such components of variation may be identified with particular features of the process or measurement system under consideration.

Thus, in determining prepackaged quantities, the apparent variation may comprise the real prepackage-to-prepackage variation in quantities, the variation in tare weights of containers, and errors in the measurement system. Again, in quantity control operations, there may be local or short-term variation, measurable by ranges or standard deviations in small samples of items taken close together from the production line, but additional variation may arise from differences in average level over time (for example, due to fluctuations in materials or temperature controls), and different machines or filling heads may operate at slightly different levels, contributing yet another source of variation.

D.11.2 It is frequently possible to measure these contributions separately by means of variance components, each identified by σ^2 (or s^2) with a subscript to indicate the source of the contribution. In many cases, the overall variation in a system may be estimated by combining these variances, generally (but not solely) in an additive manner.

Where the variable representing the measured output of the system or process is a linear combination of *independent* contributory variables, the mean and variance of the overall measure may be represented as follows: -

If $Y = a + bx_1 + cx_2 + dx_3...$ etc.

(where x_1, x_2, x_3 etc are the contributory variables; b, c, d etc are coefficients which may be positive or

negative and are often equal to 1; and a is a constant, (often zero), then

$\mu_{Y.} = a + b\mu_1 + c\mu_2 + d\mu_4$ . . etc.

(where μ's represent mean values of the relevant contributory variables)

and $\sigma_Y{}^2 = b^2\sigma_1{}^2 + c^2\sigma_2{}^2 + d^2\sigma_3{}^2...$ etc.

Expressions also exist for the case of non-independent contributory variables, and for non-linear combinations, but are not considered here.

D.11.3 As examples of application, we take three situations. In the first, prepackages are being weighed gross, and a sample of containers is also weighed so as to permit estimation of weight. We shall also suppose that the measurement error of the weighing instrument is known, and can be expressed as a standard error for an individual determination. We have the following data:

Gross weight $\mu_g = 443.5$ g $\sigma_g = 4.4$ g

Tare weights $\mu_t = 69.8$ g $\sigma_t = 1.6$ g

Weighing errors $\mu_w = 0$ g $\sigma_w = 0.5$ g

In an informal test, we may wish to estimate the true weight variation, after allowing for the effect on overall variation of the measurement error and the rather large tare variation.

Here, the overall variable gross weight, is of the form gross weight = weight + tare weight + *error,* the three terms on the right being reasonably assumed independent, we may thus write;

$\mu_g = \mu_n + \mu_t + \mu_w$, and $\sigma_g^2 = \sigma_n^2 + \sigma_t^2 + \sigma_w^2$

It so happens that for each expression, only one of the right side terms is unknown so that

$443.5 = \mu_n + 69.8 + 0$, hence $\mu_n = 373.7$ g

and $19.36 = \sigma_n^2 + 2.56 + 0.25$, hence $\sigma_n^2 = 16.55$ g^2

giving a mean weight of 373.7 g and a standard deviation of 4.1 g.

D.11.4 A packer using bottles of closely controlled capacity on a filling line, filling to approximately constant vacuity, wishes to ascertain the true mean and variation of the filled content. The data are as follows:-

Brim level capacity, $\mu_c = 528$ ml, σ_c ,=4 ml. Vacuity, μ_v ,= 22 ml, σ_v =2 ml.

In this application, the structure of the final variable filled content is;

filled contents = capacity - vacuity,

so that $\mu_f = \mu_c - \mu_v = 528 - 22 = 506$ ml

and $\sigma_f^2 = \sigma_c^2 + \sigma_v^2 = 4^2 + 2^2 = 20$ ml^2

giving a mean content of 506 ml and standard deviation close to 4.5 ml. Note that the negative sign of the coefficient (- 1) for vacuity becomes positive on squaring.

D.11.5 Finally, consider a packing process with short-term variation of prepackaged quantity represented by a standard deviation, σ_o, of 35 g. Irregular and apparently random fluctuations in the average level of the process can similarly be described by a standard deviation component of 15 g. To formulate a quantity control system, the packer needs to estimate his overall medium term variation. We have here (irrespective of the mean quantity concerned):

medium-term variance= short-term variance plus variance of random fluctuations;

i.e. $\sigma_m^2 = \sigma_o^2 + \sigma_r^2$,

Now $\sigma_o = 35$ g and $\sigma_r = 15$ g

so that $\sigma_m^2 = 1225 + 225$ g $\sigma_m^2 = 1450$ g^2

and σ_m, the estimate of the medium term standard deviation becomes 38 g approximately. In an example of this king the packer may need to add an additional component to allow for the tare weight variations.

D.12 Conclusion

This chapter presents a brief summary of the statistical concepts underlying reference tests and quantity control. It is by no means exhaustive (for example, the basis of the more complex reference tests for measuring container bottles has not been covered) and is directed towards providing an introduction to the statistical material of this Manual.

10 Annex E Quantity Control by Sampling

E.1 Introduction

E.1.1 Two basic procedures are described for the evaluation of sample data. These are the Control Chart (often termed the Shewhart Chart, after its originator, although the details of operation often differ appreciably from the original) and the Cumulative Sum Chart - usually contracted to 'Cusum' chart. Both types of procedure can, in fact, be operated without charts, and they provide the basis of many integrated systems involving weighing equipment linked to calculators or computers. The weighing equipment needs to be both legal and suitable for the purpose.

E.1.2 The procedures (control chart or Cusum) may be applied to various sample statistics in order to ensure control of both average quantity and the proportion of non-standard prepackages. Generally, this proportion is controlled indirectly by monitoring the variability of the prepackaged quantities so that where variability is a problem, sufficient allowance is provided to limit the proportion of non-standard items.

Thus the most common forms of control involve one statistic measuring location and another measuring dispersion, though direct monitoring of numbers of items below T, may also be encountered. The usual sample statistics are:-

Location	Dispersion
Sample mean:	Sample standard deviation:
Sample median:	Sample range,

Sample means are often used in conjunction with either standard deviations or ranges. Because control by medians is generally adopted in order to minimize calculation, it is unlikely to be encountered in conjunction with standard deviations, but rather with ranges or the use of original values with a check on numbers below T, over a sequence of samples.

Note that n is used here to denote the number of items in each sample drawn from the production line. Generally, quantity control is applied by means of small samples drawn fairly frequently (relative to the rate of production) rather than by infrequent large samples, so that values of n from 2 to 10 are those most often employed. Generally samples with n=3 or 5 are used.

E.1.3 Before launching a control system, it is necessary to set up target values for each parameter to be monitored, e.g. a target mean or median, and a target range, standard deviation or proportion non-standard. Data giving an indication of the capability of the process under normal operating conditions is generally necessary in order to formulate reasonable and achievable targets.

The essential parameters needed for most conventional methods of control are: --

σ_0 a measure of the short-term variation as seen in within-sample differences between items.

σ_m. the medium-term variation, which will include σ_0 but may also contain other contributions from process fluctuations between the drawing of control samples. These contributions are measured by:

σ_1 representing the between-sample element of variation separated from within sample effects.

The within-samples element, σ_1, may be estimated either from sample standard deviations (s - values, averaged and adjusted to provide s_o) or via sample ranges, averaged and adjusted by an appropriate conversion factor. To avoid computing an estimate of σ_1, (or σ_m.), the packer may choose to adopt one of the simple procedures described in this chapter, though he may tend to over-control his process thereby. Methods for avoiding this over-control are described in this chapter.

E.1.4 For application of the target-formulation principles set out in the next section, it is assumed that estimates of σ_0 or σ_m are available from past records, or from a process capability check.

E.2 Principles for Estimating the Target Quantity, Qt

E.2.1 A reasonable basis for assessing a packer's system is to judge its effectiveness against that of a reference test. In order to set uniform standards, a test involving 50 items sampled from a production group of 10,000 prepackages is adopted as a yardstick for these comparisons. There is a risk to the packer of 1 in 200 that, when he packs to exactly the nominal quantity on average, a group will fail the reference test. Many packers will wish to incorporate a small level of overfill in order to reduce this risk, and the accompanying inconvenience of rejection, rectification, re-labeling or disposal.

E.2.2 However, the packer who actually adopts the n = 50 sample size for his own test would use as a rejection criterion the quantity

$$Q_n - 2.58\,\sigma/\sqrt{50} = Q_n - 0.364\sigma$$

Using this criterion, as well as rejecting 1 in 200 groups whose average contents are, in fact, at Q_n, he would also reject 12.5% whose contents were at $Q_n - 0.182\sigma$ and 50 % whose contents are at $Q_n - 0.364\sigma$. When operating on a continuous basis, this implies that if the process average falls to Q, - 0.364σ, this situation would (on average) be detected in the time taken to produce two 'groups'. These average run lengths, together with targets formulated in accordance with the Three Rules for Packers, provide the basic criteria for evaluating or setting up control procedures.

E.2.3 So far, only average quantity has been considered. Since control of the proportion of defective units will generally be implemented by setting the target average at a suitable level to avoid excessive numbers below TU_1 similar criteria can be applied to the adequacy of control systems in respect of the second Rule. Assuming an approximate Normal distribution of packed quantities (different coefficients as multipliers for σ may be necessary in cases of known non-Normality), a target set at $TU_1 + 2\sigma$, or more, will be appropriate and the 1 in 2 and 1 in 8 rejection probabilities will then occur at points which are respectively at $TU_1+\sigma$ (1.96-0.364) and $TU_1+ \sigma$(l.96-0.182) i.e. $TU_1+ 1.596\sigma$ and $TU_1+ 1.778\sigma$.

E.2.4 For ease of implementation, and with little effect on performance, the criteria may be rounded to the first decimal place. Along with a 1 in 10,000 limiting interpretation for the occurrence of inadequate prepackages arising as 'tail' values in an otherwise acceptable distribution, the principles for target setting, based on a quantification of the Three Rules and the application of the above principles, are as follows:

E.2.5 The Three Rules for Packers, appropriately quantified for target setting and the formulation of control systems, are as follows.

- The target quantity, Qt, may not be less than the nominal quantity declared, Qn.
- Not more than 1 in 40 (2.5%)[21] prepackages may contain less than the value of TU1 appropriate to the nominal quantity.

[21] Regarding the second rule, the Directive specifies an acceptable number of prepackages below TU_1 for each of reference test sample size. The proportion of prepackages below TU_1 needs to be sufficiently small, in general it appears that not more than 2.5% below TU_1 is appropriate.

- The packer must not pack quantities below the value of TU_2, appropriate to the nominal quantity, and system design must ensure that not more than 1 in 10,000 prepackages violate TU_2 by chance.

Where the assumption of a normal distribution of prepackaged quantities is reasonable, Rules 2 and 3 become:

2(a) $Q_t > T_1+2\sigma$

3(a) $Q_t > T_1+ 3.72\sigma$

(i.e. 1.96 rounded to 2σ, and 3.72σ correspond respectively to the 1 in 40 and 1 in 10,000 fractiles)

so that Qt becomes the largest of the three quantities Qn, $TU_1+2\sigma$, $TU_2+3.72\sigma$, plus any allowances as detailed in the later sections of this chapter.

For a Normal distribution of prepackaged quantities, whether the requirement for average contents, defective units or inadequate prepackages is the most critical will depend on the magnitude of the standard deviation. For non-normal distributions, the packer will need to examine the effects of the three requirements empirically.

For a Normal distribution, the criteria are as follows:-

- For σ not greater than 0.5 TNE, $Q_t \geq Q_n$
- For σ exceeding 0.5 TNE but not exceeding TNE/1.72, Q_t,>, TU1+2σ.
- For σ exceeding TNE/1.72, Q_t >, TU2+3.72σ.

Although where $Q_t = Q_n$., the requirement for not more than 1 inadequate prepackage in 10,000 becomes critical for σ > TNE/1.86. The requirement is already provided for by $TU_1+2\sigma$ up to the point where σ equals TNE/1.72; at this "changeover" value, we have:

Q_n.-2 TNE+ 3.72σ =.Q_n -TNE + 2σ (i.e. $TU_2 + 3.72\sigma = T_1 + 2\sigma$),

so that 1.72σ =TNE, or σ =TNE/1.72.

E.2.6 In order to summaries the requirements for system characteristics, we consider an average quantity, Q, which just satisfies the most stringent of the Three Rules, so that Q, is equal to Q_n, to $TU_1 + 2\sigma$ or to $TU_2 + 3.72\sigma$, where the quantities are assumed to be Normally distributed. One or other of the following properties is then required.

a If the process average falls to Q, - 0.2a (which corresponds to 58 % of prepackaged quantities below

Q_n .where rule 1 is critical, to 3.5% below TU1 where rule 2 is critical, and to about 1 in 5,000 where rule 3 is critical), the average time for detection should be not more than ten production 'periods'; or alternatively, the probability of generating a signal for corrective action within one 'period' should be not less than 0.1..

b If the process average falls to Q, - 0.4σ (corresponding to 65 %. of prepackaged quantities below Q_n, or to 5.5%. below TU_1, or to about 1 in 2,000 below TU_2), the average detection time should be not more than two periods, or the probability of generating a signal within one period should be at least 0.5.

c The risk to the packer of 'false alarms' ' i.e. of signals produced by unfortunate samples (or runs of samples) is at the packer's discretion. In the simple systems offered in the Packers' Code, it is of the order of 1 in 500 (or an average run length of 500 periods between false alarms when the process is operating exactly at the target level).

E.2.7 Some general points need to be considered in relation to both target setting and control procedures. Firstly, the sampling levels and allowances resulting from the foregoing principles, and detailed below, apply only where the packer is entirely dependent on sampling checks, and has no other sources of information on his process performance. Other relevant information may accrue from the use of checkweighers, in ways other than those constituting their recognition as prescribed instruments, from liquid level scanners, by occasional detailed process evaluation (such as head-balance checks on multi-head machines), any of which may be accompanied by lower levels of routine sampling than is suggested for the 'sampling only' procedures. The inclusion of such auxiliary data will not generally satisfy the requirements for adequate checks using lawful instruments, but may affect their intensity or frequency. Secondly, most packers will wish to monitor against excessive overfill, as well as against illegal under filling indeed, such overfill may itself be contrary to other legal requirements, as in the case of aerosols or dutiable goods. Thirdly, the target setting procedures are based on a knowledge of the variation of contents. Where indirect estimates are involved (e.g. gross minus constant tare, volume via weight and density) allowances may be necessary for the imprecision introduced by such methods. Guidance is provided in this chapter, for certain types of allowances commonly needed.

E.3 The 'production period'

E.3.1 For effective control, the system of checks must be related to a period of operation of the production process, or to a number of prepackages. Although an on-line reference test may be applied to a one-hour production run, which in some circumstances might constitute only a few hundred prepackages, it would be unreasonable to expect the slow-speed packer to check a greater proportion of his output than the high-speed packer. Strict application of an 'hourly principle' would involve anomalous implications of this kind.

For purposes of production checks by the packer, and with no implication whatever on the selection of groups for reference testing, the period of production to which the criteria should apply, is as follows:-

(i) Where production is at the rate of 10,000 or more per hour, the 'production period' is one hour.

(ii) Where production is at a rate of less than 1,000 per hour, the 'production period' is one day or shift, generally of 8-1 0 hours' duration. Special arrangements, reached by consultation between the packer and the responsible Inspector, may be appropriate for very low production rates, e.g. less than 500 per day or shift.

(iii) For intermediate cases, the 'production period' is the time taken to produce 10,000 prepackages, i.e.
Period = (10,000) / P where P is the normal hourly production rate.

E.3.2 The total number of items checked in each production period may often comprise several samples (say k) each of the same size (n). In subsequent sections, except where k and n are both identified, it is the total of k x n items sampled in a production period that governs the level of sampling allowances required. For organizational purposes, it may be preferable to define an hourly sampling rate. From the relationships above, and from the definition of total sample size per period (kn), we have:-

hourly sampling rate=kn items		for $P \geq$ 10,000,
or	(kn) / 8 items	for $P \leq$ 1,250,
or	(knP) / 10,000	for 1,250 $\leq P \leq$ 10,000.

In some cases a sample size (n) may be established from previous practice, or from considerations of operational convenience. The packer may also choose a level of fill that represents a reasonable precaution against violation of the Three Rules, and will then wish to calculate the number of samples (h) per hour that he must take in order to implement the suggested control procedures. This will be given by

$$h = kP / 10{,}000,$$

where k is the value for samples per period appropriate to the chosen overfill (expressed as a multiple of σ) and sample size, n, taken from the master Table E.3 of this chapter.

E.4 Procedure for target setting

E.4.1 Case 1: taking a single sample, N, in each production period (noting that where a signal occurs he may need to take *retrospective* corrective action where he samples only once per period), the relationship between the allowance for sampling variations and the sample size can be simply expressed as follows:-

$Q_t \geq Q_n + \sigma (u / \sqrt{N} - 0.4)$ where $\sigma \leq 0.5$ TNE

$Q_t \geq TU_1 + \sigma (u / \sqrt{N} - 1.6)$ where $\sigma > 0.5$ TNE but $\sigma <$ TNE / 1.72

$Q_t \geq TU_2 + \sigma (u / \sqrt{N} + 3.3)$ where $\sigma >$ TNE/1.72

(Alternatively, Q_t, may be calculated using all three expressions and the highest value selected).

In these expressions, u is the standard Normal deviate associated with the desired risk of false alarm. Frequently adopted risks are 1 in 1,000 (for which u may be taken as 3), 1 in 200 (u=2.58) and 1 in 40 (u= 1.96, often rounded to 2.0).

E.4.2 Where the target quantity is specified, and the appropriate sample size is required, the expressions are:-

$N \geq (u\sigma / (Q_t - Q_n + 0.4\sigma))^2$ where $\sigma \leq 0.5$ TNE,

$N \geq (u\sigma / (Q_t - Q_n - 1.6\sigma))^2$ where $\sigma > 0.5$ TNE but $\leq$ TNE / 1.72,

$N \geq (u\sigma / (Q_t - Q_n - 3.3\sigma))^2$ where $\sigma <$ TNE / 1.72

Again N may be calculated from all three expressions and the *largest* value taken.

E.4.3 Values of u appropriate to certain simple control chart procedures are given in Table E.1. For procedures A, B and C, large shifts in process average quantity are detected very quickly, and the criterion (b) of paragraph E.2.6 is applicable. For the widely used procedure D, the probability of a signal occurring is dependent on the values obtained from two or more successive samples and an average run length criterion is more relevant than a probability specification. Further this procedure is more effective than, for example, procedure A in detecting small shifts in the process average, and a corresponding decrease in either sample size or target level (or some combination of both) is acceptable. These two features are incorporated into the values of u listed in Table E.1 for procedure D.

E.4.4 Whether used with a control chart, as simple decision rules or in connection with a calculator/computer linked system, the procedures involve corrective action being taken under the following circumstances.

Procedure A (based on the principle of a control chart with single 'Action' limit at the 1 in 1,000 point when the process is at its target level).

Action required if a sample mean or median of N items falls below $Q_t - 3\sigma / \sqrt{N}$

Procedure B (as above, but with 'Action' limit at the 1 in 200 point when the process is at its target level). Action required if the sample mean or median falls below $Q_t - 2.58\sigma / \sqrt{N}$

Procedure C (as above, but with 'Action' limit at the 1 in 40 point when the process is at its target level).

Action required if the mean or median falls below $Qt - 2\sigma / \sqrt{N}$

Procedure D (Control chart with 1 in 1,000 'Action' limit and 1 in 40 'Warning' limit).

Action required if any mean or median falls below $Qt - 3\sigma / \sqrt{N}$

or if any two successive means or medians fall below $Qt - 2\sigma / \sqrt{N}$

(the use of 'strict' 1 in 1,000 and 1 in 40 points, so that 3.09 is used in place of 3 and 1.96 in place of 2 in the above expressions, is of no practical significance).

E.4.5 The values of Z in table E.1 are then used as follows:-

(i) To formulate Q_t , given N.

For $\sigma \leq 0.5$ TNE, $Qt = Q_n + Z\sigma$

For 0.5 TNE $\leq \sigma \leq$ TNE/1.72, $Qt = TU_1 + (2 + Z)\sigma$.

For $\sigma >$ TNE/1.72, $Qt = TU_2 + (3.72 + Z)\sigma$.

(ii) To find N, given Q,.

For $\sigma <$ 0.5TNE, $Z= (Qt - Qn) / \sigma$

For 0.5TNE $> \sigma \geq$ TNE/1.72, $Z= (Qt - TU1) / \sigma - 2$

For $\sigma >$ TNE/1.72, $Z= (Qt - TU2) / \sigma - 3.72$

The resulting value of Z must be positive. Then, according to the procedure (A, B, C or D) to be operated, select the value of N corresponding to Z. Where interpolation or rounding is necessary, the integer value of N must be obtained by rounding upward.

E.4.6 Case II: taking several samples during a production period, in k sub-units, each of size n, some modifications are necessary. This method is usually better at detecting small changes than is the single-sample-per-period method discussed above. However, detection of gross changes is still rapid, and the packer should be encouraged, rather than discouraged, to pay frequent attention to the operation of the process. The requirement adopted here is that the control procedure should provide an average run length of 2k samples for a shift in process average of 0.4σ from the target, or an average run length of 8k samples for a shift of 0.2σ. Table E.2 provides average run length data for the three procedures. An additional procedure E is a standard Cusum technique, described later in this chapter. Such Cusum procedures are characterized by a decision interval (h) and a reference shift (f), the values of these parameters for procedure E being h=5 and f=0.5.

Table E.1 Sampling allowance for item control methods based on a single sample of N items per period.

Procedure	A	B	C	D
Standard normal variate	u = 3	u = 2.58	U = 2	see paragraph E.4.3
Criterion for z	3 / √N - 0.4	2.58 / √N - 0.4	2 / √N - 0.4	The smaller of 2.75 / √N - 0.4 or 1.55 / √N - 0.2
Likely values of N	Values of z for each procedure			
50	0	0	0	0
40	0.07	0	0	0.03
30	0.15	0.07	0	0.08
25	0.2	0.12	0	0.11
20	0.27	0.18	0.05	0.15
16	0.35	0.25	0.1	0.19
12	0.47	0.34	0.18	0.25
10	0.55	0.42	0.23	0.29
8	0.66	0.51	0.31	0.35
6	0.82	0.65	0.42	0.43
5	0.94	0.75	0.49	0.49
4	1.1	0.89	0.6	0.58
3	1.33	1.09	0.75	0.69

Table gives average number of samples from the onset of a shift of 'z' standard errors until the occurrence of a control signal.

E.4.7 In order to obtain the z value for a particular procedure and a total sample size (per period) N made up of k sub-units each of size n, the following method is applied:

Find z corresponding to N in Table E.1 for the appropriate procedure. Note that this step does not apply for Cusum technique.

Find z' in the appropriate column of Table E.2 corresponding to an average run length L = 8k. Divide z' by √n and subtract 0.2 to obtain z, i.e. $z = z'_{8k} / \sqrt{n} - 0.2$.

Also find z' corresponding to an average run length L = 2k. Divide by √n and subtract 0.4, i.e. $z = z'_{2k} / \sqrt{n} - 0.4$

The smallest of the three resulting z values may now be adopted for target setting, following the method of paragraphs E.4.1. and E.4.2..

E.4.8 It is recognized that the above procedure is tedious and so Table E.3 therefore provides z-values for most combinations of k and n likely to occur in practice.

Table E.2 Average run lengths at various standardized shifts from target for four control procedures

	Procedure			
Critical shift from Qt in standard errors, z	A (As in Table E.1)	B	D Cusum h= 5, f=0.5	E
0.01	741	200	556	930
0.3	288	88	196	100
0.35	248	78	167	79
0.4	215	68	142	58
0.45	186	60	121	45
0.5	161	53	103	38
0.6	122	42	76	26
0.7	93	33	57	20
0.8	72	27	41	16
0.9	56	22	33	13
1.00	44	17.5	26	10.5
1.1	35	14.4	20	9.4
1.2	28	11.9	16	8.3
1.3	22	10.0	13	7.2
1.4	18	8.4	10.6	6.5
1.5	15	7.1	8.8	5.8
1.6	12.4	6.1	7.4	5.4
1.7	10.3	5.3	6.2	5.0
1.8	8.7	4.6	5.4	4.7
1.9	7.4	4.0	4.6	4.4
2.0	6.3	3.6	4.1	4.1
2.25	4.4	2.7	3.1	3.6
2.5	3.2	2.1	2.4	3.2
2.75	2.5	1.8	2.0	2.8
3.0	2.0	1.5	1.7	2.6

E.4.9 The decision rules, analogous to those of paragraph 9.4.4 for single samples per period, are as follows. In all cases they are based on the standard error of sample means, Φ_e. In some cases, this standard error will correspond to $\sigma_0/\sqrt{n}$, and it may also be expressed in units of sample range, but the definitive form is that given in this section.

Procedure A Action value at Q_t, $-3\,\sigma_e$,. (or Q_t, $-3.09\,\sigma_e$)

Procedure B Action value at Q_t, $-2.58\,\sigma_e$.

Procedure D Action value at Q_t, $-3\,\sigma_e$, and warning value at Q_t, $-2\,\sigma_e$, (or action value at Q_t, $-3.09\,\sigma_e$ and warning value at Q_t, $-1.96\,\sigma_e$)

Procedure E Cusum scheme operated with h=5.0 (i.e. H=5a.) and f=0.5 (i e F=0.5a,).

Whilst other equally valid sets of decision rules exist, they will not be compatible with the target setting procedure indicated in these sections, and evaluation of the adequacy of such procedures from basic principles may then be necessary.

E.5 Allowances other than for sampling variation

E.5.1 In addition to an allowance (designed to protect both packer and consumer against violation of the three rules) related to sampling variation, other allowances may be appropriate either because of measurement imprecision, or to compensate for known features of his process (e.g. drifting between sampling occasions, or cyclic fluctuations) or product (e.g. desiccation, or contraction of volumetrically controlled contents on cooling after a hot filling process). The allowance for measurement errors is the most widely applicable especially when these arise from the use of an assumed constant tare weight of containers in control by gross-minus-tare estimates of contents.

E.5.2 Using the methods provided, the extent of variation on tare weight of containers needs to be established. Where the short-term standard deviation of tare weights exceeds one tenth of the appropriate value of TNE, the packer will need either to pre-tare, or to make some allowance to guard against inadvertent under filling. Denoting the standard deviation of tare containers by st, there are two methods by which a tare variation may be applied. The first is to incorporate the tare variation into the value of a used for target formulation and control procedures. Assuming (as is generally reasonable) that content varies independently of container tare, we have σ gross = $\sqrt{(\sigma^2 \text{ process} + \sigma^2 \text{ tare}]}$

and, substituting estimates s_p. and s_t, for the σ's, the value of s for control purposes becomes

$$s_p.(t) = \sqrt{s_p^2 + s_t^2}$$

The alternative, which will result in larger allowances, is to use the procedure given for checkweighers, which is to add an independent tare allowance of 0.85 s_t to the target quantity, in addition to the sampling variation allowance.

E.5.3 Where constant tare is assumed, it is also necessary that the average tare subtracted from the gross weights should correspond to the actual current stock of containers, and regular cheeks on average tare, and tare variation, must be incorporated into the control system. It may be noted that the value of sp (t) should correspond to the standard deviation of gross prepackage weights, and there is no objection to this standard deviation being estimated directly from gross weights. However, because of the need to monitor tare deliveries for both average and variation, the diligent packer will prefer to obtain information by occasional weight determination, even if destructive, on the actual contribution of tare weights to his process variation.

E.5.4 Prepackages subject to loss by evaporation of weight or volume to a material extent may be reference tested for up to seven days following the end of the day on which they were packed. The packer will need to make an allowance for any such losses that may occur during this period, and may prudently allow for losses over longer periods. There may also be situations where losses may occur during storage, which may not be regarded as desiccation, and similar allowances, usually expressed as a proportion or percentage of the packed quantity, may be applied in such cases. In practice, the packer will need to make sufficient allowance so that prepackages will not become inadequate during normal conditions of storage and distribution.

Table E.3 Master list of sampling allowance factors, Z (For procedures A, D and E of section 9.4)

n=No of items per Procedure		k=number of sub-samples per period										
sub-sample		1	2	3	4	5	6	8	10	12	16	20
2	A	-	0.84	0.70	0.61	0.54	0.47	0.35	0.27	0.21	0.13	0.07
	D	-	0.58	0.43	0.35	0.29	0.25	0.19	0.15	0.12	0.07	0.03
	E	-	0.37	0.25	0.19	0.15	0.12	0.08	0.05	0.03	0	0
3	A	-	0.65	0.53	0.46	0.37	0.31	0.21	0.15	0.10	0	0
	D	-	0.43	0.32	0.25	0.20	0.17	0.12	0.08	0.06	0	0
	E	-	0.26	0.16	0.12	0.08	0.06	0.03	0	0	0	0
4	A	1.10	0.54	0.44	0.35	0.27	0.21	0.13	0.07	0.03	0	
	D	0.58	0.35	0.25	0.19	0.15	0.12	0.07	0.03	0	0	
	E	0.42	0.20	0.12	0.08	0.05	0.03	0	0	0	0	
5	A	0.94	0.46	0.37	0.27	0.20	0.15	0.07	0			
	D	0.49	0.29	0.20	0.15	0.11	0.08	0.03	0			
	E	0.35	0.16	0.08	0.05	0.02	0	0	0			
6	A	0.82	0.40	0.31	0.21	0.15	0.10	0.03	0			
	D	0.43	0.25	0.17	0.12	0.08	0.06	0	0			
	E	0.30	0.13	0.06	0.02	0	0	0	0			
8	A	0.66	0.32	0.21	0.13	0.07	0.03	0				
	D	0.35	0.19	0.12	0.07	0.03	0	0				
	E	0.23	0.08	0.02	0	0	0	0				
10	A	0.55	0.26	0.15	0.07	0						
	D	0.29	0.15	0.08	0.03	0						
	E	0.19	0.05	0	0	0						
12	A	0.47	0.21	0.10	0	0						
	D	0.25	0.12	0.06	0	0						
	E	0.16	0.03	0	0	0						
16	A	0.35	0.13	0	0							
	D	0.19	0.07	0	0							
	E	0.11	0	0	0							
20	A	0.27	0.07									
	D	0.15	0.03	0								
	E	0.08	0	0								
25	A	0.20	0									
	D	0.11	0									
	E	0.05	0									
30	A	0.15	0									
	D	0.08	0									
	E	0.02	0									
40	A	0.07	0									
	D	0.03	0									
	E	0	0									
50	All	0										
Notes:												

Multiply s. by the appropriate factor to arrive at the necessary allowance.

Allowances are not required for sampling variation if N or kn > 50.

Guidance on what should be regarded as a production period is given above

$Q_t = Q_n + Z\sigma$ (for σ < TNE/2)

or $TU_1 + (2+Z)\sigma$ (for σ > TNE/2 but σ < TNE/1.72)

or $TU_2 + (3.72 + Z)\sigma$ (for σ > TNE/1.72).

E.5.5 Formulation of target quantities by the methods of section E.4 provides reasonable assurance of conformity of groups of 10,000 prepackages or of one shift/day's production. Except for production rates of 10,000 per hour or more, where the basis of control becomes the hour, such groups will not correspond to the one-hour period that may form the basis of an on-line reference test, and there is the risk that if a medium term-component of variation exists (s_1), one hour's production may be rejected where the process 'wanders' during this period from its longer-term average.

The packer wishing to guard against this risk may include an allowance so that such one-hour wanderings are unlikely to lead to rejection in an on-line reference test. Two possible methods are as follows.

(i) Where two or more samples are taken during each hour, the target quantity may be increased (over and above any sampling and other allowances) by an amount equal to:

$$2\, s_1 / \sqrt{g},$$

where s_1 is the between-samples component and g is the number of samples (each of two or more items) taken per hour; or

(ii) where the normal sampling rate is one sample per hour, or less frequently, the value of s_1 may again be estimated, but should be related to a special sequence of samples taken once per hour for evaluation purposes. Then, the risk of reference test failure may be set at 1 in 200 (or less) by incorporating an allowance of at least Us_1, over and above the value Q_n, $TU_1 + 2\sigma$ or $TU_2 + 3.72\ \sigma$, as appropriate, where U is calculated as:-

$U = (2.68 \sqrt{(b^2 + 0.02)} - 0.379)/b$ where $b = s_1 / s_0$

Thus,

for $\sigma \leq 0.5$ TNE $Q_t \geq Q_n + us_1$:

for σ > 0.5 TNE but $\sigma \leq$ TNE / 1.72 $Q_t \geq TU_1 + 2s_0 + us_1$:

for σ > TNE / 1.72 $Q_t \geq TU2 + 3.72s_0 + us_1$:

E.5.6 Any other allowances that are incorporated above the nominal quantity, such as 'aesthetic' allowances to fill transparent containers to a level of acceptable appearance (albeit that a lower level would meet the declaration) may be incorporated in the sampling variation allowance when calculating target values or estimating sample size frequency requirements.

E.6 Control of short-term variation or proportion of defective units

E.6.1 In addition to ensuring that the average quantity packed conforms to the declared quantity, the packer's checks need to show that the proportion of defective units does not exceed 2.5%. In most cases, this assurance can be provided by monitoring the within-sample variation by the use of either sample ranges (R) or sample standard deviations (s).

This method has the further advantage that if any change in underlying process variation is detected, account can be taken of such changes in the procedures for monitoring average quantity. Control methods for averages depend on a measure of the process variation, which should be up-dated or reviewed periodically -monitoring of short-term variation provides a convenient basis for such review.

E.6.2 An alternative to monitoring sample standard deviations or ranges is to carry out a periodic, e.g. 3-monthly, process capability check in the manner used to set up the control system.

E.7 Control procedures (1)

E.7.1 *Control charts e.g. Shewhart* The statistical basis of the control chart has been described, and an example of the calculation of control limits for sample means was given. The same principle can be applied to monitoring of other sample statistics, notably medians, ranges, standard deviations and numbers of non-standard prepackages.

The essential stages are as follows:-

(i) Define the sample statistic to be monitored.

(ii) Formulate a target value for this statistic, which corresponds to the typical value when the process is operating satisfactorily.

(iii) Using the appropriate sampling distribution (or tables compiled for the appropriate sampling distribution), obtain the control chart limits. These usually comprise an approximate 1 in 1,000 probability point (Action Value) and on approximate1 in 40 probability point (Warning Value) under the target conditions. These limits may be lower limits (as for control of process average against under fill), upper limits (as for monitoring within-sample variation against an increase that would jeopardize conformity to Packers' Rule 2 or 3, or indicate a need to revise control limits for averages), or combined upper and lower limits (as where overfill is to be avoided as well as under fill).

(iv) A chart is then drawn up embodying the target value, the limits and any other relevant information, e.g. sample size, statistics on which the limits are based, etc.

(v) Sample values are entered on the chart as they become available, and adjustment or corrective action taken when a 'signal' occurs. A 'signal', in this context, constitutes any violation of an Action Limit and in general any two successive violations of a Warning Limit.

Two special points need to be made about two-way control charts for monitoring sample averages (means or medians) against both overfill and under fill.

Firstly, there may be a band of acceptable process performance, so that a target zone (rather than a single target level) is applicable. The appropriate procedure here is to take the lower boundary of the target zone as the level from which to calculate lower control limits, and the upper boundary of the target zone as the level from which to calculate the upper limits.

Secondly, and independently of the first point, if two successive sample values violate opposite Warning Limits, the implication is that an increase in process variation may have occurred and investigation on this point is appropriate.

E.7.2 *Control limits for sample means*. The appropriate sampling distribution for calculation of control limits is, in most cases, a Normal distribution whose mean, M, is at the target quantity Qt, and whose standard error is either:

(i) $\sigma_e = \sigma_0 / \sqrt{n}$ for samples of size n where no between-samples component exists; or

(ii) $\sigma_e = /(\sigma^2_1 + \sigma^2_0/n)$ where there is a between-samples component σ_1.

In the latter case, σ_e may be estimated directly from the sample means obtained in a process capability study.

The control chart limits are set out from the target quantity:

Action Limit = Q_t-3 σ. (or Q_t-3.09σ) Warning Limit = Q_t -2 σ. (or Q_t - 1.96σ)

For upper limits, with which the Inspector is less likely to be concerned, the values are:

Action Limit = Q'_t+3σ (or Q'_t,+3.09σ) Warning Limit = Q'_t, +2σ. (or Q'_t + 1.96σ)

Note that in some cases, Q'_t may be greater than Q_t where a target zone rather than a single target value is formulated. (In all these expressions, it is necessary in practice to substitute the estimates $s_{e,}$, $s_{1,}$ s_0 for the parameters σ_e, σ_1, σ.

E.7.3 *Control limits for sample medians.* The sampling variation of sample medians is approximately 25 % greater than that for sample means. The packer who uses medians rather than means may allow for this lower precision of control in several ways, of which the simplest is to set up the target values and limits in exactly the same way as for sample means. He will then meet the average run length criteria set out in 9.2, but at the expense of rather more frequent false alarms, i.e. some over-control, as compared with the packer using sample means. In fact, the average run length at target quantity would fall from 741 (for sample means) to 122 (for medians) in the case of Action-Limit-only charts, and from 640 (for means) to 97 (for medians) in the case of Action-and-Warning procedures. As an alternative to avoid over-control, the packer may base his limits on the observed standard error of sample medians, which, as noted, may be some 25 % greater than for means but he must also increase the Z factors used for formulating targets by 25 % before incorporating them into the expressions in section E.4.

E.7.4 *Control limits for standard deviation.* The sampling distribution of the sample estimate of standard deviation is not Normal (it is, in fact, a χ or $\sqrt{\chi^2}$ distribution), To avoid tedious calculation from tables of χ^2, control limits for standard deviation are expressed as multiples of the target value of a. This will usually be obtained from a process capability study, and the table below gives the multipliers for both upper and lower limits for most sample sizes in common use. Note that where s_0 (the estimate of σ_0.) is obtained by averaging a number of small-sample standard deviations, the appropriate bias correction factor, must be applied to the average before using this table.

It will be noted that upper and lower pairs of limits are given in Table E.4, The upper limits are those of interest in respect of conformity to the average quantity principles, but some packers will wish to take (justifiable) advantage of any reduction in variation, so as to tighten control limits for sample average or to reduce target quantities.

Table E.4 Control limits for sample standard deviation

Multipliers for:	**2**	**3**	**4**	**5.**
Upper Action limit	3.29	2.63	2.33	2.15
Upper Warning limit	2.24	1.92	1.77	1.67
Lower Warning limit*	-	0.16	0.27	0.35
Lower action limit*	-	0.032	0.09	0.15

Values for other sample sizes may be obtained from tables
* Optional limits

Multiply the target short-term standard deviation by the factors given to obtain the control limits.

E.7.5 *Control limits for sample range.* The sampling distribution for sample ranges (R) is less tractable than the other sample statistics, though approximations based on χ^2 exist. However, factors for provision of control chart limits are widely available, and need merely to be multiplied by the average sample range appropriate to the target level of short-term variation.

This, in turn, may be estimated either directly from the average range (R) in samples from a capability study, or by multiplying the target short-term standard deviation by the appropriate d. factor.

The control limit factors applicable to R are listed in Table E.5 below. Range control charts are not suited to sample sizes larger than 10 (for larger sample sizes, control by sample standard deviation should be used, as described in paragraph E.7.4 above).

Multiply the target sample range, **Rt** by the factors given in the table to obtain the control limits.

Table E.5 Control limits for sample range

Sample size	2	3	4	5	6	8	10
Upper action limit	4.12	2.98	2.57	2.34	2.21	2.04	1.93
Upper warning limit	2.81	2.17	1.93	1.81	1.72	1.62	1.56
Lower warning limit *	0.04	0.18	0.29	0.37	0.42	0.5	0.54
Lower action limit *	--	0.04	0.1	0.16	0.21	0.29	0.35

Multiply the target sample range, **Rt**, by the factor in the Table to give the control limit. * Optional limits.

E.7.6 *Control limits for number defective units ('TU1 count')*. The target level for a control chart based on a count of prepackages whose contents fall below TU1 is expressed as the average number of such prepackages expected per sample when the process is at its target level. Obviously this target average count must not exceed n/40, where n is the sample size; it may well be considerably lower if the packer wishes. The sampling distribution appropriate to the T, count is the Binomial distribution with parameters n (sample size) and p (target proportion of non-standard prepackages). For ease of computation or table look-up, the Poisson distribution usually approximates this with parameter m (or sometimes symbolized by λ or a) equal to np.

Action Limits then correspond to the number of defective units at (or in most cases just beyond) the 0.999 cumulative probability of the distribution. Because the Poisson (and Binomial) distributions apply only to integer values, there are generally no exact 0.999 and 0.975 points, and the conventional limits are therefore as defined.

Alternatively, where the Action Limit is at A, the Warning Limit is set at A-1, and a signal is generated if fewer than a critical number of non-warning values appear between two warning values. This critical gap may vary from 1 (implying two successive warning values) upward, depending on the average number of non-standard items at the target quality and sample size concerned.

For counts of defective units, it must be noted that a signal is generated if the limit is reached or exceeded. For this reason, fractional limits are often used - although such numbers never appear as counts, the limits are thereby made unequivocal. Typical fractional limits are (A-0.5) and (W-0.5) in traditional charts. Limits of the traditional type may be obtained from tables of the Poisson distribution, from quality control texts, or by calculation of the terms of the Poisson distribution until the required points of the cumulative distribution are reached.

E.7.7 Where small sample sizes are used, control charts for the number of defective units are insensitive, and a better procedure is to monitor the number of non-standard prepackages over a moving 'span' of the most recent N prepackages sampled, and to investigate or adjust if this exceeds a critical value. The simple scheme for packers using original values and medians is based on this procedure where N=60 or 100, but it may be applied to other values as in Table E.6.

Target proportion of non-standard items	Values of N							
1 in 40	4	4	5	6	6	7	8	10
1 in 50	4	4	4	5	6	6	7	9
1 in 100	3	3	3	3	4	4	5	6
1 in 200	2	2	3	3	3	3	3	4
1 in 500	-	2	2	2	2	2	2	3
1 in 1,000	-	-	-	-	2	2	2	2

Critical values are 1 in 20 points of the Poisson distribution with m = np, where p is the target proportion of non-standards

E.8 Control procedures (2)

E.8.1 *Cusum Charts.* The procedures in section E.7 related to charts of the Shewhart type. This section deals with Cusum charts. The essential steps are as follows:-

(i) Define the sample statistic to be monitored.

(ii) Formulate a target value of this statistic, which corresponds to the typical value when the process is operating satisfactorily.

(iii) Prepare a suitably scaled chart for plotting the Cusum. The simplest scale convention is to decide on the 'horizontal' distance between successive plotting positions (the sample interval) and to scale the 'vertical' axis of the chart so that a distance equal to the sample interval represents approximately 2 σ_e Cusum units (for sample means, or a. R or a.σ_0 units for sample ranges or standard deviations ('a' factors are given in Tables E.7 and E.8).

(iv) Label the vertical Cusum scale outwards from a central zero, positive values upwards and negative values downwards, in units corresponding to roughly $2\sigma_e$, or to a.Rk or a.σ_0 as appropriate.

(v) Prepare a mask as in for sample means, or for ranges or standard deviations. Note that the lower part of the mask is optional for the Cusum chart for means, because it is concerned with detection of overfill, and that only the lower part of the mask is required for Cusum charts for ranges or standard deviations (Cusum charts are not recommended for detecting reductions in process variation unless more sophisticated methods are adopted).

(vi) As each sample value x (be it sample mean, range or standard deviation) becomes available the target value for the sample statistic is subtracted, and the resulting deviations are cumulated to form the Cusum. The Cusum, C, is then plotted against the sample number.

(vii) At any stage, and especially if the appearance of the chart indicates a possible change (experience will rapidly provide a guide to this) the mask is applied to the chart and adjustment or correction is required if the Cusum path cuts the limb of the mask. No action is required if the Cusum path lies within the limb(s) of the mask.

(viii) At any stage, and especially if a change has been signalled, the average value of the sample statistic over a segment from the (i + 1)th to jth values inclusive may be calculated by reading from the chart (or from the data used to plot the chart) the values of the Cusum at the ith and jth samples.

E.8.2 Cusum decision rules for average charts. Although the parameters may be varied to suit individual applications, a useful general-purpose Cusum scheme for sample means uses the parameters h= 5, f=0.5. These are multiplied by □ $_e$ (the standard error of the sample means) for construction of the mask, thus:-

$H=h\,\sigma_e = 5\,\sigma_e$; $F = f\,\sigma_e = 0.5\,\sigma_e$.

The parameter H is the Decision Interval, and F corresponds to the slope (in Cusum units per sample interval) of the limbs of the mask. The average run length performance of this scheme is given in column E of Table E.2.

E.8.3 Cusum decision rules for range charts. Where Cusum charts are used for average control, it is logical to use them also for control of variation. Table E.7 gives the scale (a), decision interval (h) and slope (f) factors for monitoring against an increase in average range. Where any increase is signalled, the current value of average range can be estimated by the method given in paragraph E.8.1 (viii) and if required can be converted to an estimate of short-term standard deviation, s., upon dividing by the appropriate dn factor.

Table E.7 Cusum factors for range charts

Sample size, n	2	3	4	5	6	8	10
Scale factor (a)	1.50	1.00	0.85	0.75	0.65	0.55	0.50
Decision interval factor (h)	2.50	1.75	1.25	1.00	0.85	0.55	0.50
Slope factor (f)	0.85	0.55	0.50	0.45	0.45	0.40	0.35

The factors are used by multiplying them by the average range at target process performance, R,, viz: scale, aR units per distance corresponding to one sample interval, decision interval H=h.R, slope of mask in chart units per sample interval, F=f.R,.

E.8.4 *Cusum decision rules for standard deviation charts*. Table E.8 gives the scale (a), decision interval (h) and slope (f) factors for monitoring against an increase in average sample standard deviation, s. Where any such increase is signalled, the current value of s can be estimated by the method of E.8.1 (viii), and multiplied by the bias adjustment b_n factor to provide s., the estimate of current short-term variation.

E.8.5 *Cusum for 'TU1, count'*. As in the case of control charts, the discrete nature of counts of non-standard items places some restrictions on the parameters for decision rules. The following schemes are offered for use in conjunction with mt, the mean or expected number of non-standard items per sample, under target conditions. The smallest value of m, listed in Table E.8 is 0.25.

This implies a minimum sample size of n= 10, where the target fraction of defective units is 2.5 %

(i.e. m_t = N x % defective units = 10 x 0.025).

The method of scaling the chart also needs to be modified to take account of the discrete nature of the observations. The following procedure is suggested:

a. calculate the average number of samples required to yield an average of one defective unit;

b. round this value up to a convenient integer, and adopt it as the horizontal (sample axis) interval for the Cusum chart. The horizontal intervals are likely to be small - a reminder that individual samples provide little information; and

c. mark off the vertical (Cusum) scale in intervals of the same length as the horizontal scale, and label as consecutive even integers upward and downward from zero.

The Cusum $\sum (x-m_1)$ is formed and plotted in the usual way, where x is the number of non-standard items in a sample (this will often be zero giving a preponderance of negative contributions to the Cusum, offset by occasional but larger positive contributions).

Table E.8 provides decision interval and slope parameters, which are used to construct masks for use with the Cusum chart.

m_t	0.25	0.32	0.4	0.5	0.64	0.8	1	1.25	1.6	2
H	1.75	3.5	4.5	3	4	6	3	4	3	5
F	0.5	0.18	0.1	0.5	0.36	0.2	1	0.75	1.4	1

(Note that the H and F values are already scaled and must not be multiplied by m_t).

For intermediate values of m_t use H, F for the next higher value of m_t in the table.

E.9 Example for Sample Control

A cheese packer is packing 200 g nominal packs. His packing line can produce 5,000 packs an hour and he wants to control the process by taking samples of 5 packs every half hour. He is plotting the average of each sample on a Shewhart chart having a target weight of 201 g, with warning limits of 198 g and 203 g, and action limits of 197 g and 204 g. He uses a non-automatic weighing machine to check the gross weight of the product, it has d = 0.1g.

From his records over the last 40 samples of size 5 the calculated overall mean is 201.2g. The standard deviation of all the 200 readings is 5.0 g and the sample standard deviation is 3.1 g. He is using a constant tare of 2.7 g, and his checks on the tare show a variability of 0.1 g.

Knowing that for a 200 g pack the TNE is 9 g, and having Table E.3 showing for this type of control z-factors for samples of size 5 as

k	1	2	3	4	5	6	8
z	0.49	0.29	0.20	0.15	0.11	0.08	0.03

where k is the number of samples taken in each production period, comment on his system.

Comments on Packers System

1. Test for stability of process
$s_p / s_o = 5.0 / 3.1 = 1.6$. As this value is not approximately 1 it indicates that the mean of the process wanders. Therefore s_p should be used to set the target quantity.

2 Measurement uncertainty

Process variation	5.0 g	
Tare variation	0.1 g	
Balance error	0.3 g	weights exceeding 200 g uses more than 2,000d, so mpe is 3d
Combined s	5.2 g	5.0 g is dominant so add to it combined other uncertainties

$\sqrt{\{(0.1)^2 + (0.3 / \sqrt{3})^2\}} = 0.2$ g

3 Sampling allowance

With a packing rate of 5,000 / hr the production period is 2 hr, so the number of samples taken from each period is 4. From Table C1, with k = 4 and n = 5, z = 0.15

4 Storage allowance

No data given to assess if this is needed.

5. Tare variability

As s_t = 0.1 g, and this is less than TNE/10 , it can be considered insignificant and so no allowance is necessary.

6 Target quantity

Q_t is the larger of Q_n = 200.0 g

$T_1 + 2\ s_p$ = 191 + 2 x 5.2 = 201.4 g

$T_2 + 3.72\ s_p$ = 182 + 3.72 x 5.2 = 201.3 g

PLUS ALLOWANCES, in this example the information suggests that only a sampling allowance is required.

This is $z \times s_p$ = 0.15 x 5.2 = 0.78 g $\approx$ 0.8 g

So the minimum target weight is 201.4 + 0.8 = 202.2 g

Control limits for the average

Warning limits 202.2 $\pm$ 1.96 x 3.1 / $\sqrt{5}$ = 199.5 and 204.9 g

Action limits 202.2 $\pm$ 3.09 x 3.1 / $\sqrt{5}$ = 197.9 and 206.5 g

7. Control limits for sample standard deviation

From Table E.4

upper warning limit = 3.1 x 1.67 = 5.2 g

upper action limit = 3.1 x 2.15 = 6.7 g

8. Control limits for sample range

Target range $= s_o \times d_n$ = 3.1 x 2.326 = 7.2 g where d_n can be obtained from statistical tables.

From Manual Table E.5

upper warning limit = 7.2 x 1.81 = 13.0 g

upper action limit = 7.2 x 2.34 = 16.8 g

From the above it can be seen that his target quantity is too low and his limits are inappropriate. Furthermore his existing upper limits are closer to the target than the lower limit, this would have the effect of depressing the mean and so is not appropriate.

Note:

Control of the variation of process is needed as a small change in variability may produce inadequates.

If the packer wanted he could add 2.7g to the minimum target weight to obtain the minimum gross target weight and use this to monitor the average.

11 Annex F Checkweighers and other automatic instruments

F.1 The place of checkweighers in the average system

F.1.1 The requirements to be met by checkweighers in order to be usable for the checks made by the packer under the Directive are dependent on the type that is used and their manner of use. There are no restrictions on the precision of measurement required in the use of checkweighers, but conditions are laid down for allowances covering tare weight variation of containers and the 'zone of indecision' of the set point.
The 'zone of indecision' is the extent, expressed in units of mass, of the zone within which the machine may take two contrary decisions with respect to the same load. The value of the zone of indecision is taken as equal to 6xstandard deviation of the accept/reject distribution.
The 'set point' is the value, in units of mass, at which the instrument is set to accept/reject product.

F.1.2 Where checkweighers are used only to reject inadequate prepackages, or to reject (or control the proportion of) non-standard prepackages, back-up checks by sampling are necessary in order to ensure conformity to average quantity requirements. Where control of average quantity can be implemented via the checkweigher, either by a 'count ratio' method or by direct calculation of averages, no back-up checks by sampling are required, but certain checks on the integrity of the system, on the validity of set points, on the zone of indecision and on tare weights (average and variation) are necessary.

F.1.3 The current OIML recommendations[22] use different terminology, in this section 'zone of indecision' is referred to in OIML documents as 'maximum permissible standard deviation' (random). The documents also refer to a setting (target) error as the 'maximum permissible mean error' (systematic)
Both of these errors contribute to the uncertainty of measurement, with the random error dependant upon the accuracy class of the instrument.

F.2 Types of checkweighers

F.2.1 Four types of checkweighers are in use. Paragraphs F.2.2 to F.2.5 list the required capabilities of these types, which are in increasing order of complexity.

F.2.2 The first type includes simple checkweighers, which are capable of assigning each prepackage passing over them to one or two or more groups, according to whether the checkweigher detects their gross weights as being above or below a set point. For use under the average system, the set points (after allowing for average and variation of tare weights, the zone of indecision, desiccation, etc) must correspond to one or more of the following:

a. the nominal quantity, Q_n;

b. the value of TU_1 appropriate to the nominal quantity;

c. the value of TU_2 appropriate to the nominal quantity.

Prepackages lighter than a chosen set point (depending on the mode of operation selected by the packer) must be separated from the main stream of the production line. Such checkweighers may also be provided with feedback to the filling machine for adjustment based on the output of the checkweigher; a set point for separating or counting overweight prepackages (according to a criterion appropriate to the circumstances) is also permissible. In subsequent sections, checkweighers of this type will be referred to as *grading checkweighers.*

[22] OIML Recommendation R51

F.2.3 The second type covers checkweighers, which are capable of rejecting prepackages at one set point, which must correspond to either TU_1 or TU_2 and also of *counting* the accepted prepackages in one of two ways:

(i) the number of prepackages classified as heavier than the target set point corresponding to Q_n plus appropriate allowances; and the total number of accepted prepackages; or

(ii) as in (i) but instead of the total number of accepted prepackages, the number of prepackages classified as heavier than the TU_1 or TU_2 set point but lighter than the target set point corresponding to Q_n plus allowances.

At intervals depending on the rate of production, the ratio of the numbers counted in these two groups is calculated and any necessary adjustment made. Facilities for filling machine adjustment, and for rejecting overweight prepackages, may be incorporated as for grading checkweighers. In subsequent sections, checkweighers of this type will be referred to as *counting* checkweighers.

F.2.4 The third type must be capable of separating from the production line prepackages sensed as being lighter than a TU1 or TU2 set point, and of processing the signals from the weighing unit so as to produce a permanent record (printed or in an electronic store) of the average quantity, measured from a pre-set tare weight, of the accepted prepackages produced in every successive hour of operation. Means must also be provided either for giving a clear visual and/or audible warning if the average over any run of half the number of prepackages normally produced in one hour falls below the target weight, or for adjusting the filling machine so as to restore the average for the whole hour's run to the larger weight. In subsequent sections checkweighers of this type will be referred to as *average recording checkweighers.*

F.2.5 The final type of checkweigher operates on the weight of the product deposited in a load receptor. Signals from the weighing unit operate a device, which can separate from the prepackaged output any prepackages below TU_1 or TU_2 (according to the mode of operating selected by the packer). The machine must also either produce a one-hour average record and the half-group warning as for average-recording checkweighers, or must be arranged to adjust the filling machine, where necessary, at intervals not exceeding half the normal hourly production, so as to restore the average for the whole hour's run to the target weight. In subsequent sections, checkweighers of this type will be referred to as *weighing checkweighers.*

F.3 Operation of checkweigher system

F.3.1 The type of checkweigher used and its method of operation are important factors affecting other aspects of the system of overall quantity control. The two main considerations are the formulation of the target quantity, Q_t' and the need for, and nature of, back-up checks by sampling in some cases. We consider the various types of checkweigher in turn, noting in each case that when considering the allowances to be incorporated, certain allowances are common not only to all the types of checkweigher, but also to other control systems. These include allowances for desiccation, other storage allowances, and any provision for drift in process average fill over time.

F.3.2 *Grading checkweighers.* The essential factor is the choice of set point. Where a packer wishes to discharge his duties by treating the nominal quantity as a minimum, a grading checkweigher with its set point corresponding to Q_n (plus allowances) provides an acceptable means of ensuring conformity to the average system, and no other checks and records are required other than those relating to the setting up and maintenance of the checkweigher. Alternatively, if the packer chooses to *reject* prepackages whose contents are below TU1 then, irrespective of the inherent variation in his process, he may formulate the target quantity using the target setting equations relating to Q_n . For processes where σ_p exceeds 0.5 TNE, this may result in smaller sampling allowances being required for the back up checks. Finally, where the reject set point corresponds to TU_2 the packer must again carry out back up checks, but even for σ_p exceeding TNE/1.72, he may adopt the less stringent criterion involving TU_1 rather than TU_2. The level of the back up cheeks is at the packer's discretion, provided that he formulates the target quantity in accordance with the methods in paragraphs E.4 and E.5 of this Manual, and that the sample size and frequency are not less than two samples each of five items twice per 'production period'.

In summary, the alternatives are listed in Table F. 1.

Table F.1 Alternative procedures for grading checkweighers

Set point	Back-up cheeks	Formulation of target quantity
Qn'	None	At packer's discretion to avoid excessive rejections
TU1'	2 x 5 per period	Q_n,+Z σ
TU2'	2 x 5 per period	TU_1 + (2+ Z)σ , σ > 0.5 TNE or Q_n + *Z* σ, $\sigma \leq$ 0.5 TNE

Notes: Production period = 1 hour for P $\geq$ 1 0,000 or 8 hours for P $\leq$ 1,250, or 10,000/P otherwise,

where P is normal hourly rate of production.

Z is sampling allowance factor of Table E.3.

F.3.3 *Counting checkweighers.*

In addition to the usual checks on set points and zone of indecision, it is necessary for the packer to evaluate the Reference Count Ratio on setting up and whenever significant changes in filling process characteristic are suspected. The method involves adjusting the filling process until the average (gross) weight of small samples drawn from the 'accepted' stream is close to but not consistently less than the target gross weight. The target set point is then established. The reference count ratio is established as follows.

(i) Draw about 100 prepackages at random from an hour's production from the 'accepted' stream, starting with grade counters set to zero.

(ii) At the end of the hour, provided that the average of the sampled prepackages is not less than the target quantity, obtain the readings from the grade counters, which may comprise: C_1, number heavier than TU_1 or TU_2, but lighter than Q_t, and C_2 number heavier than Q_t,, or C_3 total number accepted in the hour.

Then the Reference Count Ratio is either:

$R_1 = C_2/C_3$ (in the region of 0.5),

or R_2 C_2/C_1 (in the region of l.0).

The extent to which these ratios will depart from 0.5 or 1.0 will depend on the shape of the distribution of prepackaged quantities, on the filling process variation, and on the tare variation and zone of indecision of the checkweigher. As examples, consider two situations for a Normal distribution of quantities with σ =0.5 TNE, and Q_t set to exactly Q_n.

The first case assumes that there is no tare variation and the zone of indecision is zero (a perfect checkweigher). Here, 2.275 % of prepackages will be rejected at the TU2 set point and 97.725% accepted. Of all prepackages produced, 50% will be counted as being at or above the Qn set point. In a typical set of 5,000 prepackages, the counts would be:

C_1 2,386 C_2 2,500 C_3 4,886 Giving $R_1 = 0.5117$ or $R_2 = 1.0478$

Secondly, if the tare variation is exactly equal to 1/10 TNE and the checkweigher zone of indecision exactly 0.25 TNE, so that no allowances on target fill or adjustment to set points are required (see section F.5 below), the apparent fill standard deviation would become 0.5679 TNE, and the corresponding counts would be approximately:

C_1 2304 C_2 2500 C_3 4804 Giving R_1= 0. 52 or $R_2 = 1.085$

The control system must include an evaluation of the actual count ratio in each half-period, so that adjustment may be effected to ensure the correct average quantity over each full period, in addition to the checks on set points and zone of indecision.

F.3.4 *Average recording checkweighers.*

In addition to the usual checks on set points and zone of indecision, periodic checks on the integrity of the weighing/recording system are necessary, on initial setting up and when ever there are grounds for suspecting significant changes in the operating characteristic of the checkweigher. The procedure is as follows.

(i) A sample of 100 prepackages is taken at random from the "accepted" stream over the period of production time between successive recordings of average weight. The prepackages are weighed gross on a suitable and legal weighing machine (weighing to < 1/10 TNE) and the average weight calculated using the same assumed tare as is used in the checkweigher system.

(ii) If necessary, the checkweigher is adjusted so that the discrepancy between the actual average obtained from the sample and that recorded by the checkweigher does not exceed A+B+C,

where A is one unit in the last significant decimal place of the recorder or display,

B is half the zone of indecision,

and C is one-fifth of the filling process standard deviation

The contributions B and C allow, respectively, for 'weighing error' in the checkweigher, and sampling variation in the mean weights of sets of 100 prepackages sampled from the production line.

F.3.5 *weighing checkweighers.*

In this case, a test weight equal to the target quantity, Q,, is supported by the checkweigher load receptor and the average weight over 100 weighing cycles is recorded. The weighing unit is adjusted, if necessary, so that the discrepancy between the weight of the test weight and the recorded average from the checkweigher does not exceed A+B,

where A and B are defined as in paragraph F.3.4 for average recording checkweighers. Because no sampling from the production line is involved, contribution C does not appear in this expression. The test weight must be calibrated on a suitable and legal weighing machine (weighing to < 1/10 TNE).

F.4 Routine cheeks on set points and the zone of indecision

F.4.1 The test for all mandatory set points (which may comprise one or more of Q_n, TU_1 and TU2 as described in section F.2) involves the preparation of dummy packs corresponding to a 20 %. chance of rejection and an 80 %,, chance of rejection at the set point. These correspond to set point ±0.142 x zone of indecision. (Note: If 0.142 x zone of indecision is smaller than one-tenth of the TNE, use the latter figure instead.)

The test may be carried out either by passing one of each type of dummy pack over the checkweigher several times, or by preparing a number of identical dummy packs of each type. The full test procedure requires that each dummy pack be passed over the checkweigher 20 times (or that 20 dummy packs at each test weight be passed over the checkweigher). The number of rejections (R) of the underweight dummy pack is recorded, and also the number of rejections (G) of the overweight pack.

The test is acceptable if

R+G lies between l5 and 25 inclusive

and R-G lies between 7 and l7 inclusive.

If this test fails, it is repeated and it the combined R, G totals fail to satisfy the criteria

R + G between 32 and 48 inclusive,

R-G between l6 and 32 inclusive,

the checkweigher requires adjustment or attention before further use.

For routine use at the commencement of each production period (1 hour, 10,000 prepackages or 8 hours as defined in Table F. 1) a shortened procedure may be used,

- One pass of each (1 overweight, 1 underweight) test pack. If each is correctly graded, test ends. Otherwise
- four more passes at each test weight. If at least one pass at each weight results in correct grading, test ends. Otherwise complete the full set point check as above by 15 further passes at each test weight.

F.4.3 The tests for determination of the zone of indecision take four forms, depending on the type of checkweigher and whether it is tested under normal production conditions or not.

a. The method unsuitable for use under normal conditions involves a single adjustable dummy pack. This pack is passed over the checkweigher 20 times at each of a number of increments; at the lower end so that all 20 passes result in "rejection" and at the upper end so that all 20 result in "acceptance". The whole procedure is repeated another four times, giving a total of 100 passes at each increment. The proportions of rejections at each increment are then plotted against the added weight on arithmetic probability paper. If a reasonable straight-line relationship appears to apply, the set point is taken as the 50%. accept/reject level of weight increment and the zone of indecision as 6 σ, where σ is estimated as one quarter of the difference between the added weights corresponding to 2.275 % and 97.725%. rejection.

b. For use under normal production conditions, a suitable number of dummy packs are prepared to span the estimated zone of indecision. The dummy packs are inserted into the production line and the accept/reject pattern noted for 100 passes of each pack. The method of estimation of zone of indecision and set point is then as in (a) above.

c. Where (a) and (b) are impracticable and (d) is not applicable, sampling from the checkweigher output becomes the only available procedure. It is necessary to obtain at least 100 items from the reject channel, and at least a comparable number (over the same period) from the accept channel. The frequency distribution of the weights from the two channels is used in order to estimate the zone of indecision, again using a probability plot at the final stage of analysis. The test can be tedious and time-consuming where little rejection occurs (unless this can be artificially increased for a period), and the precision of assessment of the checkweigher performance is subject to the chance variations of the actual filling operation.

d. For average recording and -weighing checkweighers where individual weights can be displayed or recorded, a direct estimate of the zone of indecision is available. At gross or weights (as applicable) corresponding to each of Q_t,, TU_1 and TU_2 the average and standard deviation of 100 passes of a test weight or pack are obtained, viz:

X_a, S_a, Mean and standard deviation of Q_t test weight;

X_b, S_b, Mean and standard deviation of TU_1 test weight

X_c, S_c, Mean and standard deviation of TU_2 test weight.

The discrepancy between the X's and the corresponding true test weights (as measured by a machine reading to < 1/10 TNE) must not exceed A+ B (as defined in paragraph F. 3.4). In order to determine the zone of indecision (and hence to obtain B for the check on set point), calculate zone of indecision = $2(S_a, + S_b + S_c)$, i.e. six times the average of the three standard deviations.

F.5 Allowances for tare variation and zone of indecision

F.5.1 In calculating target quantities for use in conjunction with checkweigher systems, most of the allowances normally applicable to other forms of control remain relevant. These include allowances for desiccation, other storage losses, process drift, etc. The average value of tare weight of containers needs to be monitored so as to allow for batch-to-batch changes in tare weight, or for changes due to gain or loss in moisture of hygroscopic prepackaging materials. However, special considerations apply to allowances for tare variation in the short term, as the incorporation of the tare standard deviation into a measure of gross process variation is inapplicable.

F.5.2 The tare variation, along with the zone of indecision, comprises the measurement error of the checkweigher system. Because of the need for independent monitoring of these two contributions, separate allowances are also calculated as follows.

F.5.3 Where the zone of indecision does not exceed one quarter of the value of TNE appropriate to the quantity packed, no allowance is necessary. However, for larger zones of indecision, an allowance is made by shifting the *set point(s)* upwards so as to restore the rejection characteristics of the checkweigher to what they would be if the zone of indecision were 0.25 TNE. This is achieved by an upward shift of the set point(s) equal to 0.5 (zone of indecision) - 1/8 TNE.

Thus when the zone of indecision equals 0.25 TNE, the shift becomes zero. Negative shifts for smaller zones of indecision are not permissible.

F.5.4 To allow for tare variation, an allowance of 0.85 of the short-term tare variation (s_t,) is added to the target quantity, except where s_t, does not exceed one-tenth of the TNE, when no tare variation allowance is required. This allowance (except where the tare variation is trivial) provides for a 4 in 5 chance of rejection of prepackages whose contents are just at the set point (TU2, TU1, or Q_n.) providing the basis for separation of the product stream.

F.6 Records for checkweigher systems

F.6.1 For checkweigher systems to qualify as checks using suitable equipment, certain basic data used for setting up and maintaining the checkweigher must be recorded and made available to the Inspector. These records will vary somewhat in type according to the nature and mode of operation of the checkweigher. They are as follows.

F.6.2 On initial setting up of a checkweigher for a particular product, a record must be made of:

(a) the zone of indecision;

(b) allowance (if any) for desiccation,

(c) checks on the integrity of the system - the error of the weight seen and the accuracy of the calculations.

F.6.3 On initial setting up, and thereafter whenever it is suspected that *the filling process or tare* characteristics have changed, sufficiently to invalidate the calculation of the Q_t', TU1', or TU2' set points, a record must be made of:

(a) the filling process variability, s_o;

(b) the average tare;

(c) the tare variability, s_t;

(d) the Reference Count Ratio (R_1 or R_2).

F.6.4 On initial setting up and whenever they are re-calculated, the values of those of the following set points used in the operation of the checkweigher must be recorded:

target set point;

(a) Q_n' set point;

(b) TU1' set point;

(c) TU2' set point.

F.7 Other automatic instruments

Certain types of automatic equipment exist for the monitoring of filling levels for prepackaged liquids, and such instruments may provide a substantial element of assurance of conformity to the average system. They may therefore be incorporated in a packer's checking system, but will require back-up checks, never less stringent than those for checkweighers. The packer also needs to demonstrate the effectiveness of his system overall (except where he carries out full as defined in section 9.4 appropriate to the target fill and production rate).

F.8 Example for Checkweigher control.

A packer is producing tins of peas with a nominal quantity of 425 g, producing 3,000 tins an hour. The filling process has a variation (s) of 6g, and the tins have a tare of 50 g with a 4 g variation. The checkweigher has a zone of indecision, zoi, (random variability) of 3.6 g. What is the minimum target set point value and what other checks need to be carried out?

The TNE is 3% of 425 g = 12.75 g, this must be rounded up to the next 0.1g, i.e. to 12.8g.

The target set point is

Q_n	=	425 g,	plus allowances for
Tare weight	=	50 g,	
Tare variability	=	3.4 g,	see F.5.4 an allowance of 0.85 x variability
Storage	=	0 g,	no information given in example

Zoi	=	0.2 g	see F.5.3 an allowance of zoi/2 – TNE/8
Set point	= 478.6 g		

Monitoring of the set point, see F.4, needs to be carried out, for this set point the test blocks need to be 478.6 $\pm$ 0.142 x 3.6 g, that is 478.1 g and 479.1 g.

12 Annex G Control using Measuring Container Bottles and Templates

G.1 Measuring container bottles (MCBs) are made and checked for compliance with the requirements in Directive 75/107/EEC. These require the appropriate markings :

(a) a mark by which the manufacturer can be identified

(b) the reversed epsilon (3) at least 3 mm high

(c) its nominal capacity,

(d) an indication of the brim capacity expressed in centiliter and not followed by the symbol cl, and / or an indication of the distance in millimeter from the brim level to the filling level corresponding to the nominal capacity, followed by the symbol mm

The minimum height of the indications mentioned in c) and d) above are the same as those in paragraph A.11 above but with a minimum of 3 mm.

G.2 Directive 76/211/EEC requires that MCB are 'of the type defined in the Directive relating thereto[23], filled under the conditions prescribed in that Directive and herein'. MCB only relate to capacities of 50ml to 5 liters. The conditions relate to the volume being at 20^{o} C and that the 3 packers rules are met. The uncertainty of measurement, from both the error permitted on the MCB and the measurement of the liquid level, must be taken into account when establishing the appropriate fill height and control limits.

G.3 Unless other means of controlling the contents are employed, MCB can only be used for monitoring content fill if they are used with a certificated template. The template must bear the following inscriptions:

(a) identity mark,

(b) identity of MCBs for which it is to be used

(c) the nominal quantity being packed

(d) graduations to permit the determination of fill quantity errors

(e) the operational temperature of the liquid, and if that temperature is not 20 ^{0}C, a description of the liquid and the apparent thermal co-efficient of cubical expansion [24] by reference to which the template has been graduated.

(f) if it is to be used over a closure a statement to that effect and the identity of the closure.

G.4 The template must either be graduated in milliliters or in millimeters. In the latter instance a conversion chart must be available so that actual volume errors in the liquid fill can be determined.

G.5 The packer should take into account when setting the target fill volume the following:

(a) variation in the filling process

(b) variation in the MCBs, and

(c) possible error in determining the fill height using the template.

---

[23] Directive 75/107/EEC

[24] That is the thermal coefficient of cubical expansion of the liquid minus the thermal coefficient of cubical expansion of the material of the MCBs.

G.6 For the MCB to be suitable, for the portion of the template scale between TU2 and the fill height, 1 mm difference in liquid height should correspond to at least one-fifth of the TNE. Also in this range the meniscus should be clearly visible and there should be no distortion..

G.7 Certification of a template

G.7.1 Because of the variation in thickness of the material of which the bottles are made, in the relevant part of the neck, templates must be calibrated by reference to MCBs conforming close to the average design specification. The calibration procedure, which consist of the following:

(a) establish a calibration curve or equivalent tabulation relating to liquid levels with differences, expressed in milliliters at the intended operational temperature, from the scale mark corresponding to the nominal volume, in respect of the pattern of MCBs in question. This 'master' calibration function must be established by means of experimental measurements on not less than 10 MCBs, which have been selected as being as close as practicable to the average design height, diameter (or breadth) and capacity at the nominal fill level.

(b) templates must be constructed in accordance with the master calibration curve or tabulated function as described in a) above, against which they must be verified by an Inspector, who will test the scale marks by linear measurements, at three or more points relative to the brim or closure seating, that is the nominal volume scale mark and the two extreme scale marks. The maximum error of position allowed is $\pm$0.5 mm.

(c) The MCBs is filled with water at 20 ^{0}C to the fill level marked on it. The level is adjusted by reference to a depth gauge and should be measured centrally, inside the bottle, to the point where the end of the depth gauge just and only just touches the surface of the water.

(d) Where the template is to be used at 20 ^{0}C it is then placed in position either on the naked brim or on a suitable uniform closure, depending on which method is to be used, and the nominal volume scale mark viewed horizontally against the bottom of the water meniscus. The maximum error allowed is $\pm$ 0.5 mm.

(e) Where the template is marked with an operational temperature other than 20 ^{0}C the procedure in d) above is modified as follows as the position of the nominal volume scale mark will have been appropriately adjusted using the apparent thermal coefficient of cubical expansion for the liquid:

i. for an operational temperature below 20 ^{0}C the numerically smaller limiting value of the coefficient must be used,

ii. for an operational temperature above 20 ^{0}C the numerically larger coefficient should be used,

iii. after the MCBs is filled as in c) above, water is inserted or extracted equal in volume to the amount by which the nominal volume of the liquid with which the template is intended to be used would have increased or decreased in volume respectively if its temperature were to change from 20 ^{0}C. The nominal volume scale mark is checked (over a closure if specified), a maximum error of 0.5 mm is allowed.

G.7.2 Where the template is being used over a closure the component of variability of the measurement of the fill level attributable entirely to it must not exceed ± 1mm and must be taken into account in establishing the target quantity. To determine the variability take at least 10 normally filled and closed bottles from the line and measure on each the distance from the bottom on the MCBs to the top of the closure, remove the closure and determine the distance from the bottom of the MCBs to the brim. Subtract for each bottle the second from the first to determine the increase in height attributable to the closure. The average of these heights gives the distance the nominal volume scale mark should have been adjusted from the fill height marked on the MCBs and the standard deviation of these differences will give the component of variability due to the closure assuming the MCBs has uniform height.

G.8 Example for MCBs and Template Control

A packer is packing bottles of drink with a nominal quantity of 200 ml in a MCBs with a fill height of 10 mm. The neck diameter at the fill height is 40mm. The MCBs tolerance is 6 ml, inferring s=3 ml, and the filling process has a standard deviation of 5 ml.

The TNE is 9 ml. The MCBs cross sectional area at the fill height is □ (2^2) = 12.6 cm^2 . Therefore 1 mm change in liquid height is equivalent to 1.3 ml.

Uncertainties of measurement

Element	Value	Divisor	Multiplier	Std Uncertainty
		(distribution)	(sensitivity)	(ml)
MCBs	6 ml	2	1	3
Templet error	0.5 mm	√3	1.3	0.4
Reading error	0.5 mm	√3	1.3	0.4

As MCBs uncertainty is dominant add to it the combined uncertainties of the other elements.

Uncertainty (s) = $3 + \sqrt{\{ (0.4 / \sqrt{3})^2 + (0.4 / \sqrt{3})^2 \}}$ = 3.3 g. Combining this with the filling variation gives an overall variation in the filling/checking systems of $\sqrt{\{ (3.3)^2 + (5)^2 \}}$ = 6 ml.

Therefore the target quantity needs to be the greater of

Q_n = 200.0 ml

$TU_1 + 2\ s$ = 191 + 2 x 6 = 203.0 ml

$TU_2 + 3.72\ s$ = 182 + 3.72 x 6 = 204.3 ml

Target quantity is 204.3 assuming more than 50 items are checked within a production period, otherwise a sampling allowance (using z-factor) would also be required. The limits for the mean and range or standard deviation are calculated as in E.9.

13 Annex H Reference Tests

H.1 Nature of test

H.1.1 The two tests specified in Annex II to the Directive check for compliance with the first of the three Packer's Rules. The third Packer's Rule (there shall be no prepackage having a negative error greater than twice the tolerable negative error (that is below TU2) is not part of the reference test but during a reference test such a prepackage is treated as being a non-standard (below TU1) for the purposes of the reference test.

H.1.2 An Inspector can choose which of the tests to carry out first, and if that test fails there is no need to carry out the other test. Generally if the fill process variability is less than 0.5 TNE then the first Packer's Rule (the average rule) is going to be the most critical and that test should be performed first. Conversely if the process variability exceeds 0.5 TNE then the second Packer's Rule (for the number of non-standards) should be carried out first.

H.1.3 The Directive specifically states that destructive tests shall only be used where non-destructive test is impractical, as destructive tests are less effective than non-destructive tests. Examples would be where the tare is not constant or where templates have been used to fill MCBs. For good practice non-destructive tests should not be carried out on batches of fewer than 100 prepackages.

H.1.4 Besides the effectiveness of the statistical tests, generally checks are more effective, give a better indication as appropriateness of the quantity control system and are more likely to identify problem areas if carried out on samples drawn from batches already produced. Checking on the packing line is unlikely to find problems as the packer can continue his normal process control.

H.2 Choosing a sampling plan

H.2.1 An assessment of measurement uncertainty (see below) needs to be carried out to determine which method of testing can be used. Non-destructive sampling can only be carried out if the tare variability is sufficiently small and, for gravimetric testing of volume quantities, the error is determining the density is sufficiently small.

H.2.2 If destructive testing has to be carried out this is best done using a double sampling plan to minimize the number of prepackages that have to be opened. When sampling from storage it is best to take the sample size needed for both parts of the double sampling plan at the same time and also to take a few extra prepackages in case of breakages etc.

H.2.3 All sampling must be carried out randomly, for large batches various levels can be introduced to make sampling easier. For example the levels may represent the pallet, level on the pallet, the outer container on a level and the item within the outer container - each of these levels must be determined randomly.

H.3 Measurement

H.3.1 The requirement for errors of measurement, in the estimation of quantities, is that the total measurement error may not exceed 1/5th TNE. In order to consider the contribution of various sources of error, it is necessary to interpret this in terms of

$$\text{Standard deviation, } s_z \leq \text{TNE}/10$$

where s_z is the overall standard error of measurement, which includes the errors in the equipment being used.

H.3.2 However the overall measurement error may comprise two or more contributions where any indirect determination of contents is involved. For example, where quantity is estimated as gross-minus-tare weights, with measurement errors of s_a and s_b respectively at the two stages, (sometimes both these errors will be of similar magnitudes), then

$s_z = \sqrt{(s_a^2 + s_b^2)}$

H.3.3 More complex situations may be involved, when for example a destructive test is carried out on a tub of solid or viscous product. Here a possible method of determination of volume may be as follows.

Weigh tub and contents: W_1

Weigh tub + contents + water added to brim of tub: W_2

Weigh empty tub: W_3

Weigh tub filled with water to brim: W_4

The weight of contents is obviously $W_1 - W_3$ and the weight of an equal volume of water is given by $(W_1 + W_4) - (W_2 + W_3)$. The overall error in the two weight estimates is then:

contents; $\sqrt{(s_1^2 + s_3^2)}$ (where the subscripts refer to the respective weighing operations); equivalent volume of water; $\sqrt{(s_1^2 + s_4^2 + s_2^2 + s_3^2)}$.

H.3.4 It is the latter, involving four error contributions that provide the measurement error appropriate to a determination of quantity for a product declared by volume. In other cases, volume may be estimated by weight divided by a density estimate. Where s_w represents the total error in the weight determination for an estimated quantity, W, and s_d represents the error in determination for a density estimate, D, the overall error in volume determination is given approximately by:

$s_v = W/D \sqrt{((s_w/W)^2 + (s_d/D)^2)}$

H.3.5 Finally, where a constant tare weight is assumed for estimation via gross-tare, the actual tare variation, s_t , must be sufficiently small to ensure that

$s_z \tilde{=} (s_{tare}^2 +$ other error variance contributions) does not exceed 1 /10 th TNE.

H.3.6 An example of two of the more complex situations is appropriate. Firstly, suppose that for the product in tubs, the errors were:-

$s_1 = s_2 = s_4 = 0.5$ g

$s_3 = 0.2$g (empty tub weighed on a more precise instrument)

With $W_1 = 145$ g, $W_2 = 155$ g, $W_3 = 20$ g, $W_4 = 160$ g, for a product with $Q_n = 125$ g.

Here, TNE = 4.5 % of 125 g, i.e. 5.625 g, this is rounded up to 5.7 g

Assuming the weighing can be carried out at 20'C, the error in volumetric determinations, estimated by (145+170)-(155 + 20)= 140 g or ml, becomes

$s_z = \sqrt{(0.25 + 0.25 + 0.04 + 0.25)} = 0.89$ g,

which exceeds 1/10th TNE and does not therefore constitute a valid measurement procedure. More precise determination of the larger weights (although each individually satisfies the 1/10th TNE criterion) is required, or alternatively the test must be regarded as informal.

H.3.7 Secondly, consider an estimation of volume via gross weight minus a constant tare, divided by an estimate of density. For example, nominal one-liter quantities of an oil may be packed in plastic containers of average weight 50 g. The tare is assumed constant, but in fact the standard deviation of tare weights is (we suppose) one gram. The gross weights (typically 850 g) are determined on an instrument yielding weighing error standard deviation of 0.5 g (well within 1/10th TN E= 1.5 g).

The overall error in weight determination is thus $s_w \tilde{=} (0.5^2 + 1.0^2) = 1.118$g, at a typical weight W of 850 - 50 = 800 g.

Now suppose that a density estimate D =0.85 g/ml is obtained, subject to error 0.001. The overall error in volume measurement is now

$s_v = 800/0.85 \sqrt{((1.118/800)^2 + (0.001/0.85)^2)} = 1.72$ ml.

With TNE= 15 ml, this again just fails the 1/10th TNE criterion. Note that, if the density determination were exact, the volumetric error standard deviation (with $s_d = 0$) would become 1.315 ml; within the 1/10th TNE criterion.

H.4 Action when test fails

H.4.1 National legislation may specify what actions are possible when a batch fails a reference test and these are likely to include:

- rectification of the batch tested,
- recall of previous prepackages produced under similar circumstances,
- modification of the system to prevent further similar occurrences.

14 Annex I Special Areas

I.1 Desiccating product

I.1.1 Product that, even though in prepackages, can diminish in quantity by evaporation, whether of the product or of an ingredient, are referred to as 'desiccating products'.

I.1.2 Article 2.2 of the Directive defines a prepackage as being product in a package which has a predetermined quantity and which 'cannot be altered without the package being either opened or undergoing a perceptible modification. As the quantity of a desiccating product does alter without the package being opened or undergoing a perceptible modification it could be argued that these products cannot comply with the Directive and should not be 'e'-marked. National legislation may specify how desiccating products should be controlled

I.1.3 As desiccation tends to alter the quality of goods packers will endeavour to ensure that the change is minimized to maintain the quality of the product. A practical solution is that at the time product is put on the EEA market, whether by the packer or importer, the product shall comply with all the requirements of the Directive. Thereafter the quantity in the prepackage shall not reduce to below TU2 at any time while in the distribution chain, this should be up until any 'best before' date or similar marked on the prepackage.

I.1.4 Furthermore the packer or importer should be able to substantiate that the product, as prepackaged, is a desiccating product and ideally the prepackaging should make it clear to the consumer that the product desiccates.

Note: When the Directive is updated the time at which the 'Packer's Rules' apply needs to be clarified.

Note: Directive 76/768/EEC, article 6 requires cosmetic products to be labeled with 'the nominal content at the time of packaging'.

Note: OIML[25] recommendations implies that goods only have to be correct when leaving the packer's/importer's possession.

I.2 Drained weight

I.2.1 The Directive pre-dates the requirement for drained weights on food products and so only applies to the total content of the prepackage. The drained weight of the product must comply with the requirements of directive 2000/13/EC once acceptable test method and tolerances have been agreed.

I.3 Aerosols

I.3.1 The original Directive on aerosol dispensers requires both weight and volume of contents to be declared. The contents can be easily determined and was verifiable. Directive 80/232/EEC gave a derogation to aerosols made up in prescribed quantities and in a prescribed capacity container to be labeled only with the volume of the contents (as well as other information such as the capacity of the container). Volume declaration may be justified on safety grounds in compliance with Directive 75/324/EEC and may also obviate deceptive packaging but makes verifying the contents very difficult an expensive and probably not within the 0.2 TNE measurement uncertainty required by the directive.

Note: OIML [26] recommend that a weight declaration be made, that of the total quantity of product and propellant expelled when following the packer's directions for use. This is a verifiable declaration.

[25] OIML recommendation 79 paragraph 2
[26] OIML recommendation 79 paragraph 5.3.2

I.4 Bread

I.4.1 Bread is a desiccating product and if a packer carries out checks while the loaf is hot allowance must be made for the weight loss during cooling prior to packaging. The Packer should justify the allowance made for the checks.

I.5 Rug & knitting yarn

I.5.1 These products' weights can vary with the humidity of the environment. There are standards, which specify methods of determining weight by drying the product (to a constant weight) and then adding on a re-gain factor to take into account the natural moisture that would be found in that fibre.

I.6 Growing media for plants (compressible)

I.6.1 EEAropean Standards EN 12579, Sampling and EN 12580, Quantity Determination specify how to sample and determine the volume of growing media and soil improvers. The volume is calculated via the density and weight and the Standard specifies the equipment that should be used for this purpose.

I.6.2 Any comments or problem found with the methodology should be referred to CEN/TC 223.

I.7 Density measurement

I.7.1 Checking volume declarations via weight and density can be undertaken where the product is homogeneous. The method of density determination will depend on the type of product. Various methods for determining density and the volume of products are mentioned in an OIML publication. Their appropriateness needs to be considered in the light of ensuring that the maximum error of measurement must not exceed 0.2 TNE when carrying out a reference test

I.7.2 Carbonated product (the density of which includes dissolved carbon dioxide) pose problems, although it may be possible to mark the level of the liquid on the package (bottle), open the package and empty the product, and calibrate the volume of the package (bottle) at the mark using water of known density. A variation on this can be used to determine the volume of inhomogeneous product such as yogurt containing pieces of fruit. The brim filled package can be used as a measure if a suitable 'strike' (flat glass disc) is used to minimize the error of filling to the brim.

I.8 Poultry

I.8.1 Poultry is sometimes graded into weight bands and the nominal quantity to apply determined after the weight of the poultry. Commission Regulation 1538/90/EEC specifies tests similar to the reference test in the Directive and also weight bands that should be used. The tolerable negative errors specified in these regulations are not the same as in the directive. The following table gives them:

Nominal weight g	Tolerable negative error, g	
	carcases	cuts
< less than 1,000	25	25
1,100 - < 2,400	50	50
2,400 and more	100	50

I.8.2 The Regulations lay down single sampling plans for use in reference testing for batches of 100 prepackages or greater and requires a 100% non-destructive test on batches of less than 100. Unlike the Directive the Regulations do not permit other equivalent reference tests.

I.8.2 Should the prepackaging also bear an 'e'-mark then the batch of prepackages must also comply with the tighter criteria specified in the Directive.

I.9 Deceptive packaging

I..9.1 If the quantity in a package (the nominal quantity) is reduced by the packer it is important that the consumer is not misled in any way. Recent research has shown that consumers pay more heed to the size of the packaging than the declared nominal quantity. Section 6 of the OIML recommendation 79 endeavours to ensure that consumers are not mislead, and that packers have fair competition, by setting out the following recommendations:

'*6.1 Fill level* Packages should be filled in such a manner that a purchaser may not reasonably be misled with respect to the quantity or identity of the product it contains, taking into consideration any recognized and accepted production practices that may be necessary for the manufacturer or packer.

6.2 Package design and display Packages shall be manufactured, constructed and displayed in such a manner that a purchaser may not be reasonably misled with respect to quantity or identity of product contained therein.

6.3 Labeling If the prepackaged product is labeled on more than one location of its packaging, the information on all labels shall be equivalent and in accordance with the requirements of the Recommendation.'

I.9.2 Following these recommendations should prevent consumers from being misled and will assist fair competition between traders. Relating to the filling, one way a packer can ensure no one is mislead is to change the packaging when the nominal quantity is varied and also ensure that the minimum amount of ullage (vacant space) is present. And also to ensure that false bottomed containers are not used

I.9.3 . Another consideration for packaging is that it must meet the essential requirements stated in directive 94/62/EC on Packaging and Packaging Waste which specifies that it shall :

a) be minimal subject to safety, hygiene and acceptance for the packed product and for the consumer,

b) not be noxious and hazardous substances in packaging must be minimized in emissions, ash or leachate from incineration or landfill,

c) be recoverable through at least one of :

- material recycling
- incineration with energy recovery
- composting or biodegradable

d) reusable

I.9.4 Where additional information to the nominal quantity is stated on the label this must not confuse the consumer. The quantity must be the most prominent with other information such as gross weight, weight including immediate wrappings and maximum weight (for safe handling) must be less pronounced and accompany the nominal quantity. It is not acceptable for instance for the gross weight to be displayed without the nominal () weight.

15 Annex J Importer's Control System and Records

J.1 Duties

J.1.1 As mentioned in 5 above, an Importer for the purposes of the Directive is the person who first imports the prepackages into the EEA.

J.1.2 A suggested method of checking imported prepackages involves using a double sampling plan, which is a means of keeping testing (and hence the opening of prepackages) to a minimum. A preliminary sample is taken and subjected to certain criteria, it may pass or fail outright but, if it falls between (is referred) a second sample must be taken which will decide the matter. Where the importer samples from the consignment as it is being put to stock, or by means of moving stock to obtain a representative sample, he would be well advised to take a large enough sample to cover both stages, should he find the first stage inconclusive .In other circumstances it may be just as easy to take the samples for the two stages as they are required. The sample sizes are as follows:-

Group or consignment size	Sample sizes			Test for non-standard			Criterion for average contents k factor
	1st	2nd	Total	Accept	Refer	Reject	
100 - 500	8			0	1	2	0.599
		12	20	1	-	2	0.64
501 - 3,200	12			0	1	2	0.43
		18	30	2	-	3	0.503
over 3,200	20			1	2	3	0.297
		28	48	3	-	4	0.387

J.1.3 The above sample sizes relate to the case where the importer has little or no experience of the particular commodity from that packer. Where the importer has the necessary experience and records his sample size need only be eight irrespective of the consignment size. If however that reduced sample does not pass the full sampling plan must be used.

J 1.4 Having selected the sample and tested for tare variability and determined the density where necessary, the weight or volume as appropriate of each prepackage should be ascertained and recorded and the average and standard deviation calculated. The batch is only accepted if both the non-standard and average test is passed. For the purposes of the average test the mean can be deficient by up to k.s (the k factor multiplied by the standard deviation) as this takes into account the variability possible between samples.

J.1.5 The records of the tests should be passed to the importer to decide whether the consignment is acceptable, and if not what action should be taken (sorting or return). This action should also be recorded.