WELMEC 6.5
Issue 2
February 2012

WELMEC

European cooperation in legal metrology

Guidance on Controls by Competent Departments on "℮" marked Prepackages

WELMEC

European cooperation in legal metrology

WELMEC is a co-operation between the legal metrology authorities of the Member States of the European Union and EFTA.

This document is one of a number published by WELMEC to provide guidance to packers, importers and competent departments responsible for conformity assessment of their products.

The documents are purely advisory and do not themselves impose any restrictions or additional technical requirements beyond those contained in relevant EU Directives.

Alternative approaches may be acceptable, but the guidance provided in this document represents the considered view of WELMEC as to the best practice to be followed.

Published by:
WELMEC Secretariat

E-mail: secretary@welmec.org
Website: www.welmec.org

Contents

1 Introduction 4
2 Terms and Definitions 6
3 Duties of Competent Departments 6
4 Duties of Packers 7
5 Duties of Importers 8
Annex A Labelling of Prepackages 10
Annex B Measuring Equipment for the Packer and Importer 16
Annex C Records 17
Annex D Statistical Principles relating to Control Systems and Reference Tests 20
Annex E Quantity Control by Sampling 36
Annex F Checkweighers and other automatic instruments 61
Annex G Control using Measuring Container Bottles and Templates 66
Annex H Reference Tests 71
Annex I Special Areas 72
Annex J Importer's Control System and Records 74

1 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 This document is part of a series of documents published by WELMEC and which are primarily intended to provide guidance to all those concerned with the application of Directives 76/211/EEC and 2007/45/EC for prepacked products. They are intended to lead to a uniform interpretation and enforcement of these directives and assist in the removal of barriers to trade.

Those that have been agreed by WELMEC are published on their website at http://www.welmec.org/latest/guides.html

6.0 Introduction to WELMEC documents on prepackages
6.1 Definitions of terms
6.3 Guidance for the Harmonised Implementation of Council Directive 76/211/EEC as amended
6.4 Guide for packers and importers of '℮' marked prepacked products
6.5 Guidance on Controls by Competent Departments
6.6 Guide for recognition of procedures
6.7 Guide for Market Controls on Prepackages for Competent Departments
6.8 Guidance for the Verification of Drained Weight
6.9 Prepackages - Uncertainty of Measurement
6.10 Information on Controls on Prepacked Products
6.11 Guidance for Prepackages whose Quantity Changes after Packing

1.3 This document looks particularly at the requirements for 'e' marked prepackages specified in Directive 76/211/EEC (the Directive). This has been amended by Directive 2007/45/EC of 5 September 2007, and this latter Directive revoked Directives 75/106/EEC and 80/232/EEC. Where relevant requirements in other Directives and Regulations are also considered. The guidance in WELMEC document 6.3 is reflected in this document.

1.4 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'. The OIML recommendations are at present undergoing a revision and so the references are appropriate for the current editions of these documents[1].

[1] R79 edition 1997 and R87 edition 2004.

1.5 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.

1.6 The Directive requires the prepackages from 5 g to 10 kg or 5 ml to 10 L inclusive:

- meet the content requirements, sometimes referred to as the 3 Packer's Rules,
- meet the labelling 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.

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

a) the average quantity of product in prepackages shall not be less than the stated (nominal) quantity,
b) there shall be no more than 2.5%[2] of prepackages with a quantity of product below TU1, and
c) there shall be no prepackages with a quantity of product below TU2.

Where :

- TU1 is the nominal quantity less one tolerable negative error (TNE) and
- TU1is the nominal quantity less two TNE.

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

Nominal quantity (Q_n) in grams or millilitres	Tolerable negative error	
	As % of Q_n	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	

1.9 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 millilitre.

[2] Regarding the second rule, the Directive specifies an acceptable number of prepackages below TU1 for each reference test sample size. The proportion of prepackages below TU1 needs to be sufficiently small, in general it appears that not more than 2.5% below TU1 is appropriate.

2 Terms and Definitions

Throughout this document, and other WELMEC series 6 documents the terms used, and their definitions, are those stated in WELMEC document 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].

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 labelling 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 of product in prepackages.

3.3 Checks on 'e' marked prepackages, 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 risk using:

- 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.

3.4 Checks 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 Appendix 3 of WELMEC document 6.3.

3.5 The Directive does not preclude any other checks[4], 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.

3.6 The Directive covers prepackages labelled with a weight or volume quantity declaration from 5 g or 5 ml to 10 kg or 10 L inclusive. Domestic legislation may

[3] Directive 76/211/EEC, Annex I.5
[4] Directive 76/211/EEC, Annex I.6

cover prepackages outside these limits, or products sold by reference to length, area and number.

3.7 When the quantity of product contained in a prepackage is not measured by the packer, the Competent Department recognizes the packers procedures in the way specified in national legislation. This may also result in an approval to mark prepackages with the '℮' mark. For the methods of recognition used in Member States refer to WELMEC 6.0.

3.8 Where there have been changes in the quality control system these 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 (quantity of) product or its labelling 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 usually fully labelled and has passed all the checks in the packing process.

4.3 The packer's duties[5] are :

a) to ensure the prepackages meet the labelling and quantity requirements of the Directive,
b) either to measure the quantity of product into each package or to check the actual quantity of product,
c) where the actual contents are not measured, to carry out checks so that the quantity of goods is effectively assured.
d) to use suitable and legal measuring equipment for these purposes.
e) to keep at the disposal of the competent department documents containing the results of checks, together with any corrections and adjustments which they have shown to be necessary

4.4 Essentially the duties in 4.3.b) above permit the packer to have one of two systems. The first is to measure the quantity of product in 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 quantity of product in the prepackages. 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

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

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 contents is effectively ensured". As he is also responsible for ensuring that the prepackages meet the requirements of 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.

4.8 For the packer's procedures to give the necessary assurance they shall:
a) cover the setting up, monitoring and review of the quantity control system,
b) justify the target quantity and control limits for each product,
c) contain procedures to be followed when limits are breached,
d) 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 examples 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 of the importer at the place of first entry into the EEA,
d) 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 should state that the procedures of the packer in the exporting country have been assessed and that the controls and records guarantee compliance with the requirements of the Directive. In all cases, the evidence needs to specify the prepackages, the nominal quantity and labelling that has been assessed. The evidence should include an identification (code) used for traceability.

5.5 Checks appropriate for 5.3.c) and 5.3.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.

Annex A Labelling of Prepackages

A.1 Indication of Quantity

A.1.1 Prepackages shall be marked with the nominal quantity and an '℮' mark to show compliance with the requirements of the Directive.

A.1.2 The nominal 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[6]. Community and national legislation may require indications of volume or mass for specified categories of products or types of prepackages.

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[7].

Subsequent to the Directive, Regulation (EC) 764/2008[8] has come into force which acknowledges that products lawfully marketed in one Member State should generally be allowed to be marketed elsewhere. Where a competent department wishes to apply a national technical regulation the procedure in Chapter II of the Regulation has to be followed.

An example of when this procedure needs to be followed is as where a packer marks the weight and an 'e'-mark on packaging, to comply with domestic requirements. The ice-cream is sent to another Member State where the domestic legislation requires ice-cream to be marked with the quantity by volume. If the latter country insisted that the package needed to be marked with the volume of the ice cream then the Competent |Department would have to follow the procedure in Chapter II of Regulation (EC) 764/2008.

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 solids and liquid, or the solid part of a mixture of a solid and liquid[9].

A.1.3 Prepackaged foodstuffs that are normally sold by number are not required to be marked with an indication of volume or mass[10]. The '℮' mark does not apply to such prepackages.

[6] Legal units within the meaning of Directive 80/181/EEC.
[7] Directive 76/211/EEC, Article 4.3
[8] Regulation (EC) No 764/2008 of the European Parliament and of the Council of 9 July 2008 laying down procedures relating to the application of certain national technical rules to products lawfully marketed in another Member State
[9] OIML R79 (1997) *Labelling requirements for prepackaged products* Article 5.3.1
[10] Directive 2000/13/EC, Article 8.3

A.1.4 Prepackaged solid foodstuffs presented in a liquid medium shall be marked with an indication of drained weight in addition to an indication of weight[11]. At present the '℮' mark applies to the total weight of the foodstuff and liquid medium and not the drained weight.

A.1.5 Products that are prepackaged in aerosol form shall be marked with indications of volume and total capacity of the container[12]. The directive on aerosol dispensers[13] requires them to be marked with both the weight and volume of the product (which includes the quantity of propellant), but Directive 2007/45/EC gives a derogation from the requirement to mark a quantity by weight.[14]

Note : For aerosols, OIML[15] states: 'In the case of a product packed in a container designed to deliver the product under pressure, the statement shall declare the net quantity in mass that will be expelled when the instructions for use are followed. The propellant is included in the net quantity statement. Statements of quantity shall be the kilogram, gram or milligram.'

A.1.6 Certain prepackages of defined wines and spirits have to be made up in specified quantities. Where two or more individual prepackages of specified wines or spirits are made up in a multipack, each individual prepackage must contain one of the specified quantities listed Where a prepackage is made up of two or more items NOT intended to be sold individually then the total quantity of the prepackage must be one of the specified quantites[16].

Until the time limits specified[17], national requirements for prepackaging milk, butter, dried pasts, coffee and sugar in specified quantities must be met.

A.1.7 Community and national legislation may permit indications other than the nominal quantity[18] or for indications to be given for quantities above 10 kg or 10 L. The '℮' mark does not apply to such indications.

[11] Directive 2000/13/EC, Article 8.4
[12] Directive 2007/45/EC article 4.1
[13] Directive 75/324/EEC. Article 8.1.e
[14] Directive 2007/45/EC article 4.2
[15] OIML R79 (1997) paragraph 5.3.2
[16] Directive 2007/45/EC article 5 and Annex.
[17] Directive 2007/45/EC article 2, 11 October 2013 for sugar, 11 October 2012 for other products.
[18] For example, pre-packed fertilizers may be marked by weight or by gross weight (Directive 76/116/EEC).

A.2 Manner of Marking and Presentation

A.2.1 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:-

a) having information hidden in the fold of the packaging,
b) having markings on a clear container in a similar colour as the product contained in it.
c) having markings only on the rear of a container.

A.2.2 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:

a) the name under which the product is sold ;
b) 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
c) with respect to beverages having an alcohol strength of more than 1.2% vol., the actual alcoholic strength by volume[19].

A.2.3 'Easily legible' must take into account the size, font, orientation, 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.2.4 OIML[20] 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.

[19] Directive 2000/13/EC, clause 8
[20] OIML R 79 (1997) paragraph 5.5.2

A.3 Nominal Quantity

A.3.1 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[21].

The height of the figures in the nominal quantity 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 moulded 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.3.2 The unit of measurement shall either be written in full or abbreviated. The only permitted abbreviations are:

	For foodstuffs	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.3.3 Where a prepackage is marked with more than one indication of quantity[22], all indications shall be in close proximity to each other, shall not be more

[21] Directive 2000/13/EC, clause 8
[22] For example, an indication of weight and volume, weight and drained weight, net weight and gross weight, volume

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. The '℮' mark relates to the nominal weight or volume.

Note : OIML recommends that if the prepackaged product is labelled on more than one location of its packaging, the information on all labels shall be equivalent[23].

A.3.4 For prepackaged foodstuffs, the labelling and methods used must not be such as could mislead the purchaser to a material degree as to the quantity of foodstuffs[24].

A.3.5 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.

A.3.6 Where a prepackage is legally marketed in another member state and the competent department wishes to apply their own national technical regulation the procedure in Chapter II of Regulation (EC) 764/2008[25] has to be followed.

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.

[23] OIML R79 (1997), Article 6.3

[24] Directive 2000/13/EC, Article 2

[25] Regulation (EC) No 764/2008 of the European Parliament and of the Council of 9 July 2008 laying down procedures relating to the application of certain national technical rules to products lawfully marketed in another Member State

A.4 Prepackages bearing the "℮" mark

A.4.1 The '℮' 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.

A.4.2 Where the "℮" 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 3 mm in height.

A.4.3 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.

A.4.4 Where more than one statement is given on a multi-pack, for example '4x10 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 and it would be clearer to mark the packaging with ℮ 4x10g 40 g. Where the individual items are not appropriate for selling singularly it would be the 40 g and the marking would be clearer as ℮ 40 g 4x10 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.

A.5 Identity Mark or Inscription

A.5.1 Prepackages shall bear a mark or inscription enabling the competent departments to identify the packer or the person arranging for the packing to be done or the importer established in the Community.

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

- Markings shall be 'easily legible and visible on the prepackage in normal conditions of presentation'.
- 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..."[26].
- Other legislation may require the Country of Origin to be stated on the label.

A.5.2 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.

[26] OIML R79 (1997) paragraph 4

Annex B Measuring Equipment for the Packer and 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 The word "suitable" includes a number of conditions of use that arise from the need to limit the uncertainty of measurement. When the total uncertainty of measurement does not exceed 1/5 TNE, a packer/importer usually meets the ‘suitability’ criterion.’

B.3 All equipment should 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 should 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.

Annex C Records

C.1 Production Records

C.1.1 The Directive requires that checks be so organized to effectively ensure the quantity of the product in prepackages. 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.

C.1.2 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 Record retention

C.2.1 All records need to be retained for as long as prepackages remain in the distribution chain, with a minimum of 1 year.

C.2.2 It is up to nation legislation if and how often verification and re-verification of measuring instruments is carried out. Records of such verification and re-verification of measuring instruments shall be kept by the packer for at least two verification periods.

C.2.3 Records of the checks that packers carry out on their measuring instruments, as mentioned in WELMEC 6.6 paragraph 4.4[27] should be kept sufficiently long to show that they have met their specifications.

C.3 Validation of records

C.3.1 The Competent Department – or equivalent status in the Member State - 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 quantity of product of prepackages and comparing the average 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.

C.3.2 The records must include any corrections or adjustments made to the process and also show that they have been properly and accurately carried out.

C.3.3 The following records should be maintained, as appropriate to the process involved.

C.4 <u>Identification and specification</u>

- procedures and work instructions documents involved
- identity of people and their responsibilities

C.5 <u>Batch data</u>

- product identity
- batch identity
- batch size

[27] WELMEC 6.6 paragraph 4.4: 'Measuring instruments must be checked on a regular basis by an accepted method to verify that they meet specifications.'

- density and its variability (if applicable)
- nominal quantity
- tare and its variability (if applicable)
- other allowances

C.6 Process characteristics

- measuring instruments involved for traceability of the control system
  - equipment maintenance
  - calibration records
  - if applicable, data about official verification for in-service instruments
  - checks on equipments to ensure no drift (e.g for checkweighers, checks on set points)
- parameters related to the performance of the measuring instruments involved for traceability of the control system, e.g. :
  - for checkweighers, the zone of indecision or accuracy class, the mean error and standard deviation
  - for automatic gravimetric filling instruments, the dispersion or accuracy class, standard deviation
- parameters related to the settings of the measuring instruments involved for traceability of the control system
  - for checkweighers, target value, set points
  - for automatic gravimetric filling instruments, target value, control limits, number and peroidicity of control cycles
- Parameters for the filling process
  - Target value
  - Standard deviation
  - Action limits
  - Warning limits
  - Allowances
  - Sample size and sample frequency, when relevant (for filled prepackages, density and packing material)
  - Storage of measuring results and calculations
- Calculations

C.7 Production checks

- identification of checkpoint / packing line
- reference to product identity
- batch identity
- time and date of sampling of prepackages, density and packing material, when relevant
- number of prepackages in a sample, of prepackages, density and packing material, when relevant
- average and variance of actual quantity of product (sample data)
- average and variance of actual quantity of product (batch data)
- ratio and/or number of prepackages with a quantity of product less than TU1 and/or TU2
- corrective actions when signalled by warning limits and action limits relating to average quantity of product, TU1 and/or TU2.
- for checkweighers - checks on set points, to ensure no drift
- test of good functioning of rejection mechanism

C.8 Software used for processing of quantity control data

Although software was not involved in directive 76/211/EEC, its impact on the control system has to be taken into account to ensure right decisions about conformity of batches. There should be records to show:

- validation prior to use and after every change, records to be kept,
- issue status (e.g. date or version number) identified with the records produced from it,
- means or tools to protect against corruption and unauthorized modification.

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, that is

- Not less than TU1 (adequate)
- Less than TU1 (non-standard)
- Less than TU2 (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 U 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(0) = 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 $\pm 1.96\sigma$),
- 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 D.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 g and σ = 5 g. The value of the variable of interest is the weight corresponding to TU1 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

$$P(r) = e^{-m} \times m^{r} / r! \text{ with } m^{0}, 0! \text{ and } 1! \text{ all defined as } 1.$$

(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 $P(0) = 0.975^{30}$ (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 .5.6 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.5.7 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 non-standard 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(0)	P(1)	P(2)	P(≥ 3)
Hypergeometric	N=200, n=30, p=0.025	0.43974	0.39736	0.13800	0.02490
BinomiaI	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 $\bar{x}$ 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 $\bar{x}$ 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 $\bar{x}$ obtained from a sample of n observations on a population whose standard deviation is σ and invoking the Central Limit Theorem, it is now possible to construct a *confidence interval for* the true (but unknown) population mean μ,

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

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

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

$(\bar{x} - \mu)/ \sqrt{n} / \sigma$, but $(\bar{x} - \mu) \sqrt{n} / s$

where s has a sampling distribution known as Student ‘t’ distribution, (with n-1 degrees of freedom).

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

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,	$\bar{x}$ = 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 being 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

$H_0 : \mu \geq Q_n$ where Q_n is the nominal quantity.

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

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

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

$H_1 : \mu < Q_n$

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 :

$H_0 : p < 0.025$,

where p is the batch proportion of prepackages containing less than TU1. 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:

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

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:

$\bar{x} > Q_n - 0.640s$ Accept

and $\bar{x} < Q_n - 0.640s$ Reject for n = 20.

Here, 0.640 is in fact 2.861 / √20, or t / √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.

D.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)

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

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

Q_n = nominal quantity of the pre-prepackage;

s = standard deviation calculated on the sample,

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

where

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

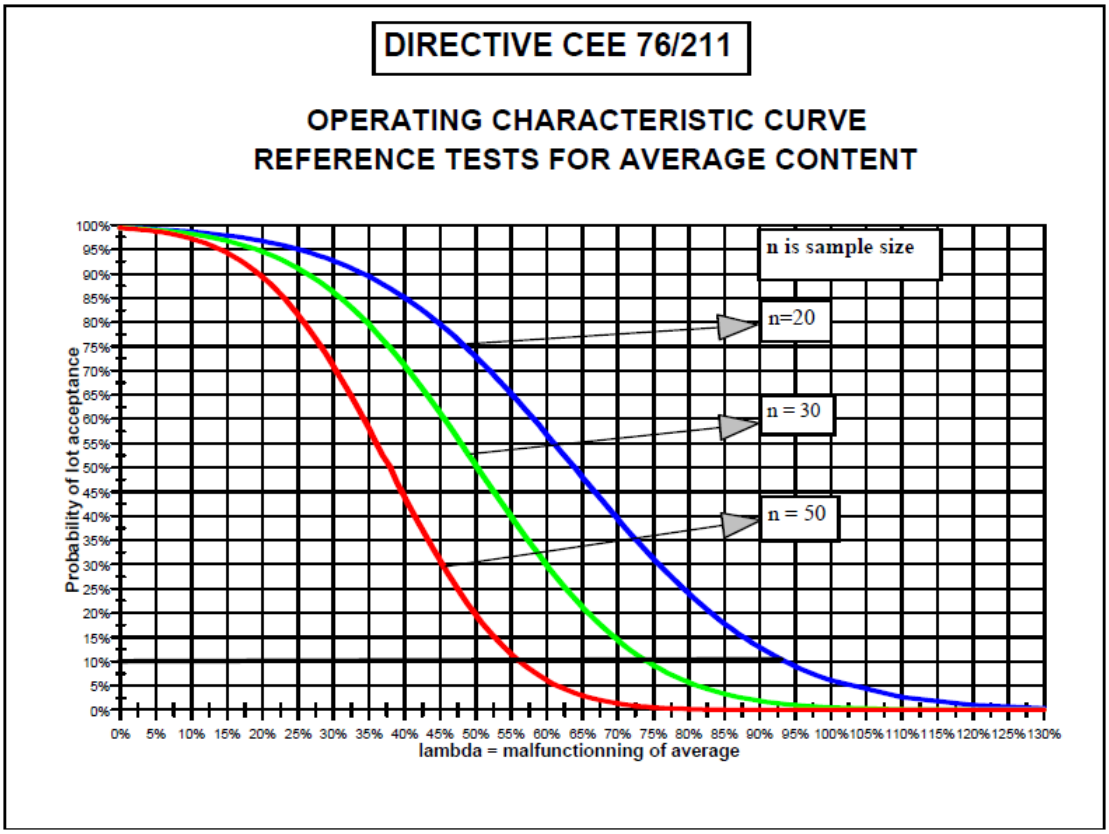
$$P_A = F\left[t_{1-\alpha} - \left(\lambda.\sqrt{n}\right)\right]$$

F : Distribution function of a Student random variable,

(1-α): Level of confidence of the test = 0,995 according the directive

α−1*t:* Fractile of order (1-α) of a Student random variable with (n-1) degrees of freedom.

The OC curve is given by the following graph:



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 =

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

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

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

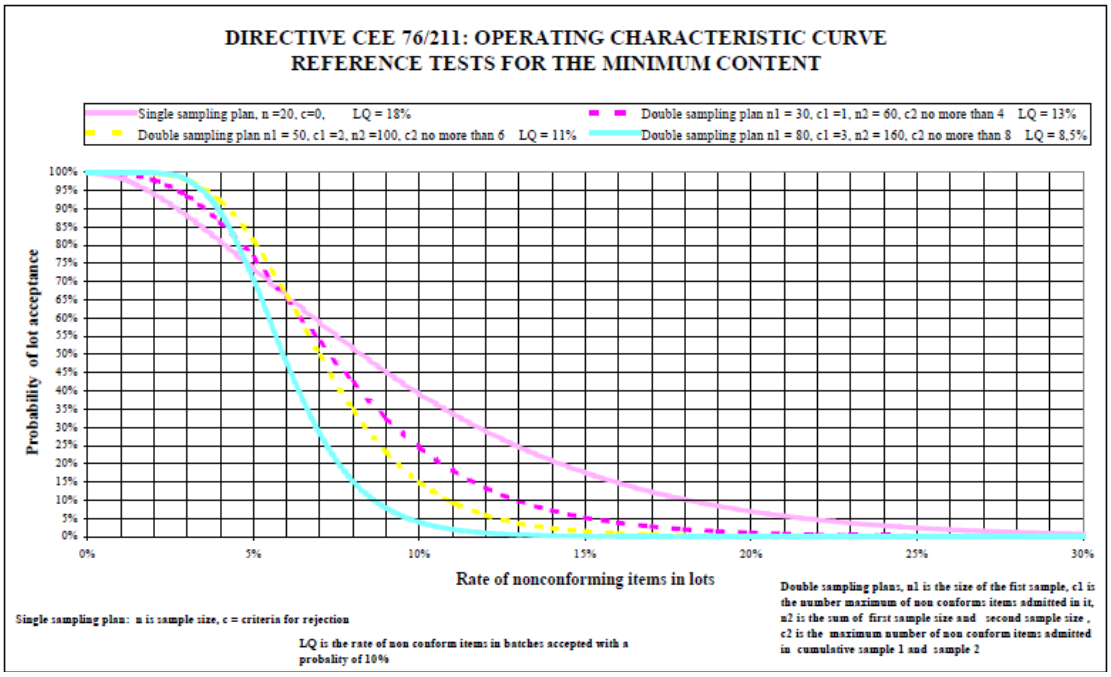
c2 = maximum number of non conform items admitted in the cumulative first and second sample;

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 ≤ r ≤ c2.



D.9 Quantity control DIRECTIVE CEE 76/211: OPERATING CHARACTERISTIC

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 $\Sigma (x - T)$ 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 $Y = a + bx_1 + cx_2 + dx_3 \ldots$ 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 $\sigma_Y^2 = b^2\sigma_1^2 + c^2\sigma_2^2 + d^2\sigma_3^2 \ldots$ 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.

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. In a process capability check, two parameters are calculated, the capability index, C_p and the performance index, P_p. If these are not the same, this reflects unwanted presence of between-sample standard deviation.

C_p = (Upper Specification Limit – Lower Specification Limit)/$6\sigma_0$
P_p = (Upper Specification Limit – Lower Specification Limit)/$6\sigma_m$
where
σ_0 = standard deviation for single values, calculated from within-sample standard deviation
σ_m = medium term standard deviation for single values calculated as the global standard deviation for one production period

E.1.5 For the calculation of C_p and P_p we need to know the specification limits. The upper specification limit is set by the packer. Only TU2 is a lower specification limit for all items, whereas TU1 and Q_N are not specification limits, but critical bounds on the distribution curve of prepackages. But for the purpose of inspection of between-sample standard deviation, both specification limits can be set arbitrarily, and only the ratio $C_p/P_p = s_m / s_0$ is calculated.

E.1.6 If between-sample standard deviation is close to zero, this ratio is 1. In most cases it will be more than 1, because the performance is usually somewhat less than the capability. The capability of a process is a measure of the potential of the process when it is working at its best. To see if the between-sample variation is significant calculate σ_m/σ_0 and if it exceeds the critical values in The E.0 the variation is significant at the 1 in 40 level,

Table E.0 Critical values[28] (1 in 40 significance) for σ_m/σ_0

No of samples, k	No of items per sample, n									
	2	3	4	5	6	8	10	12	15	20
20	-	-	-	1.083	1.067	1.048	1.038	1.031	1.024	1.0181
25	-	-	1.098	1.075	1.061	1.044	1.035	1.028	1.022	1.0164
30	-	-	1.087	1.066	1.053	1.039	1.030	1.025	1.020	1.0145
35	-	1.115	10.79	10.60	1.048	1.035	1.028	1.023	1.0179	1.0133
40	-	1.107	1.073	1.056	1.045	1.033	1.026	1.021	1.0167	1.0124
50	1.172	1.093	1.065	1.049	1.040	1.029	1.023	1.0187	1.0147	1.0109
60	1.154	1.084	1.059	1.045	1.037	1.027	1.021	1.0174	1.0138	1.0102
70	1.140	1.077	1.053	1.041	1.033	1.024	1.0190	1.0156	1.0124	1.0092
80	1.129	1.071	1.050	1.038	1.031	1.023	1.0178	1.0147	1.0116	1.0086
100	1.114	1.064	1.044	1.034	1.028	1.020	1.0161	1.0133	1.0105	1.0078

E.1.7 It should be noted that if the between-sample standard deviation is substantial, the behaviour charts will show many signals of unpredictable behaviour. When

[28] UK Department of Trade Manual of Practical Guidance for Inspectors, Issue 1 1979

this behaviour happens the standard procedure is to halt production, find the assignable cause of variation and correct it. Assignable causes are special large sources of variation, which make the process jump or vary unpredictably / uncontrolled. When all assignable causes are eliminated, only common causes are present, small random or chance causes, which make the process vary close to a mean value, or at least within limits, in a predictable / controlled manner. In this situation the packing process is predictable, and samples drawn from the batch are representative for the entire production batch.

E.1.8 When a competent department performs a reference test, by drawing random samples from a production line for prepackages in order to accept or reject the entire batch; it is important that all samples are from a predictable and stable process. Otherwise, the samples are not necessarily representative for all prepackages, which have been produced during the packing period. For the same reasons, when the packer uses sampling in a procedure for controlling his packing process, the packing process must be predictable and stable, and the packer should be able to demonstrate this. Keeping all records and charts for his production batches is an effective way to demonstrate to inspectors that the packing process is and has been predictable and that filling level is and has been according to the requirements.

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-labelling 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 9.91 % 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 $Q_n - 0.364\sigma$, 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 TU1 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 TU1 + 2σ, 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 TU1+σ (1.96-0.364) and TU1+ σ(l.96-0.182) i.e. TU1+ 1.596σ and TU1+ 1.778σ.

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:

- The target quantity, Qt, may not be less than the nominal quantity declared, Qn.
- Not more than 1 in 40 (2.5%)[29] prepackages may contain less than the value of TU1 appropriate to the nominal quantity.
- The packer must not pack quantities below the value of TU2, appropriate to the nominal quantity, and system design must ensure that no more than 1 in 10,000 prepackages violate TU2 by chance.

E.2.5 Where the assumption of a normal distribution of prepackaged quantities is reasonable, Rules 2 and 3 become:

2(a) $Q_t > T1+2\sigma$

3(a) $Q_t > T2+ 3.72\sigma$

(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, TU1+2σ, TU2+3.72σ, 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, $Q_t > TU1 + 2\sigma$.
- For σ exceeding TNE/1.72, $Q_t > TU2+3.72\sigma$.

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 TU1+2σ 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. TU2 + 3.72σ = TU1 + 2σ),

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 TU1 + 2σ or to TU2 + 3.72σ, 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, - 0.2 σ (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.

[29] Regarding the second rule, the Directive specifies an acceptable number of prepackages below TU1 for each of reference test sample size. The proportion of prepackages below TU1 needs to be sufficiently small, in general it appears that not more than 2.5% below TU1 is appropriate.

b) If the process average falls to Q, - 0.4σ (corresponding to 65 %. of prepackaged quantities below Q_n, or to 5.5%. below TU1, or to about 1 in 2,000 below TU2), 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 $1{,}250 \leq P \leq 10{,}000$.

E.3.3 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

$Q_t \geq TU1 + \sigma (u / \sqrt{N} - 1.6)$ where $\sigma > 0.5$ TNE but $\sigma <$ TNE / 1.72

$Q_t \geq TU2 + \sigma (u / \sqrt{N} + 3.3)$ 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$/1.72$, $Qt = TU1 + (2 + Z)\sigma$.

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

(ii) To find N, given Qt

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

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.

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

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

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	E Cusum h= 5, f=0.5
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 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 uncertainty is the most widely applicable. Sources of uncertainty are treated in a similar manner whether it is caused by imprecision of measurement,, variation of tare,, variation of the density or measurement uncertainty in the determination of the density of a liquid.
Allowance for measurement uncertainty is made at the packers risk and cost.

E.5.2 *Additional Comments on Tare*

Requirements on prepackages are set on net content. Net content is calculated from gross weight minus tare weight. The variation of tare is a source to measurement uncertainty in the determination of mean value of tare and consequently also to the measurement uncertainty of the net content.
The variation of the tare is dependent on the tare material and it is common known that a paper, plastic or foil tare has a very small variation and that glass and metal normally has a large variation which normally is measured and expressed as a standard deviation found from a sample.
But it is not only the standard deviation measured on short or long term that could form a problem. The average weight can change from one pallet or batch of tare to another and if the average tare weight increases without adjustment of the target (for example by changing tare weight in the weighing instrument), the sampling checks will indicate that filling level can be reduced. This can cause under filling.The measurement uncertainty of the tare can be minimized by updating the tare value more often. If a fixed value is used for the tare value, a long term variation of the tare weight will give larger measurement uncertainty.

E.5.3 Sampling of empty tare for the calculation of tare mean value should be representative for the tare in use. Besides other acceptable methods that a

packer can use, the following is considered as best practice. If the mean tare weight is updated for every production period, the tare usually has a very small contribution to the measurement uncertainty. If however a fixed value is used for tare mean value, the long term variation of tare must be recorded and continually updated. Also, for each production period, tare values must be checked to be within the long term variation. How much long term variation is acceptable depends on the allowance that is given for tare long term variation. It is recommended to update the tare value for each production period, or at least for each production batch of tare. An alternative approach is to make the producer of the tare supply every batch of tare with a certificate containing information about mean value and standard deviation.
The complete workload of recording the long term variation of tare and to demonstrate that each production period has a tare value that is within the allowance applied for tare variation, is not smaller than updating the tare mean value. It only leads to larger overfill or violation of the filling rules. Especially the first Packers Rule that requires *The actual content shall not be less on average than the nominal quantity* is vulnerable to the target setting procedure, including tare determination.

E.5.4 *Additional Comments on the accuracy of Instruments*

A guideline on good practice for choosing the accuracy of the weighing instrument is given in table 3 in Welmec Guide 6.4, which uses the premise that the Maximum Permissible Error (MPE initial) shall be less than TNE/10.

This premise is based upon the nominal content of the prepackages, and does not take into account the packing process itself. If a packer who is using process behaviour charts for controlling a packing process with very small variation, the packer could find that a choice in line with table 3 in WG 6.4, gives too large resolution in order to control the process within control limits.

If this is the case there are two possible solutions - chose another instrument with a smaller verification scale interval or look for an instrument with better resolution but with the same verification interval (many scales can be delivered with a division (d) of 1/10 e). It could be advantageous to choose an instrument with a division equal to one standard deviation of the process.

Two arguments are in favour of choosing a sufficiently high resolution:

1) If the resolution is poor compared to the natural dispersion of the proves, the control limits may be too small, giving many signals of unpredictable behaviour even though it is in fact a predictable process.
2) Alignment of the centre line of the process to the target value is much easier and faster with better resolution, because the average value can be estimated more correctly without large rounding errors (large compared to the process variation).

The directive does not require the use of a verified instrument for measurement of the tare but if a verified instrument is used remember that it has a minimum capacity which is the minimum load for which it is verified and should not be used for measurements below this. If a non verified instrument is used then it should be calibrated or checked regularly during use according to national legislation.

The tare weight is usually much smaller than the nominal weight of the product and the tare variation is usually much smaller than the variation of prepackages. For the determination of tare weight and its variation, the packer might consider using a weighing instrument with better accuracy and resolution, or he could weigh n samples of tare at the same time, effectively extending the resolution of the weighing instrument by a factor of n. The dispersion of the individual tare is estimated by: $s_{tare}(x_i) = \sqrt{n} \cdot s_{tare}(\bar{x})$

E.5.5 Prepackages subject to loss by evaporation of weight or volume, refer to I.1 for guidance.

TABLE E.3:
Master list of sampling factors, Z[30]

n = No of items per sample set	Procedure	k = number of sample sets per period*											
		1	2	3	4	5	6	8	10	12	16	20	25
2	A	—	0·84	0·70	0·61	0·54	0·47	0·35	0·27	0·21	0·13	0·07	0
	D	—	0·58	0·43	0·35	0·29	0·25	0·19	0·15	0·12	0·07	0·03	0
	E	—	0·37	0·25	0·19	0·15	0·12	0·08	0·05	0·03	0	0	0
3	A	—	0·65	0·53	0·46	0·37	0·31	0·21	0·15	0·10	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	0									
	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:
*A production period is the time taken to produce 10,000 prepackages, with a minimum of 1 hour and a maximum of one shift/day.
Multiply s_o 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 in E.5.5
$Q_1 = Q_a + Z\sigma$ (for $\sigma \leq$ TNE/2)
or T1 + (2 + Z)σ (for $\sigma >$ TNE/2 but $\sigma \leq$ TNE/1·72
or T2 + (3·72 + Z)σ (for $\sigma >$ TNE/1·72
Procedure A = Shewhart Control with Action Limit 1 in 1000.
Procedure D = Shewhart Control with Warning Limit 1 in 40 and Action Limit 1 in 1000.
Procedure E = Cusum Control with h = 5, f = 0·5 (as per BS:5703).

[30] Taken from Table 9.3 of the UK Department of Trade "Manual of Practical Guidance for Inspectors" Issue 1, 1980, ISBN 0 11 512501 9

E.5.6 The reference test carried out at the end of the production line will be based on samples from a one-hour period. This means that if the process 'wanders' slowly during the packing period, the packer might have succeeded fulfilling all rules for his entire production period, but the result of the reference test might still be rejection if the one-hour period for the reference test coincides with a period with a lower mean value. If the packer wants to guard against this kind of problem, there are three parameters to play with, sampling frequency (k), number of items in the sample (n) and additional overfill. Sampling frequency (k) should be high enough to detect a significant shift in the mean value for every one-hour period as a minimum, every half-hour if the packers also wants to overfill on the rest of the one-hour period to restore the mean value for the complete one-hour period. Increasing the number of items (n) in each sample is a means of increasing the sensitivity of the detection of a shift in the mean value of the process.

E.5.7 A pragmatic approach is to take into account that predictable processes can have a slowly shift of the mean value of up to 1.5σ and so overfill by up to 1.5σ if sampling frequency is low compared to a one-hour period, and to reduce this allowance if sampling frequency is higher.

Another approach is to give allowance to compensate for low sampling frequency (k) and low sensitivity (n). Table E.3 is used to calculate an allowance for different choices of k and n, but the packer should bear in mind that this is based on the complete production period and does not take into account that the reference test is carried out on a one-hour basis.

Welmec Guide 6.6, item 4.5.2, gives an alternative formula to calculate the minimum number of samples, N = kn, from the actual overfill:

$$N \geq \left(\frac{t_{n-1,0.995}\, s}{Q_n + Overfill - Rejection\ limit} \right)^2$$

This formula can also be used the other way round, to calculate an adequate overfill from the number of samples drawn from the process, N = kn.

$$Overfill \geq t_{n-1,0.995} s \left(\frac{1}{\sqrt{N}} - \frac{1}{\sqrt{n_{cd}}} \right)$$

n_{cd} = number of items sampled in a reference test carried out by the competent department.
We see that if $N \geq n_{cd}$, no overfill is needed. But if $N < n_{cd}$, an overfill proportional to the standard deviation is needed because of low sampling.

E.5.8 A reference test carried out on stock will be done on 10 000 items no matter of the production rate. If the reference test accepts or rejects the reduced lot, the complete batch is accepted or rejected. If the 10 000 items in the reference test represents a period shorter than one hour, detection of a possible shift of the mean value becomes time critical. A general guideline is to have at least 3 samples during a one-hour period.

E.5.9 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.5.10 Summing up calculation of total allowance

$$Total\ allowance = a_1 + \sqrt{a_2^2 + a_3^2}$$

Allowance a_1 is a linear shift of the entire distribution so that the critical points on the distribution are acceptable for all three filling rules, Qn, TU1 and TU2.

Allowance a_2 is the measurement uncertainty. Combination of all independent uncertainty sources is carried out by sum of variances.

Allowance a_3 is "an uncertainty" due to low sampling, a slow drift of the mean value is not detected until one of the detection limits are violated, see also average run lengths. Allowance a_3 is established by three alternative procedures:
a) Master table E.3 (from the choice of sampling regime with detection rules or Cusum, □ and N = kn, decides which fill factor z to use (allowance$_3$ = z□).
b) Overfill formula from E.5.7 connecting overfill and number of samples, N, when taking into account the rejection limit and the number of items sampled in the reference test carried out by the competent department.
c) A pragmatic approach is to overfill with up to 1.5 □ to guard against what might happen between samples.

In a_3, also 'aesthetic' allowances can be included.

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. Other detection rules are 4 out of 5 values in a sequence outside 1σ limit, or 8 values in a sequence on one side of the target line. Also other detection rules may be used, but simultaneous use of many detection rules will increase the number of false alarms.

(vi) Centring of the process to the target.

The control and warning limits are drawn in the control charts as parallel lines to the target line. The centre line of the process is the grand average of all measurement values, and this line needs to be aligned with the target value by adjustment of the filling level. Such alignment procedure might include larger number of items in each sample (n) and/or higher sampling frequency (k) and the simultaneous use of all four detection rules mentioned above.

(vii) Monitoring the process by process behaviour charts

After having centred the process, monitoring the process by process behaviour charts will show when the process is producing predictably and when there is a reason to make corrections, using only two detection rules at 3σ control limits and 2σ warning limits.

E.7.2 Two detection rules concerning the action limits and the warning limits provide indications (signals) when the behaviour of the process is not predictable / out of control and there is a reason to act on the process to bring it back to the state

of predictable behaviour. When no signals show up in the charts, there is only natural variation present in the process, and the process should be continued. Both detection rules are balanced to give a reasonable trade-off between the risk of producing false alarms (type 2 error) and having no signal when there actually is a reason to act (type 1 error).

E.7.3 *Control limits for sample means or for individual values*
Control charts for subgroup means, $\bar{x}$, or charts for individual values, x, monitor the location of the process compared to a target value.
There are several ways of calculating 3σ control limits for the sample means charts, ($\bar{x}$ chart). The most convenient way is to use standard tables, see for example tables in “Understanding Statistical Process Control”, 3rd edition by Donald J. Wheeler and David S. Chambers or “Introduction to Statistical Quality Control” issue 5e by Douglas C Montgomery.

E.7.3.1 For a predictable process, 3σ control limits can be calculated from different measures of dispersion: the average of moving ranges, the average of subgroup ranges, the average of within subgroup standard deviations or the average of root mean square deviations. In any case, samples are drawn from a free running process with no adjustments during the sampling. The average of many values gives a robust estimate of the control limits, even though the process is unpredictable. Control limits will only be somewhat inflated, if just a few samples are affected by assignable causes. The averaging effectively suppresses a few large numbers among many (k is more than ten).

E.7.3.2 Table E.4 and E.5 can be used to calculate control limits for sample means, $\bar{x}$, or for individual values, x, based on for example k = 10 or 20 samples from a free running process. For the sake of example, sample size is n = 4, average within-sample standard deviation: $\bar{s} = 0{,}92\ g$ and average within-sample range is $\bar{R} = 2.09\ g$

Table E.4. Factors for calculation of control limits for mean values, $\bar{x}$ (UCL and LCL), for individual values, x (UNPL and LNPL), and for standard deviation, s (USDL and LSDL) based on average within-sample standard deviation, $\bar{s}$, for samples of n items.

	Factor for control limits of $\bar{x}$, UCL and LCL	Factor for control limits of x, $UNPL$ and $LNPL$	Factor for Lower Limit of s, $LSDL$	Factor for Upper Limit of s, $USDL$
n	A3	E3	B3	B4
2	2.659	3.760	-	3.267
3	1.954	3.385	-	2.568
4	1.628	3.256	-	2.266
5	1.427	3.191	-	2.089
6	1.287	3.153	0.030	1.970

E.7.3.3 From table E.4 we calculate 3σ Upper Control Limit, $UCL = \bar{\bar{X}} + A3 \cdot \bar{s}$ and 3σ Lower Control Limit, $LCL = \bar{\bar{X}} - A3 \cdot \bar{s}$ around the target value, the centre line, $\bar{\bar{X}}$.

$UCL = \bar{\bar{X}} + 1.628 \cdot 0.92\ g = \bar{\bar{X}} + 1.5\ g$

$LCL = \bar{\bar{X}} - 1.628 \cdot 0.92\ g = \bar{\bar{X}} - 1.5\ g$

Warning limits for the means of sample subgroups with n = 4 are at ± 2σ, which is ± 1.0 g.

For a behaviour chart for individual values (x-chart), 3σ limits are called Upper Natural[31] Process Limit and Lower Natural Process Limit. $UNPL = \bar{\bar{X}} + E3 \cdot \bar{s}$, $LNPL = \bar{\bar{X}} - E3 \cdot \bar{s}$

$UNPL = \bar{\bar{X}} + 3.256 \cdot 0.92\ g = \bar{\bar{X}} + 3.0\ g$

$LNPL = \bar{\bar{X}} - 3.256 \cdot 0.92\ g = \bar{\bar{X}} - 3.0\ g$

Warning limits for individual values are at ± 2σ, which is ± 2.0 g.

It should not be a surprise that the 3σ control limit for individual values is twice the value as the 3σ control limit for the mean value of n = 4 items in a sample, because

$sigma(x) = \sqrt{n} \cdot sigma(\bar{x})$.

Table E.5. Factors for calculation of control limits for mean values, $\bar{x}$ (UCL and LCL), for individual values, x (UNPL and LNPL), and for range, R (URL and LRL), based on average range, $\bar{R}$, for samples of n items.

	Factor for control limits, $\bar{x}$, UCL and LCL	Factor for control limits, x, $UNPL$ and $LNPL$	Factor for Lower Range Limit, LRL	Factor for Upper Range Limit, URL
n	A2	E2	D3	D4
2	1.880	2.660	-	3.268
3	1.023	1.772	-	2.574
4	0.729	1.457	-	2.282
5	0.577	1.290	-	2.114
6	0.483	1.184	-	2.004

E.7.3.4 From table E.5 we calculate 3σ Upper Control Limit, $UCL = \bar{\bar{X}} + A2 \cdot \bar{R}$ and 3σ Lower Control Limit, $LCL = \bar{\bar{X}} - A2 \cdot \bar{R}$ around the target value, the centre line, $\bar{\bar{X}}$.

[31] 'Natural' to indicate that it makes sense to draw specification limits in the same chart as the individual values for direct comparison, specification limits in charts for mean values are not useful in the same way. Also, the capability index and the performance index for a process should be calculated for individual values, not for means.

$$UCL = \bar{\bar{X}} + 0.729 \cdot 2.09\ g = \bar{\bar{X}} + 1.5\ g$$
$$LCL = \bar{\bar{X}} - 0.729 \cdot 2.09\ g = \bar{\bar{X}} - 1.5\ g$$

Warning limits for the means of sample subgroups with n = 4 are at ± 2σ, which is ± 1.0 g.

For a behaviour chart for individual values (x-chart), 3σ limits are calculated:
$$UNPL = \bar{\bar{X}} + E2 \cdot \bar{R} = \bar{\bar{X}} + 1.457 \cdot 2.09\ g = \bar{\bar{X}} + 3.0\ g$$
$$LNPL = \bar{\bar{X}} - E2 \cdot \bar{R} = \bar{\bar{X}} - 1.457 \cdot 2.09\ g = \bar{\bar{X}} - 3.0\ g$$

Warning limits for individual values are at ± 2σ around the target value, which is ± 2.0 g.

E.7.3.5 It should not be a surprise that the 3σ control limits for $\bar{x}$, and 3σ control limits for individual values, x, are the same as when we used table E.4 to calculate similar control limits. We should reach the same result irrespective of the type of statistics we use for estimation of control limits. We have used the average dispersion value of a number of samples (k > 10) for our calculations. Also the median value of statistics will give a robust estimate of the dispersion, and they can also be used for calculation of control limits for the location ($\bar{x}$ or x) or for the dispersion (R, s, R_{median} or s_{median}), using different tables.

E.7.4 *Control limits for standard deviation.*

A subgroup standard deviation chart (s-chart) monitors the variation of the process. From table E.4, control limits are calculated from the estimated average within-sample standard deviation of for example k = 10 or 20 samples (with n = 4): $\bar{s} = 0.92\ g$

Upper Standard Deviation Limit, $USDL = B4 \cdot \bar{s}$, and Lower Standard Deviation Limit, $LSDL = B3 \cdot \bar{s}$, on each side of the central line, $\bar{s}$.

$USDL = B4 \cdot \bar{s} = 2.266 \cdot 0.92\ g = 2.08\ g$
$LSDL = B3 \cdot \bar{s} = 0$ There is no lower limit for the standard deviation when n = 4.

E.7.5 *Control limits for sample range.*
A subgroup range chart or the moving range chart of individual values (R-charts) also monitors the variation of the process. From table E.5, control limits are calculated from the estimated average range of for example k = 10 or 20 samples (with n = 4): $\bar{R} = 2.09\ g$

Upper Range Limit, $URL = D4 \cdot \bar{R}$, and Lower Range Limit, $LRL = D3 \cdot \bar{R}$, on each side of the central line, $\bar{R}$.

$URL = D4 \cdot \bar{R} = 2.266 \cdot 2.09\ g = 6.83\ g$
$LRL = D3 \cdot \bar{R} = 0$ There is no lower limit for the range when n = 4.

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 $2\sigma_e$ Cusum units (for sample means, or aR or $a\sigma_0$ 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 aRk or $a\sigma_0$ 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 j^{th} 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 σ_e (the standard error of the sample means) for construction of the mask, thus:-

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

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.6 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 d_n factor.

Table E.6 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, $\bar{R}$ that is to say: scale, a $\bar{R}$ units per distance corresponding to one sample interval, decision interval H=h $\bar{R}$ slope of mask in chart units per sample interval, F=f. $\bar{R}$

E.8.4 *Cusum decision rules for standard deviation charts*. Table E.7 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_o, 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.7 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 $\Sigma (x-m_1)$ 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.7 provides decision interval and slope parameters, which are used to construct masks for use with the Cusum chart

m_t	25	32	4	5	64	8		125	6	
H	75	5	5							
F	5	18	1	5	36	2		75	4	

(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 milk packer is packing 1 000 ml nominal packs. The packing line can produce 5 000 packs an hour, and the packer wants to control the process by taking samples of n = 4 packs every half hour.
The apparent density in air of low-fat milk is ρ = 1.033 g/ml and he uses a class III non-automatic weighing instrument (NAWI) with verification scale interval e = 1 g and scale division, (d) of 1 g.
The cardboard containers have a mean value of $\bar{t} =$ 27.0 g, with standard deviation of the mean, $s_{\bar{t}} =$ 0.2 g.
The packer is plotting a Shewhart chart for sample means ($\bar{x}$ chart) as well as a chart for the sample range (R-chart). Previous records show that the average of sample ranges for a free-running process is $\bar{R} =$ 2.09 g, and the average of within-sample standard deviations is $\bar{s} =$ 0.92 g.

Using the appropriate bias factor the standard deviation of the filling process and tare is estimated as 1 g.

Comments on Packers System

1. Target setting by calculation of allowance

Three types of allowances are normally taken into account:
Allowance, a_1 when Packers' rule 2 or 3 is the most critical (if Packers rule 1 is most critical, allowance$_1$ = 0).
Allowance, a_2 when the sampling is less than during a reference test.
Allowance, a_3 for the measurement uncertainty.

Total allowance is calculated by:

$$Total\ allowance = a_1 + \sqrt{a_2^2 + a_3^2}$$

Allowance,a_1 for the critical filling rule:
The target, Q_t, for this process has to be the larger of Q_n, TU1 + 1.96s and TU2 + 3.72s.This allowance will take care of the situation where Packers rule 2 or 3 is more critical than rule 1.

For each filling rule, Q_n, TU1 = (Q_n – TNE) and TU2 = (Q_n – 2 TNE) must be considered with respect to the critical point of the distribution.
Filling rule no 1: $Q_{t1} - Q_n > 0$ (centre point of the distribution > Q_n, allowance$_1$ = 0)
Filling rule no 2: $Q_{t2} - (Q_n - TNE) > 1.96s$ (< 2.5 % of the distribution is below 1.96σ)
Filling rule no 3: $Q_{t3} - (Q_n - 2\ TNE) > 3.72s$ (< 1 out of 10 000 items is below 3.72σ)

The standard deviation of the mean value of n = 4 items is found from s = 1 / √4 = 0.5 g and used for the calculation of control limits.

Calculation of Q_t for the three filling rules and conclusion:

$$Q_{t1} = Q_n \cdot \rho = 1000\ ml \cdot 1.033 \frac{g}{ml} = 1033.0\ g$$
$$Q_{t2} = TU1 \cdot \rho + 1.96\ s(x) = 985\ ml \cdot 1.033 \frac{g}{ml} + 1.96 \cdot \sqrt{4} \cdot 0.508\ g = 1019.5\ g$$
$$Q_{t3} = TU2 \cdot \rho + 3.72\ s(x) = 970\ ml \cdot 1.033 \frac{g}{ml} + 3.72 \cdot \sqrt{4} \cdot 0.508\ g = 1005.8\ g$$

Conclusion: Q_{t1} is the largest, therefore rule 1 is the most critical and allowance$_1$ = 0 g

Allowance a_2 for sampling
With a packing rate of 5 000 items per hour, the time for producing 10 000 items (referred to as 'production period' in Table E.3 above) is 2 hours. The number of samples taken from each period is k = 5, because the first sample is taken at the start of the period, on every half hour and the last sample at the end of the period. Each sample consists of n = 4 items so the number of sampled items (nk = 20). Because 20 items is low compared to the number of samples in a reference test, an allowance$_2$ from table E.3 is calculated: Allowance$_2$ = z · σ.

We use two detection rules:
Detection rule no 1: 1 sample outside control limits at 3σ
Detection rule no 2: 2 out of 3 samples outside warning limits at 2σ

Procedure A: Use only detection rule no 1, z = 0.27 for k = 5 and n = 4.
Procedure D: Use both detection rules simultaneously, z = 0.15 for k = 5 and n = 4.

For procedure A, allowance a_2 = 0.27 · 0.508 g = 0.14 g
For procedure D, allowance a_2 = 0.15 · 0.508 g = 0.08 g

We use procedure D, so the allowance a_2 = 0.08 g

c) Allowance a_3 for measurement uncertainty

According to WELMEC Guide 6.9, the sources of measurement uncertainties are:

1) measuring net weight of the product,
2) measuring the weight of the tare
3) measuring the density of the liquid

1) The sources of uncertainty measuring the weight of one prepack on the weighing instrument are:
- the MPE in service (mpe = 2 g at 1 kg),
- the resolution (d) at the measurement point,
- the error at zero ,

all treated as rectangular distributions.

Standard measurement uncertainty for the gross weight is u_g:

$$u_g = \sqrt{\left(\frac{mpes}{\sqrt{3}}\right)^2 + \left(\frac{d}{2\sqrt{3}}\right)^2 + \left(\frac{d}{2\sqrt{3}}\right)^2} = \sqrt{\left(\frac{2}{\sqrt{3}}\right)^2 + \left(\frac{1}{2\sqrt{3}}\right)^2 + \left(\frac{1}{2\sqrt{3}}\right)^2} = 1.22g$$

2) If the same weighing instrument is used for determination of individual tare, or the mean value of n = 10 tare in the same weighing operation, the same formula is applied for determination of measurement uncertainty. But now the mpe is 1 g (at 27 g for individual tare and at 270 g for n = 10 tare). The standard deviation of the mean value from measurements of 10 empty tare is given in the example, $s_{tare}(\bar{x}) = 0.2\ g$.

Standard measurement uncertainty for the tare weight is u_t:

$$u_t = \sqrt{\left(\frac{mpes}{\sqrt{3}}\right)^2 + 2\left(\frac{d}{2\sqrt{3}}\right)^2 + \left(\frac{s_t}{\sqrt{n_t}}\right)^2} = \sqrt{\left(\frac{1}{\sqrt{3}}\right)^2 + 2\left(\frac{1}{2\sqrt{3}}\right)^2 + (0.2)^2} = 0.73\ g$$

3) Standard uncertainty for density, ρ, is estimated to 0.0005 g/ml. The standard measurement uncertainty of the target mass because of uncertainty in density is consequently:

$$u(m) = V \cdot u(\rho) = 1000\ ml \cdot 0.0005 \frac{g}{ml} = 0.5\ g$$

Combined standard measurement uncertainty is:

$u_c = \sqrt{(u_G)^2 + (u_{\bar{T}})^2 + (u_\rho)^2} = \sqrt{(1.22)^2 + (0.73)^2 + (0.5)^2} = 1.51\ g$

If allowance due to measurement uncertainty is given at 1σ level (68% confidence level), u_c, we get:

Allowance$_3$ = u $_c$= $1.51\ g$

Calculating total allowance

$Total\ allowance = a_1 + \sqrt{a_2{}^2 + a_3{}^2} = 0 + \sqrt{1.51^2 + 0.08^2} = 1.51\ g$

The target Q_t is calculated as:

$Q_t = Q_n \cdot$ density + tare + total allowance = 1000 ml · 1.033 g/ml + 27.0 g + 1.51 g = 1061.51 g

This example shows that the measurement uncertainty caused by the use of a verified class III NAWI with e = d = 1 is the major source to the total allowance.

A reduction of the measurement uncertainty could be achieved if the packer instead used a NAWI with e = d = 0.5 g. This would reduce the u_g from 1.22 g to 0.89 g.
And if instead a calibrated scale (like the one used in Guide 6.9) was used to measure the tare material the u_t could have been reduced from 0.73 g to 0.5 g causing the allowance$_3$ to be reduced from 1.51 g to 1.02 g.

It is also possible for the packer to calibrate his weighing instrument with calibrated weights. By doing this on a regularly basis it will be possible to bring the uncertainty caused by the NAWI down. The example in Guide 6.9 shows that a reduction of the uncertainty with approximately 50 % will be possible. This is of course dependent on the class of weights used and the procedure of calibration.

2. Control and warning limits for the $\bar{x}$-chart

For the $\bar{x}$-bar chart, we find A2 in table E.5, for calculation of control limits.

$UCL = Target + A2 \cdot \bar{R} = 1061.51\ g + 0.729 \cdot 2.09\ g = 1063.03\ g \approx 1063.0\ g$

$LCL = Target - A2 \cdot \bar{R} = 1061.51\ g - 0.729 \cdot 2.09\ g = 1059.99\ g \approx 1060.0\ g$

$3\sigma = A2 \cdot \bar{R} = 0.729 \cdot 2.09\ g = 1.524\ g$ giving 1σ = 0.508 g & 2σ = 1.02 g
1σ is used above for calculation of allowance$_1$. 2σ are used for calculation of the warning limits for the sample mean (± 1.02 g around the target):

$$UWL = 1061.51\ g + 1.02\ g = 1062.53\ g \approx 1062.5\ g$$
$$LWL = 1061.51\ g - 1.02\ g = 1060.49\ g \approx 1060.5\ g$$

Control and warning limits are drawn on the chart as symmetric lines around the target line. The centre line for the process, $\bar{\bar{X}}$, needs to be aligned with the target line.

3. Control limits for the R-chart

For the R-chart, we find D4 and D3 in table E.5, for calculation of control limits.

$$URL = D4 \cdot \bar{R} = 2.282 \cdot 2.09\ g = 4.77\ g$$
$$LRL = D3 \cdot \bar{R} = 0 \cdot 2.09\ g = 0\ g$$

For the alternative s-chart, we use B4 and B3 in table E.4 for calculation of control limits:

$$USDL = B4 \cdot \bar{s} = 2.266 \cdot 0.92\ g = 2.08\ g$$
$$LSDL = B3 \cdot \bar{s} = 0 \cdot 0.92\ g = 0\ g$$

Annex F Checkweighers and other automatic instruments

F.1 General

WELMEC 6.4 provides guidance for packers when choosing the instrument to be installed in a packing line.

This annex describes conditions applicable for instruments when installed in a packing line.

F.1.1 An overview of different specifications applicable to installed checkweighers

Automatic checkweighers are now a category of instruments included in the Measuring Instrument Directive 2004/22/EC[32] (designated under acronym "MID"). Among solutions recognised to fulfil the essential requirements of MID, OIML R51/2006 has been defined as the normative document for checkweighers ([33]).

A transition period between national certification and MID certification of ten years exists, that is until 30 October 2016, national certificates will remain valid. A lot of instruments will therefore still be subject to older national certification systems based on the following possibilities:

- OIML R51/1996. A majority of instruments are in line with metrological and technical requirements according to OIML R51/1996 which was recognised as a basis for applying the WELMEC type approval agreement ([34]).
- OIML R51/1985. A lot of instruments are still in line with an older version of R51, that is R51/1985, the requirements of which are equivalent to these of a repealed EEC directive, that is Directive 78/1031/EEC which was transposed in several EEC countries.
- National particular requirements. Some instruments in some Members States are in use, which are in line with some national requirements which are only known in the states where these requirements apply.

F.1.2 A quick summary of involved characteristics for the purpose of this guide and scope of this annex

- OIML R51/1996 is based on metrological accuracy classes designated by X(x) defining tolerances for mean value and standard deviation.
- MID (and also OIML R51/2006) defines accuracy classes designated by XI(x), XII(x), XIII(x) and XIV(x)([35]). The difference with R51/1996 is that for mean value, tolerances are a more detailed.
- Directive 78/1031/EEC was based on the concept of "Zone of indecision" defined as a value of the interval inside which the decision of the checkweigher is undetermined.
  With this concept you have :
  * The nominal zone of indecision "Un" marked on the instrument and
  * the actual zone of indecision "Ua" the conventional value of which is 6 times the standard deviation for the load being tested. This value is directly calculated from tests performed on the checkweigher.

F.1.3 The tolerances in service are the following :

* $Ua \leq Un$ ($Ua \leq 0.8\ Un$ for initial verification) and
* Sorting error $\leq 0.5\ Un$ (the same applies for initial verification)

- For all other national cases, no WELMEC guidance can be provided.

[32] MID Annex MI006, chapter 2
([33]) Communication of the Commission n°2006/C 269/01
([34]) Welmec guide 9
([35]) Although MID designates the class as "XIV(x)" and OIML R51/2006 designates the class as "XIIII(x)", these two designations are equivalent.

From this point, Annex F only deals with checkweighers according to OIML R51/1996 and MID (with normative document OIML R51/2006).

F.2 Maximum permissible errors applicable to checkweighers

F.2.1 Values for coefficient x

- In R51/1996, classes are designated as X(x). The factor x has the form 1×10^k, 2×10^k or 5×10^k, where k is a negative whole number or zero. R51/1996 doesn't define any maximum value for x, so it is up to national legislation to define the maximum value for x.
- In MID, additional classes have been created, i.e XI, XII and XIV. The factor x shall follow the following rules :
  * for classes XI and XII, $x < 1$
  * for class XIII, $x \leq 1$
  * for class XIV, $x = 2$

  Note : R51/2006 doesn't define any maximum value for x but in this case, MID applies, so the rule is that 2 is the maximum value for x.

WELMEC WG6 recommends that the following classes or better are used for prepackages :

* R51/1996 : X(x) having (x) ≤ 1
* MID : XIII(x) ; in this case, (x) ≤ 1

However, this doesn't automatically imply that there is no need to compensate for other parameters (e.g tare variability, density, …).

F.2.2 Maximum permissible values for the mean error (systematic error)

The tables below show the maximum permissible mean error according to R51/1996 and MID for initial verification.

Table F.1 : R51/1996 - Maximum permissible mean error

Net load (m) expressed in verification scale intervals (e)		Maximum permissible mean error for class X(x) instrument
Where x ≤ 1	Where x > 1	
0 < m ≤ 500 500 < m ≤ 2000 2000 < m ≤ 10000	0 < m ≤ 50 50 < m ≤ 200 200 < m ≤ 1000	± 0.5 e ± 1 e ± 1.5 e

Table F.2 : MID - Maximum permissible mean error

Net load (m) in verification scale intervals (e)				Maximum permissible mean error
XI	XII	XIII	XIV	
0 < m ≤ 50000 50000 < m ≤ 200000 200000 < m ≤ 1000000	0 < m ≤ 5000 5000 < m ≤ 20000 20000 < m ≤ 100000	0 < m ≤ 500 500 < m ≤ 2000 2000 < m ≤ 10000	0 < m ≤ 50 50 < m ≤ 200 200 < m ≤ 1000	± 0.5 e ± 1 e ± 1.5 e

Note 1 : the maximum permissible mean error is valid whatever the value of factor x is.

Note 2 : both R51/1996 and R51/2006 define the maximum permissible mean error as twice the maximum permissible mean error at initial verification. But this could differ according to national legislation.

F.2.3 Maximum permissible values for the standard deviation (random error)

The table below shows the maximum permissible standard deviation according to R51/1996 and MID for initial verification.

Table F.3 : R51/1996 and MID-Maximum permissible standard deviation

Net load (m)	Maximum permissible standard deviation (as percentage of m or in grams) for class X(1) instrument
m ≤ 50	0.48 %
50 < m ≤ 100	0.24 g
100 < m ≤ 200	0.24 %
200 < m ≤ 300	0.48 g
300 < m ≤ 500	0.16 %
500 < m ≤ 1000	0.8 g
1000 < m ≤ 10000	0.08 %
10000 < m ≤ 15000	8 g
15000 < m	0.053 %

Note 1 : the maximum permissible standard deviation has to be multiplied by factor x.

Note 2 : both R51/1996 and R51/2006 define the maximum permissible standard deviation as given in the following table F.4. But this could differ according to national legislation.

Table F.4 : R51/1996 and MID-Maximum permissible standard deviation in service (also to be multiplied by factor x)

Net load (m)	Maximum permissible standard deviation (as percentage of m or in grams) for class X(1) instrument
m ≤ 50	0.6 %
50 < m ≤ 100	0.3 g
100 < m ≤ 200	0.3 %
200 < m ≤ 300	0.6 g
300 < m ≤ 500	0.2 %
500 < m ≤ 1000	1.0 g
1000 < m ≤ 10000	0.1 %
10000 < m ≤ 15000	10 g
15000 < m	0.067 %

F.3 Operation of checkweigher system

When a “legal” checkweigher is used in a packing line for both ensuring conformity to average quantity requirements and individual quantity requirements, no additional back-up checks by sampling (e.g. statistical control) are necessary.

If only one aspect (average quantity or number below TU1 and/or TU2) is performed, additional checking is necessary for the other requirements (e.g. back-up checks by sampling).

National legislation could give different requirements.

However, the checkweigher shall be under control and supervision. This paragraph gives guidance for some aspects of checkweighers in use as :

- checkweighers not adjusting the filling process upstream from them
- checkweighers having parameters to be set according to the conditions of the packing line or according to the product or packing material
- controls of the checkweigher by the packer

Records of data and results of the control system are described in Annex C.

F.3.1 Use of checkweighers

A checkweigher shall present the average quantity of product and the number/percentage below TU1 and TU2 of at least every production hour. The packers procedures will include requirements to consider this information and act on it if necessary. It will be very difficult to monitor each hour's production from an accumulated average (over a period longer than one hour).

F.3.2 Setting parameters for a checkweigher

Where applicable, desiccation, special storage conditions, tare, density and the standard deviation (or zone of indecision) of the checkweigher are factors which shall be taken into consideration.

When the packing line produces a high number of defective units with a quantity close to TU1(when the standard deviation of the packing process is greater than ½ TNE), then this should be dealt with to avoid the checkweigher not being able to handle the prepackages according to the directive. This is due to the fact that the random error (standard deviation) will not permit to a checkweigher flooded with prepackages that should be rejected to make the right decision.

Possibilities to handle such eventuality are :

- to set the acceptable proportion of prepackages under TU1 to 0% or
- to raise the average target value for those instruments where the limits are automatically linked to the average target or
- to have a limit set to a value greater than TU1 without changing the average target value (where the instrument allows this).

F.3.3 *Monitoring a checkweigher*

A checkweigher may be sensitive to the environment in which it operates or may go wrong. Therefore, the control system of a packer shall include checks of the checkweigher in such a way that anomalies are detected as soon as possible after having occurred.

Whatever checks are consider necessary, WELMEC WG6 recommends that a special check to include the accuracy of the weighing operation and the correct functioning of the reject mechanism shall be performed at an appropppriate period. This could mean 'daily' for basic checks and less frequent for more comprehensive/metrological checks. Packers also need to take into account the risks involved.

Checking weighing operation

A verified/calibrated NAWI with a verification scale interval (e) smaller than (ideally less than 0.2) the interval of the AWI is needed to be able to check the weighing operation. The check shall always be done in normal weighing operation (i.e dynamically if the instrument operates in dynamic mode) with prepackages identical to those produced on the packing line (because of the buoyancy in air and weight distribution characteristics). The check should show that the error of the AWI is less than the maximum permissible error in-service. Also see WELMEC 6.4, paragraph 6.3.2.

Checking reject mechanism

The reject mechanism should be tested for proper function. This means that the test has to show that the rejected prepackage is taken away from the packing line.
The test should be done with a prepackage that has an actual quantity 3 standard deviations less than the value of the reject points set, which could be at TU2, TU1 and/or the nominal quantity which is being packed at the moment of test.

F.3.4 Most checkweighers are linked to servers through data transmission systems to which they transfer information about the weighings.

Some of these systems do not have the capacity to keep or retrieve information about every single prepackage and in that case the records should consist at least the statistics from the checkweigher for every hour's of production.

F.4 Example for checkweigher control

A packer is producing tins of peas with a nominal quantity of 425 g.

The filling process has a standard deviation of 8.6 g.

The tins have an average tare weight of 15 g with a 0.7 g standard deviation.

What is the minimum value of the target set point?

The TNE is 3 % of 425 g = 12.75, this must be rounded up to 12.8 g.

The standard deviation of the tare may be disregarded when it is < 1/10 TNE

Storage variability may also be disregarded as product does not lose weight when packed in tins.

The target set point is Q_n = 425 g.

The set points at TU1 and TU2 need to take into account the zone of indecision of the checkweigher, so:

the T1 set point should be at TU1 + 2*Ua, and

the T2 set point should be at TU2 + 3.72*Ua

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 centilitre and not followed by the symbol cl, and / or an indication of the distance in millimetre 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[36],filled under the conditions prescribed in that Directive and herein'. MCB only relate to capacities of 50ml to 5 litres. The conditions relate to the volume being at 20° 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° C, a description of the liquid and the apparent thermal coefficient of cubical expansion[37] 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 millilitres or in millimetres. In the latter instance a conversion chart must be available so that actual volume errors in the liquid fill can be determined.

[36] Directive 75/107/EEC

[37] 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.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.

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 millilitres 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 ± 0.5 mm.

(c) The MCBs is filled with water at 20° 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° 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 ± 0.5 mm.

(e) Where the template is marked with an operational temperature other than 20° 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° C the numerically smaller limiting value of the coefficient must be used,

ii. for an operational temperature above 20° 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° 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

G.8.1 A packer is packing bottles of drink with a nominal quantity of 200 ml in an MCB with a neck diameter of 25 mm at the fill height. The average volume of the bottles is 200.3 ml and the standard deviation is 1.0 ml. The filling process has a standard deviation of 5 ml and TNE is 9 ml.
The MCB's cross sectional area at the fill height is $\pi * r^2 = 3.14 * 1.25^2 = 4.91$ cm^2. Therefore 1 mm change in liquid height is equivalent to 0.49 ml.

Uncertainties of measurement

Element	Value	Divisor	Multiplier	Std. Uncertainty
		(distribution)	(sensitivity)	(ml)
MCB	4*1 ml = 4 ml	2	1	2
Template error	0.5 mm = 0.245 ml	√3	1.3	0.18
Reading error	0.5 mm = 0.245 ml	√3	1.3	0.18

G.8.2 Since the uncertainty from the MCB is dominant (more than 10 times the contribution from the template) normally only the uncertainty from the MCB is used. But in this example we will show the calculation.

Uncertainty = $\sqrt{2^2 + 0.18^2 + 0.18^2}$ = 2.02 ml

Combining this with the filling variation gives an overall variation of $\sqrt{2.02^2 + 5^2}$ = 5.39 ml

The target quantity needs to be the greater of:

Q_n + K = 200 + (-0.3) = 199.7 ml
TU1 +2s + K = 191 + 2 * 5.39 + (-0.3) = 201.5 ml
TU2 + 3.72s + K = 182 + 3.72 * 5.39 + (-0.3) = 201.8 ml

K = Q_n – average volume of bottles = 200.0 – 200.3 = -0.3 ml

The target quantity is 201.8 ml

It is assumed that more than 50 items are checked within a production period otherwise a sampling allowance (using z-factor) would also be required.

G.8.3 If the packer does not have access to reliable information about the average volume and the standard deviation of the bottles used then the 'Value' for the MCB in the table above should be changed to the tolerance for the specific bottle size from the directive.
If the tolerance is changed to the corresponding value for a 200 ml MCB (6 ml) in the example above, then the calculation is:

Uncertainties of measurement

Element	Value	Divisor	Multiplier	Std. Uncertainty
		(distribution)	(sensitivity)	(ml)
MCB	6 ml	2	1	3
Template error	0.5 mm = 0.245 ml	√3	1.3	0.18
Reading error	0.5 mm = 0.245 ml	√3	1.3	0.18

G.8.4 Since the uncertainty from the MCB is dominant (more than 10 times the contribution from the template) normally only the uncertainty from the MCB is used. But in this example we will show the calculation.

Uncertainty = $\sqrt{3^2 + 0.18^2 + 0.18^2}$ = 3.01 ml

Combining this with the filling variation gives an overall variation of $\sqrt{3.01^2 + 5^2}$ = 5.84 ml

The target quantity needs to be the greater of:

$Q_n + K = 200 + (0) = 200.0$ ml

TU1 +2s + K = 191 + 2 * 5.84 + (0) = 202.68 ml

TU2 + 3.72s + K = 182 + 3.72 * 5.84 + (0) = 203.7 ml

$K = Q_n$ – average volume of bottles = 200.0 – 200.0 = 0 ml

The target quantity is 203.7 ml

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 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.1.3 Besides the effectiveness of the statistical tests, generally checks are more effective, give a good 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. However checking on the packing line may provide other information to evaluate the appropriateness of the packers procedures.

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

For information on uncertainty of measurement for prepackages, refer to WELMEC guide 6.9.

H.4 Action when test fails

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.

Annex I Special Areas

I.1 Prepackages that change quantity after packing.

I.1.1 In order that packers have to meet same requirements, WELMEC WG6 recommends that Competent Departments apply the Directive's requirements for these products whose quantity changes after packing as follows:

a) That the prepackages shall meet the three quantity requirements at time the prepackages have passed the quantity checks specified in the packer's or importer's[38] quantity control system, and so are ready for placing on the market, and

b) be able to demonstrate this from records, and

c) No prepackages shall have a deficiency greater than twice the tolerable negative error anywhere in the distribution chain.

I.1.2 In order to preserve the quality of the product, the packer or importer should provides organisations in the distribution chain with the necessary information as to the storage and handling that needs to be observed.

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. WELMEC document 6.8 provides guidance on a test method and tolerances to apply.

I.3 Bread

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.4 Rug & knitting yarn

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.5 Growing media for plants (compressible)

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

I.6 Density measurement

I.6.1 Checking volume declarations via weight and density can be undertaken. 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 G14. 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, see WELMEC 6.9.

[38] 'importer' shall mean any natural or legal person established within the Community who places a product from a third country on the Community market; Decision 768/2008/EC, article R1.5

I.6.2 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.

Annex J Importer's Control System and Records

J.1 Duties

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.2 Checks

J.2.1 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	sizes			non-standard			Criterion for average contents k factor
		nd	otal	Accept	efer	Reject	
100 - 500				0		2	0.599
		2	0	1		2	0.640
501 - 3,200				0		2	0.430
		8	0	2		3	0.503
over 3,200				1		3	0.297
		8	8	3		4	0.387

J.2.2 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.2.3 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.3 Records

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.