INTERNATIONAL STANDARD

ISO 16820

Second edition
2019-10

Sensory analysis — Methodology — Sequential analysis

Analyse sensorielle — Méthodologie — Analyse séquentielle

Reference number
ISO 16820:2019(E)

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 34, *Food products*, Subcommittee SC 12, *Sensory analysis*.

This second edition cancels and replaces the first edition (ISO 16820:2004), of which it constitutes a minor revision. Information on using the Thurstonian δ approach and a citation for a new Reference [7] has been added in 5.1.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Sensory analysis — Methodology — Sequential analysis

1 Scope

This document specifies a procedure for statistically analysing data from forced-choice sensory discrimination tests, such as the triangle, duo-trio, 3-AFC, 2-AFC, in which after every trial of the discrimination test the decision can be made to stop testing and declare a difference, to stop testing and declare no difference, or to continue testing.

The sequential method often allows for a decision to be made after fewer trials of the discrimination test than would be required by conventional approaches that use predetermined numbers of assessments.

The method is effective for:

a) determining that either:

   1) a perceptible difference results; or

   2) a perceptible difference does not result when, for example, a change is made in ingredients, processing, packaging, handling or storage;

b) selecting, training and monitoring assessors.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 5492, *Sensory analysis — Vocabulary*

3 Terms, definitions and symbols

3.1 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO 5492 and the following apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1.1
alpha-risk
α-risk
probability of concluding that a perceptible difference exists when one does not

Note 1 to entry: This is also known as Type I error, significance level or false positive rate.

3.1.2
beta-risk
β-risk
probability of concluding that no perceptible difference exists when one does

Note 1 to entry: This is also known as Type II error or false negative rate.

3.1.3
sensitivity
general term used to summarize the performance characteristics of the test

Note 1 to entry: In statistical terms, the sensitivity of the test is defined by the values of α, β and p_d.

3.2 Symbols

p_0 probability of a correct response when no perceptible difference exists

p_d proportion of assessments in which a perceptible difference is detected between the two products

p_1 probability of a correct response when a perceptible difference does exist

4 Principle

The type of discrimination test (triangle, duo-trio, etc.) is chosen. The sensitivity of the test is defined by selecting values for α, β and p_d.

The boundaries of the decision regions are computed based on α, β, p_0 and p_1. After every trial of the discrimination test, the total number of correct responses [for the panel, see Clause 1 a), or per assessor, see Clause 1 b)] is compared to the decision boundaries to determine:

— if testing can be stopped and a difference can be declared;

— if testing can be stopped and no difference can be declared;

— if testing should continue.

5 Procedure

5.1 Construct a graph, as in Figure A.1, which illustrates the boundaries of the decision regions based on α, β, p_0 and p_1, as follows.

a) α and β are chosen based on the risks the researcher is willing to take of obtaining a false positive or a false-negative result, respectively. α is the probability of declaring that a difference exists when the true probability of a correct response is p_0. β is the probability of failing to declare that a difference exists when the true probability of a correct response is p_1 ($p_1 > p_0$).

b) p_0 is the probability of a correct response when no perceptible difference exists (i.e. the probability of a correct guess). The value of p_0 depends on the discrimination test being used:

1) for the triangle and the 3-AFC tests, $p_0 = 1/3$;

2) for the duo-trio and the 2-AFC tests, $p_0 = 1/2$.

c) p_1 is the probability of a correct response when a perceptible difference does exist. The value of p_1 depends on p_d:

1) for the triangle and 3-AFC tests, $p_1 = p_d + \left(\frac{1-p_d}{3}\right)$;

2) for the duo-trio and 2-AFC tests, $p_1 = p_d + \left(\frac{1-p_d}{2}\right)$;

3) researchers who use the Thurstonian δ approach to measure the magnitude of the sensory difference between two products can use the conversion tables (from δ to p_d and from p_d to δ)

presented in Reference [7] to choose the value of p_d that corresponds to the chosen value of δ for the test method being used.

d) The lines that form the boundaries of the decision regions are calculated as:

$$\text{lower line: } d_0 = \frac{\lg(\beta) - \lg(1-\alpha) - n \times \lg(1-p_1) + n \times \lg(1-p_0)}{\lg(p_1) - \lg(p_0) - \lg(1-p_1) + \lg(1-p_0)}$$

$$\text{upper line: } d_1 = \frac{\lg(1-\beta) - \lg(\alpha) - n \times \lg(1-p_1) + n \times \lg(1-p_0)}{\lg(p_1) - \lg(p_0) - \lg(1-p_1) + \lg(1-p_0)}$$

where α, β, p_0 and p_1 are as defined above, and n is the number of trials of the test.

NOTE The distance between the two lines depends on $p_1 - p_0$.

5.2 After each trial of the discrimination test, plot the total number of correct responses (on the vertical axis) versus the number of trials (on the horizontal axis):

— if the total number of correct responses falls between the lower and upper lines on the chart, then continue testing by conducting another trial;

— if the total number of correct responses falls above the upper line on the chart, then stop testing and conclude that a perceptible difference exists (at the α-level of significance);

— if the total number of correct responses falls below the lower line on the chart, then stop testing and conclude that no meaningful difference exists [i.e. there is less than a $(1 - \beta)$ probability that the true probability of a correct response is as high as p_1].

Annex A
(informative)

Examples

A.1 Example 1 — Sequential analysis of a series of triangle tests —Acceptance versus rejection of two trainees on a panel

A.1.1 Background

A sensory analyst wishes to base the decision to accept or reject two trainees on the panel on their performance in triangle tests using a typical pair of products. Each trainee receives a series of triangle tests. Intervals between tests are kept long enough to avoid sensory fatigue.

A.1.2 Test design

The number of trials required to accept or reject a trainee is determined by sequential analysis using a graph as shown in Figure A.1. To position the boundaries of the decision regions (i.e. the two lines in Figure A.1), assign a value to each of the four parameters, α, β, p_0 and p_1. In the triangle test, p_0 = 1/3 (i.e. the probability of a correct guess, p_d = 0). Usually the minimum acceptable rate of detection is set at p_d = 50 %, which makes:

$$p_1 = 0{,}50 + (1 - 0{,}50)\left(\frac{1}{3}\right) = \frac{2}{3}$$

If it is desired to reduce the number of trials to reach a decision, lower the minimum acceptable rate of detection, e.g. to p_d = 40 %, which makes:

$$p_1 = 0{,}40 + (1 - 0{,}40)\left(\frac{1}{3}\right) = 0{,}60 \text{ , etc.}$$

NOTE In this example, the definition of p_d is not the proportion of the population of assessors who can distinguish the samples but rather the proportion of trials in which a single assessor actually distinguishes the samples.

The analyst chooses the following values for the parameters:

— α = 0,05 is the probability of selecting an unacceptable trainee;

— β = 0,10 is the probability of rejecting an acceptable trainee;

— p_0 = 1/3 is the maximum unacceptable ability (i.e. the null hypothesis *p*-value of the triangle test);

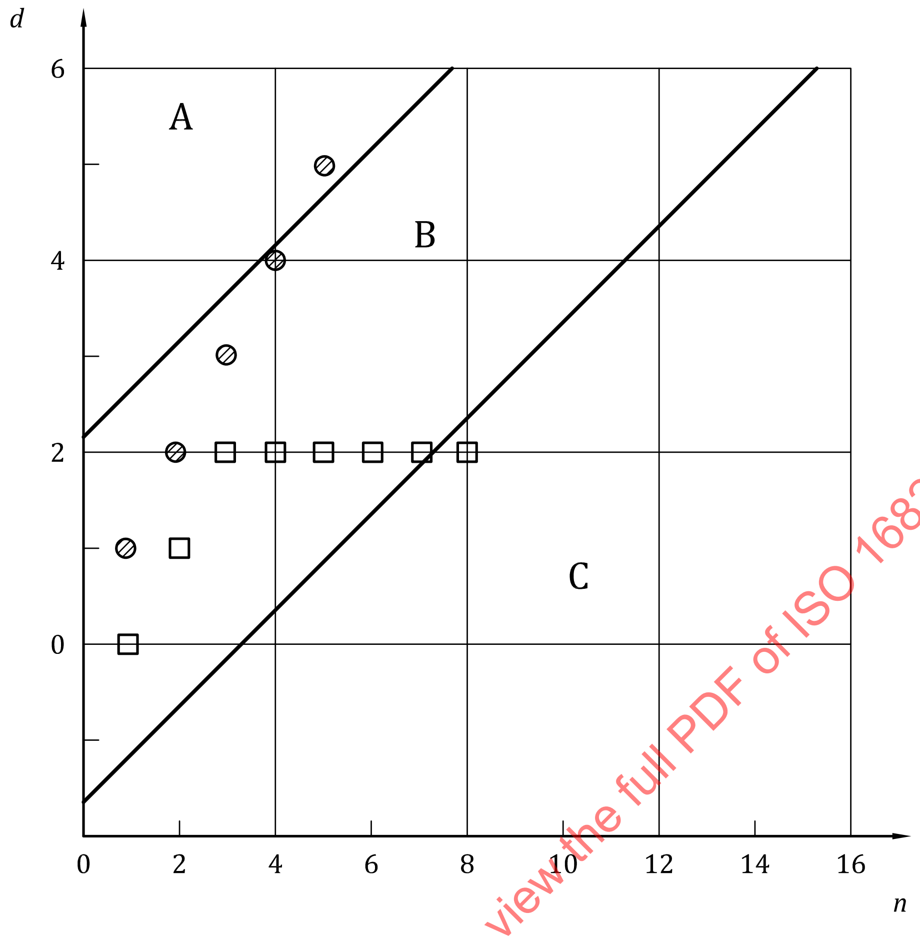
— p_0 = 2/3 is the minimum acceptable ability (i.e. the probability that the odd sample will be detected when p_d = 0,50).

A.1.3 Analysis and interpretation of results

As each triangle is completed, the results are entered in the diagram in Figure A.1 as follows. Enter the result of the first trial, if correct, as (x, y) = (1,1) and, if incorrect, as (x, y) = (1,0). For each succeeding trial, increase the value of x by 1 and increase y by 1 for a correct response, or increase x by 1 and y by 0 for an incorrect response. Continue testing until a plotted point touches or crosses either of the decision boundaries. Draw the indicated conclusion (i.e. accept or reject the trainee).

Trainee A is correct in all tests and is accepted after five trials. Trainee B fails in the first triangle, succeeds in triangles 2 and 3, but then fails on every subsequent triangle and is rejected after the eighth trial.

Parameters of the test:	$\alpha = 0{,}05$ $p_0 = \frac{1}{3}$	$\beta = 0{,}10$ $p_1 = \frac{2}{3}$
Boundary lines:	Lower: $d_0 = \frac{\lg(\beta) - \lg(1-\alpha) - n \times \lg(1-p_1) + n \times \lg(1-p_0)}{\lg(p_1) - \lg(p_0) - \lg(1-p_1) + \lg(1-p_0)}$ Lower: $d_0 = \frac{\lg(0{,}10) - \lg(1-0{,}05) - n \times \lg\left[1-(2/3)\right] + n \times \lg\left[1-(1/3)\right]}{\lg(2/3) - \lg(1/3) - \lg\left[1-(2/3)\right] + \lg\left[1-(1/3)\right]}$ Lower: $d_0 = -1{,}624 + 0{,}5\,n$	
	Upper: $d_1 = \frac{\lg(1-\beta) - \lg(\alpha) - n \times \lg(1-p_1) + n \times \lg(1-p_0)}{\lg(p_1) - \lg(p_0) - \lg(1-p_1) + \lg(1-p_0)}$ Upper: $d_1 = \frac{\lg(1-0{,}10) - \lg(0{,}05) - n \times \lg\left[1-(2/3)\right] + n \times \lg\left[1-(1/3)\right]}{\lg(2/3) - \lg(1/3) - \lg\left[1-(2/3)\right] + \lg\left[1-(1/3)\right]}$ Upper: $d_1 = 2{,}085 + 0{,}5\,n$	



Key

⊘	Trainee A	A	acceptance region
□	Trainee B	B	continue testing region
n	number of trials	C	rejection region
d	number of correct responses		

NOTE At the fifth trial, Trainee A passes out of the "continue testing region" and into the "acceptance region". At the eighth trial, Trainee B passes out of the "continue testing region" and into the "rejection region".

Figure A.1 — Use of sequential analysis in triangle tests — Example 1: Selection of two trainees

A.2 Example 2 — Sequential analysis of a series of duo-trio tests — Warmed-over flavour (WOF) in beef patties during storage

A.2.1 Background

A manufacturer's quality control panel has detected warmed-over flavour (WOF) in beef patties refrigerated for 5 days and then reheated. The project leader wants to set a reasonable maximum for the number of days the beef patties may be refrigerated.

A.2.2 Test design

Preliminary evaluations have shown that beef patties stored for 5 days exhibit strong WOF while patties stored for 1 day exhibit no WOF. The sensory analyst chooses to run a series of duo-trio tests involving patties that have been stored for 1, 3 and 5 days. Each of the stored patties will be compared to a freshly grilled control sample that has undergone no refrigerated storage.

The three sample pairs (control versus 1-day storage, control versus 3-day storage and control versus 5-day storage) are presented in separate duo-trio tests. The freshly grilled control is used as a constant reference in all three tests. As each assessor completes an evaluation, the result is added to the previous responses and the cumulative number of correct responses is plotted as illustrated in Figure A.2. The test series continues until the stored sample is declared either similar to or different from the freshly grilled control.

The analyst chooses the following values for the parameters:

— α = 0,10 is the probability of accepting an unacceptable patty;

— β = 0,10 is the probability of rejecting an acceptable patty;

— p_0 = 0,50 is the probability of a correct response when no perceptible difference exists (i.e. the null hypothesis p-value of the duo-trio test);

— p_1 = 0,70 is the probability that the stored sample is selected as being different from the control when p_d = 0,40 (i.e. p_1 = 0,40 + (1 − 0,40) (0,50) = 0,70).

A.2.3 Analysis and interpretation of results

Given the analyst's choices for α, β, p_0 and p_1, the formulae for the lines that form the boundaries of the decision regions, presented in Clause 5, are: $d_0 = -2{,}59 + 0{,}60n$ and $d_1 = 2{,}59 + 0{,}60n$. The lines are plotted in Figure A.2 along with the cumulative number of correct responses (see Table A.1) from the tests on the three stored samples.

Parameters of the test:	α = 0,10 β = 0,10 p_0 = 0,50 p_1 = 0,70
Boundary lines:	Lower: $d_0 = \dfrac{\lg(\beta)-\lg(1-\alpha)-n\times\lg(1-p_1)+n\times\lg(1-p_0)}{\lg(p_1)-\lg(p_0)-\lg(1-p_1)+\lg(1-p_0)}$ Lower: $d_0 = \dfrac{\lg(0{,}10)-\lg(1-0{,}10)-n\times\lg(1-0{,}70)+n\times\lg(1-0{,}50)}{\lg(0{,}70)-\lg(0{,}50)-\lg(1-0{,}70)+\lg(1-0{,}50)}$ Lower: $d_0 = -2{,}59+0{,}60\,n$
	Upper: $d_1 = \dfrac{\lg(1-\beta)-\lg(\alpha)-n\times\lg(1-p_1)+n\times\lg(1-p_0)}{\lg(p_1)-\lg(p_0)-\lg(1-p_1)+\lg(1-p_0)}$ Upper: $d_1 = \dfrac{\lg(1-0{,}10)-\lg(0{,}10)-n\times\lg(1-0{,}70)+n\times\lg(1-0{,}50)}{\lg(0{,}70)-(0{,}50)-\lg(1-0{,}70)+\lg(1-0{,}50)}$ Upper: $d_1 = 2{,}59+0{,}60\,n$

Table A.1 — Results obtained in Example 2 — Sequential analysis of a series of duo-trio tests — Warmed-over flavour (WOF) in beef patties during storage

Assessor	Test A (1-day)		Test B (3-day)		Test C (5-day)	
	Results	Count	Results	Count	Results	Count
1	I	0	I	0	C	1
2	I	0	C	1	C	2
3	I	0	I	1	C	3
4	C	1	C	2	C	4
5	I	1	I	2	I	4
6	C	2	C	3	C	5
7	I	2	I	3	C	6
8	C	3	C	4	C	7
9	I	3	C	5	I	7
10	C	4	C	6	C	8
11	I	4	C	7	C	9
12			I	7	C	10
13			C	8		
14			C	9		
15			C	10		
16			C	11		
17			I	11		
18			I	11		
19			C	12		
20			C	13		
21			I	13		
22			I	13		
23			I	13		
24			C	14		
25			I	14		
26			C	15		
27			C	16		
28			C	17		
29			C	18		
30			C	19		

NOTE Results: I = incorrect, C = correct; Count = cumulative number of correct results.

The sample stored for 1 day is declared to be similar to the freshly grilled control after 11 trials of the duo-trio test. The sample stored for 5 days is declared to be different from the control after 12 trials. The sample stored for 3 days cannot be declared similar to, or different from, the control after 30 trials (see Figure A.2). The analyst reports the findings to the project leader with the recommendation that 3-day storage be accepted as a tentative specification until additional testing can be completed to obtain definitive results for that storage period also.