INTERNATIONAL STANDARD

ISO 5495

Third edition
2005-11-15

Sensory analysis — Methodology — Paired comparison test

Analyse sensorielle — Méthodologie — Essai de comparaison par paires

Reference number
ISO 5495:2005(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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

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.

ISO 5495 was prepared by Technical Committee ISO/TC 34, *Food products*, Subcommittee SC 12, *Sensory analysis*.

This third edition cancels and replaces the second edition (ISO 5495:1983), which has been technically revised.

Sensory analysis — Methodology — Paired comparison test

1 Scope

This International Standard describes a procedure for determining whether there exists a perceptible sensory difference or a similarity between samples of two products concerning the intensity of a sensory attribute. This test is sometimes also referred to as a directional difference test or a 2-AFC test (Alternative Forced Choice). In fact, the paired comparison test is a forced choice test between two alternatives.

NOTE The paired comparison test is the simplest existing classification test since it concerns only two samples.

The method is applicable whether a difference exists in a single sensory attribute or in several, which means that it enables determination of whether there exists a perceptible difference concerning a given attribute, and the specification of the direction of difference, but it does not give any indication of the extent of that difference. The absence of difference for the attribute under study does not signify that there does not exist any difference between the two products.

This method is only applicable if the products are relatively homogeneous.

The method is effective

a) for determining

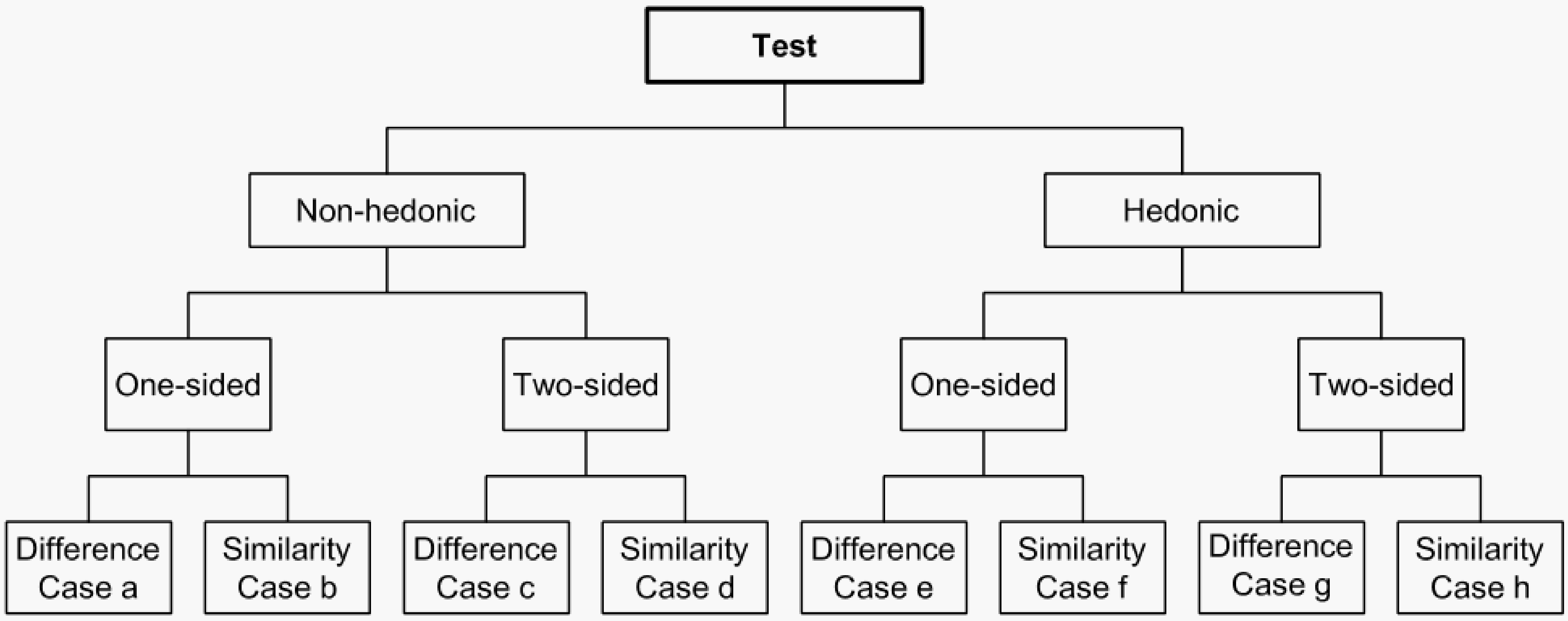
— whether a perceptible difference exists (paired difference test), or

— whether no perceptible difference exists (paired similarity test) when, for example, modifications are made to ingredients, processing, packaging, handling or storage operations, or

b) for selecting, training and monitoring assessors.

It is necessary to know, prior to carrying out the test, whether the test is a one-sided test (the test supervisor knows a priori the direction of the difference, and the alternative hypothesis corresponds to the existence of a difference in the expected direction) or a two-sided test (the test supervisor does not have any a priori knowledge concerning the direction of the difference, and the alternative hypothesis corresponds to the existence of a difference in one direction or the other).

The paired test can also be used in order to compare two products in terms of preference. The different cases of use of the paired test are summarized in Figure 1.



NOTE Only non-hedonic tests are dealt with in this International Standard.

Figure 1 — Possible different cases of use of the paired comparison test

EXAMPLE 1 (Case a) The production of a biscuit has been modified in order to render it more crisp. It is desired to check whether this increase is perceptible. Therefore it is necessary to try to highlight a difference to see whether the new product is perceived as being crispier than the usual product (control).

EXAMPLE 2 (Case b) A manufacturer knows that the product may contain traces of an ingredient which imparts an off-flavour to the product. He therefore wishes to determine the maximum acceptable quantity so that the flavour difference with a reference product without this ingredient is barely perceptible and therefore without any regrettable consequences.

EXAMPLE 3 (Case c) It is desired to produce a new soup and to compare two ingredients which will provide the salty flavour. For cost-intensive reasons, the ingredient which, at the same concentration, will provide the strongest salty flavour is sought. Therefore it is necessary to try to highlight a difference. It is not known a priori which ingredient will produce the strongest salty flavour.

EXAMPLE 4 (Case d) A manufacturer of plastics used, in particular, by car manufacturers for dashboards is seeking, for economic reasons, to replace the usual lubricant by a new one, but does not wish that the new plastics formula be perceived as presenting less or more surface slip than the usual one. It is a question of determining whether, for a same concentration, the new lubricant provides the same “surface slip” level as the usual product. It is necessary to show that both lubricants are similar in terms of “surface slip”, but it is not known a priori which lubricant can produce the highest surface slip characteristics.

2 Normative references

The following referenced documents are indispensable for the application 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 1992, *Sensory analysis — Vocabulary*

ISO 6658:1985, *Sensory analysis — Methodology — General guidance*

ISO 8586-1:1993, *Sensory analysis — General guidance for the selection, training and monitoring of assessors — Part 1: Selected assessors*

ISO 8586-2:1994, *Sensory analysis — General guidance for the selection, training and monitoring of assessors — Part 2: Experts*

ISO 8589:1988, *Sensory analysis — General guidance for the design of test rooms*

3 Terms and definitions

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

3.1
α (alpha) risk
probability of concluding that a perceptible difference exists when one does not exist

NOTE This is also called a type I error, significance level or false-positive rate.

3.2
β (beta) risk
probability of concluding that no perceptible difference exists when one does exist

NOTE This is also called a type II error or false-negative rate.

3.3
difference
situation in which samples can be distinguished based on their sensory attributes

NOTE The proportion of assessments during which a perceptible difference is detected between the products for the sensory attribute under study is given by the symbol p_d.

3.4
one-sided test
test in which the test supervisor has a priori knowledge concerning the direction of difference

NOTE The null hypothesis is H_0, the products are not different; the proportion of correct responses observed, p, is equal to 1/2. The alternative hypothesis is H_1, $p > 1/2$.

3.5
two-sided test
test in which the test supervisor does not have any a priori knowledge concerning the direction of difference

NOTE The null hypothesis is H_0, the products are not different; the proportion of responses observed for one of the samples, p, is equal to 1/2. The alternative hypothesis is H_1, $p \neq 1/2$.

3.6
correct responses
expected responses
number of assessors, in the case of a one-sided test, having selected the sample expected by the test supervisor to be the most intense for the sensory attribute under study

3.7
consensual responses
highest value, in the case of a one-sided test, of the number of assessors having selected sample A and those having selected sample B

NOTE This is calculated as above since there are not any correct responses.

3.8
product
material to be evaluated

3.9
sample
unit of product prepared, presented and evaluated during the course of the test

3.10
sensitivity
general term employed to summarize the performance characteristics of the test

NOTE In statistical terms, the sensitivity of the test is defined by the values of α, β and p_d.

3.11
similarity
situation in which any perceptible differences between the samples are so small that the products can be used interchangeably

4 Principle

The number of assessors is chosen on the basis of the sensitivity desired for the test (See 6.2 and the footnote that accompanies Tables A.4 and A.5).

The assessors receive a set of two samples (i.e. a pair). They designate the sample which they consider to be the most intense regarding the sensory attribute under consideration, even if this choice is based only on a guess.

NOTE One of the samples may be a control.

The number of times that each sample is selected is counted and the significance is determined by reference to a statistical table, taking into consideration the results obtained for the expected sample (one-sided test) or the highest number of responses obtained for either of the samples (two-sided test).

5 General test conditions

5.1 Define the objective of the test in a clear way to determine if the attempt is to be a one-sided or a two-sided test, if it is a difference or similarity test, and which is the most appropriate sensitivity.

5.2 Carry out the test under conditions that prevent all communication among assessors until the evaluations have been completed, using facilities and booths complying with ISO 8589.

5.3 Prepare the samples out of sight of the assessors and in an identical manner for each one of them; i.e. same apparatus, same vessels.

5.4 Assessors shall not be able to draw any conclusions regarding the intensity of the attribute from the manner in which the samples are presented to them. For example, for a tactile test, any differences in appearance shall be avoided. Mask all colour differences if the test objective does not concern the colour by using light filters and/or subdued lighting. The samples may also be presented successively and non-simultaneously in the case of slight differences in appearance.

5.5 Code the samples or the vessels containing the samples in a uniform manner, preferably using 3-digit numbers chosen at random for each test. Each pair is composed of two samples, each with a different code. Preferably, different codes should be used for each assessor during a session. However, the two same codes may be used for all assessors within a test, provided that each code is used only once per assessor during a test session (e.g. if several paired tests on different products are being conducted during the same session).

5.6 The quantity or volume served shall be identical for the two samples constituting each pair, just as that of all the other samples in a series of tests on a given type of product. The quantity or volume to be assessed can be imposed. If it is not, it should however be specified to the assessors to take quantities or volumes that are always similar whatever the sample.

5.7 The temperature of the samples constituting each pair shall be identical just as that of all the other samples in a series of tests on a given type of product. It is preferable to present the samples at the temperature at which the product is generally consumed.

5.8 The assessors shall be told whether or not they have to follow a special protocol in order to assess the products (e.g. whether or not to swallow the samples for a taste test, or carry out a specific gesture for a tactile test) or whether they are free to do as they please. In this latter case, they should be requested to proceed in the same manner for all the samples.

5.9 During the test sessions, avoid giving information about product identity, expected treatment effects or individual performance until all tests are completed.

6 Assessors

6.1 Qualification

All assessors should possess the same level of qualification, this level being chosen on the basis of the test objective (see ISO 8586-1 and ISO 8586-2). Experience and familiarity with the product can increase the performance of an assessor and can consequently increase the likelihood of finding a significant difference. Monitoring the performance of assessors over time may prove to be useful for increased sensitivity.

All assessors shall be familiar with the mechanisms of the paired test (the scoresheet, the task and the evaluation procedure). In addition, assessors shall be capable of recognising the sensory attribute on which the test is based. This attribute shall be defined verbally, by means of a reference substance or by presenting a few samples having different levels of intensity for the attribute under examination.

6.2 Number of assessors

Choose the number of assessors so as to obtain the level of sensitivity required for the test (see Table A.4 for a one-sided test and Table A.5 for a two-sided test). The use of a large number of assessors increases the likelihood of detecting small differences between the products. However, in practice, the number of assessors is often determined by material conditions (e.g. duration of the experiment, number of available assessors, quantity of product). When conducting a difference test, the number of assessors is typically approximately 24 to 30. When conducting a similarity test, about twice as many assessors (i.e. approximately 60) are required for equivalent sensitivity. When testing for similarity, evaluations should not be replicated by the same assessors. For a difference test, replications may be considered but should still be avoided whenever possible. However, if replicate evaluations are required in order to produce a sufficient total number of evaluations, every effort should be made to have each assessor perform the same number of replicate evaluations. For example, if only 10 assessors are available, have each assessor perform three paired tests in order to obtain a total of 30 evaluations.

NOTE Analysing three evaluations performed by 10 assessors as 30 independent evaluations is not valid when testing for similarity using Table A.3. However, the difference test using Tables A.1 and A.2 is valid even when replicate evaluations are performed [5], [6]. Some recent publications [1], [2] on replicated discrimination tests suggest alternative approaches for analysing replicated evaluations.

7 Procedure

7.1 Prepare the worksheets and scoresheets (see Figures B.1, B.2 and B.3) prior to conducting the test so as to use an equal number of the two possible presentation sequences of both products, A and B.

7.2 Present the two samples constituting a pair successively or simultaneously (see 5.4). In the case of simultaneous presentation, arrange the two samples in the same manner for each assessor (in line from left to right, in line from the bottom up, etc.). The assessors shall examine the two samples constituting the pair in the order indicated in the scoresheet, but assessors are generally authorized to make repeated evaluations of each sample if so wished (if, of course, the nature of the product allows for repeated evaluations).

7.3 Provision should be made for one scoresheet per pair of samples. If an assessor is to perform more than one test during the course of a session, collect the completed scoresheet and the unused samples prior to serving the subsequent pair. The assessor can neither go back to any of the previous samples, nor modify his/her verdict concerning any of the previous tests.

7.4 Do not ask any questions about preference, acceptance or degree of difference following the selection of the most intense sample. The selection the assessor has just made may bias the response to any additional questions. Responses to such questions may be obtained through separate tests concerning preference, acceptance, degree of difference, etc. (see ISO 6658). A “Comments” section requesting the reasons for the choice may be included for the assessors' remarks.

7.5 The paired test is a “forced choice” procedure; assessors are not allowed to choose the “no difference” option. An assessor who detects no difference between the samples should be instructed to select one of the samples and to indicate that the selection was only a guess in the “Comments” section of the scoresheet.

8 Analysis and interpretation of results

8.1 When testing for a difference

8.1.1 Case of a one-sided test

Use Table A.1 to analyse the data obtained from a paired test. If the number of correct responses is greater than or equal to the number given in Table A.1 (corresponding to the number of assessors and to the α-risk level chosen for the test), conclude that a perceptible difference exists between the samples (see B.1).

If desired, calculate a confidence interval on the proportion of the population able to distinguish the samples. This method is described in B.5.

No conclusion should be drawn for maximum numbers of correct responses under $n/2$.

8.1.2 Case of a two-sided test

Use Table A.2 to analyse the data obtained from a paired test. If the number of consensual responses is greater than or equal to the number given in Table A.2 (corresponding to the number of assessors and to the α-risk level chosen for the test), conclude that a perceptible difference exists between the samples (see B.3).

If desired, calculate a confidence interval on the proportion of the population able to distinguish the samples. This method is described in B.5.

8.2 When testing for similarity [1)]

8.2.1 Case of a one-sided test

Use Table A.3 to analyse the data obtained from a paired test. If the number of correct responses is less than or equal to the number given in Table A.3 (corresponding to the number of assessors, to the β-risk level and to the value of p_d chosen for the test), conclude that no meaningful difference exists between the samples (see B.2). If the results are to be compared from one test to another, then the same value of p_d should be chosen for all tests.

If desired, calculate a confidence interval on the proportion of the population able to distinguish the samples. This method is described in B.5.

No conclusion should be drawn for maximum numbers of correct responses under $n/2$.

---

1) In this International Standard, “similar” does not mean “identical”. This term signifies rather that the two products are sufficiently alike to be used interchangeably. It is impossible to prove that two products are identical. However, it can be demonstrated that any difference that does exist between two products is so minor as to have no practical significance.

8.2.2 Case of a two-sided test

Use Table A.3 to analyse the data obtained from a paired test. If the number of consensual responses is less than or equal to the number given in Table A.3 (corresponding to the number of assessors, to the β-risk level and to the value of p_d chosen for the test), conclude that no meaningful difference exists between the samples (see B.4). If the results are to be compared from one test to another, then the same value of p_d should be chosen for all tests.

If desired, calculate a confidence interval on the proportion of the population able to distinguish the samples. This method is described in B.5.

9 Report

Report the test objective, the results and the conclusions. It is recommended to add the following additional information:

— the purpose of the test and the nature of the treatment being studied;

— the complete identification of the samples: origin, method of preparation, quantity, shape, storage prior to testing, quantity served, temperature (the information concerning the sample should indicate that all storage, handling and preparation operations have been carried out so as to yield samples that differ only due to the variable of interest, if any);

— the number of assessors, the number of correct or consensual responses and the result of the statistical evaluation (including the values of α, β and p_d used for the test);

— the assessors: experience (in sensory testing, with the product, with the test samples), age and gender (see ISO 8586-1 and ISO 8586-2);

— any specific information and recommendations given to the assessors in connection with the test, in particular in the case where a precise definition and reference samples illustrating the attribute under test and/or a test protocol have been indicated to the assessors;

— the test environment: test facility used, simultaneous or sequential presentation, whether the identity of the samples was disclosed after the test and if so, in what manner;

— the location and date of test, and the name of the panel leader.

10 Precision and bias

Because the results of sensory discrimination tests are dependent on individual sensitivities, it is impossible to make a general statement regarding the reproducibility of the results that is applicable to all populations of assessors. The precision regarding a particular population of assessors increases as the size of the panel increases, as well as with the training and exposure to the product.

As a "forced-choice" procedure is used, the results obtained by this method are bias-free, provided that the precautions given in Clause 7 are fully observed.

Annex A
(normative)

Tables

A.1 Determination of perceptible difference or similarity

See Tables A.1 to A.3.

Table A.1 — Minimum number of correct responses required to conclude that a perceptible difference exists, based on a one-sided paired test [2), 3)]

n	α					n	α				
	0,20	0,10	0,05	0,01	0,001		0,20	0,10	0,05	0,01	0,001
10	7	8	9	10	10	36	22	23	24	26	28
11	8	9	9	10	11	37	22	23	24	27	29
12	8	9	10	11	12	38	23	24	25	27	29
13	9	10	10	12	13	39	23	24	26	28	30
14	10	10	11	12	13	40	24	25	26	28	31
15	10	11	12	13	14						
						44	26	27	28	31	33
16	11	12	12	14	15	48	28	29	31	33	36
17	11	12	13	14	16	52	30	32	33	35	38
18	12	13	13	15	16	56	32	34	35	38	40
19	12	13	14	15	17	60	34	36	37	40	43
20	13	14	15	16	18						
						64	36	38	40	42	45
21	13	14	15	17	18	68	38	40	42	45	48
22	14	15	16	17	19	72	41	42	44	47	50
23	15	16	16	18	20	76	43	45	46	49	52
24	15	16	17	19	20	80	45	47	48	51	55
25	16	17	18	19	21						
						84	47	49	51	54	57
26	16	17	18	20	22	88	49	51	53	56	59
27	17	18	19	20	22	92	51	53	55	58	62
28	17	18	19	21	23	96	53	55	57	60	64
29	18	19	20	22	24	100	55	57	59	63	66
30	18	20	20	22	24						
						104	57	60	61	65	69
31	19	20	21	23	25	108	59	62	64	67	71
32	19	21	22	24	26	112	61	64	66	69	73
33	20	21	22	24	26	116	64	66	68	71	76
34	20	22	23	25	27	120	66	68	70	74	78
35	21	22	23	25	27						

NOTE 1 The values in the table are exact because they are based on the binomial distribution. For values of n not included in the table, an approximation of the missing entries may be obtained in the following manner: Minimum number of responses (x) equals the nearest whole number greater than $x = (n+1)/2 + z\sqrt{0,25\,n}$, where z varies as a function of the significance level as follows: 0,84 for α = 0,20; 1,28 for α = 0,10; 1,64 for α = 0,05; 2,33 for α = 0,01; 3,09 for α = 0,001.

NOTE 2 The values of $n < 18$ are usually not recommended for paired difference tests.

---

2) The values given in this table have been calculated from the exact formula of the binomial distribution for parameter p = 0,5 with n replications thanks to the SAS software developed in Reference [4].

3) The values correspond to the minimum number of correct responses required for significance at the stated α-level (i.e. column) for the corresponding number of assessors, n (i.e. row). Reject the "no difference" affirmation if the number of correct responses is greater than or equal to the value in the table.

Table A.2 — Minimum number of consensual responses required to conclude that a perceptible difference exists, based on a two-sided paired test [2), 3)]

n	α					n	α				
	0,20	0,10	0,05	0,01	0,001		0,20	0,10	0,05	0,01	0,001
10	8	9	9	10		36	23	24	25	27	29
11	9	9	10	11	11	37	23	24	25	27	29
12	9	10	10	11	12	38	24	25	26	28	30
13	10	10	11	12	13	39	24	26	27	28	31
14	10	11	12	13	14	40	25	26	27	29	31
15	11	12	12	13	14						
						44	27	28	29	31	34
16	12	12	13	14	15	48	29	31	32	34	36
17	12	13	13	15	16	52	32	33	34	36	39
18	13	13	14	15	17	56	34	35	36	39	41
19	13	14	15	16	17	60	36	37	39	41	44
20	14	15	15	17	18						
						64	38	40	41	43	46
21	14	15	16	17	19	68	40	42	43	46	48
22	15	16	17	18	19	72	42	44	45	48	51
23	16	16	17	19	20	76	45	46	48	50	53
24	16	17	18	19	21	80	47	48	50	52	56
25	17	18	18	20	21						
						84	49	51	52	55	58
26	17	18	19	20	22	88	51	53	54	57	60
27	18	19	20	21	23	92	53	55	56	59	63
28	18	19	20	22	23	96	55	57	59	62	65
29	19	20	21	22	24	100	57	59	61	64	67
30	20	20	21	23	25						
						104	60	61	63	66	70
31	20	21	22	24	25	108	62	64	65	68	72
32	21	22	23	24	26	112	64	66	67	71	74
33	21	22	23	25	27	116	66	68	70	73	77
34	22	23	24	25	27	120	68	70	72	75	79
35	22	23	24	26	28						

NOTE 1 The values in the table are exact because they are based on the binomial distribution. For values of n not included in the table, an approximation of the missing entries may be obtained in the following manner: Minimum number of responses (x) is the nearest whole number greater than $x = (n+1)/2 + z\sqrt{0,25\,n}$, where z varies as a function of the significance level as follows: 1,28 for α = 0,20; 1,64 for α = 0,10; 1,96 for α = 0,05; 2,58 for α = 0,01; 3,29 for α = 0,001.

NOTE 2 The values of n < 18 are usually not recommended for paired difference tests.

Table A.3 — Maximum number of correct or consensual responses required to conclude that two samples are similar, based on a paired test [4), 5)]

n	β	p_d					n	β	p_d				
		10 %	20 %	30 %	40 %	50 %			10 %	20 %	30 %	40 %	50 %
18	0,001	—	—	—	—	—	60	0,001	—	—		—	33
	0,01	—	—	—	—	—		0,01	—	—	—	33	36
	0,05	—	—	—	—	9		0,05	—	—	32	35	38
	0,10	—	—	—	9	10		0,10	—	30	33	36	40
	0,20	—	—	9	10	11		0,20	—	32	35	38	41
24	0,001	—	—	—	—	—	66	0,001	—	—	—	—	37
	0,01	—	—	—	—	12		0,01	—	—	33	36	40
	0,05	—	—	—	12	13		0,05	—	—	35	39	43
	0,10	—	—	12	13	14		0,10	—	34	37	40	44
	0,20	—	—	13	14	15		0,20	—	35	39	42	46
30	0,001	—	—	—	—	—	72	0,001	—	—	—	37	40
	0,01	—	—	—	—	16		0,01	—	—	36	40	44
	0,05	—	—	—	16	17		0,05	—	—	39	43	47
	0,10	—	—	15	17	18		0,10	—	37	41	44	48
	0,20	—	15	16	18	20		0,20	—	39	42	46	50
36	0,001	—	—	—	—	—	78	0,001	—	—	—	40	44
	0,01	—	—	—	18	20		0,01	—	—	40	44	48
	0,05	—	—	18	20	22		0,05	—	39	43	47	51
	0,10	—	—	19	21	23		0,10	—	40	44	48	53
	0,20	—	18	20	22	24		0,20	—	42	46	50	54
42	0,001	—	—	—	—	21	84	0,001	—	—	—	44	48
	0,01	—	—	—	21	24		0,01	—	—	43	48	53
	0,05	—	—	21	23	26		0,05	—	42	46	51	55
	0,10	—	—	22	25	27		0,10	—	44	48	52	57
	0,20	—	22	24	26	28		0,20	—	46	50	54	59
48	0,001	—	—	—	—	25	90	0,001	—	—	—	48	53
	0,01	—	—	—	25	28		0,01	—	—	47	52	57
	0,05	—	—	25	27	30		0,05	—	45	50	55	60
	0,10	—	—	26	28	31		0,10	—	47	52	56	61
	0,20	—	25	27	30	33		0,20	45	49	54	58	63
54	0,001	—	—	—	—	29	96	0,001	—	—	—	52	57
	0,01	—	—	—	29	32		0,01	—	—	50	56	61
	0,05	—	—	28	31	34		0,05	—	49	54	59	64
	0,10	—	27	30	32	35		0,10	—	50	55	60	66
	0,20	—	28	31	34	37		0,20	48	53	58	62	68

4) The values given in this table have been obtained thanks to the program based on the calculation of confidence intervals from the exact formula of the binomial distribution, developed in Reference [7].

5) The values correspond to the maximum number of correct or consensual responses required for "similarity" at the chosen levels of p_d, β and n. Accept the "no difference" assumption at the $100(1 - \beta)$ % level of confidence if the number of correct or consensual responses is less than or equal to the value in the table.

Table A.3 (*continued*)

n	β	p_d					n	β	p_d				
		10 %	20 %	30 %	40 %	50 %			10 %	20 %	30 %	40 %	50 %
102	0,001	—	—	—	55	61	120	0,001	—	—	61	67	73
	0,01	—	—	54	59	65		0,01	—	—	65	71	78
	0,05	—	52	57	63	68		0,05	—	62	68	75	81
	0,10	—	54	59	64	70		0,10	—	64	70	77	83
	0,20	51	56	61	67	72		0,20	60	67	73	79	85
108	0,001	—	—	54	59	65	126	0,001	—		64	70	77
	0,01	—	—	57	63	69		0,01	—	—	68	75	82
	0,05	—	55	61	67	72		0,05	—	66	72	79	85
	0,10	—	57	63	68	74		0,10	—	68	74	81	87
	0,20	54	60	65	71	76		0,20	64	70	76	83	89
114	0,001	—	—	57	63	69	132	0,001	—	—	67	74	81
	0,01	—	—	61	67	73		0,01	—	65	72	79	86
	0,05	—	59	65	71	77		0,05	—	69	76	83	90
	0,10	—	61	67	72	79		0,10	—	71	78	85	92
	0,20	57	63	69	75	81		0,20	67	73	80	87	94

NOTE 1 The values in the table are exact because they are based on the binomial distribution. For the values of n not included in the table, compute the 100(1 – β) % upper confidence limit for p_d, as follows:

$$\left[2\left(x/n\right)-1\right]+2\times z_{\beta}\sqrt{\left(nx-x^{2}\right)/n^{3}}$$

where x is the number of correct or consensual responses, n the number of assessors and z_β varies as follows: 0,84 for β = 0,20; 1,28 for β = 0,10; 1,64 for β = 0,05; 2,33 for β = 0,01; 3,09 for β = 0,001. If the computed value is lower than the preselected limit for p_d, then declare the samples similar at the β significance level.

NOTE 2 The values of n < 30 are usually not recommended for paired similarity tests.

NOTE 3 The values corresponding to numbers of correct responses under n/2 are not mentioned in this table. They are coded by the sign —.

A.2 Statistical approach for the determination of the number of assessors on the basis of Tables A.4 (one-sided test) and A.5 (two-sided test)

The statistical sensitivity of the test depends on three values: the α-risk, the β-risk and the maximum authorized proportion of "distinguishers" p_d[6]. Prior to conducting the test, select the values for α, β and p_d using the following guidelines.

As a general rule, a statistically significant result for an α-risk:

— between 10 % and 5 % (0,10 to 0,05) indicates slight evidence that a difference was apparent;

— between 5 % and 1 % (0,05 to 0,01) indicates moderate evidence that a difference was apparent;

— between 1 % and 0,1 % (0,01 to 0,001) indicates strong evidence that a difference was apparent; and

— below 0,1 % (< 0,001) indicates very strong evidence that a difference was apparent.

6) In this International Standard, the probability of a correct response, p_c, is modelled as $p_c = 1 \times p_d + (1/2) \times (1 - p_d)$, where p_d is the proportion of the entire population of assessors able to distinguish between the two samples. A psychometrical model of the assessor's decision-making process, such as the Thurstone-Ura model (Reference [3]), could also be applied in the case of the paired test.

For β-risks, the strength of the evidence that a difference was not apparent is assessed using the same criteria as those specified above (replacing “was apparent” by “was not apparent”).

The maximum authorized proportion of “distinguishers”, p_d, falls into three ranges:

— $p_d < 25$ % represents small values;

— 25 % $< p_d <$ 35 % represents medium sized values; and

— $p_d > 35$ % represents big values.

Choose the number of assessors so as to obtain the level of sensitivity required by the test. Identify in Table A.4 the section corresponding to the selected value of p_d and the column corresponding to the selected value of β. The minimum required number of assessors is therefore located in the row corresponding to the selected value of α. Alternatively, Table A.4 can be used to develop a set of values for p_d, α and β that provide acceptable sensitivity while maintaining the number of assessors within practical limits. This approach is presented in detail in Reference [4].

Table A.4 — Number of assessors required for a one-sided paired test [7), 8)]

α		β					
		0,50	0,20	0,10	0,05	0,01	0,001
0,50		— [a]	—	—	9	22	33
0,20		—	12	19	26	39	58
0,10	p_d = 50 %	—	19	26	33	48	70
0,05		13	23	33	42	58	82
0,01		35	40	50	59	80	107
0,001		38	61	71	83	107	140
0,50		—	—	9	20	33	55
0,20		—	19	30	39	60	94
0,10	p_d = 40 %	14	28	39	53	79	113
0,05		18	37	53	67	93	132
0,01		35	64	80	96	130	174
0,001		61	95	117	135	176	228
0,50		—	—	23	33	59	108
0,20		—	32	49	68	110	166
0,10	p_d = 30 %	21	53	72	96	145	208
0,05		30	69	93	119	173	243
0,01		64	112	143	174	235	319
0,001		107	172	210	246	318	412
0,50		—	23	45	67	133	237
0,20		21	77	112	158	253	384
0,10	p_d = 20 %	46	115	168	214	322	471
0,05		71	158	213	268	392	554
0,01		141	252	325	391	535	726
0,001		241	386	479	556	731	944
0,50		—	75	167	271	539	951
0,20		81	294	451	618	1 006	1 555
0,10	p_d = 10 %	170	461	658	861	1 310	1 905
0,05		281	620	866	1 092	1 583	2 237
0,01		550	1 007	1 301	1 582	2 170	2 927
0,001		961	1 551	1 908	2 248	2 937	3 812

[a] The empty boxes correspond to cases which do not present any practical interest (high values for α and β taking into account the selected value of p_d)

7) The values given in this table have been taken from Reference [4] or computed from the exact formula of the binomial law for parameter p = 0,5 with n responses thanks to the SAS software developed in that reference.

8) The values correspond to the minimum number of assessors required to perform a paired test with a specified level of sensitivity determined by the values of p_d, α and β. Identify in the table the section corresponding to the selected value of p_d and the column corresponding to the selected value of β. Read the minimum number of assessors from the row corresponding to the selected value of α.

Table A.5 — Number of assessors required for a two-sided paired test [8), 9)]

α		β					
		0,50	0,20	0,10	0,05	0,01	0,001
0,50	p_d = 50 %	— [a]	—	—	23	33	52
0,20		—	19	26	33	48	70
0,10		—	23	33	42	58	82
0,05		17	30	42	49	67	92
0,01		26	44	57	66	87	117
0,001		42	66	78	90	117	149
0,50	p_d = 40 %	—	—	25	33	54	86
0,20		—	28	39	53	79	113
0,10		18	37	53	67	93	132
0,05		25	49	65	79	110	149
0,01		44	73	92	108	144	191
0,001		48	102	126	147	188	240
0,50	p_d = 30 %	—	29	44	63	98	156
0,20		21	53	72	96	145	208
0,10		30	69	93	119	173	243
0,05		44	90	114	145	199	276
0,01		73	131	164	195	261	345
0,001		121	188	229	267	342	440
0,50	p_d = 20 %	—	63	98	135	230	352
0,20		46	115	168	214	322	471
0,10		71	158	213	268	392	554
0,05		101	199	263	327	455	635
0,01		171	291	373	446	596	796
0,001		276	425	520	604	781	1 010
0,50	p_d = 10 %	—	240	393	543	910	1 423
0,20		170	461	658	861	1 310	1 905
0,10		281	620	866	1 092	1 583	2 237
0,05		390	801	1 055	1 302	1 833	2 544
0,01		670	1 167	1 493	1 782	2 408	3 203
0,001		1 090	1 707	2 094	2 440	3 152	4 063

[a] The empty boxes correspond to cases which do not present any practical interest (high values for α and β taking into account the selected value of p_d).

9) The values given in this table have been computed from the exact formula of the binomial law for parameter p = 0,5 with n responses thanks to the SAS software developed in Reference [4].

Annex B
(informative)

Examples

B.1 Example 1 — One-sided paired test to confirm that a difference exists concerning the intensity of an attribute between two samples

B.1.1 Context

Following some remarks made by consumers, some technological modifications have been made in order to produce a crispier biscuit that the usual product. Before proceeding to a larger scale preference test involving consumers, the development department wishes to ascertain that the technological modifications have provided the desired effect. It wishes to limit the risk of concluding in favour of a difference that does not exist. On the other hand, since it has the possibility of making other technological modifications, it is ready to accept a high risk of not detecting a difference which exists.

B.1.2 Test objective

This is to confirm that the new product is indeed crispier. It is therefore a case for a one-sided test.

B.1.3 Number of assessors

In order to prevent the development department from wrongly concluding in favour of a difference which would not exist, the sensory analysis supervisor proposes an α threshold of 0,05, a percentage of assessors detecting the difference p_{d} equal to 30 % and a β of 0,50. It therefore consults Table A.4 and finds that at least 30 assessors are required.

B.1.4 Conducting the test

Thirty plates with a biscuit “A” (control) and 30 plates with a biscuit “B” (prototype) are coded with unique random numbers. For 15 assessors, the products are presented in the order AB, for the 15 others in the order BA. A specimen scoresheet is shown in Figure B.1.

B.1.5 Analysis and interpretation of results

Twenty-one assessors designate sample B as being crispier. Referring to Table A.1 in the row corresponding to n = 30 and in the column α = 0,05, it can be seen that 20 responses in the expected direction suffice to declare the 2 samples significantly different.

B.1.6 Report and conclusions

The sensory analyst reports that the prototype appeared to be crispier for the panel (n = 30, x = 21) at a 5 % significance level. Biscuits may therefore be manufactured with the new process for preference tests with consumers.

Paired test

Name: .... Assessor code:.................................. Date: ....

Instructions:

Taste the two samples beginning with the one on your left. Indicate the code of the sample which is the crispiest in the space below. If you are not sure, take a guess; you can indicate under the heading "Comments" that it is a guess.

The crispiest sample is: ....

Possible comments: ....

....

....

Figure B.1 — Scoresheet for Example 1

B.2 Example 2 — One-sided paired test to confirm whether two samples are similar concerning a given attribute

B.2.1 Context

A manufacturer knows that the product may contain traces of an ingredient which imparts a herbaceous off-flavour to the product. He therefore wishes to determine the maximum acceptable quantity so that the flavour difference with a reference product without this ingredient (T) is barely perceptible and therefore without any regrettable consequence.

B.2.2 Test objective

This is to determine the maximum acceptable quantity of the ingredient so that the herbaceous flavour difference with a reference product without this ingredient is barely perceptible and therefore without any regrettable consequence.

B.2.3 Number of assessors

The manufacturer wishes to be reasonably sure of the specifications concerning the permissible quantity of the ingredient responsible for a herbaceous off-flavour. Thus, in this test, the risk of not detecting a difference in herbaceous flavour (β) has to be kept as low as possible. The α-risk of wrongly concluding in favour of the existence of a difference which would not exist is of lesser importance, since it would only lead to a more conservative specification. β is therefore fixed at 0,05, α at 0,50 and the percentage of assessors detecting the difference p_d is fixed at 20 %. The manufacturer therefore consults Table A.4 and finds that at least 67 assessors are required. However, on consulting Table A.3, it is noted that for the selected values of β and p_d, a minimum number of 78 assessors are required in order to be able to use the table (in the cases below 78, the maximum numbers of proposed correct or consensual responses are below chance, i.e. $n/2$, and therefore do not figure in the table). The manufacturer therefore decides to recruit 78 assessors.

B.2.4 Conducting the test

Taking into account preliminary tests and previous knowledge, a target concentration C is defined. The two solutions are prepared and each one is divided up into 78 plastic cups coded with unique random numbers. For 39 assessors, the products are presented in the order TC, for the 39 others in the order CT. A specimen scoresheet is shown in Figure B.1 but, in this case, it is a question of comparing the samples with regard to the herbaceous characteristic.

B.2.5 Analysis and interpretation of results

Seventy-eight assessors take part in the test. Forty-one assessors designate sample C as being more herbaceous. In Table A.3, it can be seen that for n = 78 and for a p_d of 20 % and a β of 0,05, the maximum value for concluding that the samples are similar is equal to 39. As the number of correct responses obtained during the test is above this value, the analyst notes that it is not possible to conclude in favour of a similarity between the samples.

B.2.6 Report and conclusions

The sensory analyst notes in the report that concentration C is too strong to be accepted as a tolerated permissible value and recommends to test a slightly lower concentration.

B.3 Example 3 — Two-sided paired test to confirm that a difference exists concerning the intensity of an attribute between two samples

B.3.1 Context

A soup manufacturer wishes to determine which of two sodium-based ingredients provides the most intense salty flavour. This ingredient will be selected from the formulation of a new production as it may be used at a weaker concentration and will be less expensive (both products costing the same price per kilogram). If no significant difference appears, other ingredients will be tested.

B.3.2 Test objective

This is to determine which of two ingredients provides the most intense salty flavour for a same concentration. It is therefore a case for a two-sided test.

B.3.3 Number of assessors

The sensory analyst wishes to be 95 % certain that a high proportion of assessors will be able to perceive a difference. α is therefore fixed at 0,05 and p_d at 50 %. However, to conclude wrongly that no difference exists will result in additional costs since other ingredients will have to be tested. Consequently, the sensory analyst fixes β at 0,10. On consulting Table A.5, it is found that at least 42 assessors are required. It is therefore decided to recruit 44.

B.3.4 Conducting the test

Two batches of soup A and B are prepared, their sole difference being the ingredient that provides the salty flavour. The two samples are presented hot in ceramic bowls coded with unique random numbers. For 22 assessors, the soups are presented in the order AB, for the 22 others in the order BA. A specimen scoresheet is shown in Figure B.2.

B.3.5 Analysis and interpretation of results

Forty-four assessors take part in the test. Thirty-two designate sample A as being saltier and 12 designate sample B as being saltier. Referring to Table A.2 in the row corresponding to n = 44 and in the column α = 0,05, it can be seen that the highest of the two numbers obtained for A or B must be greater than or equal to 29 in order to declare the 2 samples significantly different. Table A.2 also shows that a value greater than or equal to 31 means that a significant difference at the 1 % threshold can be concluded.

B.3.6 Report and conclusions

The sensory analyst notes in the report that ingredient A is perceived as being saltier than ingredient B at the 1 % threshold. This ingredient A is therefore adopted for the future production.

Paired test

Test code: 845-2003

Assessor No. 14 Name:.................................... Date:.............................

Instructions:

Taste the two samples from left to right. Select the saltiest sample and identify it by placing a cross in the corresponding box.

Comments:

842 ☐ 376 ☐

....

....

If you wish to make a comment concerning the reasons for your choice or the characteristics of the samples, you can insert it under the "Comments" heading.

Figure B.2 — Scoresheet for Example 3

B.4 Example 4 — Two-sided paired test to confirm whether two samples are similar concerning a given attribute

B.4.1 Context

A manufacturer of plastics used in particular by car manufacturers for dashboards is seeking, for economic reasons, to replace the usual lubricant by a new one, but does not wish that the new plastics formula be perceived as presenting less or more surface slip than the usual one.

B.4.2 Test objective

This is to determine whether, for a same concentration, the new lubricant provides the same "surface slip" level as the usual product.

B.4.3 Number of assessors

The manufacturer wishes to be reasonably certain that the new ingredient provides the same level of perceived "surface slip" as the usual lubricant. Thus, in this test, the risk of not detecting a difference in surface slip must be kept as low as possible. The α-risk of wrongly concluding in favour of the existence of a difference which would not exist is nevertheless of importance, because it would lead to keeping the usual more expensive lubricant. β is therefore fixed at 0,05, α at 0,10 and the percentage of assessors detecting the difference p_d is fixed at 30 %. The manufacturer therefore consults Table A.5 and finds that at least 119 assessors are required. In order to balance out the orders of presentation, it is decided to recruit 120 assessors.

B.4.4 Conducting the test

Two batches of plastic sheets A and B are prepared, their sole difference being the lubricant which provides the surface slip characteristic. The two sheets are presented in boxes coded with unique random numbers. For 60 assessors the samples are presented in the order AB, for the 60 others in the order BA. A specimen scoresheet is shown in Figure B.3.

B.4.5 Analysis and interpretation of results

The 120 recruited assessors take part in the test. Of these, 67 assessors designate sample A as presenting more surface slip and 53 designate sample B. Referring to Table A.3 in the section corresponding to n = 120, the row corresponding to a β of 0,05 and in the column p_d = 30 %, the value 68 is found. The highest of the two numbers obtained (67) being less than this value (68), it can therefore be concluded that a similarity exists between the two samples at the 95 % confidence interval and at the value of p_d selected for the test.

B.4.6 Report and conclusions

The sensory analyst reports that the new lubricant provides a level of surface slip similar to the usual lubricant and can therefore be chosen for the future production.

Paired test

Name: .... Assessor code: .... Date: ....

Instructions:

You are requested to compare the 2 samples from the **surface slip** characteristic standpoint. A sheet with surface slip is one which does not present any resistance to the tangential movement of the hand over its surface. Please touch both products, beginning with the one on your left, and observing the following protocol: movement of the flat of the hand, applying a slight pressure, from left to right.

Select the sample with the most surface slip and identify it by placing a cross in the corresponding box.

192 ☐ 526 ☐

Figure B.3 — Scoresheet for Example 4

B.5 Example 5 — Confidence intervals for paired tests

B.5.1 Context

If desired, analysts can calculate a confidence interval on the proportion of assessors actually able to differentiate the 2 samples (i.e. corrected for chance). The calculations, based here on approximation by normal distribution, are as follows, where x is the number of responses in the expected direction (one-sided test) or number of consensual responses (two-sided test) and n is the total number of responses:

— p_c (proportion of expected or consensual responses) = x/n

— $\hat{p}_d$ (proportion of distinguishers) = $2\,p_c - 1$

— s_d (standard error of $\hat{p}_d$) = $2\sqrt{p_c(1-p_c)/n} = 2\sqrt{\left(n \times x - x^2\right)/n^3}$

— upper confidence limit = $\hat{p}_d + z_\alpha s_d$

— lower confidence limit = $\hat{p}_d - z_\alpha s_d$

where z_α is the critical value of the standardized centred normal distribution given in the table below:

Level	Test	
	z_α one-sided	z_α two-sided
80 %	0,84	1,28
90 %	1,28	1,64
95 %	1,64	1,96
99 %	2,33	2,58
99,9 %	3,10	3,29

B.5.2 Analysis and interpretation of results

In Example 3 (two-sided paired difference test), the data were as follows: n = 44, x = 32, α = 0,05 and β = 0,10. It follows from this that

— $p_c = x/n$ = 32/44 = 0,73

— $\hat{p}_d$ (proportion of distinguishers) = 2 p_c – 1 = 2 × 0,73 – 1 = 0,45

— s_d (standard error of $\hat{p}_d$) = $2\sqrt{\left(n \times x - x^2\right)/n^3} = 2 \times \sqrt{\left(32 \times 44 - 32^2\right)/44^3}$ = 0,134

— 95 % upper confidence limit = $\hat{p}_d + z_\alpha s_d$ = 0,45 + 1,96 × 0,134 = 0,71

— 95 % lower confidence limit = $\hat{p}_d - z_\alpha s_d$ = 0,45 – 1,96 × 0,134 = 0,19

The sensory analyst can therefore be 95 % certain that at least 19 % of the population is capable of distinguishing the samples. This result concords with the conclusion given in Example 3, indicating sample A as being saltier, since it shows that the confidence interval does not contain the null value and goes beyond the limit initially fixed for the test (p_d at 50 %).

In Example 4 (two-sided paired similarity test), the data were as follows: n = 120, x = 67, α = 0,10 and β = 0,05. It follows from this that

— $p_c = x/n$ = 67/120 = 0,56

— $\hat{p}_d$ (proportion of distinguishers) = 2 p_c – 1 = 2 × 0,56 – 1 = 0,12

— s_d (standard error of $\hat{p}_d$) = $2\sqrt{\left(n \times x - x^2\right)/n^3} = 2 \times \sqrt{\left(120 \times 67 - 67^2\right)/120^3}$ = 0,09

— 95 % upper confidence limit = $\hat{p}_d + z_\alpha s_d$ = 0,12 + 1,96 × 0,09 = 0,29

— 95 % lower confidence limit = $\hat{p}_d - z_\alpha s_d$ = 0,12 – 1,96 × 0,09 = – 0,06

The sensory analyst can therefore be 95 % certain that the actual proportion of the population capable of distinguishing the samples is not over 29 %.

Taken together, the confidence intervals authorise a 5 % error for the upper and lower limits, therefore the sensory analyst can be 90 % certain that the proportion of distinguishers is between 0 % and 29 % of the population. This is below the initially fixed limits (p_d of 30 %). The analyst can therefore conclude that there is no perceptible difference between the samples.

Bibliography

[1] ENNIS, D.M. and BI, J. The Beta-Binomial Model: Accounting for Inter-trial Variation in Replicated Difference and Preference Tests. *Journal of Sensory Studies*, **13**, No. 4, 1998, pp. 389-412

[2] BROCKHOFF, P.B. and SCHLICH, P. Handling replications in discrimination tests. *Food Quality and Preference,* **9**, No. 5, 1998, pp. 303-312

[3] FRIJTERS, J.E.R. Three-Stimulus Procedure in Olfactory Psycholophysics: An Experimental Comparison of Thurstone-Ura and Three-Alternative Forced-Choice Models of Signal Detection Theory. *Perception and Psychophysics*, **28**, No. 5, 1980, pp. 390-397

[4] SCHLICH, P. Risk Tables for Discrimination Tests. *Food Quality and Preference*, **4**, 1993, pp. 141-151

[5] KÜNERT, J. and MEYNERS, M. On the triangle test with replications. *Food Quality and Preference*, **10**, 1999

[6] KUNERT, J. On repeated difference testing. *Food Quality and Preference*, **12**, 2001, pp 385-391

[7] MACRAE, A. Confidence intervals for the triangular test. *Food Quality and Preference*, **6**, 1995, pp. 61-67

[8] ISO 3534-1, *Statistics — Vocabulary and symbols — Part 1: Probability and general statistical terms*

ICS 67.240

Price based on 21 pages

INTERNATIONAL STANDARD

ISO 5077

Second edition
2007-02-15

Textiles — Determination of dimensional change in washing and drying

Textiles — Détermination des variations dimensionnelles au lavage et au séchage domestiques

---

Reference number
ISO 5077:2007(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

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.

ISO 5077 was prepared by Technical Committee ISO/TC 38, *Textiles*, Subcommittee SC 2, *Cleansing, finishing and water resistance tests*.

This second edition cancels and replaces the first edition (ISO 5077:1984), which has been technically revised.

Textiles — Determination of dimensional change in washing and drying

1 Scope

This International Standard specifies a method for the determination of the dimensional change of fabrics, garments or other textile articles when subjected to an appropriate combination of specified washing and drying procedures.

In the case of textile articles or deformable materials, it is necessary to exercise all possible caution in the interpretation of the results.

2 Normative references

The following referenced documents are indispensable for the application 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 139, *Textiles — Standard atmospheres for conditioning and testing*

ISO 3759, *Textiles — Preparation, marking and measuring of fabric specimens and garments in tests for determination of dimensional change*

ISO 6330, *Textiles — Domestic washing and drying procedures for textile testing*

3 Principle

The specimen is conditioned in the specified standard atmosphere and measured before subjection to the appropriate washing and drying procedures. After drying, conditioning and remeasuring of the specimen, the changes in dimensions are calculated.

4 Apparatus and reagents

Use apparatus and reagents as specified in ISO 3759 and ISO 6330.

5 Atmospheric conditions

The atmospheric conditions required for conditioning and testing are specified in ISO 139.

6 Test specimens

6.1 The selection, dimensions, marking and measuring of test specimens are specified in ISO 3759.

6.2 When possible, three specimens from each sample should be used. One or two specimens may be used when insufficient sample is available.

7 Procedure

7.1 Determine the original length and width dimensions, as appropriate, after the specimens have been conditioned and measured according to the procedure specified in ISO 139 and ISO 3759.

7.2 Wash and dry the specimens according to one of the procedures specified in ISO 6330, as agreed between the interested parties.

7.3 After washing and drying, condition and measure the specimens and calculate the dimensional change of the specimens according to the procedure specified in ISO 3759.

8 Expression of results

8.1 Calculate the mean changes in dimensions in both the length and width directions in accordance with the arrangement in ISO 3759 as follows:

$$\frac{x_\text{t} - x_\text{o}}{x_\text{o}} \times 100$$

where

x_o is the original dimension;

x_t is the dimension measured after treatment.

Record the changes in measurement separately as a percentage of the corresponding original value.

8.2 Express the average dimensional changes to the nearest 0,5 %.

8.3 State whether the dimension has decreased (shrinkage) by means of a minus sign ($-$) or increased (extension) by means of a plus sign ($+$).

9 Test report

The test report shall specify the following:

a) the number and year of this International Standard;

b) the number of specimens washed and dried;

c) the procedure used for washing and drying from ISO 6330;

d) for fabric specimens, the average dimensional change in the length (warp or wale) and the average dimensional change in the width (weft or course) to the nearest 0,5 %;

e) for garments, the description, make and size of the garment tested;

f) for garments, an adequate description of each measuring position and the average dimensional change to the nearest 0,5 % at each position for each garment tested.