E

TWC/22/8
ORIGINAL: English
DATE: May 24, 2004

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS
GENEVA

TECHNICAL WORKING PARTY ON AUTOMATION AND COMPUTER PROGRAMS

Twenty-Second Session
Tsukuba, Japan, June 14 to 17, 2004

ASSESSMENT OF DISTINCTNESS FOR SEGREGATING CHARACTERISTICS

Document prepared by experts from France and United Kingdom

ASSESSMENT OF DISTINCTNESS FOR SEGREGATING CHARACTERISTICS

Authors: Vincent Gensollen (FR), Sylvain Grégoire (FR), Sally Watson (UK)

1. At the twenty-first session of the TWC, held in Tjele, Denmark, from June 10 to 13, 2003, the TWC agreed that a revised version of document TWC/21/2 would be prepared for the twenty-second session of the TWC to be held in Japan in 2004. The revision should include a comparison of Chi-square and Fisher exact tests. This document presents the Chi-square test and the Fisher exact test. A brief comparison is also done.

2. Within a variety, variability can occur due to genetic and/or environmental variation. This is the case in particular for cross-pollinated, including synthetic, varieties. For such varieties the expression of a characteristic should be recorded using more than one observation. In general, records are taken from a number of individual plants (see “TGP/9”).

3. For quantitative characteristics the statistical method recommended by UPOV for the assessment of distinctness is COYD[1] analysis, which takes into account variation between years (for details see TGP/9 Draft 1, section 5.3.2). UPOV has proposed several statistical methods to deal with uniformity in measured quantitative characteristics. One method which takes into account variation between years is the COYU[2] method (for details see TGP/10.3.1).

4. For qualitative characteristics and pseudo-qualitative characteristics two cases are possible (see Annex 1 for examples).

5. In the first case, all the plants of a given variety express the same state of expression for a characteristic, i.e. they are homogeneous in their expression of the characteristic. For example, all plants of a variety express the same state of the characteristic “Sex of plant” (e.g. dioecious female (1), dioecious male (2), monoecious unisexual (3), or monoecious hermaphrodite (4)). The ploidy level is another example of this kind of characteristic (e.g. diploid (2), tetraploid (4), or hexaploid (6)).

6. In this first case, the characteristic could be used as a grouping characteristic. In cross-pollinated, including synthetic, varieties these characteristics are rather rare but very useful; UPOV recommends the assessment of uniformity by the off-type procedure in order to prevent breeders from creating varieties which are heterogeneous in these characteristics. Distinctness between two varieties can be established when they express different states in such a characteristic.

7. In the second case, plants of a given variety can have different states of expression for a characteristic, i.e. they are heterogeneous in their expression of the characteristic. These characteristics are important for distinctness purposes because the frequency of plants expressing the different states in a variety can be very consistent and so helpful in determining distinctness. For example, in Lucerne the frequency of occurrence of plants with the different states of the “flower color” characteristic (white or yellow (1), violet (2), very dark violet (3), variegated (4)) is used to show distinctness between varieties.

---

[1] COYD : Combined Over Years Distinctness
[2] COYU : Combined Over Years Uniformity

8. In this second case, the Chi-square or the Fisher exact tests can be used to assess distinctness (SNEDECOR, G.W.; COCHRAN W. (1937); KANJI G. K. (1993)). The Chi-square or the Fisher exact tests compare the frequencies of plants expressing the different states of the characteristic in different varieties (see Annex 2). Such characteristics cannot be assessed for uniformity.

9. For assessing distinctness in pair-wise comparisons, the differences between varieties must be significant in either 2 out of 2 or 2 out of 3 successive cycles of examination. Significance is assessed by using the Chi-square or the Fisher exact tests at a UPOV recommended significance probability for the crop. The differences must also be of the same sign in the 2 cycles that have significant differences, e.g. variety A must have consistently more plants with variegated flowers than variety B.

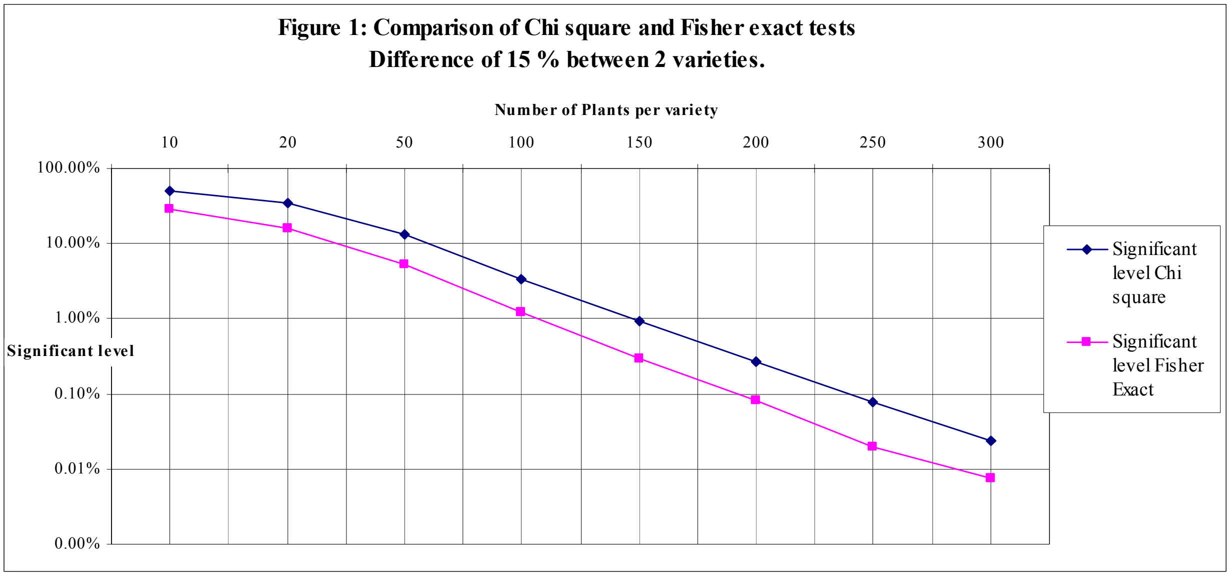
10. The main differences between the Chi-square and the Fisher exact tests are as follows:

- The two tests measure the difference between the observed data and the expected data. The Fisher exact test works in exactly the same way as the Chi-square test for independence, however, the Chi-square test gives only an estimate of the significance probability whereas the Fisher exact test is exact.

- The classical Chi-square estimate might not be very reliable where the number of plants in each column or row are very uneven or has a low value (less than five) in one of the cells. In that case the Fisher exact test is a good alternative for the Chi-square test. This could be a great advantage when the number of plants is small like in the DUS spaced plant trials (60 plants, in general).

- With a large number of plants or cases, the Chi-square test is preferred as the Fisher exact test is difficult to calculate. It is not advisable to use the Fisher exact test for which the number of cases is larger than 300 in a 2*3 table, 150 in a 2*4 table or 75 in a 2*5 table. In these cases the Chi-square test is rather robust, even in extreme cases (Uitenbroek, D.G. (1997)).

- The number of plants needed by the Fisher Exact test is lower than for the Chi-square test to reach a certain significant level. That is shown in the following figure. For example, the one percent significant level is reached at around 100 plants with the Fisher exact test whereas 150 plants are needed with the Chi-square test. This example comes from a comparison of 2 varieties with a difference of 15 % between the 2 classes (more examples will be presented at the TWC meeting).



11. When there are just two states (e.g. presence or absence), a characteristic can be analysed for distinctness either by using the Chi-square test or the Fisher exact test, as described above, or, they may be summarized for the variety quantitatively as a percentage presence. In the latter approach, the COYD criterion is applied to the variety-by-years table of the percentage presence of the characteristic (for details see TGP/9 Draft 1, Section 5.3.2).

12. Such characteristics cannot be assessed for uniformity. The COYU analysis, recommended by UPOV for uniformity, needs a quantitative note for each plant to perform the test. The reason is that the test is based on the standard deviation calculated for each variety in each replicate. Here we have for each variety and each replicate just the number of plants with a qualitative characteristic (for example the number of plants with inflorescences or without inflorescences). in other words, we have only one datum for each variety in each replicate and therefore we cannot calculate a standard deviation.

Remark: *When data are transformed into states of expression for the purpose of variety description, characteristics might appear to be quantitative. However, the statistical analysis is not made at this level of process. For example in the characteristic "Tendency to form inflorescences in year of sowing" the range of expressions is divided into: absent or very weak (1), weak (3), medium (5), strong (7), or very strong (9). The statistical analysis is made on the number of plants with inflorescences or without inflorescences.*

ANNEX I

List of Qualitative Characteristics and Descriptions of how they can be Used to Assess Distinctness and Uniformity

Name of characteristic	Distinctness			Uniformity		
	States for assess-ment	Description (states of expression)	Type of scale	Unit of assess-ment	Description (states of expression)	Type of scale
Sex of plant	1 2 3 4	dioecious female dioecious male monoecious unisexual monoecious hermaphrodite	nominal scale qualitative data [3]	True-type Off-type	Number of plants belonging to the variety Number of off-types	nominal scale qualitative data
Ploidy	2 4 6	diploid tetraploid hexaploid	nominal scale qualitative data [3]	True-type Off-type	Number of plants belonging to the variety Number of off-types	nominal scale qualitative data
Flower colour for Lucerne varieties	1 2 3 4	white or yellow violet very dark violet variegated	combination of ordinal and nominal scale qualitative data [4]		It is not possible to assess uniformity	
Tendency to form inflores-cences in year of sowing	1 9	absent present	nominal scale qualitative data [5]		It is not possible to assess uniformity	
Resistance to *Xanthomonas translucens (campestris) pv graminis* for Ryegrass varieties	1 9	dead plant living plant	nominal scale qualitative data [5]		It is not possible to assess uniformity	

[Annex II follows]

[3] Distinctness occurs when varieties express different states of expression.
[4] Distinctness can be assessed by applying the Chi-square test.
[5] Distinctness can be assessed by applying the Chi-square test or the COYD criterion on percentage.

ANNEX II

Examples of Application of the Chi-square Test to a Qualitative Characteristic

1. The Chi-square test can be applied to RxC tables where, for example, each of the R rows relates to one of the R different varieties which are being compared (in the case of DUS studies this is usually two, as all pair comparisons are computed), and each of the C columns relates to one of the C different states of expression of the characteristic. The value in the cell for variety *i* and state *j* is the observed frequency of plants in variety *i* which express state *j*. It should be noted that the roles of the rows and columns could equally well be reversed so that rows represented states and columns varieties.

2. The principle of the test is to compare the observed frequencies with the frequencies that would be expected if all R varieties, on average, had the same frequency of occurrence of the different states. This is the Null Hypothesis.

3. A Chi-square statistic is calculated as follows:-

$$\chi^2 = \sum_{i=1}^{R}\sum_{j=1}^{C}\frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

where:

O_{ij} denotes the observed number of plants (frequency) in variety *i* expressing state *j* of the characteristic.

E_{ij} denotes the expected number of plants (frequency) in variety *i* expressing state *j*.

E_{ij} is calculated as follows:-

Equation [1]

$$E_{ij} = \frac{\left(\text{Total number of plants for variety } i\right)\times\left(\text{Total number of plants expressing state } j\right)}{\text{Total number of plants for all varieties}}$$

4. The value of the Chi-square statistic is compared with critical values from Chi-square distribution tables on $(R-1)(C-1)$ degrees of freedom.

5. For a given significance probability, if the Chi-square statistic is less than the critical value, the varieties are declared to be not distinct, at that significance level, for the given characteristic. If the Chi-square statistic is greater than the critical value, then the varieties are declared to be distinct, at that significance level, for the given characteristic.

Example 1: Use of the Chi-square test to assess distinctness between two varieties of Rye-grass for the characteristic Resistance to *Xanthomonas translucens* pv *graminis,* using Excel software:

Observed frequencies of plants

	Variety X	Variety Y	*Total number of plants*
Number of dead plants	36	74	*110*
Number of living plants	67	29	*96*
Total number of plants	*103*	*103*	*206*

Expected frequencies of plants if the Null Hypothesis is true

	Variety X	Variety Y	*Total number of plants*
Number of dead plants	55	55	*110*
Number of living plants	48	48	*96*
Total number of plants	*103*	*103*	*206*

Chi-square statistic value = **28,17**

6. Of the 103 plants observed in Variety X, 36 plants were found to be dead, whilst for variety Y, 74 plants were found to be dead out of 103 plants observed. The difference between variety X and variety Y is significant at the 0.1% level. The above table shows the observed frequencies and the Chi-square statistic value. The expected frequencies are calculated using equation [1] above.

7. The table below, taken from an Excel worksheet, shows how each of the 4 cells contributes to the Chi-square statistic value. The Chi-square statistic value is the sum of the 4 contributions.

Example of Chi-square computation using Excel

OBSERVED

	variety X	variety Y	sum
dead	36	74	110
alive	67	29	96
sum	103	103	206

EXPECTED

	variety X	variety Y	sum
dead	55	55	110
alive	48	48	96
sum	103	103	206

Chi-square statistic = sum for each cell of (observed - expected)*(observed - expected)/ expected

Cell	**formula**	**value**
dead X	=(36-55)*(36-55)/55	6.56
alive X	=(67-48)*(67-48)/48	7.52
dead Y	=(74-55)*(74-55)/55	6.56
alive Y	=(29-48)*(29-48)/48	7.52
	(sum of 4 cells) **Chi-square**	28.17

degrees of freedom = (nR-1)*(nC-1) = (2-1)*(2-1) = 1
nR = number of rows = 2
nC = number of columns = 2

Null Hypothesis: on average the two varieties have the same frequencies

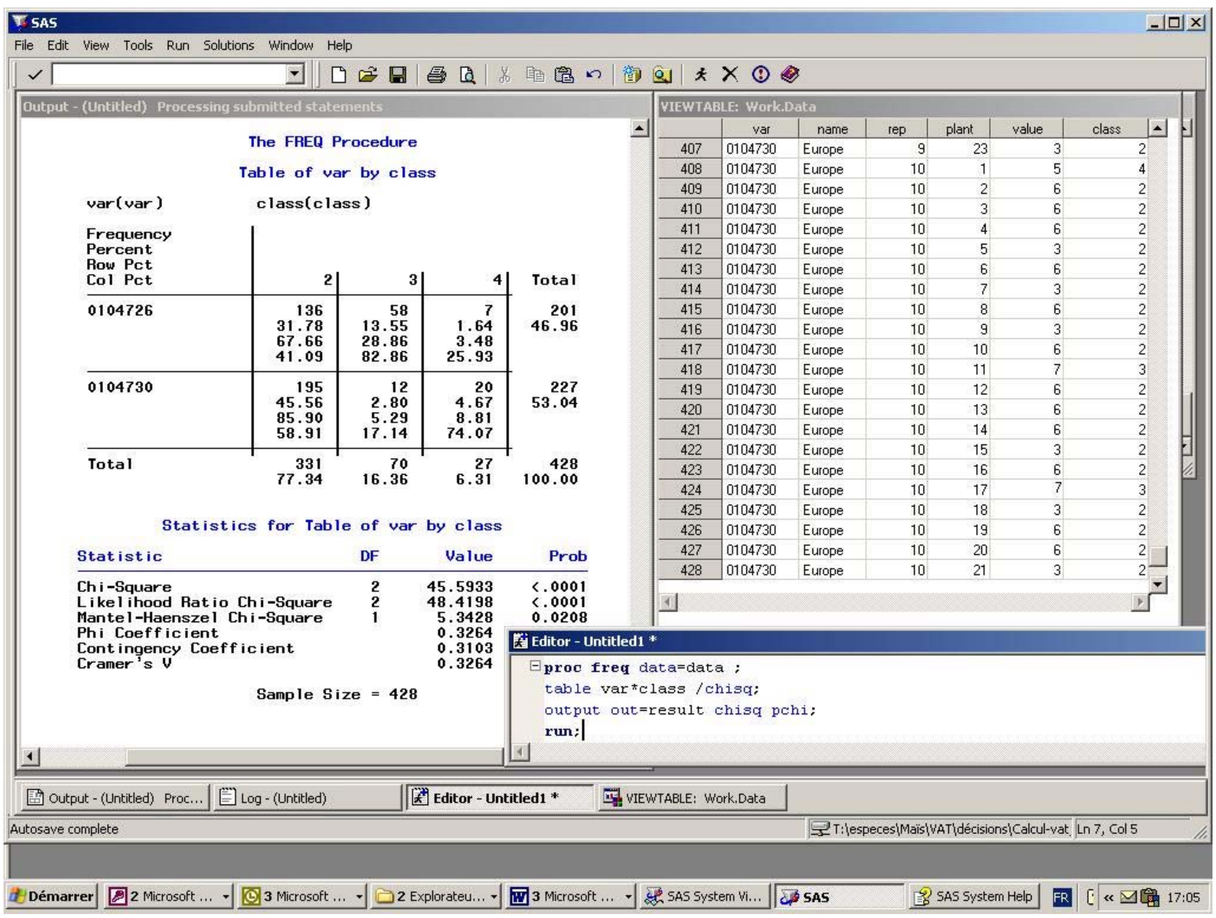
8. The Chi-square statistic is calculated as 28.17 on 1 degree of freedom. The 5%, 1% and 0.1% critical values of the Chi-square distribution on 1 degree of freedom are 3.841, 6.635 and 10.83 respectively. Hence Variety X and Variety Y are distinct at the 0.1% significance level.

Example 2: Output from Chi-square tests using SAS software to assess distinctness, between varieties of Lucerne, for the characteristic Flower Color.

9. Trials comprise candidate and reference varieties. Each candidate variety is tested in a pair-wise comparison with all candidates and all reference varieties.

10. In the screen copy below is an example in which the following are shown:

- a simple SAS program to obtain a Chi-square test from a dataset *(see editor window)*
- a partial view of the dataset, where "var" is the variety code (0104726 for Derby variety and 104730 for Europe variety) and "class" is the 4 states of the "flower colour" characteristic (white or yellow (1), violet (2), very dark violet (3), variegated (4)). "Var" and "class" are used to obtain the rows and columns of the table of frequencies *(see Viewtable window).*
- a default output for a Chi-square test which shows the frequencies and the Chi-square value (45.59) for a comparison of Derby and Europe on 10 replicates of at least 20 plants per replicate. In this example 201 plants has been observed for Derby and 227 plants for Europe.*(see output window)*



11. In general, it is recommended that the total number of plants in each column or row in a table of frequencies to be used in a Chi-square test, should have more than 5 of the total number of plants observed. Grouping of characteristic states can be done in order to avoid a situation where a column (or row) has less than 5 of the total number of plants observed. In this case the

test procedure is as before but uses a table with fewer columns or rows and consequently there are fewer degrees of freedom for the Chi-square test.

12. The value of the Chi-square statistic depends linearly on the number of plants observed. This means that, as is usually the case with other statistical tests, if the sample size is increased, smaller differences in percentages between 2 varieties can be shown to be “significant”.

ANNEX III

Examples of Application of the Fisher Exact Test to a Qualitative Characteristic

1. It is assumed in this case that the Fisher exact test is applied to a 2 x 2 table where, for example, each of the 2 rows relates to one of the 2 different varieties which are being compared (in the case of DUS studies this is usually two, as all pair comparisons are computed), and each of the 2 columns relates to one of the 2 different states of expression of the characteristic. In this way a 2 x 2 Contingency table can be built up:

	Class 1	Class 2	Total
Sample 1	a	b	a + b
Sample 2	c	d	c + d
Total	a + c	b + c	n = a + b + c + d

2. The Fisher exact test is calculated as follows (KANJI G. K. 1993):

Equation [2]

$$\sum p = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!} \sum_i \frac{1}{a_i! b_i! c_i! d_i!}$$

where the summation is over all i possible 2 x 2 schemes with a cell frequency equal to or smaller than the smallest experimental frequency, keeping the row and column totals fixed as above:
If $\sum p$ is less than the significance level chosen, we may reject the null hypothesis of independence between samples and classes, that is to say that the two samples have been drawn from one common population.

Example 3: Use of the Fisher exact test to assess distinctness between two varieties of Ryegrass for the characteristic Resistance to *Xanthomonas translucens* pv *graminis,* using:

Observed frequencies of plants

	Variety X	Variety Y	*Total number of plants*
Number of dead plants	36	74	*110*
Number of living plants	67	29	*96*
Total number of plants	*103*	*103*	*206*

Fisher exact statistic value > 0.01%

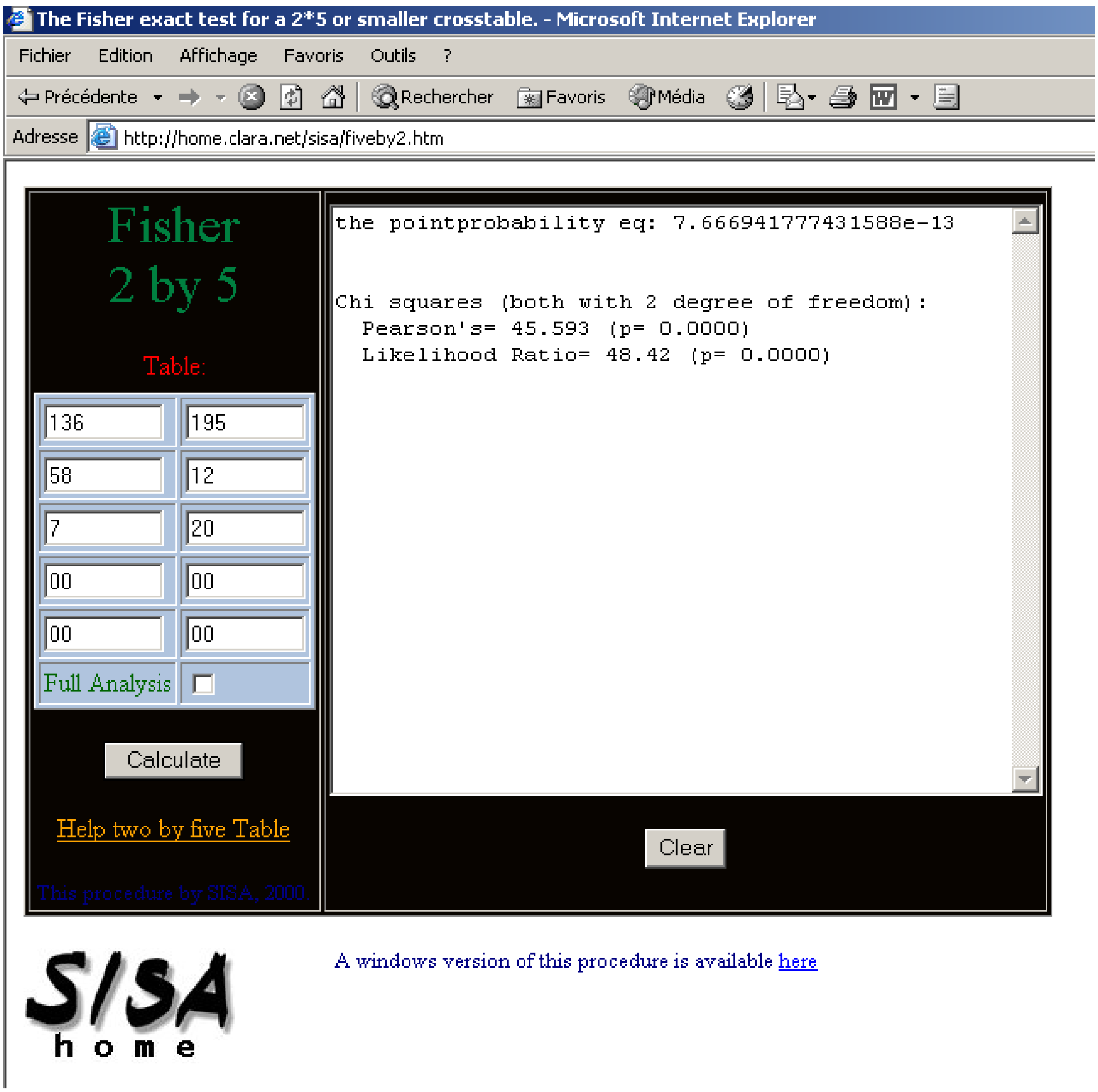
3. Of the 103 plants observed in Variety X, 36 plants were found to be dead, whilst for variety Y, 74 plants were found to be dead out of 103 plants observed. The difference

between variety X and variety Y is significant at the 0.1% level. The above table shows the observed frequencies. The significant probability is calculated with the equation [2].

Example 4: Output from Fisher exact tests using SISA software to assess distinctness, between varieties of Lucerne, for the characteristic Flower Color.

4. Trials comprise 2 varieties tested in a pair-wise comparison. The Frequencies used are the same as those used in the example 2 of the Chi-square test.

5. The screen copy below is issued from the SISA software on line:



REFERENCES

Snedecor, G.W.; Cochran W. (1937); "Statistical methods". *Iowa State University Press.*
Kanji G. K. (1993); "100 Statistical tests". *SAGE Publications.*

Uitenbroek, D.G. (1997) "SISA-Binomial", Available:
http://home.clara.net/sisa/binomial.htm. (Accessed: 2004, January 1).

[End of document]