



TWC/32/21

ORIGINAL: English

DATE: May 21, 2014

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

Geneva

TECHNICAL WORKING PARTY ON AUTOMATION AND COMPUTER PROGRAMS

Thirty-Second Session

Helsinki, Finland, from June 3 to 6, 2014

REVISION OF DOCUMENT TGP/8: PART II: SELECTED TECHNIQUES USED IN DUS EXAMINATION,
NEW SECTION: STATISTICAL METHODS FOR VISUALLY OBSERVED CHARACTERISTICS

Document prepared by the Office of the Union

Disclaimer: this document does not represent UPOV policies or guidance

1. The purpose of this document is to report on developments concerning a possible New Section: "Statistical Methods for Visually Observed Characteristics" to be introduced in document TGP/8: Part II: Techniques Used in DUS Examination, in a future revision of document TGP/8.

2. The following abbreviations are used in this document:

TC:	Technical Committee
TC-EDC:	Enlarged Editorial Committee
TWA:	Technical Working Party for Agricultural Crops
TWC:	Technical Working Party on Automation and Computer Programs
TWF:	Technical Working Party for Fruit Crops
TWO:	Technical Working Party for Ornamental Plants and Forest Trees
TWPs:	Technical Working Parties
TWV:	Technical Working Party for Vegetables

BACKGROUND

3. The TC, at its forty-ninth session, held in Geneva, from March 18 to 20, 2013, agreed that it would not be appropriate to continue the development of a section on "Statistical Methods for Visually Observed Characteristics", unless new guidance was provided beyond the methods already provided in document TGP/8. In that regard, it requested the TWC to clarify if it proposed to modify an existing method or provide a new additional method (see document TC/49/41 "Report on the conclusions", paragraph 72).

COMMENTS BY THE TECHNICAL WORKING PARTIES IN 2013

4. At their sessions in 2013, the TWO, TWF, TWV, TWC and TWA considered documents TWO/46/23, TWF/44/23, TWV/47/23, TWC/31/23 Rev. and TWA/42/23 Rev., respectively.

5. The TWC agreed that the method proposed in Annex II to document TC/49/32 "Revision of document TGP/8: Part II: Techniques Used in DUS Examination, New Section: Statistical Methods for Visually Observed Characteristics" (see document TWO/46/23 "Revision of Document TGP/8: Part II: Statistical Methods for Visually Observed Characteristics") was new and considered that it had advantages over the Chi-square test already provided in document TGP/8 for multinomial distributed data, such as visually

observed characteristics, whereas COYD for normally distributed data is not suitable for multinomial distributed data (see document TWC/31/32 "Report", paragraph 53).

6. The TWC agreed that it would be beneficial to further develop the method for multinomial data and to compare the decisions made using the two methods – the COYD method for normally distributed data and the Chi-square test – based on real data from Finland and the United Kingdom (Timothy, Red Clover and Meadow Fescue: growth habit). The TWC also noted that Finland planned to use the new method for multinomial data, once it had been established and potentially also the United Kingdom (see document TWC/31/32 "Report", paragraphs 54 and 55).

7. The TWA agreed with the TWC that it would be beneficial to further develop the method for multinomial data and to compare the decisions made using the two methods based on real data from Finland and the United Kingdom (see document TWA/42/31 "Report", paragraph 61).

8. The TWA noted that the experts from the Netherlands and Germany had the intention to use the new method for multinomial data, once it had been established (see document TWA/42/31, paragraph 62).

DEVELOPMENTS IN 2014

9. The TC, at its fiftieth session, agreed to the development of a new method for multinomial distributed data. The TC invited the TWC to compare the new method for multinomial distributed data and the Chi-square test, as set out in document TC/50/28, paragraph 10, and requested the TWC to identify a suitable expert to draft the document (see document TC/50/36, paragraph 65).

10. The Annex to this document presents a comparison of the results on distinctness decision between the COYD method for ordinal characteristics and Chi-square test, prepared by Mr. Sami Markkanen (Finland), Chairperson of the TWC, to be considered by the TWC at its thirty-second session.

11. *The TWC is invited to:*

(a) note the developments concerning a possible New Section: "Statistical Methods for Visually Observed Characteristics" to be introduced in document TGP/8: Part II: Techniques Used in DUS Examination, in a future revision of document TGP/8; and

(b) consider a comparison of the results on distinctness decision between the COYD method for ordinal characteristics and Chi-square test, prepared by an expert from Finland.

[Annex follows]

**REVISION OF DOCUMENT TGP/8: PART II: TECHNIQUES USED IN DUS EXAMINATION,
NEW SECTION: STATISTICAL METHODS FOR VISUALLY OBSERVED CHARACTERISTICS**

**A COMPARISON OF THE RESULTS ON DISTINCTNESS DECISION BETWEEN THE COYD METHOD
FOR ORDINAL CHARACTERISTICS AND χ^2 -TEST**

Introduction

1. During its 31st meeting in 2013, the TWC agreed that it would be beneficial to further develop the method for multinomial data and to compare the decisions made using the two methods, χ^2 -test and COYD method for multinomial characteristics, based on real data from Finland and the United Kingdom (Timothy, Red Clover and Meadow Fescue: growth habit). (See report TWC/31/32 page 7.)
2. A Comparison of the results of the COYD method for ordinal characteristics and χ^2 -test on distinctness decision was made using the same Meadow fescue growth habit data from Finland. The idea of the comparison is to consider if the COYD for ordinal characteristics separates more variety pairs than the χ^2 -test. As an expert from Denmark stated in the Memorandum (TC/50/28, Annex, page 2.) 'The χ^2 -test does not depend on the scale of measurements, so data recorded on the nominal scale and ordinal scale are treated the same way and because the χ^2 -test ignores the ordering of notes on the ordinal scale. The proposed new method for characteristics recorded on the ordinal scale takes this ordering into account. The proposed method is therefore expected to be more effective if the data are recorded on the ordinal scale than if they are recorded on the nominal scale.'
3. Introduction to different types of data and scale levels, including ordinal scaled quantitative data, can be found in the revision document for TGP/8 'Data to be recorded' (latest version TC/50/5 Annex II). Detailed analysis of COYD method for ordinal characteristics by expert from Denmark in TC/49/32 Annex II, pages 4 to 10). Pearson's chi-square test is explained in TGP/8/1 Part II, page 78.
4. The characteristic 'Plant: growth habit at inflorescence emergence' (TG/39/8 Meadow fescue (*Festuca pratensis* Huds and Tall fescue *F.arundinacea* Schreb.) is a visually observed characteristic TG/39/8 explains the characteristic 'The growth habit should be assessed visually from the attitude of the leaves of the plant as a whole. The angle formed by the imaginary line through the region of greatest leaf density and the vertical should be used.' . The observations for this data were done from single plants and the observer gave each individual a note.

Criteria for distinctness in the χ^2 -test

5. The p-value used in the χ^2 -test was 0,05. Yates correction was not used, because the amount of classes in the comparison was always over two.
6. The order of direction of the data was checked before distinctness decision, i.e. the growth habit of the candidate has to be constantly more erect or more prostrate than the compared reference in at least two of the three years used in the analysis. If the data compared between variety pair had different directions in different years, the result was not distinct even though the calculated p-values were under 0,05 in both of the years.
7. The recommended criteria for χ^2 -test was used (Ranta et al. 1994). Therefore 20% of the calculated expected frequencies shouldn't go under 5 and the expected frequencies should be over 1. Due to this, some of the classes had to be fused together. It was usual to have four to three classes in the analysis, because otherwise these criteria would not have met. Especially more extreme classes 1 to 3 and 6 to 9 had only few observations (see TC/49/32, Annex II, page 8).
8. The analyses for χ^2 -test were done using Excel software for Windows.

Results and conclusions

9. Candidate A could be separated from 6 reference varieties with χ^2 -test (varieties F,H,K,P,W and 1). Candidate B was separated from 3 reference varieties (F,P and 1). COYD method for ordinal characteristics separated respectively 11 reference varieties from Candidate A (varieties E,F,H,K,N,P,U,V,Z,1 and 3) and 10 reference varieties from Candidate B (varieties E,F,H,K,N,P,U,V,Z and 3). In average, the COYD method separated 20% more reference varieties than the χ^2 -test. For Candidate A, all the reference varieties separated by χ^2 -test except one (candidate W) were separated also by COYD method. For Candidate B there were also one reference variety (candidate 1) separated only by χ^2 -test.

10. Problem of the analysis of growth habit data with χ^2 -test is the low number of individuals in some of the classes. In 14 cases the p-values in comparisons with Candidate A were under 0,05, but the expected frequencies didn't fulfill the requirements (either over 20% of the expected frequencies were under 5 or some of the values were under 1). Comparisons with Candidate B showed 5 similar situations (marked as * in the Table 1. in the Annex). This low number of observations in some classes can lead into situation where the candidate variety can't be stated as distinct, because the requirements of the statistical analysis are not fulfilled, though the compared varieties could be distinct.
11. The comparison of the results of COYD method for ordinal characteristics and χ^2 -test for meadow fescue growth habit data showed that COYD method for ordinal characteristics can separate more varieties and therefore the use of COYD method with ordinal characteristic would enhance decisions on distinctness.
12. It would be useful to have same type of comparison between COYD method for multinomial characteristics and χ^2 -test with other species and characteristics.

References

- Ranta, E., Rita, H. & Kouki J. 1994. Biometria. Tilastotiedettä ekologeille. Yliopistopaino, Helsinki.
- UPOV. TGP/8/1. Trial Design and Techniques Used in the Examination of Distinctness, Uniformity and Stability. 2010
- UPOV TG/39/8 Guidelines for the conduct of tests for distinctness, uniformity and stability. Meadow fescue (*Festuca pratensis* Huds.), Tall fescue (*Festuca arundinacea* Schreb.). Geneva 2002.
- TC/50/28 REVISION OF DOCUMENT TGP/8: PART II: SELECTED TECHNIQUES USED IN DUS EXAMINATION, NEW SECTION: STATISTICAL METHODS FOR VISUALLY OBSERVED CHARACTERISTICS, January 30, 2014.
- TC/49/32 REVISION OF DOCUMENT TGP/8: PART II: SELECTED TECHNIQUES USED IN DUS EXAMINATION, NEW SECTION: STATISTICAL METHODS FOR VISUALLY OBSERVED CHARACTERISTICS, February 4, 2013.

Annex. Table 1. P-values for variety pair comparisons and information of distinctness by χ^2 -test and COYD for ordinal characteristics.

Ref. Variety	CANDIDATE A			Distinct by χ^2 -test	Distinct by COYD	CANDIDATE B			Distinct by χ^2 -test	Distinct by COYD
	2010	2011	2012			2010	2011	2012		
cand A	-	-	-	<i>no</i>	<i>no</i>	0,02(*)	0,38	0,31	<i>no</i>	<i>no</i>
cand B	0,02(*)	0,53	0,31	<i>no</i>	<i>no</i>	-	-	-	<i>no</i>	<i>no</i>
C	0,68	0,16	0,86	<i>no</i>	<i>no</i>	0,31	0,12	0,67	<i>no</i>	<i>no</i>
D	0,24	0,04(*)	0,06	<i>no</i>	<i>no</i>	0,25	0,74	0,88	<i>no</i>	<i>no</i>
E	0,003	0,07	0,07	<i>no</i>	D	0,0003	0,46	0,09	<i>no</i>	D
F	0,04(*)	0,0001	0,002	D	D	0,74	0,002	0,005	D	D
G	0,01	0,64	0,06	<i>no</i>	<i>no</i>	0,14	0,80	0,02	<i>no</i>	<i>no</i>
H	0,00002	0,0003(*)	0,03	D	D	0,0006(*)	0,16	0,01	<i>no</i>	D
I	0,40	0,77	0,85	<i>no</i>	<i>no</i>	0,01	0,33	0,66	<i>no</i>	<i>no</i>
J	0,34	0,21	0,16	<i>no</i>	<i>no</i>	0,01	0,17	0,68	<i>no</i>	<i>no</i>
K	0,13	0,001	0,04	D	D	0,43	0,09	0,07	<i>no</i>	D
L	0,14	0,40	0,27	<i>no</i>	<i>no</i>	0,15	0,76	0,65	<i>no</i>	<i>no</i>
M	0,18	0,33	0,21	<i>no</i>	<i>no</i>	0,39	0,07	0,95	<i>no</i>	<i>no</i>
N	0,09	0,0005	0,07	<i>no</i>	D	0,28	0,04(*)	0,03	<i>no</i>	D
O	0,007 d	0,005(*)	0,02 d	<i>no</i>	<i>no</i>	0,02	0,65	0,26	<i>no</i>	<i>no</i>
P	0,001(*)	0,0004	0,01	D	D	0,001	0,09	0,002	D	D
Q	0,01	0,51	0,15	<i>no</i>	<i>no</i>	0,03	0,42	0,48	<i>no</i>	<i>no</i>
R	0,26	0,54	0,08	<i>no</i>	<i>no</i>	0,53	0,42	0,17	<i>no</i>	<i>no</i>
S	0,007(*)	0,15	0,16	<i>no</i>	<i>no</i>	0,03	0,24	0,78	<i>no</i>	<i>no</i>
T	0,22	0,001	0,85	<i>no</i>	<i>no</i>	0,46	0,46	0,69	<i>no</i>	<i>no</i>
U	0,0008	0,01(*)	0,08	<i>no</i>	D	0,007	0,58	0,18	<i>no</i>	D
V	0,30	0,004(*)	0,40	<i>no</i>	D	0,66	0,39	0,06	<i>no</i>	D
W	0,15	0,03	0,04	D	<i>no</i>	0,24	0,22	0,13	<i>no</i>	<i>no</i>
X	0,02(*)	0,009 (*)	0,13	<i>no</i>	<i>no</i>	0,01(*)	0,67	0,45	<i>no</i>	<i>no</i>
Y	0,47	0,35	0,14	<i>no</i>	<i>no</i>	0,20	0,63	0,82	<i>no</i>	<i>no</i>
Z	0,04(*)	0,02(*)	0,04	<i>no</i>	D	0,01(*)	0,37	0,01	<i>no</i>	D
1	0,004	0,0001	0,02	D	D	0,02	0,14	0,03	D	<i>no</i>
2	0,39	0,15	0,14	<i>no</i>	<i>no</i>	0,39	0,43	0,22	<i>no</i>	<i>no</i>
3	0,32	0,22	0,10	<i>no</i>	D	0,04	0,32	0,72	<i>no</i>	D
4	0,17	0,01	0,09	<i>no</i>	<i>no</i>	0,13	0,47	0,46	<i>no</i>	<i>no</i>
5	0,05(*)	0,27	0,02	<i>no</i>	<i>no</i>	0,73	0,17	0,47	<i>no</i>	<i>no</i>

Explanations for the table

(*) p-values which were under 0,05, but over 20 % of the the expected frequencies were under 5 or one or more of the expected frequencies were below 1

d the direction of the difference between varieties was not constant between years

highlighted p-values in shaded cells are p-values which separated varieties,

D is for distinct