E

TWC/24/9
ORIGINAL: English
DATE: June 9, 2006

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS
GENEVA

TECHNICAL WORKING PARTY ON AUTOMATION AND COMPUTER PROGRAMS

Twenty-Fourth Session
Nairobi, June 19 to 22, 2006

A RATIONALE FOR ELIMINATION OF REFERENCE VARIETIES WHEN COYD IS USED - WORK IN PROGRESS

Document prepared by an expert from the United Kingdom

INTRODUCTION

1. After the first year of testing, results are reviewed in order to eliminate reference varieties that are clearly distinct from the candidates. When COYD is used to assess distinctness in a characteristic, no formal mechanism has yet been described to inform such early decisions on distinctness.

2. In this paper, a possible approach is described. This work is at an early stage of development and is presented for discussion.

Objective

3. To estimate the chance that a candidate will be distinct on the 2-year COYD criterion from a reference variety after the first year of test. In order to judge the variability associated with measurements in a particular character we need to have past data. If the probability is suitably high, the candidate is declared distinct from that reference variety and does not need to be compared in the second year.

Mathematical details

4. Two varieties, A and B, are tested in two years, labelled 1 and 2. Mean measurements, x_{ij}, are made in the character of interest for each variety, i, and year, j. Let the difference d_j in year j, be given by:

$$d_j = x_{Aj} - x_{Bj}$$

and assume that it is normally distributed. Also let the COYD difference after two years be $D=(d_1+ d_2)/2$. The COYD criterion says that variety A and variety B should be considered distinct if

$$\left| \frac{d_1 + d_2}{2 s_{12}} \right| \geq t_{1-p/2, \nu_{12}}$$

where $t_{1-p/2, \nu_{12}}$ denotes the 1-p/2 quantile of the student t-distribution with ν_{12} degrees of freedom and s_{12} is the square root of the residual variance for the two year COYD analysis of variance (with year and variety effects removed).

5. We wish to estimate the probability p_D that A and B will be considered distinct after two years of tests, given the first year result, d_1, and the historical data, $\underline{x}$, i.e.

$$\Pr\left(\left| \frac{d_1 + d_2}{2 s_{12}} \right| \geq t_{1-p/2, \nu_{12}} \middle| d_1, \underline{x} \right) = p_D. \qquad (1)$$

Conditional on d_1 being known, under the null hypothesis of no difference between varieties, d_2 has a Normal distribution with mean d_1 and standard deviation sqrt(2)* σ_{12}. If we assume that the true value, σ, which in fact is to be estimated from the historical data, were also known and that s_{12}^2 divided by σ^2/ν_{12}, is chi-squared distributed with degrees of freedom ν_{12}, then equation (1) could then be estimated from a non-central t distribution (D*sqrt(2)/s_{12} has non-centrality parameter sqrt(2)* d_1/ s_{12}).

6. In truth, exact knowledge of σ is an approximating assumption: nevertheless it is instructive to calculate the threshold values for d_1 under this assumption for given target values for p_D.

Application

7. The assessment of characteristic "stipule length", for field pea, is used as an example. The historical data x is based on UK semi-leafless pea data from 1995-2004. COYD is used with a probability level of 2%. It is assumed that the number of degrees of freedom in the current 2-year test is large so that D is approximately normally distributed in this example.

8. The long-term 2% LSD for a 2 year test based on the 10 years of historical data is 10.64 mm (note that the data ranges from 45.0 mm to 121.5 mm). For comparison the long-term 2% LSD for a single year test is 15.04 mm. The table below gives the approximate required thresholds for the first year difference d_1 to obtain a p_D probability of being distinct after the second year of tests. Note that the threshold for p_D is the same as the long-term LSD.

p_D	d_1 threshold
99.9%	±20.63
99%	±18.16
98%	±17.28
95%	±15.95
90%	±14.78
80%	±13.36
50%	±10.64

NOTES

1. A choice of level of p_D needs to be made.
2. A Bayesian approach to this problem could allow the use of information from sources other than the first year's experiment, e.g. from the Technical Questionnaire. It should be possible to represent the expert's judgement on the quality of this information.
3. In this first step of development, two approximations have been used to simplify the mathematics and to facilitate discussion of the approach. It is intended to investigate robustness to deviations from the normality assumption, and how to allow for imperfect knowledge of σ_{12}. The implicit assumption that σ is applicable to all years and all varieties also needs to be verified.
4. An alternative approach might be to use the ideas of sequential testing.

[End of document]