E

TGP/8.1 Draft 1
ORIGINAL: English
DATE: May 29, 2002

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

Associated Document
to the
General Introduction to the Examination
of Distinctness, Uniformity and Stability and the
Development of Harmonized Descriptions of New Varieties of Plants (document TG/1/3)

DOCUMENT TGP/8

"USE OF STATISTICAL PROCEDURES IN

DISTINCTNESS, UNIFORMITY AND STABILITY TESTING"

Section TGP/8.1: Introduction

Document prepared by experts from France and the Netherlands

to be considered by the

Technical Working Party on Automation and Computer Programs (TWC), at its twentieth session to be held in Texcoco, Mexico, from June 17 to 20, 2002

SECTION 8.1

INTRODUCTION

1. In this introduction we will address the key-elements of statistics for variety testing. Then, elementary statistical notions are addressed followed by an explanation of the principles of experimental design and some related issues. After that, an introduction is given on the nature of characteristics we observe and the way we record them, together with the statistical consequences this may have. Next there is a paragraph on validation of data, combining statistical ideas and ideas on the nature of the characteristic. Finally the spectrum of statistical tools that are presently recommended by UPOV will be covered.

2. The aim of UPOV, as laid down in the Convention (1991 Act), is to protect new varieties after examination. It aims at achieving standardized variety descriptions according to a harmonized method. The examination or testing seeks to determine if the new variety complies with the requirements so that it will be granted protection.

3. The requirements are formulated in observable characteristics that will show distinctness between varieties, uniformity within a variety and are stable over the years for all varieties. It is thus generally referred to as DUS-testing. The basic question is: can we tell apart a new variety from the standing acknowledged ones in some reliable way (discrimination). The nature of the object of interest, a plant variety, makes it susceptible to all kinds of external influences, e.g. growing conditions like latitude, climate, soil, irrigation etc., which cause variation between plants in the expression of the characteristic. There are also aspects that may lead to variation (or the lack of it) in the observed characteristic of a given variety for reasons that are internal e.g. the way it is propagated (cross-pollinated vs. cloned), seed/bulb/tuber size, maternal effects etc. We must also bear in mind that observing a characteristic leads to yet another source of variation, the measurement error. All these elements of variation of our observations must be taken into account in a model for proper analysis that is agreed by UPOV. These topics are pursued later on.

4. So each observation contains a mixture of all these sources of variation and it is good practice to have some basic notion of their magnitudes (if possible). At the end of the run this will determine the reliability of the measurements and that is what statistics is all about.

5. The first key element of experimentation in a statistical context is replication. This can either be used for determining the number of off-types among say sixty plants in just one experiment or for measuring several characteristics on say 20 separate plants per plot in 3-replicated blocks for 3 consecutive years. In all cases, the presence of replication gives us a handle for hypothesis testing: is the new variety different from the rest by some observed characteristic that exceeds the noise that is basic to the whole experiment? This noise, present in all observations (=experimental error), acts as a yard-stick to judge differences in the experiment (distinctness). When the characteristic is less consistent the difference for a decision needs to be accordingly greater. Moreover the observations on the replications can be used as a uniformity measure.

6. The second key element is randomization for each experiment. This ensures that all unknown sources of variation have an equal chance of acting upon the varieties tested, and no accidental systematic effect is recorded due to e.g. germination, planting, growing, harvesting, recording, etc. In those cases where specific ordering is required to observe a characteristic

(e.g. comparison of color or architecture) one should be aware that this is only (reluctantly) admissible in one replication, and one should be careful with conclusions on other measurements. This brings our focus to a case where some grouping is required because otherwise competition would influence the observation, e.g. early varieties would hamper the development of late ones. In those cases it is crucial to randomize within and between groups of the experiment. Experiments are carried out over the years and groups should then be arranged differently, to avoid systematic effects.

7. Proper randomization makes our conclusions valid/predictive for all similar experiments, and not just for one specific experiment, the results of which might have been caused by some unknown systematic effects, e.g. ordering of varieties, some fertility trend/spot in the soil or recording time during the day. Randomization also prevents us from working with a small experimental error that can not be reproduced, e.g. the close comparison of specific varieties or when all short varieties are lined up in the field. Randomization also enhances objective recording of the characteristic.

8. It is nevertheless common practice in UPOV to plant varieties which show little difference between them, side by side, or in the same area. This allows better evaluation by the experts during the different visits and observations through direct comparison on the field for difficult cases. As already stated, this does not invalidate the statistical reliability too much if it is confined to one replication in a blocked trial, or if different orderings are used in different blocks and in different trials. One should be aware that the yardstick used for discrimination on the other characteristics is not as reliable in this case, as the arrangement can only be based on one or two characteristic(s). Close scrutiny in one characteristic may introduce systematic or interaction effects (bias) for others, so care is needed in the interpretation.

9. The third key element is that for the purposes of UPOV the main experimental treatment is variety. Other major external sources of variation are not the object of the study e.g. latitude and soil. The data are observed on plants in good growing conditions. In those cases where (random) experimental conditions are thought to be relevant for the characteristic performance of the varieties (often year-by-year variation, e.g. climatic influences), this is compensated by repeating the experiment in several years or growing cycles. A new variety can thus be compared with the others using a different criterion that includes this extra source of variation (variety-by-year interaction). Another important reason that a new variety should be tested in more than one growing cycle is to comply with the stability criterion, even though a defect could express itself within one cycle. UPOV looks for consistent results, this implies the use of results from more than one trial for a decision. Consistency means that the general ranking between varieties should not change a lot between trials. Characteristics which do have such an effect on variety ranking are not considered as an element for distinctness by UPOV.

10. A fourth key element is the specific set of considerations that holds for a crop. There can be no general set of experiments and/or characteristics given, that will fulfil the UPOV requirements for DUS-testing. It will depend on the crop and the considerations are diverse, but general information is provided in this document. For most crops, the characteristics and requirements are defined in the Test Guidelines. But sometimes other characteristics can be used as a complement for the ‘agreed’ characteristics. Observations can be made at all different stages of development of the crop, so it is imperative that all aspects of recording a characteristic are described properly and exhaustively to ensure that they can be compared in the long run but also understood by a novice.

11. During or at the end of the study, the data, on the same set of characteristics between varieties, are used by the experts of the crop for DUS testing. The use of and the need for computations may differ considerably. In some cases the notes recorded and the knowledge of the expert are sufficient, while in other cases there is a need to compute a large set of data from more than one sowing in order to obtain objective values on which to base the final expert decision.

12. Finally, looking at the very great range of methods and software available to compute data, one should be aware that for a given method some possibilities/models/options will be in agreement with the UPOV philosophy and some others not. UPOV provides assistance in the use of computational methods which are in accordance with the UPOV philosophy of DUS testing. These methods are described in UPOV documents, and where possible, software is made available. A major advantage of using computations and statistics is that it is less subjective than the sole judgement of the expert. It provides data, additional quality control of the data, graphical representations, suggested decisions based on probabilities, etc., which can be understood, shared and used in a similar way by people from different countries. Harmonization between countries is one of the goals of UPOV.

13. General guidance on what is usually adopted, or what should be avoided, is also available through participation in the work of the UPOV Technical Working Parties.

[End of document]