E

TWC/23/9
ORIGINAL: English
DATE: May 17, 2005

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

TECHNICAL WORKING PARTY ON AUTOMATION AND COMPUTER PROGRAMS

Twenty-Third Session
Ottawa, June 13 to 16, 2005

A COMPARISON OF COYU AND A METHOD BASED ON BENNETT'S TEST FOR COEFFICIENTS OF VARIATION

Document prepared by experts from Poland

A COMPARISON OF COYU AND A NEW METHOD BASED ON BENNETT'S TEST FOR COEFFICIENTS OF VARIATION

Bogna Zawieja[1] , Wiesław Pilarczyk[1,2]

[1]Department of Mathematical and Statistical Methods, Agricultural University, Poznań, Wojska Polskiego 28, 60-637 Poznań
[2]The Research Centre for Cultivar Testing, 63-022 Słupia Wielka

Summary

1. In this paper, the decisions concerning uniformity of varieties based on COYU and a new method based on the application of Bennett's test for coefficients of variation are compared. In COYU, the standard deviations of varieties are compared. To avoid difficulties caused by different levels of expressions of characteristics under investigation, the varieties are ranged according to their mean values and comparisons are made within groups of varieties characterized by similar mean values. With the Bennett's test, the whole group of varieties can be compared because difficulties caused by different levels of expressions are overcome by comparing coefficient of variability instead of standard deviations. All considerations are illustrated with trial data concerning winter rye varieties.

Introduction

2. In uniformity testing, the homogeneity of plants within varieties is tested. It means that differences between plants within a variety should not exceed a specified threshold. The method used for checking uniformity depends both on the features of propagation of the variety and on the type of characteristic (i.e. quantitative, qualitative, or pseudo qualitative). For self-pollinated and vegetatively propagated varieties, it is possible to count the number of off-type plants and compare with the maximum number of off-types allowed. In the case of quantitative characteristics in cross-pollinated varieties, the standard deviation of a new variety can be compared with the average of standard deviations of already registered varieties.

3. In this paper, the COYU method, based on comparisons of standard deviations of new and established varieties, is compared with a method in which the coefficients of variations form the base for comparisons. This new method utilizes the Bennett's (1976) test for testing homogeneity of coefficients of variations. In our opinion this new method is better because it is independent of the level of expression of characteristics involved.

Trial Data

4. The calculations were made using the three-year trial DUS data on winter rye. These trials were conducted at the Experimental Station for Cultivar Testing at Słupia Wielka, Poland, in the period 1999-2001. All three trials were planned and were conducted in randomized complete block design with three replicates. The eight quantitative characteristics, among all those observed, were tested in these experiments. The characteristic codes are: 31 - plant height, 32 - length between upper node and ear, 33 - length of ear, 10 - length of blade of leaf next to flag leaf, 11 - width of blade of leaf next to flag leaf, 51 - number of spikelets, 52 - length of rachis. The coding is in accordance with that used in the

official DUS test. A total of 73, 83 and 75 varieties were compared in the years 1999, 2000 and 2001 respectively.

The Methods

5. Let us start by recalling the approach used in the COYU method for single trials for quantitative characteristics. The single DUS trial is the trial carried out at one place within a particular year of testing (or a particular growing season). Usually, many varieties of a given species are simultaneously tested in one trial. These experiments are usually conducted in randomized complete block design with two or three replicates of treatments and 10-20 measurements taken from each plot.

COYU: Step 1 (Weatherup, 1992)

6. All characteristics are checked one by one. For every characteristic, the standard deviation of the new variety is compared with the average of standard deviations of varieties used for comparing (sometimes called the "reference collection").

7. The threshold value for standard deviation of the new variety is calculated using the formula

$$C = \bar{s} + s_w t^{tab}_{0.02;w-1}, \qquad (1)$$

where $s_w = \sqrt{\dfrac{\sum_{i=1}^{w} (s_i - \bar{s})^2}{w-1}}$,

$t^{tab}_{0.02;w-1}$ is the critical value of *t*-Student's distribution at the significance level $\alpha = 0.02$ for *w*-1 degrees of freedom. The average of standard deviations $\bar{s}$ is calculated using the formula

$$\bar{s}^2 = \frac{\left(\sum_{i=1}^{w} s_i^2\right)}{w} \qquad (2)$$

Method based on Bennett's test for coefficients of variation

8. To introduce it, let $\{x_{ij}\}$ $(i = 1,...,v; j = 1,...,n_i)$ represent $n = \sum n_i$ independent observations (for one characteristic) from *v* normal populations (varieties) $N(\mu_i,\sigma_i^2)$. The coefficient of variation for *i*-th variety is defined traditionally as

$$\zeta_i = \sigma_i / \mu_i \qquad \text{where } \mu_i > 0. \qquad (3)$$

The subset of varieties is treated as uniform, if no varieties have distinct coefficients of variation. So the hypothesis concerning homogeneity of these coefficients

$$H_0: \quad \zeta_1 = ... = \zeta_v (= \zeta, \text{ say}) \qquad (4)$$

is tested against the alternative that some coefficients (at least one) are different.

Let $$\psi_i = \frac{\zeta_i^2}{1+\zeta_i^2}, \text{ and } y_i = \frac{n_i z_i^2}{1+z_i^2} \qquad (5)$$

where: $z_i = s_i / \bar{x}_i$ is the empirical coefficient of variation for the *i*-th variety (calculated from the sample), s_i^2 is the sample variance and $\bar{x}_i$ is the mean value of all n_i measurements of *i*-th variety. Iglewicz et al. (1968 and 1970) have shown that McKay's (1932) transformed variable y_i / ψ_i have approximate χ^2 distribution with (n_i-1) d.f. and that variates y_i have Gamma distribution (Pitman, 1939). It is necessary to mention that this approximation assumes that the probability of a negative z_i value ($\bar{x}_i < 0$) is negligible. Therefore in formula (3) we assume that $\mu_i > 0$. Then hypothesis (4) is equivalent to hypothesis that

$$H_0: \quad \psi_1 = ... = \psi_v (= \psi, \text{say})$$

for the *v* independent Gamma variates $y_1 ..., y_v$. So the appropriate statistic for H_0 is given by the formula

$$2Z = (n-v)\log\left(\frac{\sum_i y_i}{n-v}\right) - \sum_i (n_i - 1)\log\left(\frac{y_i}{n_i - 1}\right)$$

And with respect to (5) we have

$$2Z = (n-v)\log\left(\frac{\sum_i \frac{n_i z_i^2}{1+z_i^2}}{n-v}\right) - \sum_i (n_i - 1)\log\left(\frac{\frac{n_i z_i^2}{1+z_i^2}}{n_i - 1}\right), \qquad (6)$$

which is approximately distributed as χ^2 with $(v-1)$ degrees of freedom. This method we will call the Bennett's method (Bennett, 1976).

9. In fact, in the COYU method for testing uniformity, the standard deviation of the new variety is compared with the average of standard deviations calculated from standard deviations of eight (ten) varieties having similar mean values. To achieve that, all varieties are ranged according to their mean values, with the four neighbouring varieties with lower means and the four neighboring varieties with higher means from the reference collection identified for each variety. We used here a new approach to check uniformity in two versions. One of these relies on calculating the statistic (6) for ten reference varieties (for which averages are closest to the average of the new variety) for each new variety in turn. In all cases, when the set of neighbours with lower (higher) mean values contains less than five varieties, the set of neighbours with higher (lower) mean values is increased, so every new variety is compared with ten varieties from the reference collection for this particular variety. In the second version, the statistic (6) is calculated using all listed (and – of course – included in experiment) varieties, testing each new variety in turn.

The Results

10. The uniformity of the new varieties of winter rye was checked for all seven quantitative characteristics separately for the three years of testing. The COYU method was applied first. The computer program UNIF (Weatherup, 1992) was used. The two new methods were then applied to the same data.

11. In Table 1, the numbers of varieties declared as uniform (U) or not uniform (NU), at significance level $\alpha = 0.05$, are given. The results are presented in two-by-two tables for the COYU method and the Bennett's method. The results are given separately for each of two versions of the Bennett's method (for ten nearest and for all varieties in the reference collection) and for each of the characteristics included. In general, results for these two approaches are similar. Nevertheless, in the majority of cases, the Bennett's method indicated that slightly more varieties are uniform than by the COYU method. Next, the summary decisions (numbers of uniform and non-uniform varieties) were compared using the exact χ^2 Fisher test (Kendall M.G., Buckland W.R.,1986). These results are presented in Table 2. In Table 2, the empirical α-levels for rejecting the hypothesis of independence for each year (jointly for all characteristics) are given. There are no differences between the two versions of the new method.

12. All empirical α-levels for 1999 and 2000 are greater than 0.01. So for these two years, the decisions concerning uniformity of varieties are independent of the method applied.

13. On the other hand, in 2001, according to the new approach, all varieties were declared uniform with respect to all characteristics while according to the COYU criterion there were 30 cases of lack of uniformity. These results are significantly different.

Table 1. Two-by-two table of frequency of varieties

	Charac-teristic	no 31		no 32		no 33		no 41		no 42		no 51		no 52	
	COYU method	U	NU	U	NU	U	NU	U	NU	U	NU	U	NU	U	NU
Year: 1999															
Bennett (10 closest varieties)	U	38	2	37	4	39	0	38	3	39	2	39	0	39	1
	NU	0	1	0	0	2	0	0	0	0	0	2	0	0	1
Bennett (whole set of varieties)	U	38	3	37	4	39	0	38	3	39	2	39	0	39	2
	NU	0	0	0	0	2	0	0	0	0	0	2	0	0	0
Year: 2000															
Bennett (10 closest varieties)	U	41	5	47	1	50	3	52	1	50	3	52	1	47	1
	NU	3	4	4	1	0	0	0	0	0	0	0	0	5	0
Bennett (whole set of varieties)	U	41	2	49	2	50	3	51	1	50	3	52	1	42	0
	NU	3	7	2	0	0	0	1	0	0	0	0	0	10	1
Year: 2001															
Bennett (10 closest varieties)	U	43	6	41	8	47	2	45	4	49	0	43	6	45	4
	NU	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bennett (whole set of varieties)	U	43	6	41	8	47	2	45	4	49	0	43	6	45	4
	NU	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 2. The empirical significance levels of exact Fisher χ^2 test for comparing different methods of testing of uniformity of varieties

year	1999			2000			2001		
method	U	NU	α emp.	U	NU	α emp.	U	NU	α emp.
COYU	273	14	0.055	351	20	0.368	313	30	< 0.001
Bennett 10	281	6		354	17		343	0	
COYU	273	14	0.014	351	20	0.321	313	30	<0.001
Bennett whole set	283	4		347	24		343	0	
Bennett 10	281	6	0.376	354	17	0.168	313	30	----
Bennett whole set	283	4		347	24		313	30	

Comments and Conclusions

14. The analysis of three year data from winter rye trials showed that the Bennett's test for coefficients of variation gives similar results to those given by the COYU approach. Only in a few cases did the new approach accept some varieties as uniform while according to COYU criterion these would have been rejected as not uniform. However, because both the COYU criterion and the Bennett's method are based on approximate tests – they ought to be precisely compared using both generated and real data.

References

1. Bennett, B.M., 1976, On an approximate test for homogeneity of coefficients of variation, in: Contributions to applied statistics (ed. W.I. Ziegler). Birkhäuser Verlag: 169-171.

2. Iglewicz, B. and Meyers, R. H. [1970]. Comparison of approximations of the percentage points of the sample coefficient of variation. Technometrics 12,1676-169.

3. Kendall, M.G., Buckland W.R., 1986, Słownik Terminów Statystycznych, PWE Warszawa.

4. McKay, A.T. [1932], Distribution of the coefficient of variation and the extended *t* distribution. J.R. Statist. Sec. 95, 695-698.

5. Pilarczyk, W. [1993], Badanie odrębności, wyrównania i trwałości odmian I. Zakres stosowania metod statystycznych w badaniach OWT. XXIII Colloquium Metodologiczne z Agrobiometrii, PAN, str. 74-85.

6. Pitman, E. J. [1939], Tests of hypotheses concerning location and scale parameters. Biometrika 31,200-215.

7. Weatherup, S.T.C. [1992], Distinctness, Uniformity and Stability trial (DUST) analysis system. User manual. Department of Agriculture for Northern Ireland Biometrics Division. Belfast BT9 5PX.

[End of document]