Colloquium Biometricum 40
2010, 53–61

COMPARISON OF UNIFORMITY DECISIONS BASED ON COYU AND BENNETT'S METHODS – SIMULATED DATA

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

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

Summary

Uniformity decisions concerning new varieties of plants are based both on quantitative characteristics and on qualitative characteristics. Decision rules for qualitative characteristics (usually "qualitative" is equivalent with "visually assessed") are rather simple. Namely for every new variety the number of non-typical plants in a fixed sample size is counted and if it is larger than the threshold value (established by crop-experts), the variety is treated as non-uniform. More complicated procedure is applied for quantitative characteristics. Decisions are based on comparisons of standard deviation of candidate variety with average value of standard deviations of so called reference varieties. A special procedure called COYU (combined over years uniformity) was elaborated by member states of UPOV (International Union for Protection of New Varieties of Plants) for this purpose, Talbot (2000). The COYU method is – to some degree – an officially promoted method. But some other methods are still under consideration. One of such methods uses the Bennett test for coefficients of variation. The details of this new approach are given in paper by Zawieja and Pilarczyk (2005, 2006, 2007) and by Zawieja, Pilarczyk and Kowalczyk (2009). Some comparisons of uniformity decisions concerning winter wheat and oilseed rape varieties based on COYU and Bennett's test are also included in mentioned papers. During the annual session of Technical Working Party on Automation and Computer Programs (held in Alexandria, Virginia in June 2009) it was suggested to compare decisions on uniformity of varieties using simulated data based on real measurements. So in the present paper this problem is reconsidered using real data for oilseed varieties (reference set) and simulated data (candidate varieties).

Key words and phrases: Bennett's method, COYU method, DUS testing, oilseed rape, simulation, uniformity

Classification AMS 2010: 62K99, 62P10

1. Introduction

Fulfilling conditions of distinctness, uniformity and stability (DUS) are necessary requirements for every new variety to be registered and placed in national (recommended) list of varieties. Distinctness means that new variety must be distinguishable from every known variety (variety of so-called common knowledge) for at least one characteristic. Decisions are usually taken after two or three years of trialling. In such trials all varieties that are potentially indistinguishable from candidate varieties are tested together. As already mentioned, distinctness from any other variety for just one characteristic is sufficient to fulfil requirement of distinctness. On the other hand, degree of uniformity of new variety must be not worse that uniformity of all varieties used for comparison for all considered (observed) characteristics. An officially adopted method - in countries associated in UPOV (International Union for the protection of New Varieties of Plants*)* - for checking uniformity (COYU) suffers from some disadvantages, Kristensen and Roberts (2009). In papers by Zawieja and Pilarczyk (2005, 2006, 2007) an alternative approach to testing uniformity has been proposed. Namely instead of testing equality of – adjusted by moving average method – standard deviations (COYU approach), the new method that uses Bennett's test for equality of coefficients of variations was applied. The Bennett's test is much simpler and can potentially be used instead of COYU method. In general (at least for winter wheat data and oilseed rape data) the decisions concerning uniformity of candidate varieties were statistically equivalent. For testing equivalence of decisions the McNemar (1947) test was applied. Because number of candidate varieties was (and usually is) small, the McNemar test was based on relatively small sample size. During discussion on these methods at the annual meeting of Technical Working Party (TWC) of UPOV held in Alexandria in 2009, it was suggested to use the larger set of candidate varieties using simulated data. The official policy within UPOV is that new method can be approved (and officially promoted)when old and new methods support concordant decisions on uniformity of varieties. So, the aim of this paper is the comparison of COYU supported decisions concerning uniformity of varieties with the same decisions supported by Bennett's test. The method of simulation and results of comparisons of considered method are presented in next paragraphs.

2. Data

The data from DUS trials on oilseed rape performed at experimental station Słupia Wielka in the period 2006-2008 form the basis of investigations. Only data for varieties already registered are used. Because the aim of this research was comparison of decisions concerning uniformity supported by COYU and by Bennett's test, there was no necessity of use of all characteristics observed. Then, one characteristic – the plant height – was chosen. For every of analysed periods, namely 2006-2007, 2007-2008 and 2006-2008, the data for candidate varieties were generated using method as follows:

1) using 60 measurements for every variety (two plots $\times$ 30 plants observed), the minimum, x_{min}, and maximum value, x_{max}, of real variety mean and minimum and maximum values of standard deviations were calculated, s_{min}, s_{max};

2) starting from (rounded) x_{min}, the values for "candidate" varieties were formed using formula

$$x_i = x_{min} + (i\text{-}1)d, \quad i=1,2,3,....$$

where values x_i were generated as far as x_{max} was reached;

3) every value x_i was associated with all values of standard deviations generated as follows

$$s_j = s_{min} + (j\text{-}1)s, \quad j = 1,2,3,....$$

where the s_j were generated so far as s_{max} was reached.

The values of d and s were chosen in a way that guarantee the reasonable number of "candidate" varieties.

For the period 2006-2007, there were 66 established varieties (forming so-called reference set) and 187 candidate (simulated) varieties. Similarly for the period 2007-2008, there were 57 established and 272 simulated varieties and finally, for the period 2006-2008, 72 and 238 such varieties. Uniformity of every "candidate" variety was tested using the methods given bellow.

3. Method

Each candidate variety was tested using COYU (combined over year uniformity) method and Bennett's test. The method similar to that described by Zawieja, Pilarczyk and Kowalczyk (2009) was used to compare decisions concerning uniformity. The COYU method uses average values of within-plot standard deviations as a measure of uniformity. These values are next ln (natural logarithm) transformed, and “adjusted” using moving average approach. Adjusted values are compared with similar values received for the reference set varieties. All details of COYU approach can be found in a paper by Talbot (2000).

In a Bennett’s approach the coefficients of variations are used as a measure of uniformity. Equality of coefficient of variation of candidate variety and a subset of coefficients of variation of reference set varieties is a criterion of acceptance of candidate variety as uniform. It can be applied when all coefficients of variation are not higher than 0.3 (Forkman, 2006 Iglewicz and Meyers, 1970). In our case this condition was always fulfilled. The subset of reference set varieties was formed in similar way as in COYU approach, namely varieties with closest mean values were taken. More details on Bennett’s test are given in a paper by Zawieja and Pilarczyk (2006).

The decisions concerning uniformity of candidate varieties supported by the two considered methods are compared using two-way contingency table (Table 1).

The COYU and Bennett’s methods were applied at the same significance level. The $n_{11}+n_{22}$ denote the number of unanimous decisions why $n_{12}+n_{21}$ denotes the number of contradictory decisions.

Table 1. Two-way contingency table for decisions on uniformity of candidate varieties using two methods

Method		Bennett’s test	
	decision	uniform	not uniform
COYU approach	uniform	n_{11}	n_{12}
	not uniform	n_{21}	n_{22}

The two methods are fully concordant if they support exactly the same decisions concerning uniformity of tested varieties. It means that in that case $n_{12} + n_{21} = n$, where n denotes the total number of decisions.

There are several methods for testing degree of concordance of decisions with use such data. In a paper by Zawieja and Pilarczyk (2006) the Fisher exact test was used to find out if there is an association between decisions, why in a paper by Zawieja and Pilarczyk (2007) the McNemar test was used to test if the hypothesis that probabilities of contradictory decisions p_{21} and p_{12} are equal can be accepted or not.

Here the “odds ratio” *OR* (Rudas, 1998, Uebersax 2006) is applied as a measure of association between decisions. Odds ratio is calculated as

$$OR = \frac{n_{11} \cdot n_{22}}{n_{12} \cdot n_{21}}.$$

Large value of *OR* indicates association between methods. The statistical significance of lack of association can be tested using statistics Z_0 of the form

$$Z_0 = \frac{\ln(OR)}{\sigma_{\ln(OR)}},$$

where $\sigma_{\ln(OR)} = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}$. The Z_0 statistics has asymptotic normal distribution. Coefficient *OR* can be easily transformed to the Yule coefficient of association *Q* (Yule and Kendall, 1966), using formula

$$Q = \frac{OR - 1}{OR + 1}.$$

This coefficient is interpreted similarly to the coefficient of correlation. $Q = 0$ means lack of association between methods, value close to 1 means high agreement. To have additional characterisation of association, the probability *p* of agreement was also calculated according to the formula

$$p = \frac{n_{11} + n_{22}}{n},$$

where *n* denotes the total number of candidate varieties.

4. Results

The COYU method and the corrected Bennett's test (Shafer and Sulivan, 1986) were applied for three sets of data generated according to above described method (data for candidate varieties). The data for reference varieties were taken from real experiments performed at the experimental station Słupia Wielka. The COYU analysis was performed with use of DUST package of Weatherup (1992). For Bennett's test the EXCEL spreadsheet was utilized. The results for two years data concerning period 2006-2007 are given in Table 2 (testing at significance level $\alpha = 0.002$) and in Table 3 (significance level 0.02).

Table 2. Decisions on uniformity of candidate varieties for data from the period 2006-2007, $\alpha = 0.002$

Method		Bennett's test	
COYU approach	decision	uniform	not uniform
	uniform	187	0
	not uniform	0	0

Table 3. Decisions on uniformity of candidate varieties for data from the period 2006-2007, $\alpha = 0.02$

Method		Bennett's test	
COYU approach	decision	uniform	not uniform
	uniform	178	0
	not uniform	9	0

When testing was performed at the level $\alpha = 0.002$, the two methods accepted all varieties as uniform (full agreement between methods, p = 100%). But when testing at 0.02 level (Table 3), 9 of candidate varieties were rejected as not uniform by Bennett's test but were accepted as uniform by COYU. The probability of agreement between methods equals to 95.2%. For data in Tables 2 and 3, the odds ratio *OR* can not be calculated as either n_{12} or n_{21} (or both) are zero.

The results for the 2007-2008 are presented in Table 4 ($\alpha = 0.002$) and in Table 5 ($\alpha = 0.02$). The probability of agreement is equal to 100% (when testing at 0.002 level) and 94,5% (when testing at 0.02 level). Again other measures of agreement (*OR* and *Q*) can not be calculated for results in Table 4. For results given in Table 5, these measures of agreement are respectively $OR = 48{,}32$, $Q = 0.959$ (the value of $Z_0 = 6.438$ is higher than critical value $Z_{0.01} = 2.576$).

Table 4. Decisions on uniformity of candidate varieties for data from the period 2007-2008, $\alpha = 0.002$

Method		Bennett's test	
COYU approach	decision	uniform	not uniform
	uniform	272	0
	not uniform	0	0

Table 5. Decisions on uniformity of candidate varieties for data from the period 2007-2008, $\alpha = 0.02$

Method		Bennett's test	
COYU approach	decision	uniform	not uniform
	uniform	246	7
	not uniform	8	11

The results for three years period (2006-2008) are presented in Table 6 (for α = 0.002) and in Table 7 (α = 0.02). When testing was performed at α = 0.002 level, there was 219 (=217+2) concordant decisions concerning uniformity and respectively 19 (=18+1) contradictory decisions. It means that probability of agreement is p = 92.0%.

Table 6. Decisions on uniformity of candidate varieties for data from the period 2006-2008, α = 0.002

Method		Bennett's test	
COYU approach	decision	uniform	not uniform
	uniform	217	18
	not uniform	1	2

The other measures of agreement are equal respectively OR = 24.111, Q = 0,920 and Z_0 = 2,548. For results given in Table 7, the following values can be easily obtained p = 82.8%, OR = 19.667, Q = 0.903, Z_0 = 5.596 (again Z_0 much higher than critical values at 0.05 and 0.01 levels).

Table 7. Decisions on uniformity of candidate varieties for data from the period 2006-2008, α = 0.02

Method		Bennett's test	
COYU approach	decision	uniform	not uniform
	uniform	177	36
	not uniform	5	20

5. Discussion and conclusions

In a papers by Zawieja and Pilarczyk (2005, 2006) it has been shown that the COYU method and the Bennett's test applied to real data concerning winter rye varieties did not differ statistically. It was observed that the Bennett's method was slightly more tolerant than COYU method but that statistically (at α = 0.01 level) these two methods gave the same decisions. In paper by Zawieja and others (2009), using real oilseed rape data, it has been shown that again these two method did not differ statistically but for oilseed rape the method the COYU was slightly more tolerant. In all previous investigation there were very limited numbers of candidate varieties.

The results obtained here (with use mixture of real and simulated data) showed that in some cases these two methods of testing varietal uniformity did not differ (results for years 2006-2007). In some other cases (results for periods 2007-2008 and 2006-2008) there existed meaningful differences in decisions, as the Bennett's test rejected more candidate varieties. Detailed inspection of analysed data indicated that in all cases the Bennett's test rejected varieties with small mean values and high standard deviations (with large coefficients of variation). COYU method was - for part of such varieties - more tolerant.

References

Forkman J.(2006). *Statistical inference for the coefficient of variation in normally distributed data*. Research Report 2006, 2, Centre of Biostochatics Swedish University of Agricultural Sciences.

Iglewicz B., Meyers R. H. (1970). Comparison of approximations of the percentage points of the sample coefficient of variation. *Technometrics* 12, 166–169.

Kristensen K., Roberts A. (2009). Potential approaches to improving COYU. UPOV Geneva. *TWC*/27/15, 1–8.

McNemar Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 12, 153–157.

Rudas T. (1998). *Odds Ratios in the Analysis of Contingency Tables*. Thousand Oaks, CA Sage Publ.

Shafer N. J., Sulivan J. A. (1986). A simulation study of test for the equality of the coefficient of variation. *Communications in Statistics – Simulation and Computation* 15, 681–698.

Talbot M. (2000). The Combined-Over-Years Distinctness and Uniformity criteria. UPOV, *TWC*/18/10, Geneva.

Uebersax J. (2006). Odds Ratio and Yule's Q. http://www.john-uebersax.com/stat/odds.htm.

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.

Yule G. U. (1912). On the methods of measuring association between two attributes. *Journal of the Royal Statistical Society* 75, 579–652.

Yule G. U., Kendall M. G. (1966). *Wstęp do teorii statystyki*. PWN.

Zawieja B., Pilarczyk W. (2005). The comparison of traditional UPOV uniformity criterion and new approach based on Bennett's test for coefficients of variation. *Colloquium Biometryczne* 35, 155–163.

Zawieja B., Pilarczyk W. (2006). The comparison of decisions on uniformity of rye varieties based on COYU approach and Bennett's test. *Colloquium Biometricum* 36, 225–233.

Zawieja B., Pilarczyk W. (2007). Further comparison of decisions concerning uniformity of rye varieties based on COYU approach and on Bennett's test. *Colloquium Biometricum* 37, 71–76.

Zawieja B., Pilarczyk W., Kowalczyk B. (2009). The comparison of uniformity decisions based on COYU and Bennett's method – oilseed rape data. *Colloquium Biometricum* 39, 170–176.

PORÓWNANIE DECYZJI O WYRÓWNANIU ODMIAN PODJĘTYCH PO ZASTOSOWANIU METODY COYU I TESTU BENNETTA – DANE SYMULOWANE

Streszczenie

Decyzje dotyczące wyrównania nowych odmian (odmian „kandydatów”) przed ich zarejestrowaniem dotyczą zarówno cech ilościowych jak i jakościowych. W przypadku cech jakościowych reguły decyzyjne są bardzo proste. Mianowicie w próbie o ustalonej wielkości obserwuje się liczbę roślin nietypowych i jeśli frakcja takich roślin przekracza pewną wartość progową, odmianę uznaje się za niespełniającą warunku wyrównania. Bardziej złożoną procedurę stosuje się w przypadku cech ilościowych. Najogólniej, przy analizie cech ilościowych, porównuje się odchylenia standardowe (obliczone na podstawie próby 60 roślin) odmiany-kandydata ze średnim odchyleniem standardowym z pewnej liczby – specjalnie wybranych – odmian zarejestrowanych. Oficjalnie zalecaną procedurą w krajach stowarzyszonych w organizacji UPOV jest tzw. procedura COYU. Jedną z innych metod możliwych do zastosowania jest wykorzystanie testu Bennetta, w którym bada się jednorodność współczynników zmienności. Szczegółowy opis metody Bennetta został podany w opracowaniach Zawieji i Pilarczyka (2005, 2006 i 2007). W tym opracowaniu porównane są decyzje dotyczące wyrównania odmian podjęte po zastosowaniu metody COYU (combined over years uniformity) i testu Bennetta. Dla odmian wzorcowych wykorzystano dane rzeczywiste z lat 2006-2008 dotyczące rzepaku ozimego z doświadczeń przeprowadzonych w stacji doświadczalnej oceny odmian w Słupi Wielkiej. Z powodu małej liczby odmian-kandydatów dane dla nich zostały wygenerowane. Decyzje podejmowane na podstawie obu metod nie różnią się istotnie od siebie. Jednakże, w kilku przypadkach metoda Bennetta okazała się nieco bardziej restrykcyjna.

Słowa kluczowe: badania OWT, metoda Bennetta, metoda COYU, rzepak ozimy, symulacja, wyrównanie odmian

Klasyfikacja AMS 2010: 62K99, 62P10