Heredity **72** (1994) 141–145

Received 3 June 1993

Missing observations in the analysis of stability

HANS-PETER PIEPHO
University of Kassel, Faculty of Agriculture, Steinstrasse 19, 37213 Witzenhausen, Germany

In crop variety testing it is frequently of interest to estimate measures of yielding stability or phenotypic stability. The common procedures for stability analysis require a balanced two-way table of genotypes and environments, in which all cells are filled. Frequently, however, empirical data sets are unbalanced due to missing observations. This paper explores methods to estimate stability when some cells are empty. A set of wheat data is used to exemplify these methods.

Keywords: Genotype–environment interaction, method of moments, MINQUE, reliability, stability variance.

Introduction

Yielding stability as a selection trait in plant breeding programmes and evaluation trials is constantly gaining importance over yielding ability. Various statistical concepts and measures of stability have been proposed (see reviews by Lin *et al.*, 1986; Westcott, 1986; Becker & Léon, 1988). Most of them are based on a two-way table of genotypes and environments, where 'environments' may be different locations or different years or both, depending on the scope of the analysis. Usually such data exhibit genotype–environment interaction, which makes selection of high yielding genotypes a difficult task. A genotype interacting strongly with environments may outperform most genotypes in some environments while being at a disadvantage in other environments. The larger the genotype–environment interaction of a genotype the less stable, i.e. the less predictable, is its performance in different environments. If one adheres to this concept of stability it is desirable to minimize genotype–environment interaction. Common measures for this type of stability are the ecovalence (Wricke, 1965), the stability variance (Shukla, 1972) and the non-parametric measures suggested by Nassar & Hühn (1987) and Hühn & Nassar (1989). For these measures to be computable it is necessary that all cells in the data set be filled, i.e. that each genotype be grown in the same set of environments. Rather frequently, however, one is faced with unbalanced data. If data sets from several locations are combined, some genotypes may not have been tested in all sites. Similarly, if yield tests of different years are accumulated, genotypes are typically tested in many but not all of the years. Genotypes change from year to year as new genotypes become available and older ones become obsolete. The purpose of this paper is to explore ways to estimate stability in cases where the data are unbalanced. We will confine our attention to Shukla's stability variance.

Estimation methods

For the statistical analysis it is common to assume the following model:

$$y_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ij} (i = 1,\ldots,K; j = 1,\ldots,N),$$

where y_{ij} = mean yield of genotype i in environment j, μ = grand mean (fixed), α_i = effect of genotype i (fixed), β_j = effect of environment j (random), $(\alpha\beta)_{ij}$ = interaction of genotype i with environment j (random) and e_{ij} = mean error of genotype i in environment j (random).

It is assumed that random effects β_j, $(\alpha\beta)_{ij}$, and e_{ij} are independently normally distributed with variances $\mathrm{var}[\beta_j] = \sigma^2_\beta$, $\mathrm{var}[(\alpha\beta)_{ij}] = \sigma'^2_i$, and $\mathrm{var}[e_{ij}] = \sigma^2_0$, respectively (Shukla, 1972). The assumption of homogeneous error variances is reasonable if the test design is the same for all environments. In accordance with the concept of stability used in this paper, genotype–environment interaction variance is allowed to differ among genotypes. Maximum stability of a genotype is attained if the interaction variance $\sigma'^2_i = 0$. The larger σ'^2_i, the less stable is the corresponding genotype.

In the model for means y_{ij} we cannot distinguish interaction from error. It is solely possible to estimate $\tau_{ij} = (\alpha\beta)_{ij} + e_{ij}$, and hence to estimate the variance $\sigma^2_i = \mathrm{var}(\tau_{ij})$ of genotype i. σ^2_i is the stability variance

introduced by Shukla (1972). We see that

$$\sigma_i^2 = \sigma_i'^2 + \sigma_0^2 .$$

So if the assumption of homogeneous error variance is correct, the genotype rank order given by the stability variance σ_i^2 will exactly equal that given by the interaction variance $\sigma_i'^2$. The most stable genotype, i.e. the genotype with the smallest interaction variance $\sigma_i'^2$, will then also have the smallest stability variance σ_i^2, the genotype with the second smallest $\sigma_i'^2$ will have the second smallest σ_i^2, and so on.

Shukla (1972) gave the following estimate of the stability variance:

$$\tilde{\sigma}_i^2 = \frac{KW_i}{(K-2)(N-1)} - \frac{\sum_{s=1}^{K} W_S}{(K-1)(K-2)(N-1)}, \qquad (1)$$

where

$W_i = \Sigma_j (y_{ij} - \bar{y}_{i\cdot} - \bar{y}_{\cdot j} + \bar{y}_{\cdot\cdot})^2$ with

$\bar{y}_{i\cdot} = \Sigma_j y_{ij}/N$, $\bar{y}_{\cdot j} = \Sigma_i y_{ij}/K$, and $\bar{y}_{\cdot\cdot} = \Sigma_{ij} y_{ij}/KN$.

This estimate can be shown to be identical to Grubbs' estimate of the variance in errors of measurement (Piepho, 1992). For the first genotype Grubbs' estimate is given by:

$$\tilde{\sigma}_1^2 = (K-1)^{-1}\left[\sum_{r=2}^{K} V_{1-r}^2 - (K-2)^{-1} \sum_{1<s<r}^{K} V_{s-r}^2\right],$$

where

$V_{s-r}^2 = [\Sigma_j x_{srj}^2 - (\Sigma_j x_{srj})^2/N]/(N-1)$

with $x_{srj} = y_{sj} - y_{rj}$ (Grubbs, 1948).

Estimates for the other genotypes are obtained by an obvious rotation of subscripts. Grubbs estimate is based on the method of moments, in which sample moments are equated to population moments (Jaech, 1985). We have

$$E[V_{s-r}^2] = \sigma_s^2 + \sigma_r^2 \quad (s = 1,\ldots,K-1; r > s).$$

In order to estimate σ_i^2, $E[V_{s-r}^2]$ is replaced by V_{s-r}^2. The system of equations to be solved for σ_i^2, is then given by

$$V_{s-r}^2 = \sigma_s^2 + \sigma_r^2 \quad (s = 1,\ldots,K-1; r > s).$$

There are $K(K-1)/2$ different equations in K unknowns, so that for $K > 3$ there are more equations than there are unknowns. Grubbs' estimates are the least squares solutions of these equations (Jaech, 1985). Formally the system of equations can be represented as

$$\boldsymbol{Q}\boldsymbol{\sigma} = \boldsymbol{V} \qquad (2)$$

where $\boldsymbol{\sigma}$ is a K dimensional vector of σ_i^2, $\boldsymbol{V}$ is a $K(K-1)/2$ dimensional vector of V_{s-r}^2's, and $\boldsymbol{Q}$ is a $K(K-1)/2 \times K$ matrix with elements 0 and 1, that picks the appropriate σ_i^2's. $\boldsymbol{Q}'\boldsymbol{Q}$ has full rank and can thus be inverted. The solution of eqn 2 is

$$\tilde{\boldsymbol{\sigma}} = (\boldsymbol{Q}'\boldsymbol{Q})^{-1}\boldsymbol{Q}'\boldsymbol{V}. \qquad (3)$$

Grubbs' estimates are unbiased, which is seen by taking expectations on both sides of eqn 3 and inserting eqn 2 on the right hand side:

$$E[\tilde{\boldsymbol{\sigma}}] = E[\boldsymbol{Q}'\boldsymbol{Q})^{-1}\boldsymbol{Q}'\boldsymbol{V}] = E[\boldsymbol{Q}'\boldsymbol{Q})^{-1}\boldsymbol{Q}'\boldsymbol{Q}\boldsymbol{\sigma}] = E[\boldsymbol{I}\boldsymbol{\sigma}] = \boldsymbol{\sigma}.$$

The method of moments may also be employed when some data are missing. For two genotypes s and r we can compute V_{s-r}^2 as long as they are grown together in at least two environments. If this were the case we will say that the two genotypes s and r are connected. To obtain a unique solution of eqn 2, we require that there be at least K connected pairs of genotypes as we need at least as many equations as there are unknowns. Also, each genotype must be connected to at least one other genotype. Thus, the method may break down in some instances when very many data are missing.

In the unbalanced case, another estimate of σ_i^2 can be obtained by the MINQUE principle of estimation (Rao, 1970). MINQUE stands for *MI*nimum *N*orm *Q*uadratic *U*nbiased *E*stimation or *E*stimator, depending on context. It is noted that for balanced data Shukla's estimator is a MINQUE of σ_i^2 (Shukla, 1972). Rao (1970) provides a computational procedure for MINQUE in the general case, which can be used in data sets with empty cells.

We write the linear model for the two-way classification in the form

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{b} + \boldsymbol{\tau}, \qquad (4)$$

where $\boldsymbol{Y}$ is the vector of observations y, $\boldsymbol{b}$ is the parameter vector of main effects, $\boldsymbol{\tau}$ is a vector of τ-effects and $\boldsymbol{X}$ is the design matrix. Denote by n the number of filled cells, by $\boldsymbol{M} = \{m_{pq}\}$ $(p, q = 1,\ldots,n)$ the projection matrix $\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}'$, by $\boldsymbol{t}$ the vector of squares of the residuals $(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-}\boldsymbol{X}')\boldsymbol{Y}$, and by $\boldsymbol{\theta}$ the vector of variances θ_p^2. $(\boldsymbol{X}'\boldsymbol{X})^{-}$ stands for a g-inverse of $\boldsymbol{X}'\boldsymbol{X}$. Note that the dimension of $\boldsymbol{\theta}$ is equal to n, the number of filled cells. For a cell p, we set $\theta_p^2 = \sigma_i^2$, where i is the subscript of the genotype in that cell. Now define $\boldsymbol{F} = \{m_{pq}^2\}$ and consider the equations given by $\boldsymbol{F}\boldsymbol{\theta} = \boldsymbol{t}$. Without loss of generality, let the first n_1 variances of $\boldsymbol{\theta}$ be equal to σ_1^2, the second n_2 to $\sigma_2^2,\ldots$. Then add up the first n_1 equations of $\boldsymbol{F}\boldsymbol{\theta} = \boldsymbol{t}$ to obtain the first equations in σ_1^2, $\sigma_2^2,\ldots$. Similarly add up the next n_2 equations to obtain the second equation in σ_1^2, $\sigma_2^2,\ldots$. Let the reduced equations be denoted by $\boldsymbol{G}\boldsymbol{\sigma} = \boldsymbol{w}$, where $\boldsymbol{\sigma}$ is the vector of different variances σ_i^2 (all assumed

different). Lemma 6 in Rao (1970) states that the MINQUE of $\sigma_1^2,\ldots,\sigma_K^2$ are solutions to the equation $\boldsymbol{G\sigma = w}$ provided $\boldsymbol{G}$ is non-singular. Again, if very many cells are missing, $\boldsymbol{G}$ may become singular so that the stability variance is not estimable.

Variability of estimates

With the methods outlined in the foregoing section it is possible to obtain estimates of the stability variance σ_i^2, even when there are only few observations for genotype i. Such estimates may be subject to considerable sampling variation, thus making the estimates rather unreliable. It is therefore useful to study the sampling variance of Grubbs' estimate and of the MINQUE.

With $\boldsymbol{R}=(\boldsymbol{Q}'\boldsymbol{Q})^{-1}\boldsymbol{Q}'$ the vector of Grubbs' estimate can be expressed as $\boldsymbol{RV}$, which has variance $\boldsymbol{RC_VR}'$, where $\boldsymbol{C_V}$ is the variance covariance matrix of $\boldsymbol{V}$. The elements of $\boldsymbol{C_V}$ are obtained by observing that

$$\text{var}(V_{s-r}^2)=\frac{2\sigma_r^4}{n_{sr}-1}+\frac{2\sigma_s^4}{n_{sr}-1}+\frac{4\sigma_r^2\sigma_s^2}{n_{sr}-1},$$

$$\text{cov}(V_{s\ r}^2,\ V_{s-t}^2)$$

$$=\frac{(n_{sr}-1)(n_{st}-1)(2n_{rt}+n_{rt}^2-n_{sr}n_{st})+2n_{rt}(n_{rt}-1)}{n_{sr}(n_{sr}-1)n_{st}(n_{st}-1)}\sigma_s^4,$$

and $\quad \text{cov}(V_{s-r}^2,\ V_{u-t}^2)=0,$

where n_{sr} is the number of environments in which genotypes s and r were grown together. For all cells filled we have $n_{sr}=N$ for all pairs (r, s), in which case it is easier to use the variance formula given by Grubbs (1948). For $i=1$ one has

$$\text{var}\sigma_1^2=\frac{2\sigma^4 1}{(N-1)}+\frac{4\left[\sum_{r=2}^{K}\sigma_1^2\sigma_r^2+\sum_{s=2}^{K-1}\sum_{r>s}\sigma_s^2\sigma_r^2/(K-2)^2\right]}{(N-1)(K-1)^2} \tag{5}$$

The variances for the other genotypes may be found by rotating the subscripts. The formula in eqn 5 gives the same result as the diagonal elements of $\boldsymbol{RC_VR}'$ in the balanced case. For balanced data, Grubbs estimate and MINQUE are identical, and hence the variance in eqn 5 is also the variance of the MINQUE in eqn 1.

The variance of the MINQUE in the unbalanced case can be derived from the results given in Rao (1972). To do so we write the model in eqn 4 in a different form by partitioning the vector of residuals $\boldsymbol{\tau}$. We can write:

$$\boldsymbol{\tau = Z_1\tau_1 + Z_2\tau_2 + \ldots + Z_K\tau_K},$$

where $\boldsymbol{\tau_i}$ is the vector of residuals of the i-th genotype and $\boldsymbol{Z_i}$ is an appropriate design matrix with elements 0 and 1. The model in eqn 4 can now be written as

$$\boldsymbol{Y = Xb + Z_1\tau_1 + Z_2\tau_2 + \ldots + Z_K\tau_K}.$$

Generally, the MINQUE is a quadratic form given by $\boldsymbol{Y'AY}$. For the r-th genotype we find from the results given in Rao (1970, 1972) that

$$\boldsymbol{A_r}=\sum_{s=1}^{K} g^{sr}\boldsymbol{MV_sM},$$

where $\boldsymbol{G^{-1}}=\{g^{sr}\}$ and $\boldsymbol{V_s}=\boldsymbol{Z_sZ_s}$. Now define $\boldsymbol{V}=\sigma_1^2\boldsymbol{V_1}+\sigma_2^2\boldsymbol{V_2}+\ldots+\sigma_K^2\boldsymbol{V_K}$. If residuals τ_{ij} are normally distributed, the MINQUE $\boldsymbol{Y'A_rY}$ has variance $2\text{tr}(\boldsymbol{A_rVA_rV})$ (Rao, 1972). Again, in the balanced case, this leads to the same results as eqn 5.

For all formulae in this section to be computable we need to know the 'true' values of σ_i^2. As these are usually unknown, it is intuitive to use estimates of σ_i^2 instead. This approach was suggested by Jaech (1985) for computation of the variance of Grubbs' estimate by eqn 5. Using estimates in place of true values must be considered a rough approximation and consequently the variance estimates thus obtained should serve as a rough guide only.

Biological example

To demonstrate the two methods (method of moments and MINQUE) for unbalanced data, we employ an example that has been used by Graybill (1954), Han (1969), Shukla (1972), Levy (1975), Ellenberg (1977), Snee (1982) and Mudholkar & Sarkar (1992) to illustrate their proposed procedures for studying non-homogeneous variances in a balanced two-way layout. The data set comprises 4 wheat varieties grown in 13 environments (Table 1). For the purpose of demonstration, unbalancedness is created by deleting one or more observations. Computations were carried out on a personal computer using SAS/IML software.

For the balanced data in Table 1 (example 1), the estimates obtained by Grubbs' method of moments and by MINQUE are identical. We find

$$\sigma_1^2=145.97,\ \sigma_2^2=-14.14,\ \sigma_3^2=75.15 \text{ and } \sigma_4^2=18.25.$$

The estimates for genotype 2 are negative. Negative estimates are not uncommon with MINQUE and the method of moments. They are a reflection of the large sampling variation often associated with variance component estimates (Searle *et al.*, 1992). For practical purposes negative estimates may be set equal to zero.

Now we delete the value of variety 3 grown in environment 1 (19.47) (example 2). Results for this

Table 1 Wheat yield data

Environment	Genotype			
	1	2	3	4
1	43.60	24.05	19.47	19.41
2	40.40	21.76	16.61	23.84
3	18.08	14.19	16.60	16.08
4	19.57	18.61	17.78	18.29
5	45.20	29.33	20.19	30.08
6	25.78	25.60	23.31	27.04
7	55.20	38.77	21.15	39.95
8	55.32	34.19	18.56	25.12
9	19.79	21.65	23.31	22.45
10	46.24	31.52	22.48	29.28
11	14.88	15.68	19.79	22.56
12	7.52	4.69	20.53	22.08
13	41.17	32.59	29.25	43.95

unbalanced data set are given in Table 2. It is observed that MINQUE and method of moments no longer yield identical estimates. The estimates do not deviate much from the balanced case and the rank order of genotyes has not altered. So if we were to select the most stable genotype, the result would be the same as in example 1.

We now consider the data in Table 3, which were obtained from Table 1 by deleting many observations (example 3). The results are displayed in Table 4. We see that the estimates differ considerably from examples 1 and 2. Moreover, method of moments and MINQUE yield rather discordant estimates. The MINQUE displays the same rank order as in the balanced data set, while the method of moments leads to a reversed ranking. This is indicative of high sampling variation in the estimates. The estimates of σ_4^2 in particular do not merit much confidence, as shown by the large variance estimates shown in Table 5 (27868.76 and 33655.55). This is expected because there are only two observations for genotype 4.

Generally, the high variability of estimates shown in Table 5 suggests that in all examples the stability estimates are not very accurate and should thus be viewed with some caution.

Table 2 Estimates of σ_s^2 for by method of moments and by MINQUE for unbalanced data (example 2)

	Genotype			
	1	2	3	4
Method of moments	148.36	−13.73	81.11	14.45
MINQUE	146.68	−15.71	82.93	19.11

Table 3 Wheat yield data with many empty cells

	Genotype			
	1	2	3	4
Environment				
1	43.60	24.05	19.47	.
2	40.40	21.76	16.61	.
3	.	14.19	16.60	.
4	.	18.61	17.78	.
5	.	.	20.19	.
6	.	.	23.31	.
7	55.20	.	.	39.95
8	55.32	.	.	25.12
9	19.79	21.65	.	.
10	46.24	31.52	.	.
11	14.88	15.68	.	.
12	7.52	4.69	.	.
13	41.17	32.59	.	.

Table 4 Estimates of σ_i^2 for by method of moments and by MINQUE for unbalanced data (example 3)

	Genotype			
	1	2	3	4
Method of moments	34.54	46.94	−34.48	77.21
MINQUE	96.70	−12.94	30.77	15.05

Concluding remark

It has been shown that the stability variance can be estimated in data sets with missing cells. The two methods discussed in this paper may fail, however, if too many observations are missing. In this case one may try to form subsets in which some σ_i^2 are estimable.

For the method of moments, we have described the conditions to be met in order that all σ_i^2 are estimable. These are basically conditions for the non-singularity of the matrix $\boldsymbol{Q}'\boldsymbol{Q}$. Similarly, the estimability of the MINQUE's depends on whether or not the matrix $\boldsymbol{G}$ is singular. An investigation of general conditions under which $\boldsymbol{G}$ is non-singular seems worthwhile but is beyond the scope of this paper. Some hints may be found in Rao (1970), who discusses two sufficient conditions for the non-singularity of $\boldsymbol{F}$. It is conjectured that the conditions for $\boldsymbol{G}$ are similar to those for $\boldsymbol{Q}'\boldsymbol{Q}$.

Table 5 Estimates† of variances for method of moments estimate and MINQUE

	Genotype			
	1	2	3	4
Example 1				
Method of moments	4069.33	138.94	1423.07	306.53
MINQUE	4069.33	138.94	1423.07	306.53
Example 2				
Method of moments	3329.49	51.28	3006.49	183.83
MINQUE	4161.34	160.27	1836.98	241.01
Example 3				
Method of moments	2892.07	142.29	1699.03	27868.76
MINQUE	8678.86	4258.53	3601.85	33655.55

†In the computation of sampling variances negative Grubbs' estimates and MINQUE's were set to zero.

If the data sets are sparse, stability estimates should be used with caution as sampling errors may be quite high. Variance estimates as proposed in this paper may be a helpful guide in judging the reliability of stability estimates.

References

BECKER, H. C. AND LEON, J. 1988. Stability analysis in plant breeding. *Plant Breeding,* **101,** 1–23.

ELLENBERG, J. H. 1977. The joint distribution of the standardized row sums of squares from a balanced two-way layout. *J. Am. Stat. Ass.*, **72**, 407–411.

GRAYBILL, F. A. 1954. Variance heterogeneity in a randomized block design. *Biometrics*, **10**, 516–520.

GRUBBS, F. E. 1948. On estimation of precision of measuring instruments and product variability. *J. Am. Stat. Ass.*, **43**, 243–264.

HAN, C. P. 1969. Testing the homogeneity of variances in a two-way classification. *Biometrics*, **25**, 153–158.

HÜHN, M. AND NASSAR, R. 1989. On tests of significance for nonparametric measures of phenotypic stability. *Biometrics*, **45**, 997–1000.

JAECH, J. L. 1985. *Statistical Analysis of Measurement Errors.* Wiley, New York.

LEVY, K. J. 1975. A multiple range procedure for correlated variances in a two-way classification. *Biometrics*, **31**, 243–246.

LIN, C. S., BINNS, M. R. AND LEVKOVITCH, L. P. 1986. Stability analysis: where do we stand? *Crop Sci.*, **26**, 894–900.

MUDHOLKAR, G. S. AND SARKAR, I. C. 1992. Testing homoscedascity in a two-way table. *Biometrics*, **48**, 883–888.

NASSAR, R. AND HÜHN, M. 1987. Studies on estimation of phenotypic stability: tests of significance for nonparametric measures of phenotypic stability. *Biometrics*, **43**, 45–53.

PIEPHO, H. P. 1992. *Vergleichende Untersuchungen der statistischen Eigenschaften verschiedener Stabilitatsmasse mit Anwendungen auf Hafer, Winterraps, Ackerbohnen sowie Futter- und Zuckerrüben.* Doctoral Thesis, Kiel.

RAO, C. R. 1970. Estimation of heteroscedastic variances in linear models. *J. Am. Stat. Ass.*, **65**, 161–172.

RAO, C. R. 1972. Estimations of variance and covariance components in linear models. *J. Am. Stat. Ass.*, **67**, 112–115.

SHUKLA, G. K. 1972. Some statistical aspects of partitioning genotype–environmental components of variability. *Heredity*, **29**, 237–245.

SEARLE, S. R., CASSELLA, G. AND MCCULLOCH, C. E. 1992. *Variance Components.* Wiley, New York.

SNEE, R. D. 1982. Nonadditivity in a two-way classification: is it interaction or nonhomogeneous variance? *J. Am. Stat. Assoc.*, **77**, 515–519.

WESTCOTT, B. 1986. Some methods of analysing genotype–environment interactions. *Heredity*, **56**, 243–253.

WRICKE, G. 1965. Die Erfassung der Weschselwirkungen zwischen Genotyp und Umwelt bei quantitativen Eigenschaften. *Zeitschrift für Pflanzenzüchtung*, **53**, 266–343.