Kaizô I. Beltrão and Djalma G.C. Pessoa, Instituto de Matemática Pura e Aplicada

1. Introduction

In this paper some estimators of treatment effects and of strata variances are analyzed, under the framework of General Balance (G.B.) introduced by Nelder (1965 a,b). We deal with the case where T, the treatment space, is balanced with respect to all strata in the block structure.

Under (G.B.) the data covariance matrix is determined by the block structure and the strata variances $\{\xi_\alpha\}$. Assuming $\{\xi_\alpha\}$ known, the overall BLUE of the treatment effects (or contrasts of it) uses information from all effective strata. It is computed by combining the BLUE in each stratum, the latter not depending on $\{\xi_\alpha\}$, with weights proportional to the inverse of the $\{\xi_\alpha\}$. Alternative estimators of treatment effects can be obtained by plugging in the expression of the overall BLUE different estimators of $\{\xi_\alpha\}$.

In Section 2 we present the needed notation and in Section 3 we describe several estimators of $\{\xi_\alpha\}$, the strata variances. The following estimators are considered: the straightforward one using only the stratum residuals; Yates estimator; Nelder's estimator; and a "leave-one-stratum-out" estimator. In Section 4 we derive the distribution of some of those estimators, in the case of two effective strata (e.g., BIB, Youden Squares) and under normality assumptions for the data. In Section 5 we report the results of a Monte-Carlo simulation comparing the estimators in a BIB design.

2. Notations and Basic Notions

Assume we have a random vector of observations $y = (y_i)_{i\varepsilon I}$, taking values in a vector space $D = R^I$ and indexed by a set I with cardinality $|I| = n$. Consider in D the usual Euclidean norm.

The models to be considered are linear, $\tau = EY \varepsilon T$, where $T \subseteq D$ is a linear subspace of D. Following Nelder (1965) we consider an orthogonal block structure given by a partition $D = \oplus S_\alpha$, or equivalently by the decomposition $I = \Sigma S_\alpha$, where S_α is, for each α, the orthogonal projection on S_α and $\{S_\alpha\}$ satisfy $S_\alpha = S_\alpha' = S_\alpha^2$, $\forall_\alpha$ and $S_\alpha S_\beta = S_\beta S_\alpha = 0$ if $\alpha \neq \beta$.

The dispersion matrix of Y will be assumed to have the form $DY = V = \Sigma\xi_\alpha S_\alpha$, where $\xi_\alpha \geq 0$ for all α. Situations where the above model arises are described in Houtman and Speed (1983), from whom we borrow the notation and terminology used here. Moreover we call *strata* the subspaces $\{S_\alpha\}$ and, since $DS_\alpha y = \xi_\alpha S_\alpha$, we call $\{\xi_\alpha\}$ strata variances. In order to relate T and $\{\xi_\alpha\}$ we consider the property of General Balance (G.B.).

Definition. A design with block structure $\{S_\alpha\}$ is generally balanced (G.B.) with respect to the partition $T = \oplus_\beta T_\beta$ if there exists $\{\lambda_{\alpha\beta}\}$ such that for all α $TS_\alpha T = \Sigma_\beta\lambda_\beta T_\beta$, where T is the projection on to T and for each β, T_β is the orthogonal projection on T_β. In particular in this paper, we deal with the case where the whole space T is a eigenspace of $TS_\alpha T$, i.e., $TS_\alpha T = \lambda_\alpha T$, $\forall_\alpha$.

If a design satisfies the G. B. property and the $\{\xi_\alpha\}$ are known (up to a constant) it is very simple to get the BLUE of τ and of contrasts $\langle t,\tau\rangle$, $\langle t,l\rangle = 0$, t e T, by simply combining their BLUEs in different strata. The reduced data $S_\alpha y$ to the stratum S_α satisfies

$$ES_\alpha Y = S_\alpha \tau \ \varepsilon \ S_\alpha T \text{ and } DS_\alpha Y = \xi_\alpha S_\alpha,$$

and the usual least-squares analysis can be used (S_α is the identity in S_α). The least-squares estimator $\hat{\tau}_\alpha$ of τ within the stratum S_α satisfies the normal equation

$$TS_\alpha T\tau = TS_a y$$

and is given by $\hat{\tau}_\alpha = \lambda_\alpha^{-1}TS_\alpha y$.

Under (G.B.) the overall BLUE of τ is as follows

$$\hat{\tau} = Uy = \Sigma\omega_\alpha\hat{\tau}_\alpha$$

where $\omega_\alpha = \xi_\alpha^{-1}\lambda_\alpha/\Sigma(\xi_\alpha^{-1}\lambda_\alpha)$ is the weight for the α stratum. The BLUE of a contrast $\langle t,\tau\rangle$, t ε T and $\langle t,l\rangle = 0$, is then

$$\langle t,\hat{\tau}\rangle = \Sigma\omega_\alpha\lambda_\alpha^{-1}\langle t,TS_\alpha Y\rangle$$

$$= \Sigma\omega_\alpha\langle t,\hat{\tau}_\alpha\rangle .$$

Now, observe that $cov(\langle t,TS_\alpha Y\rangle,\langle t,TS_{\alpha'}Y\rangle) = 0$ if $\alpha \neq \alpha'$ and $var(\langle t,\tau_\alpha\rangle) = \lambda_\alpha^{-1}\xi_\alpha||t||^2$, therefore the overall BLUE of $\langle t,\tau\rangle$, in the G.B. case, is a combination of the BLUEs of $\langle t,\tau\rangle$ in each stratum, with weight proportional to the inverse of its variance.

In order to compute the appropriate weights we need to estimate the $\{\xi_\alpha\}$. The ANOVA table within the stratum α allows, in some cases, an easy way of estimating ξ_α (see Table I).

3. Estimators of Strata Variances

In this section we present different estimators of $\{\xi_\alpha\}$. They are all obtained by equating the expected value of a residual squared norm to its observed value. Consider first the simplest estimator, using the residual sum of squares within each stratum, when $d_\alpha = \dim S_\alpha - \dim T \neq 0$. This estimate is defined as

$$\hat{\xi}_{\alpha,s} = \frac{||S_\alpha Y||^2 - \frac{1}{\lambda_\alpha}||TS_\alpha Y||^2}{\dim S_\alpha - \dim T}$$

and is not feasible for several commonly used designs such as the symmetric BIBD and some types of lattices where $d_\alpha = 0$. Yates (1939) proposed for the BIBD design an estimator for the interblock variance based upon the interblock sum of squares eliminating the treatment. This estimator can be extended to the multistrata case considering the sum of squares for a given stratum $S = S_\gamma$ eliminating T, denoted by $SS_{S/T}$ and given by

$$SS_{S/T} = SS_{T/S} + SS_S - SS_T$$

$$= \| \sum_{\alpha \neq \gamma} \frac{1}{\lambda_\alpha} S_\alpha T S_\alpha y \|^2 + \| S_\gamma y \|^2 - \| Ty \|^2 \tag{1}$$

with expectation

$$E(SS_{S/T}) = \xi_\gamma(\dim S_\gamma - \lambda_\gamma \dim T) + \sum_{\alpha \neq \gamma} \xi_\alpha (1 - \lambda_\alpha) \dim T \,. \tag{2}$$

If we have unbiased estimators for all the $\{\xi_\alpha\}$ but one, say ξ_γ, with $d_\gamma = 0$, we can define from (2) an unbiased estimator of ξ_γ

$$\hat{\xi}_{\gamma,Y} = \frac{SS_{S/T} - [\sum_{\alpha \neq \beta} \hat{\xi}_\alpha (1-\lambda_\alpha)]\dim T}{\dim S_\gamma - \lambda_\gamma \dim T} \tag{3}$$

where $\hat{\xi}_\alpha$ is an unbiased estimator of ξ_α. If $d_\alpha = 0$ in more than one stratum, we can write a linear system of equations similar to (3) and get explicit estimators for all the $\hat{\xi}_\alpha$'s with $d_\alpha = 0$.

In the first estimator described, the number of degrees of freedom (d.f.) left to estimate ξ_α within S_α is $d_\alpha = \dim S_\alpha - \dim T$. In Yates' estimator we use $\dim S - \lambda_\gamma \dim T$ d.f. by subtracting only a fraction of the d.f. of the sum of squares due to treatments. Nelder's approach (1968) instead subtracts from $\dim S_\gamma$ the fraction $\omega_\alpha \dim T$ of the d.f. due to T, corresponding to the amount of information on T in S_α. For this, he uses the "actual" residual in S_α which is $S_\alpha(I-U)y$.

The residual norm squared considered is then $\| S_\alpha(I-U)y \|^2$, with expectation

$$E \| S_\alpha (I-U) Y \|^2 = d'_\alpha \xi_\alpha \tag{4}$$

where

$$d'_\alpha = d_\alpha + (1-\omega_\alpha)\dim T = \dim S_\alpha - \omega_\alpha \dim T$$

since ω_α depends on ξ_α, the equation

$$\| S_\alpha (I-U) y \|^2 = d'_\alpha \xi_\alpha \tag{5}$$

can be solved only iteratively. It is worth noticing that equation (5) defines the likelihood equation for ξ_α based upon $\| (I-T)y \|^2$ under the assumption that Y has a multivariate normal distribution (for details see Patterson and Thompson (1971)).

Another iterative estimator for the variance of the γ stratum can be obtained using the information on τ from all but the γ stratum, and which we call the leave-one-out stratum estimator. The BLUE of τ based upon $(I-S_\gamma)y$, with known $\{\xi_\alpha\}$, $\alpha \neq \gamma$, is

$$\hat{\tau}_\gamma = Q_\gamma y = \Sigma \omega'_\alpha (1/\lambda_\alpha) T S_\alpha y$$

where $\omega'_\alpha = (\lambda_\alpha/\xi_\alpha)/\Sigma(\lambda_\alpha/\xi_\alpha)$ $\forall$ α with all the summations defined for $\alpha \neq \gamma$.

In this case, the residual under consideration is $S_\gamma(I-Q_\gamma)y$, with the expected squared norm

$$E \| S_\gamma (I-Q_\gamma) y \|^2 = \xi_\gamma [\dim S_\gamma + \omega'_\gamma \dim T] .$$

The estimator can be computed by iteratively solving

$$\| S_\gamma (I-Q_\gamma) y \|^2 = \xi_\gamma (\dim S_\gamma + \omega'_\gamma \dim T) . \tag{6}$$

Since to estimate ξ_γ we use the BLUE of τ, computed by excluding the information in S_γ, the "leave-one-out stratum estimator" will work better when estimating the ξ_γ with greater value.

4. Variances of Estimators of ξ_α. Two Strata Case.

In this section we assume that Y has a multivariate normal distribution with mean vector $\tau = 0$ and dispersion matrix $V = \Sigma \xi_\alpha S_\alpha$.

To get the distribution of the variance strata estimators we apply the following theorem (see Box (1954)).

Theorem. If Y has multivariate normal distribution with mean vector $\tau = 0$ and dispersion matrix V, the real quadratic form $Q = Y'MY$ is distributed as a linear combination of independent Chi-squares with one d.f. $(\chi_{(1)}^2)$ and weights given by the eigenvalues of VM.

Using the above theorem it is immediate to show that $\hat{\xi}_{\alpha,S}$ as defined by equation (1) is distributed as a $\xi_\alpha/d\chi_{(d)}^2$, which has expectation ξ_α and variance $2\xi_\alpha^2/d$, (with $d = d_\alpha$). Using the regular within stratum estimator for ξ_2, Yates interblock variance estimator is

$$\hat{\xi}_{1,Y} = \frac{SS_{S/T} - (1-\lambda_2)\dim T\, \hat{\xi}_{2,Y}}{\dim S_1 - \lambda_1 \dim T}$$

It is easy to check that $SS_{S/T}$ and $\hat{\xi}_{2,S}$ are independent and $SS_{S/T}$ is distributed as

$$\frac{(\xi_1\lambda_2+\xi_2\lambda_1)}{\dim S_1 - \lambda_1 \dim T} \chi^2_{(\dim T)} + \frac{\xi_1}{\dim S_1 - \lambda_1 \dim T} \chi^2_{(d_1)} ,$$

and from this we can compute the variance of $\hat{\xi}_{1,Y}$.

Considering that Nelder's estimator is defined as a solution of (5), which can only be solved iteratively, its variance would be hard to compute. It is possible however to get the distribution of the actual residual $\| S_\alpha(I-U)Y \|^2$ and this is

$$\xi_\alpha [\chi^2_{(d_\alpha)} + (1-\omega_\alpha) \chi^2_{(\dim T)}] , \tag{7}$$

which could be used to approximate the distribution of $\hat{\xi}_{\alpha,N}$. Nelder (1968) checked empirically, through a Monte-Carlo study the approximation of this distribution by $\xi_\alpha \chi_{d'}{}^2/d'$ which is actually a first moment approximation to (7) (where $d' = d_\alpha$).

The leave-one-stratum-out estimators when only two effective strata exist have closed forms

$$\hat{\xi}_{1,L}=\frac{||S_1(y-\hat{\tau}_1)||^2\dim S_2-||S_2(y-\hat{\tau}_1)||^2\frac{\lambda_1}{\lambda_2}\dim T}{\dim S_1\dim S_1-\dim^2 T}$$

and

$$\hat{\xi}_{2,L}=\frac{||S_2(y-\hat{\tau}_2)||^2\dim S_1-||S_1(y-\hat{\tau}_1)||\ \frac{\lambda_2}{\lambda_1}\dim T}{\dim S_1\dim S_2-\mathrm{do},^1 T}\ .$$

It is straightforward to check that $((\xi_1/\lambda_1)S_1-(\xi_2/\lambda_2)S_2)T$, T, $(S_1-(1/\lambda_1)S_1TS_1)D$ are the common eigenspaces (with eigenvalue 0 for T) and that the numerators are distributed, respectively, as:

$$(\xi_1+\frac{\lambda_1}{\lambda_2}\xi_2)d_2\chi^2_{(\dim T)}$$
$$+\xi_1\dim S_2\chi^2_{(d_1)}-t\frac{\lambda_1}{\lambda_2}\xi_2\chi^2_{(d_2)}$$

and

$$(\xi_2+\frac{\lambda_2}{\lambda_1}\xi_1)d_1\chi^2_{(\dim T)}$$
$$+\xi_2\dim S_1\chi^2_{(d_2)}-t\frac{\lambda_2}{\lambda_1}\xi_1\chi^2_{(d_1)}\ .$$

5. Empirical Study of Estimators of ξ_1 and ξ_2

The performance of the estimators discussed above was assessed by a Monte-Carlo study. The BIBD exemplified by Cochran and Cox (1957) and used by Nelder (1968) was the design chosen for the simulation. It is a BIBD with 15 blocks of two plots each, 6 treatments with 5 replications and efficiency factor $\lambda = 3./5$. One thousand pseudomultinormal vectors of dimension 30 were generated by using IMSL subroutine GGNSM for three different ξ_2/ξ_1 ratios. For each one of these 1000 vectors all the competing estimators were computed. For Nelder's iterative estimator, a 50 iterations limit and a tolerance of 10^{-4} were adopted as the stopping rule. Typical numbers of iterations until convergence were around 3 to 5, few of them being over 10 (less than 5%). Table II displays a summary of the results obtained. The values (1,1), (3,1) and (5,1) were used for the pair (ξ_1,ξ_2), where ξ_1 is the inter-block and ξ_2 the intra-block variance. In this table, the first value in each cell is the average of the respective estimator in all runs followed by its variance. The value in parenthesis is the computed theoretical variance of the estimator, whenever it had a closed form. For Nelder's estimator the corresponding figures are what the authors believe to be a lower bound for the variances. The estimators empirical averages were, as expected, always close to its true value. The empirical median bias was 1% and in 20 out of 24 cases had absolute value less than 2%. In most cells the differences between theoretical and empirical variance is of the order of two decimals points. The within stratum estimator of ξ_2 seems to do not worse than any other and is easy to compute. The leave-one-out, as expected, does the worse for the ξ_1/ξ_2 ratios considered. The within stratum estimator of ξ_1 did not perform well as indicated both by theoretical and empirical variances and the others presented similar performances. From the theoretical variances computed, the leave-one-out does slight better than Yates' for estimating ξ_1. These considerations suggest, for the design studied, that one should use either Yates estimator or a combination of within stratum to estimate ξ_2 and leave-one-out to estimate ξ_1. If we rearrange the same five replications of six treatments in a 6/5 block structure (instead of the 15/2), we will get a symmetric BIB. In this case (symmetric BIB) the within stratum estimator for ξ_1 is not defined ($\dim S_1 = \dim T$) and it is easy to see that all the other estimators considered are exactly the same.

The exact variance of $\hat{\xi}_2$ in the symmetric BIB is

$$\frac{2\xi_1^2}{\dim T}+2[(\frac{\lambda_1\xi_2}{\lambda_2})^2(\frac{1}{d_2}+\frac{1}{\dim T})+\frac{2\xi_1\xi_2\lambda_1}{\lambda_2\dim T}]\ .$$

The "lower bound" suggested for the variance of Nelder's estimator is the leading term of this expression. For the 6/5 block structure considered and for $\xi_1 = 1$, the remaining terms amount to only $.0009\ \xi_2^2 + .0333\ \xi_2$.

6. References

Box, G. E. P. (1954) - Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, II. Effect of Inequality of Variance and Correlations Between Errors in the Two-way Classification, *Annals of Mathematical Statistics,* 25, 290-302.

Cochran, W. G. and Cox, G. M. (1957) - *Experimental Designs*, Second Edition Wiley, New York.

Houtman, A. M., and Speed, T. P. (1983) - Balance in Designed Experiments with Orthogonal Block Structure, *Annals of Statistics,* 11, 1609-1085.

Nelder, J. A. (1965a) - The Analysis of Randomized Experiments with Orthogonal Block Structure, *I.* Block Structure and the Null Analysis of Variance. *Proceedings of the Royal Society* (London), Ser. A. 273, 147-162.

Nelder, J. A. (1965b) - The Analysis of Randomized Experiments with Orthogonal Block Structure, II. Treatment Structure and the General Analysis of Variance. *Proceedings of the Royal Society* (London), Ser. A. 273, 163-178.

Nelder, J. A. (1968) - The Combination of of Informations in Generally Balanced Designs. *Journal of the Royal Statistical Society*. Ser. B, 30, 303-311.

Patterson, H. D. and Thompson, E. R. (1971)- Recovery of Inter-Block Information When Block Sizes are Unequal. *Biometrika,* 58, 545-554.

Yates, F. (1939) - The Recovery of Inter Block Information in Balanced Incomplete Block Designs. *Annals of Eugenics,* 10, 317-325.

Table I. ANOVA table within stratum α

Source	d.f.	SS	E(MS)
Treatment	$\dim \mathcal{T}$	$\lambda_\alpha^{-1}\|\|TS_\alpha y\|\|^2$	$\xi_\alpha + \frac{\lambda_\alpha}{\dim \mathcal{T}}\|\|\tau\|\|^2$
Residual	$d_\alpha = \dim \mathcal{S}_\alpha - \dim \mathcal{T}$	$\|\|S_\alpha y\|\|^2 - \lambda_\alpha^{-1}\|\|TS_\alpha y\|\|^2$	ξ_α (if $d_\alpha > 0$)
Total	$\dim \mathcal{S}_\alpha$	$\|\|S_\alpha y\|\|^2$	

Table II

		Within Stratum		Yates		Nelder		Leave-One-Out	
		ξ_1	ξ_2	ξ_1	ξ_2	ξ_1	ξ_2	ξ_1	ξ_2
	$\hat{\mu}$	.998	1.023	1.019	1.026	1.015	1.036	1.021	1.057
$\xi_1 = 1$	$\hat{\sigma}^2$	.227	.204	.250	.204	.231	.201	.261	.327
	$\hat{\sigma}^2$	(.222)	(.200)	(.206)	(.200)	(.150)	(.150)	(.206)	(.295)
	$\hat{\mu}$	3.008	.977	3.036	.977	3.038	.983	3.038	1.018
$\xi_1 = 2$	$\hat{\sigma}^2$	1.934	.183	1.405	.183	1.420	.195	1.404	1.099
	$\hat{\sigma}^2$	(2.000)	(.200)	(1.472)	(.200)	(1.297)	(.171)	(1.464)	(1.097)
	$\hat{\mu}$	5.015	.977	5.052	.977	5.060	.982	5.055	1.031
$\xi_1 = 5$	$\hat{\sigma}^2$	5.342	.184	3.719	.184	3.701	.199	3.691	2.569
	$\hat{\sigma}^2$	(5.555)	(.200)	(3.939)	(.200)	(3.584)	(.180)	(3.903)	(2.497