

**Evaluating the Efficiency of
Blocking without Assuming Compound Symmetry**

by

Myra L. Samuels, George Casella and George P. McCabe

March 1990

Abstract

In analyzing a randomized blocks experiment, an investigator may wish to evaluate the effect of blocking on the efficiency of treatment comparisons. We have recently (Samuels, Casella, and McCabe 1990) discussed this topic in the context of a mixed linear model with covariance matrix assumed compound symmetric. The present note suggests a simple approach, valid for an arbitrary covariance matrix, to testing and estimating the effect of blocking on the efficiency of an experiment.

KEY WORDS: Randomized blocks; Mixed model; Analysis of variance; Linear models; Compound symmetry.

Evaluating the Efficiency of Blocking without Assuming Compound Symmetry

Myra L. Samuels, George Casella, and George P. McCabe

March 1990

1. Introduction

For a randomized blocks design, it is often appropriate to model blocks as the random factor in a mixed linear model. We have argued recently (Samuels, Casella, and McCabe 1990) that in this case it is sensible to test and estimate the impact of blocking on the efficiency of treatment comparisons. Our recent discussion was limited, however, to the case where the covariance matrix of the observations is compound symmetric. The purpose of the present note is to suggest that in many applications the assumption of compound symmetry is neither desirable nor necessary, and to point out that the effect of blocking on efficiency can be tested and estimated without it.

Let Y_{ijk} denote the k th observation on the i th treatment in the j th block, $i = 1, \dots, I$; $j = 1, \dots, J$; $k = 1, \dots, K$. We assume that the random vectors $\{\bar{Y}_{1j}, \dots, \bar{Y}_{Ij}\}'$, $j = 1, \dots, J$ are independently and identically distributed (iid) as the vector $(Y_1, \dots, Y_I)'$, which in turn is multivariate normal with mean vector $(\mu_1, \dots, \mu_I)'$ and covariance matrix Σ . (An overbar and a dot denote averaging over a subscript.) The covariance matrix Σ is compound symmetric if $\text{Var}(Y_i)$ is the same for all i and $\text{Cov}(Y_i, Y_{i'})$ is the same for all $i \neq i'$.

We consider the case of two treatments, $I = 2$, in Section 2 and extensions to the case $I > 2$ in Section 3. In Section 4 we discuss the implications of assuming compound symmetry. We summarize our conclusions in Section 5.

In Section 2 and most of Section 3 we will assume that $K = 1$, and will drop the third subscript and write the observations as $\{Y_{ij}\}$. This restriction is for notational simplicity only, and will not limit the applicability of the results.

2. The Case of Two Treatments

For $I = 2$, compound symmetry of Σ means simply that the variances are homogeneous, i.e.,

$$\text{Var}(Y_1) = \text{Var}(Y_2). \quad (2.1)$$

To compare the treatment means, the usual procedures are the paired t test of $H_0: \mu_1 = \mu_2$ and the associated confidence interval for $(\mu_1 - \mu_2)$. These procedures are based on the pairwise differences U_j defined by $U_j = Y_{1j} - Y_{2j}$ and are valid regardless of whether (2.1) holds.

2.1. An Efficiency Parameter

The precision of inference on $(\mu_1 - \mu_2)$ is determined by the variance of $(Y_1 - Y_2)$. For a non-blocked design, Y_1 and Y_2 are modeled as independent random variables, so that

$$\text{Var}_N(Y_1 - Y_2) = \text{Var}(Y_1) + \text{Var}(Y_2)$$

whereas for the blocked design we have

$$\text{Var}_B(Y_1 - Y_2) = \text{Var}(Y_1) + \text{Var}(Y_2) - 2 \text{Cov}(Y_1, Y_2).$$

The relative efficiency* of the blocked design is the ratio λ of these two variances:

$$\lambda = \frac{\text{Var}_N(Y_1 - Y_2)}{\text{Var}_B(Y_1 - Y_2)} = \frac{\text{Var}(Y_1) + \text{Var}(Y_2)}{\text{Var}(Y_1) + \text{Var}(Y_2) - 2 \text{Cov}(Y_1, Y_2)} \quad (2.2)$$

To obtain the same variance as a blocked design with J observations on each treatment, λJ observations on each treatment would be required in a non-blocked design. If the covariance between Y_1 and Y_2 is negative (a situation unlikely to be encountered in most applications, but possible in principle), then blocking results in a loss, rather than a gain, of efficiency.

Let

$$\tilde{\rho} = \frac{\text{Cov}(Y_1, Y_2)}{\frac{1}{2}[\text{Var}(Y_1) + \text{Var}(Y_2)]}. \quad (2.3)$$

* Strictly speaking, λ is the asymptotic relative efficiency, since it does not account for the loss in degrees of freedom due to blocking.

Then we can write (2.2) as

$$\lambda = \frac{1}{1 - \tilde{\rho}}. \quad (2.4)$$

Note the similarity between $\tilde{\rho}$ and the Pearson correlation coefficient

$$\rho = \frac{\text{Cov}(Y_1, Y_2)}{[\text{Var}(Y_1) \text{Var}(Y_2)]^{\frac{1}{2}}}. \quad (2.5)$$

The coefficients $\tilde{\rho}$ and ρ have the same sign. However, $|\tilde{\rho}| \leq |\rho|$, with equality if (2.1) holds. (The latter facts follow because the denominator of $\tilde{\rho}$ is the arithmetic mean of the variances, whereas the denominator of ρ is the geometric mean.)

2.2 Testing $H_0: \tilde{\rho} = 0$

From Equation (2.4) it is clear that the hypothesis that blocking has no effect on efficiency can be expressed as

$$H_0: \tilde{\rho} = 0 \quad (2.6)$$

or, equivalently,

$$H_0: \text{Cov}(Y_1, Y_2) = 0 \quad (2.7)$$

A standard test of (2.7) is based on the Pearson product-moment correlation

$$r = \frac{\sum_j (Y_{1j} - \bar{Y}_{1.})(Y_{2j} - \bar{Y}_{2.})}{[(\sum_j (Y_{1j} - \bar{Y}_{1.})^2 \sum_j (Y_{2j} - \bar{Y}_{2.})^2)]^{\frac{1}{2}}}. \quad (2.8)$$

We have under (2.7) the familiar result that

$$r \sqrt{\frac{J-2}{1-r^2}} \sim t_{J-2}, \quad (2.9)$$

where t_n is a t random variable with n degrees of freedom. Against a two-sided alternative the test based on (2.9) is the likelihood ratio test (LRT).

If it is assumed that $\text{Var}(Y_1) = \text{Var}(Y_2)$, then the LRT of (2.7) is not the t test (2.9) but rather the F test based on the fact that under (2.7)

$$\frac{\text{MS}(\text{Blocks})}{\text{MS}(\text{T*B})} \sim F_{J-1, J-1} \quad (2.10)$$

where $MS(\text{Blocks}) = 2K(J-1)^{-1} \sum_j (\bar{Y}_{.j} - \bar{Y}_{..})^2$ and $MS(\text{T*B}) = K(J-1)^{-1} \sum_{i,j} (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2$ are the usual mean squares for blocks and for treatment by block interaction, and $F_{m,n}$ denotes an F random variable with m and n degrees of freedom. However, if $\text{Var}(Y_1) \neq \text{Var}(Y_2)$ then $MS(\text{Blocks})$ and $MS(\text{T*B})$ are not independent and (2.10) fails.

Thus, we find that under the variance homogeneity assumption (2.1) there are two tests, one based on (2.9) and one based on (2.10), that are valid for testing the efficiency effect of blocking. If in fact (2.1) holds, then the F test based on (2.10) is more powerful than the t test based on (2.9) (the F test is uniformly most powerful unbiased, as can be shown by applying Theorems 1 and 2 of Chapter 5 of Lehmann 1986, and the t test is also unbiased and so cannot be more powerful). However, computations suggest that the difference in power is very small, as indicated in Table 1. We would argue that the advantage of the t test, namely, that it does not depend on the assumption of equal variances, would usually outweigh its slight potential deficiency in power, and that the t test therefore should generally be preferred to the F test.

Table 1 goes here

2.3. Estimation of $\tilde{\rho}$

Estimation of the gain (or, perhaps, loss) of efficiency due to blocking can help an investigator decide whether the blocking was cost effective (it may have been troublesome or expensive), or whether it should be omitted in future studies in similar experimental settings. We consider the estimation of the block effect parameter $\tilde{\rho}$, which is related to the relative efficiency λ through Equation (2.4).

A natural estimator of $\tilde{\rho}$ is

$$\tilde{r} = \frac{\sum_j (Y_{1j} - \bar{Y}_{1.})(Y_{2j} - \bar{Y}_{2.})}{\frac{1}{2}[\sum_j (Y_{1j} - \bar{Y}_{1.})^2 + \sum_j (Y_{2j} - \bar{Y}_{2.})^2]}. \quad (2.11)$$

This statistic is the maximum likelihood estimator of $\tilde{\rho}$, whether or not the variance homogeneity condition (2.1) is assumed to hold (Kristof 1963, Mehta and Gurland 1969). Furthermore, it is easy to show that

$$\tilde{r} = \frac{MS(\text{Blocks}) - MS(\text{T*B})}{MS(\text{Blocks}) + MS(\text{T*B})},$$

so that \tilde{r} is easily calculated from the mean squares in the ANOVA table. The statistic \tilde{r} is sometimes called an intraclass correlation coefficient or a reliability coefficient; more often, however, these names are given to a somewhat different statistic (see Winer 1971, pp. 286–287; Snedecor and Cochran 1989, p. 243).

A confidence interval for $\tilde{\rho}$ can be obtained from results of Kristof (1972), who showed that

$$A \frac{\tilde{r} - \tilde{\rho}}{\sqrt{1 - \tilde{\rho}^2}} \sim t_{J-2}, \quad (2.12)$$

where

$$A = \frac{r}{\tilde{r}} \sqrt{\frac{J-2}{1-r^2}}$$

and r is given by (2.8) and \tilde{r} by (2.11). It follows that $100(1 - \alpha)\%$ confidence limits for $\tilde{\rho}$ are given by

$$\frac{\tilde{r} \pm B\sqrt{1 - \tilde{r}^2 + B^2}}{1 + B^2}, \quad (2.13)$$

where $B = A^{-1}t_{1-\alpha/2; J-2}$ and $t_{p;n}$ is the p th percentile of a t_n distribution. Confidence limits for λ are $1/(1-L)$ and $1/(1-U)$ where L and U are the lower and upper confidence limits for $\tilde{\rho}$. The limits (2.13) are equivalent to a family of tests, of which the t test (2.9) is a member. The results of Kristof (1972) also yield confidence limits for $\tilde{\rho}$ under the variance homogeneity condition (2.1); the limits are given in Samuels, Casella and McCabe (1990). Kristof (1972) shows that the confidence interval assuming (2.1) has shorter expected length than the interval (2.13); however, the difference in expected length is very small.

2.4. Example

The following example illustrates the testing and estimation of $\tilde{\rho}$.

Example 1. For a study of environmental influences on brain anatomy, 24 young male rats, in littermate pairs, were randomly allocated to an “enriched” environment, with toys and companions, or an “impoverished” environment (see Rosenzweig 1972, and Freedman, Pisani and Purves 1978, pp. 451–452). The observation Y was the weight of the cerebral cortex after one month. The data are shown in Table 2. For these data, the (two-tailed) P -value for the test (2.9) is $P = .037$. The Pearson correlation is $r = .60$. The point estimate of $\tilde{\rho}$ is $\tilde{r} = .54$ and of λ is $(1 - \tilde{r})^{-1} = 2.2$. Thus, we estimate that it would have

required $2.2 \times 24 \approx 53$ animals to achieve the same precision with independent groups of animals rather than littermate pairs. A 95% confidence interval for $\tilde{\rho}$ is (.039, .823); the corresponding interval for λ is (1.04, 5.65).

Table 2 goes here

3. More than Two Treatments

If Σ is not assumed compound symmetric, the analysis of the randomized blocks design for $I > 2$ can be complicated; even the usual “ F statistic” for treatments does not in general have an F distribution.

3.1. A Simple Approach

A simple approach which avoids the complexity of the full multivariate problem is to concentrate on individual contrasts (of course, this requires attention to the problem of multiple inference if several contrasts are investigated).

Consider a contrast

$$\xi = \sum_i c_i \mu_i \quad (3.1)$$

where

$$\sum_i c_i = 0.$$

The usual estimator of ξ , namely,

$$X = \sum_i c_i \bar{Y}_{i.},$$

can also be written as

$$X = \bar{U} = J^{-1} \sum_j U_j \quad (3.2)$$

where

$$U_j = \sum_i c_i Y_{ij}. \quad (3.3)$$

The investigator who prefers to avoid a compound symmetry assumption can base tests and estimates of ξ on the statistics \bar{U} and

$$SS_U = \sum_j (U_j - \bar{U})^2. \quad (3.4)$$

(This approach is similar to that commonly recommended for some “repeated measures” designs (e.g., Kirk 1982, p. 543; Winer 1971, p. 540); in fact, it is equivalent, if each “within-subjects” factor has only two levels.)

Now suppose the investigator wishes to assess the effect of blocking on the efficiency of the analysis based on (3.2) and (3.4). This is easily accomplished, because a simple shift of viewpoint will convert the problem to the form discussed in Section 2.

For a given contrast ξ defined by (3.1), with corresponding $\{U_j\}$ defined by (3.3), let $S^+(S^-)$ denote the set of indices i for which c_i is positive (negative). Then U_j can be written as

$$U_j = Y_{1j}^* - Y_{2j}^* \quad (3.5)$$

where

$$Y_{1j}^* = \sum_{i \in S^+} |c_i| Y_{ij}$$

$$Y_{2j}^* = \sum_{i \in S^-} |c_i| Y_{ij}$$

and, correspondingly, ξ can be written as

$$\xi = E(Y_{1j}^*) - E(Y_{2j}^*).$$

The discussion of Section 2 can now be carried over, with Y_{1j}^* and Y_{2j}^* playing the roles of Y_{1j} and Y_{2j} . The relation (2.3) defines a block effect parameter relevant to the given contrast ξ , and (2.9), (2.11) and (2.13) give the corresponding test, point estimate, and confidence interval.

The following is an illustration.

Example 2. In a study of the effect of caffeine on muscle metabolism, volunteers underwent four exercise tests: on separate occasions in random time order, subjects exercised either their arms or their legs after ingesting either a caffeine or a placebo capsule. A metabolic variable Y was measured on each occasion. Using the labeling

Y_1 : arm, placebo

Y_2 : arm, caffeine

Y_3 : leg, placebo

Y_4 : leg, caffeine

the contrast of primary interest is a measure of the main effect of caffeine, namely,

$$\xi = \frac{1}{2}(\mu_1 - \mu_2) + \frac{1}{2}(\mu_3 - \mu_4).$$

Correspondingly, one can define for each subject the variables

$$Y_1^* = \frac{1}{2}(Y_1 + Y_3)$$

$$Y_2^* = \frac{1}{2}(Y_2 + Y_4)$$

and one can test and estimate ξ using \bar{U} and SS_U , with U_j defined by (3.5). To evaluate the impact of blocking on the efficiency of this analysis, one can apply (2.8), (2.9), (2.11), and (2.13) to Y_1^* and Y_2^* in place of Y_1 and Y_2 .

3.2. Testing all Correlations Simultaneously

For each contrast ξ , the approach of the preceding section defines a parameter, say $\tilde{\rho}_\xi$, which describes the impact of blocking on the efficiency with which ξ can be tested or estimated. If Σ is compound symmetric, then $\tilde{\rho}_\xi$ is the same for any ξ (see Samuels, Casella, and McCabe 1990). If Σ is arbitrary, however, then $\tilde{\rho}_\xi$ may vary in magnitude, and even in sign, for various ξ . It is even possible for $\tilde{\rho}_\xi$ to be zero for $(I - 1)$ orthogonal ξ , and yet for some other ξ to have $\tilde{\rho}_\xi$ nonzero.

The null hypothesis

$$H_0: \text{Cov}(Y_i, Y_{i'}) = 0 \text{ for all } i \neq i' \quad (3.6)$$

is equivalent to asserting that $\tilde{\rho}_\xi$ is zero for all ξ . The LRT of (3.6), which is based on the determinant of the sample correlation matrix, has been studied by several authors; a union-intersection test of (3.6) has also been obtained (see Seber 1984, p. 92, and references therein). It seems apparent, however, that any test of (3.6) against a nondirectional alternative, which perforce must test $I(I - 1)/2$ parameters simultaneously, would represent a less powerful approach than limiting attention, as we have suggested, to a small number of contrasts ξ selected for their substantive interest.

3.3. The Case $K > 1$

Thus far in Sections 2 and 3 we have assumed no replication ($K = 1$). We now return to the case of general K . The entire discussion of Sections 2 and 3 extends immediately, by simply identifying Y_{ij} of Sections 2 and 3 with \bar{Y}_{ij} of the general case.

4. Implications of the Compound Symmetry Assumption

A common model for a randomized blocks design is a linear model with treatments fixed, blocks random, and no treatment-block interaction. In such a model, the covariance matrix of the observations is automatically compound symmetric.

In our earlier paper (Samuels, Casella, and McCabe 1990) we argued that, when blocks are regarded as a random factor, it is preferable to have a view of blocked designs which does not depend upon treatment-block additivity; however, in that paper we did assume compound symmetry.

Compound symmetry is quite a strong assumption. It does not preclude interaction, but it severely restricts the sort of interaction which may be present. In particular, compound symmetry requires that interactions be uncorrelated with the main effect of blocks. To make this statement precise, let us decompose Y_{ijk} as

$$Y_{ijk} = W_{ij} + e_{ijk} \quad (4.1)$$

where the e_{ijk} are iid random variables with mean zero. Thus, W_{ij} is the “true” response of Block j to Treatment i ; one can visualize W_{ij} as $\lim_{K \rightarrow \infty} \bar{Y}_{ij}$. Assume that the random vectors $(W_{1j}, \dots, W_{Ij})'$, $j = 1, \dots, J$, are iid as the vector $(W_1, \dots, W_I)'$.

It is not difficult to show that, in the context of (4.1), compound symmetry of Σ implies that

$$\text{Cov}[(W_i - W_{i'}), \bar{W}] = 0 \quad \text{for all } i \neq i' \quad (4.2)$$

where

$$\bar{W} = I^{-1} \sum_i W_i.$$

In a model with no treatment-block interaction, the difference between responses to two treatments would be the same for all blocks, so that the random variables $(W_i - W_{i'})$ would

have zero variance; the condition (4.2) permits the difference between responses to vary, but only in a manner which is uncorrelated with the average response.

One source of violation of (4.2), and therefore of the compound symmetry assumption, would be a multiplicative treatment effect. Suppose, for example, that $I = 2$ and

$$W_2 = \alpha W_1 \tag{4.3}$$

for some constant $\alpha \neq 1$. Clearly, (4.3) contradicts (4.2).

Another kind of violation of (4.2) can arise, not from the *nature* of the treatment effect, but simply from its *distribution*, as illustrated by the following example.

Example 3. As in Example 1, suppose Y_1 and Y_2 are cortex weights; then W_1 and W_2 are the (hypothetical) mean weights that would result if *many* rats from a given litter were reared in Environment 1 or Environment 2. In the source population of litters, litters vary in their “potential brain size” \bar{W} and in their “sensitivity” $W_1 - W_2$ to the environmental manipulation. Suppose that potential brain size is normally distributed with mean 680 mg and standard deviation 20 mg, and that sensitivity is normally distributed with mean 30 mg and standard deviation 25 mg. Will such a population satisfy the assumption of compound symmetry? The answer is Yes if sensitivity is distributed independently of potential brain size, and No if it is not.

5. Conclusions

In many ANOVA settings, the assumption of homogeneity of variance might be regarded as a necessary evil. We have shown that, in the context of a randomized blocks design with blocks regarded as a random factor, the analogous assumption, of compound symmetry of the intrablock covariance matrix, is not innocuous and is often not necessary.

We saw in Section 4 that compound symmetry stringently limits the kinds of interaction that are permitted in the model. In Section 2 we considered the case of two treatments, where compound symmetry (which reduces to variance homogeneity) is not required for validity of the usual test for treatments. For this case we suggest using the parameter $\tilde{\rho}$ to quantify the effect of blocks on the efficiency of treatment comparisons; we have described tests and estimates of $\tilde{\rho}$ which do not require homogeneity of variances. In Section 3 we

showed how, by focusing on contrasts of interest, the ideas of Section 2 can be applied to experiments with more than two treatments.

References

- David, F.N. (1938). *Tables of the Ordinates and Probability Integral of the Distribution of the Correlation Coefficient in Small Samples*. Cambridge: Cambridge University Press.
- Freedman, David, Pisani, Robert, and Purves, Roger (1978). *Statistics*, New York: W. W. Norton.
- Kirk, Roger E. (1982). *Experimental Design: Procedures for the Behavioral Sciences* (2nd ed.), Monterey, California: Brooks/Cole.
- Kristof, Walter (1963), "The Statistical Theory of Stepped-up Reliability Coefficients When a Test Has Been Divided into Several Equivalent Parts," *Psychometrika*, **28**, 221–238.
- Kristof, Walter (1972), "On a Statistic Arising in Testing Correlation," *Psychometrika*, **37**, 377–384.
- Lehmann, E.L. (1986). *Testing Statistical Hypotheses, Second Edition*, New York: Wiley.
- Mehta, J. S. and Gurland, John (1969). "Some Properties and an Application of a Statistic Arising in Testing Correlation," *Annals of Mathematical Statistics*, **40**, 1736–1745.
- Rosenzweig, M. R., Bennett, E. L., and Diamond, M. R. (1972). "Brain Changes in Response to Experience," *Scientific American*, **226**, 22–29.
- Salib, N. M. (1985). "The Effect of Caffeine on the Respiratory Exchange Ratio of Separate Submaximal Arms and Legs Exercise of Middle Distance Runners", Master's Thesis, Purdue University, W. Lafayette, IN.
- Samuels, Myra L., Casella, George, and McCabe, George P. (1990). "Interpreting Blocks and Random Factors", submitted to *Journal of the American Statistical Association*.
- Seber, G. A. F. (1984). *Multivariate Observations*, New York: John Wiley.
- Snedecor, George W. and Cochran, William G. (1989). *Statistical Methods* (8th ed.),

Ames, Iowa: The Iowa State University Press.

Winer, B. J. (1971). *Statistical Principles in Experimental Design* (2nd ed.), New York: McGraw-Hill.

Table 1. Power of Tests of Correlation Based on (2.9) and (2.10) (one-tailed tests at $\alpha = .05$). Calculations assume equal variances.*

<i>J</i>	Test	ρ			
		.2	.4	.6	.8
10	(2.9)	.138	.322	.624	.928
	(2.10)	.139	.326	.631	.931
15	(2.9)	.178	.452	.806	.989
	(2.10)	.178	.454	.808	.989
20	(2.9)	.214	.561	.904	.998
	(2.10)	.215	.563	.904	.998
25	(2.9)	.249	.652	.953	1.000
	(2.10)	.249	.653	.954	1.000
50	(2.9)	.405	.902	.999	1.000
	(2.10)	.405	.902	.999	1.000

* Power of the test based on (2.9) was determined by interpolation in the tables of David (1938). Power of the test based on (2.10) was determined from the central F distribution using the fact that $[(1 - \rho)(1 + \tilde{r})]/[(1 + \rho)(1 - \tilde{r})] \sim F_{J-1, J-1}$ (see Kristof 1972).

Table 2. Cortex Weights of Rats Reared in Two Environments

Litter	Cortex Weight (mg)	
	Environment 1 Enriched	Environment 2 Impoverished
1	690	668
2	701	667
3	685	647
4	751	693
5	647	635
6	647	644
7	720	665
8	718	689
9	718	642
10	696	673
11	658	675
12	680	641
Mean	692.6	661.6
SD	31.8	19.5