

The dilemma of negative analysis of variance estimators of intraclass correlation

C. S. Wang¹, B. S. Yandell², and J. J. Rutledge¹

¹ Department of Meat and Animal Science, 1675 Observatory Drive, University of Wisconsin, Madison, Wisconsin 53706, USA

² Departments of Horticulture and of Statistics, 1210 West Dayton Street, University of Wisconsin, Madison, Wisconsin 53706, USA

Received December 10, 1991; Accepted February 26, 1992

Communicated by E. J. Eisen

Summary. At least two common practices exist when a negative variance component estimate is obtained, either setting it to zero or not reporting the estimate. The consequences of these practices are investigated in the context of the intraclass correlation estimation in terms of bias, variance and mean squared error (MSE). For the one-way analysis of variance random effects model and its extension to the common correlation model, we compare five estimators: analysis of variance (ANOVA), concentrated ANOVA, truncated ANOVA and two maximum likelihood-like (ML) estimators. For the balanced case, the exact bias and MSE are calculated via numerical integration of the exact sample distributions, while a Monte Carlo simulation study is conducted for the unbalanced case. The results indicate that the ANOVA estimator performs well except for designs with family size $n = 2$. The two ML estimators are generally poor, and the concentrated and truncated ANOVA estimators have some advantages over the ANOVA in terms of MSE. However, the large biases may make the concentrated and truncated ANOVA estimators objectionable when intraclass correlation (ρ) is small. Bias should be a concern when a pooled estimate is obtained from the literature since $\rho < 0.05$ in many genetic studies.

Key words: Intraclass correlation – Negative estimate – Restricted maximum likelihood

Introduction

Intraclass correlation is a measure of the degree of resemblance among groups, which are often family members, in

genetic studies. Since Fisher (1918, 1925) introduced the concept of the intraclass correlation, its estimation and the properties of its estimators have been extensively studied and documented (e.g., Fisher 1925; Snedecor and Cochran 1980). Many applications of the intraclass correlation are found in genetic studies (Kempthorne 1957; Falconer 1989). For example, if randomly chosen males are mated to several randomly chosen females with one offspring per mating then the intraclass correlation coefficient is defined as the correlation between sire-family members.

A commonly used method for correlation estimation is the analysis of variance (ANOVA) for balanced data or its analog for unbalanced data (Henderson 1953). The properties of the ANOVA estimator are well known. The estimator is biased downward (Ginsburg 1973; Ponzoni and James 1978), but the bias is usually small. Large-sample variance formulae of the ANOVA estimator have been derived for both balanced data (Fisher 1925; Osborne and Paterson 1952) and unbalanced data (Smith 1956; Swiger et al. 1964). Negative estimates of correlation can occur due to sampling errors even if the model is correct (Leone and Nelson 1966). The probability of obtaining a negative estimate depends on both the design and the size of the correlation. For a small correlation and small samples, the probability can be substantial. In animal genetic studies, intraclass correlation is often used to estimate heritability, which is non-negative by definition. Two common practices arise when a negative estimate is obtained. Some practitioners simply treat a negative estimate as zero while others may not report it to avoid embarrassment (Searle 1971). Strictly speaking, all these and others practices yield different estimators of the intraclass correlation. To the best of our knowledge, there have been no systematic studies to quantify the effects of these practices. Therefore, the basic objective of the pres-

ent study is to compare the ANOVA estimator and its two common variants, along with two maximum likelihood-like estimators in terms of bias, variance and mean squared error (MSE) via numerical integrations using the exact distributions of the estimators for the balanced case, and via Monte Carlo simulations for the unbalanced case in small samples. This approach is in accordance with Searle's (1988) call for assessing the small-sample properties of estimators of variance components in a general unbalanced mixed model setting by use of numerical and simulation means because analytical studies are either difficult or else impossible.

Models and estimators of the intraclass correlation

First consider the one-way random effects ANOVA model:

$$y_{ij} = \mu + a_i + e_{ij} \quad (i = 1, \dots, s, \text{ and } j = 1, \dots, n_i) \quad (1)$$

where $a_i \sim \text{iid}(0, \sigma_a^2)$, $e_{ij} \sim \text{iid}(0, \sigma_e^2)$, $\sigma_y^2 = \sigma_a^2 + \sigma_e^2$, and $E(y_{ij}) = \mu$. Usually, a_i and e_{ij} are assumed to follow a normal distribution.

The vector (n_1, n_2, \dots, n_s) is the n -pattern of subgroup sizes, and the data are unbalanced whenever the n_i are not all equal. In the argument which follows we shall refer to a_i as families without loss of generality.

The intraclass correlation is defined as (Fisher 1925; Snedecor and Cochran 1980):

$$\rho(y_{ij}, y_{ik}) = \text{Cov}(y_{ij}, y_{ik}) / \sigma_{y_{ij}} \sigma_{y_{ik}} = \sigma_a^2 / (\sigma_a^2 + \sigma_e^2), \quad (2)$$

where j and k are different members within a family. The limit of ρ are from 0 to 1, for $\sigma_a^2 \geq 0$.

A broader definition of the intraclass correlation coefficient is as follows. Data y_{ij} are distributed about a common mean with a common variance, σ_y^2 , and y_{ij} in the same family have a common correlation coefficient, ρ , called the intraclass correlation coefficient (Snedecor and Cochran 1980). This alternative from of the one-way random model is also called a common correlation model (CCM) (Donner and Koval 1980). This model is more general in that it permits negative values of ρ . Estimates lie in the range $-1/(n-1)$ to 1.

Both the ANOVA and CCM models can be formulated in a multivariate setting, in which the relationship between the two models can be more clearly seen. For both models, let $\mathbf{Y}_i = (y_{i1}, y_{i2}, \dots, y_{in_i})'$ denote a vector of measurements taken on the i^{th} family, $i = 1, 2, \dots, s$. Further, we assume that $\mathbf{Y}_i \sim \text{iid}(\boldsymbol{\mu}_i, \Sigma_i)$, where $\boldsymbol{\mu}_i = (\mu, \mu, \dots, \mu)'$ is a vector of length n_i . The covariance matrix for the ANOVA model is Σ_{A_i} , of order $n_i \times n_i$, where $\Sigma_{A_i}^{jj} = \sigma_y^2$ and $\Sigma_{A_i}^{jk} = \sigma_a^2$ for $j \neq k$. The intraclass correlation is $\rho = \sigma_a^2 / \sigma_y^2$. The covariance matrix of the CCM is Σ_{C_i} , where $\Sigma_{C_i}^{jj} = \sigma_y^2$ and $\Sigma_{C_i}^{jk} = \rho \sigma_y^2$ for $j \neq k$.

The two models differ in terms of the covariance structure and, consequently, in terms of the lower limit of ρ . The ANOVA model specifies the lower limit of ρ to be zero, while the CCM permits negative values of ρ . In fact the ANOVA is a special case of the CCM with $\rho = \sigma_a^2 / \sigma_y^2$. This difference has implications when the ML estimator is considered.

Historically, the ANOVA estimator of the correlation from both ANOVA and CCM models is closely tied to the analysis of variance. For both models, the standard analysis of variance is given in Table 1, in which $k = \frac{N - \sum n_i^2 / N}{s - 1}$ and $N = \sum_{i=1}^s n_i$, the total number of observations. When $n_i = n$ for all i , $k = n$, and the balanced case is obtained (Snedecor and Cochran 1980; Fisher 1925).

The ANOVA estimator of ρ for both models is

$$\hat{\rho}_A = \hat{\sigma}_a^2 / (\hat{\sigma}_a^2 + \hat{\sigma}_e^2) = (\text{MSA} - \text{MSE}) / [\text{MSA} + (k - 1)\text{MSE}], \quad (3)$$

where $\hat{\sigma}_e^2 = \text{MSE}$ and $\hat{\sigma}_a^2 = (\text{MSA} - \text{MSE})/k$. The limits of $\hat{\rho}_A$ are from $-1/(n-1)$ to 1. The negative lower limit guarantees the existence of some negative estimates of the estimator. The probability of obtaining a negative estimate was derived by Leone and Nelson (1966) and Gill and Jensen (1968). The common practice of setting a negative $\hat{\rho}_A$ to zero is equivalent to equating the negative side of the sampling distribution of $\hat{\rho}_A$ to zero. Thus, we define an estimator, a 'concentrated estimator', due to setting negative estimates to zero as

$$\hat{\rho}_{AC} = \max\{0, \hat{\rho}_A\}. \quad (4)$$

The practice of not reporting the negative estimates is essentially a truncation of the left tail at zero. So, an estimator, say a 'truncated estimator' due to not reporting negative estimates, is defined as

$$\hat{\rho}_{AT} = \hat{\rho}_A \quad \text{if } \hat{\rho}_A \geq 0, \text{ undefined otherwise.} \quad (5)$$

Similarly, we define the ML-like estimator from the CCM as

$$\hat{\rho}_{CM} = \frac{\text{MSA}(s-1)/s - \text{MSE}}{\text{MSA}(s-1)/s + (k-1)\text{MSE}}, \quad (6)$$

and the ML-like estimator from the ANOVA model as

$$\hat{\rho}_{AM} = \max\{0, \hat{\rho}_{CM}\}. \quad (7)$$

Table 1. Analysis of variance

| Source | df | MS | EMS | |
|------------------|---------|-----|-------------------------------|----------------------------|
| | | | Under CCM | Under ANOVA |
| Between families | $s - 1$ | MSA | $\sigma_y^2(1 + (k - 1)\rho)$ | $\sigma_e^2 + k\sigma_a^2$ |
| Within families | $N - s$ | MSE | $\sigma_y^2(1 - \rho)$ | σ_e^2 |

It is well known that in the balanced case the ANOVA estimator ($\hat{\varrho}_A$) is the restricted maximum likelihood (REML) estimator derived from the CCM, while $\hat{\varrho}_{AC}$ is the REML estimator from the ANOVA model, due to the fact that the variance component estimators are REML estimators. $\hat{\varrho}_{CM}$ is the ML estimator from the CCM in balanced data (Rosner et al. 1977). Similarly, $\hat{\varrho}_{AM}$ is the ML estimator from the ANOVA model in the balanced case, since the corresponding variance component estimators are the ML estimators (Herbach 1959), and $\hat{\sigma}_a^2 \geq 0$ by definition. For unbalanced data $\hat{\varrho}_A$ and $\hat{\varrho}_{AC}$ are not REML, and $\hat{\varrho}_{CM}$ and $\hat{\varrho}_{AM}$ are not ML. We call them REML-like and ML-like estimators, respectively. These estimators differ basically in two ways. Differences of estimators from different models ($\hat{\varrho}_A$ vs $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{CM}$ vs $\hat{\varrho}_{AM}$) are caused by concentrations, whereas differences between ML and REML arise because REML takes into account the loss in degrees of freedom associated with the estimation of the fixed effects.

Distributions of estimators

The distributions of the estimators are both parameter-dependent and design-dependent. In this study, due to analytical intractability, no efforts were devoted to deriving the first and second moment statistics of the estimators. Instead, numerical calculations were used to examine some of the properties of the estimators. As pointed out earlier, $\hat{\varrho}_{AC}$ concentrates the left tail of the distribution of $\hat{\varrho}_A$ at zero, as $\hat{\varrho}_{CA}$ to $\hat{\varrho}_{CM}$, while $\hat{\varrho}_{AT}$ cuts off the tail of $\hat{\varrho}_A$ below zero. Relationships between sampling distributions of the estimators are shown numerically in Fig. 1 for a design of 20 families each having five family members and $\varrho = 0.10$ as an example. These differences in distributions should result in differences in bias, variance and the MSE of the estimators.

The exact density function for $\hat{\varrho}_A$ in the balanced case was provided by Donner and Koval (1983) as follows:

$$f_A(r) = C(1-r)^{[s(n-1)-2]/2} [1+(n-1)r]^{(s-3)/2} Q^{-(sn-1)/2},$$

$$\frac{-1}{n-1} < r < 1 \quad (8)$$

where

$$C = n \left[\frac{s-1}{1+(n-1)\varrho} \right]^{(s-1)/2} \left[\frac{s(n-1)}{1-\varrho} \right]^{[s(n-1)]/2}$$

$$\Gamma\left(\frac{sn-1}{2}\right) \left[\Gamma\left(\frac{s-1}{2}\right) \Gamma\left(\frac{s(n-1)}{2}\right) \right]^{-1},$$

$Q = (s-1)[1+(n-1)r][1+(n-1)\varrho]^{-1} + s(n-1)(1-r)(1-\varrho)^{-1}$ and $\Gamma(\alpha)$ is gamma function of α . $\hat{\varrho}_{AC}$ results from the concentration of $\hat{\varrho}_A$. Consequently, the density

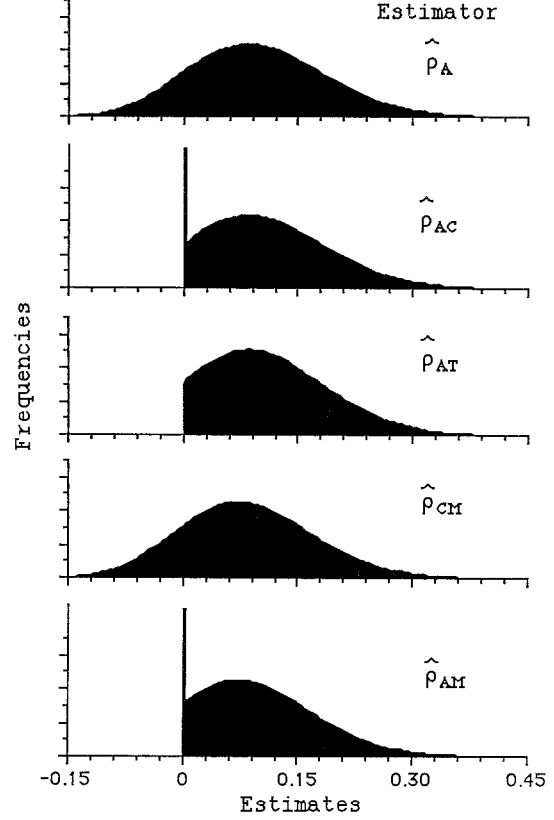


Fig. 1. Distributions of estimators generated from the exact sample distributions of the estimators. The example is taken as a design of 20 families each having five family members and $\varrho = 0.10$

function of $\hat{\varrho}_{AC}$ becomes

$$g_{AC}(r) = f(r), \quad 0 < r < 1 \quad (9)$$

$$= \int_{-1/(n-1)}^0 f(r) dr, \quad r = 0.$$

The density function of $\hat{\varrho}_{AT}$, due to the truncation of $\hat{\varrho}_A$, appears as

$$h_{AT}(r) = f(r) \Big/ \int_0^1 f(r) dr, \quad 0 < r < 1. \quad (10)$$

This is a conditional density, and the condition is the existence of $\hat{\varrho}_{AT}$, because $\hat{\varrho}_{AT}$ is not defined while $\hat{\varrho}_A$ is negative. By a similar approach we derive the probability density functions for the ML from the CCM, and for the ML from the ANOVA model ($\hat{\varrho}_{AM}$), in the balanced case as:

$$\varphi_{CM}(r) = [(s-1)/s]^{s(n-1)/2} (Q/R)^{(sn-1)/2} f(r),$$

$$-1/(n-1) < r < 1, \quad (11)$$

and

$$\eta_{AM}(r) = \varphi(r), \quad 0 < r < 1 \quad (12)$$

$$= \int_{-1/(n-1)}^0 \varphi(r) dr, \quad r = 0$$

respectively, where $R = (s-1)[1 + (n-1)r][1 + (n-1)\rho]^{-1}$ + $(s-1)(n-1)(1-r)(1-\rho)^{-1}$, and Q is the same as in (8).

Numerical integration and Monte Carlo simulation

Numerical integrations were performed for balanced data using the exact distributions of the estimators. We considered many design combinations for $N = 50, 100$ and 200 . For example, design combinations ($s \times n$) for $N = 100$ were $2 \times 50, 4 \times 25, 5 \times 20, 10 \times 10, 20 \times 5, 25 \times 4$ and 50×2 . Limits of $\hat{\varrho}_A$ and $\hat{\varrho}_{CM}$ are from $-1/(n-1)$ to 1 , and from 0 to 1 for $\hat{\varrho}_{AC}, \hat{\varrho}_{AT}$ and $\hat{\varrho}_{AM}$. For comparison purposes, intraclass correlations ranged from 0 to 1 in increments of 0.01 . For each design-parameter combination and each estimator, we calculated exact bias, variance and MSE. Integration subroutines were adapted from IMSL (IMSL, Inc., 1989). All calculations were in double-precision Fortran.

For unbalanced data, we conducted a simulation study using four levels of ρ ($0.025, 0.1, 0.2,$ and 0.5) and two sample sizes ($N = 50$ and 100). The three family sizes (n) within N chosen were $2, 5$ and 10 respectively. Two n -patterns were simulated for each of the above combinations: the unbalanced but fixed (Table 2) and the Poisson random distribution. In total, 12 design combinations or 48 design-parameter combinations were considered. The estimated biases, variances and mean squared errors of five estimators were calculated based on 1000 independent simulation runs conducted for each of 48 combinations. The one-way random effects ANOVA model (1) was used to generate the data. Normality of family and error effects was assumed throughout the study. The simulation programs were written in S language (Becker et al. 1988).

As mentioned before $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AM}$ are REML and ML estimators, respectively, from the ANOVA model in the balanced case. It is of interest to compare the performances of $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AM}$ with their counterparts, the true REML and ML estimators, in unbalanced data. For this purpose, a parallel simulation study was carried out with a similar design as the above. Simulation runs conducted

Table 2. The n -patterns used in the simulation studies

| Codes | n -patterns |
|----------------|--|
| P ₁ | 1, 1, 1, 1, 3, 3, 3, 3, 5, 5, 5, 5, 7, 7, 7, 7, 9, 9, 9, 9 |
| P ₂ | 6, 6, 6, 10, 10, 10, 10, 14, 14, 14 |
| P ₃ | 1, 1 2, 2 3, 3 |
| P ₄ | 3, 3, 3, 5, 5, 5, 5, 7, 7, 7 |
| P ₅ | 6, 6, 10, 14, 14 |
| P ₆ | 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2 3, 3, 3, 3, 3, 3, 3, 3 |

for each combination were 300 , instead of 1000 , because of computing constraints. The simulation was done using SAS and its variance component procedures (SAS Institute Inc., 1985). We do not repeat here the definitions of the REML and ML procedures of SAS, because they are well understood.

Results and discussion

In the first three subsections which follow, the exact results of five estimators computed from their exact sampling densities are presented in terms of bias, variance and mean squared error, respectively, in the balanced case. The detailed results are only given for $N = 50$ designs, since patterns of differences between estimators for $N = 100$ and 200 designs resembled that of $N = 50$ designs, though as N increased, differences between estimators became smaller. There are in total ten pair-wise comparisons for five estimators. We emphasize four presumably meaningful comparisons ($\hat{\varrho}_{CM}$ vs $\hat{\varrho}_{AM}$, $\hat{\varrho}_A$ vs $\hat{\varrho}_{AC}$, $\hat{\varrho}_A$ vs $\hat{\varrho}_{AT}$ and $\hat{\varrho}_{CM}$ vs $\hat{\varrho}_A$) here.

In the fourth subsection, the results generated from the simulation with unbalanced data are presented, and comparisons are made with the corresponding exact results of balanced data. The fifth and last subsection emphasizes comparing the performances of $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AM}$ with their counterparts, REML and ML estimators in both balanced and unbalanced cases.

Biases of estimators

Figure 2 gives the exact biases of five estimators for $N = 50$ designs based on numerical integration. Biases are both negative and positive and the horizontal line in the figure indicates the zero bias. The criterion used for comparison here is absolute bias. As mentioned earlier, $\hat{\varrho}_A$ was biased downward whereas $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AT}$, which concentrate and truncate the distribution of $\hat{\varrho}_A$, introduce upward biases. Some positive bias cancelled out with the negative bias. The remaining bias could be either positive or negative, and could be either greater or smaller than that of $\hat{\varrho}_A$, depending on the severity of truncation and concentration. For small n and small ρ , biases of $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AT}$ were huge. Another interesting point regarding $\hat{\varrho}_A$ and $\hat{\varrho}_{CM}$ was as follows: $\hat{\varrho}_A$ had negative bias, and $\hat{\varrho}_{CM}$ failed to take account of the loss in degrees of freedom associated with the estimation of fixed effects, introducing further downward bias. So the absolute bias of $\hat{\varrho}_{CM}$ was always larger than that of $\hat{\varrho}_A$. The difference in bias of these two estimators was given by Fisher (1925) and Wang et al. (1991). Furthermore, $\hat{\varrho}_{AM}$ concentrated the distribution of $\hat{\varrho}_{CM}$ producing upward bias. The tradeoff of downward versus upward bias resulted in a complicated pattern of bias comparison be-

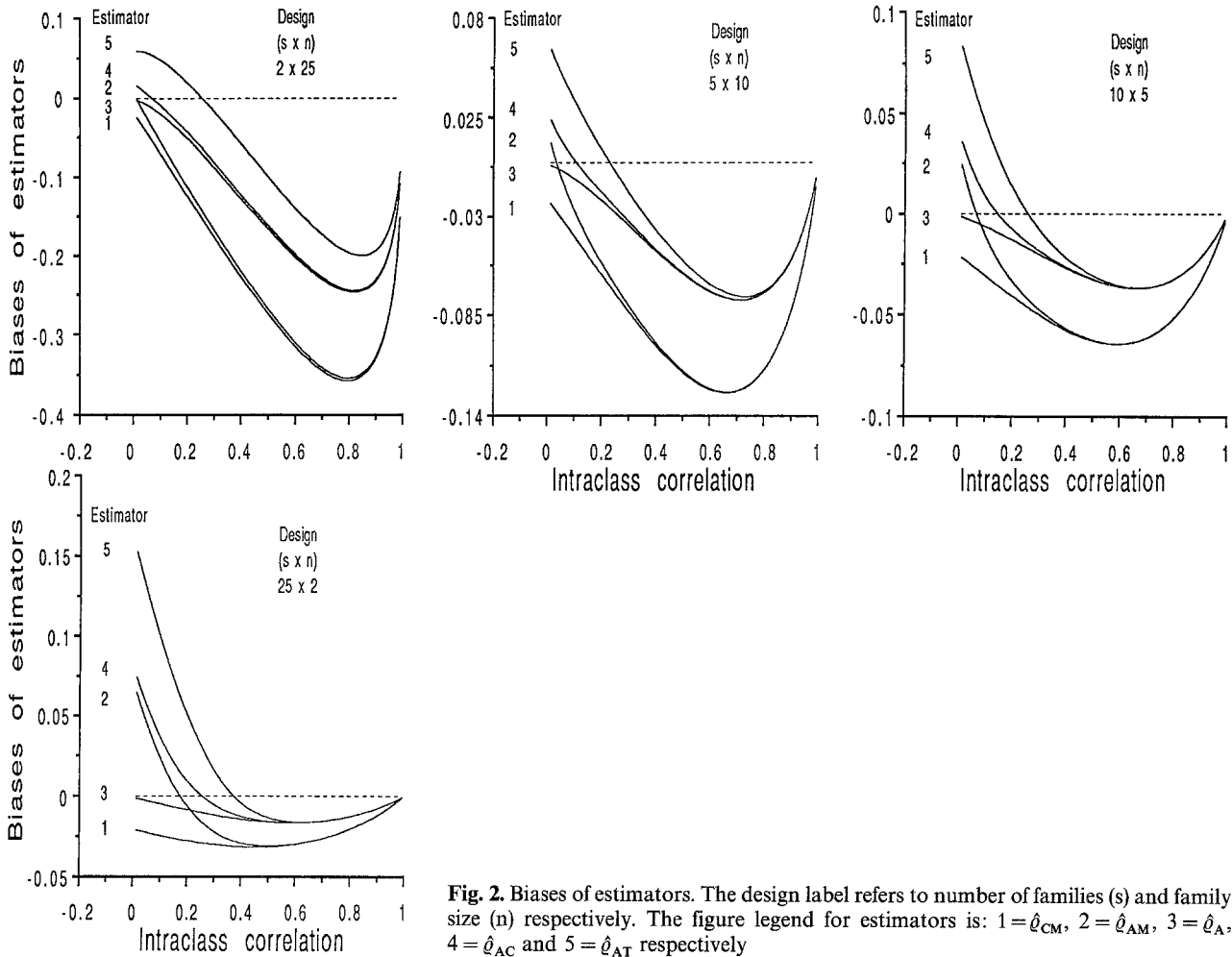


Fig. 2. Biases of estimators. The design label refers to number of families (s) and family size (n) respectively. The figure legend for estimators is: 1 = $\hat{\varrho}_{CM}$, 2 = $\hat{\varrho}_{AM}$, 3 = $\hat{\varrho}_A$, 4 = $\hat{\varrho}_{AC}$ and 5 = $\hat{\varrho}_{AT}$ respectively

tween the two ML estimators (with a similar phenomenon for $\hat{\varrho}_A$ vs $\hat{\varrho}_{AC}$ and $\hat{\varrho}_A$ vs $\hat{\varrho}_{AT}$). These non-uniform patterns of comparison between estimators imply that turning points or ‘watersheds’ must exist so that the order of comparisons change with respect to ϱ . The precise level of such watersheds would then depend on the design. Under fixed N , watersheds increased as s increased. Watersheds decreased as sample sizes (N) increased with either fixed n or fixed s . This confusing picture became clearer as ϱ increased. The basic difference in terms of bias for large ϱ was between two classes of estimators, the two ML estimators vs the three ANOVA estimators (Fig. 2). For small ϱ , particularly for $\varrho < 0.10$, a common situation in genetic studies, relative bias, defined as bias over parametric value, of $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AT}$ could be as large as 15–1000%. The bias of $\hat{\varrho}_{CM}$ was large for small s . In general, from a bias perspective, $\hat{\varrho}_A$ was preferred with a relative bias usually $< 10\%$.

Variances of estimators

Figure 3 shows variances of five estimators given $N = 50$. Variances of $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AT}$ are uniformly not greater than

that of $\hat{\varrho}_A$, and similarly for $\hat{\varrho}_{AM}$ compared with $\hat{\varrho}_{CM}$, since truncation and concentration reduce variance. The above variance differences quickly diminish as ϱ increases. Furthermore the variance of $\hat{\varrho}_{AC}$ is definitely not smaller than that of $\hat{\varrho}_{AT}$, since truncation of a distribution reduces more variance than concentration.

The comparison of the variances of $\hat{\varrho}_{CM}$ and $\hat{\varrho}_A$, namely ML and EML estimators from the CCM, yielded a conclusion of non-uniform patterns. In other words, watersheds existed in this variance comparison. For $N = 2$ designs, watersheds were around 0.01–0.02, and watersheds increased to 0.88 as n increased. This implied that for small n the variance of the $\hat{\varrho}_{CM}$ was greater than, or equal to, that of $\hat{\varrho}_A$, and vice versa for large n designs. But the difference was small except for $s = 2$ design settings, an unlikely situation in practice. A similar situation existed for the variance comparison between $\hat{\varrho}_{AM}$ and $\hat{\varrho}_{AC}$.

Mean squared errors of estimators

As discussed earlier, all five estimators are biased. In such situations, a global criterion, such as MSE, is usually used

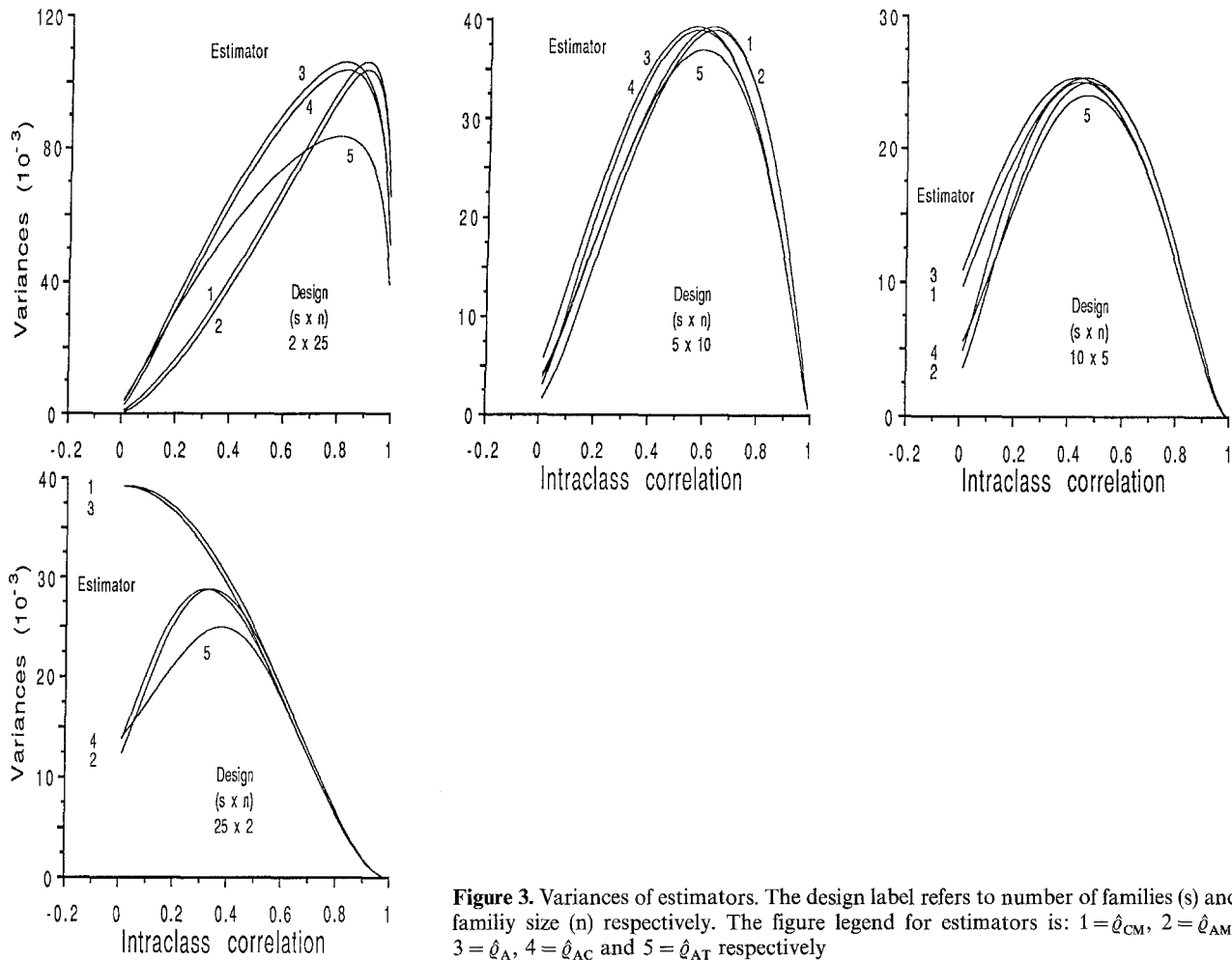


Figure 3. Variances of estimators. The design label refers to number of families (s) and family size (n) respectively. The figure legend for estimators is: 1 = $\hat{\varrho}_{CM}$, 2 = $\hat{\varrho}_A$, 3 = $\hat{\varrho}_A$, 4 = $\hat{\varrho}_{AC}$ and 5 = $\hat{\varrho}_{AT}$ respectively

to compare performances of estimators. The numerical results are presented in Fig. 4. It was clear that the MSE of $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AM}$ were always smaller than that of $\hat{\varrho}_A$ and $\hat{\varrho}_{CM}$, respectively. The reason was that $\hat{\varrho}_{AC}$ (or $\hat{\varrho}_{AM}$) concentrated the sampling distribution, reducing the variance, and either increasing or decreasing the bias. In the case of increasing bias (decreasing bias leads to decreasing MSE), variance reduction was more than enough to offset the squared bias increase, hence the MSE decreased. The differences for both comparisons were large for small n. The situation for the MSE comparison between $\hat{\varrho}_A$ and $\hat{\varrho}_{AT}$ was similar. However, the increased squared bias in the lower range of ϱ was larger than the variance reduction. Consequently, watersheds were created, making the comparison between $\hat{\varrho}_A$ and $\hat{\varrho}_{AT}$ non-uniform.

Watersheds were also observed for MSE comparisons of $\hat{\varrho}_{CM}$ vs $\hat{\varrho}_A$ and $\hat{\varrho}_{AM}$ vs $\hat{\varrho}_{AC}$, and ranged from 0 to 0.31. For example, for design 5 x 10, the MSE of $\hat{\varrho}_{CM}$ was greater than that of $\hat{\varrho}_A$ if $\varrho \geq 0.24$, and vice versa for $\varrho < 0.24$. A similar phenomenon was observed by Rosner et al. (1977) and O'Neill et al. (1987) in comparing perfor-

mances of interclass correlation estimators in terms of MSE.

Finally we may approximately rank the estimators in terms of MSE in descending order as $\hat{\varrho}_{CM}$, $\hat{\varrho}_{AM}$, $\hat{\varrho}_A$, $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AT}$, except for n = 2 designs (Fig. 4). In the small n and small ϱ case, two concentrated estimators ($\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AM}$) had some advantages over their counterparts ($\hat{\varrho}_A$ and $\hat{\varrho}_{CM}$).

A comparison of performances of five estimators in the unbalanced case with that in the balanced case

The patterns of comparisons among five estimators in the unbalanced data were similar to that in the balanced data, as shown in Figs. 2–4. For this reason, we do not give extensive tables for all design combinations. Instead, only the results from one design combination (s = 20 and p₁ n-pattern) are presented in Table 3, along with the exact results from numerical integration, in terms of relative bias (RB), variance and mean squared error. RB is the bias of an estimator divided by its true value.

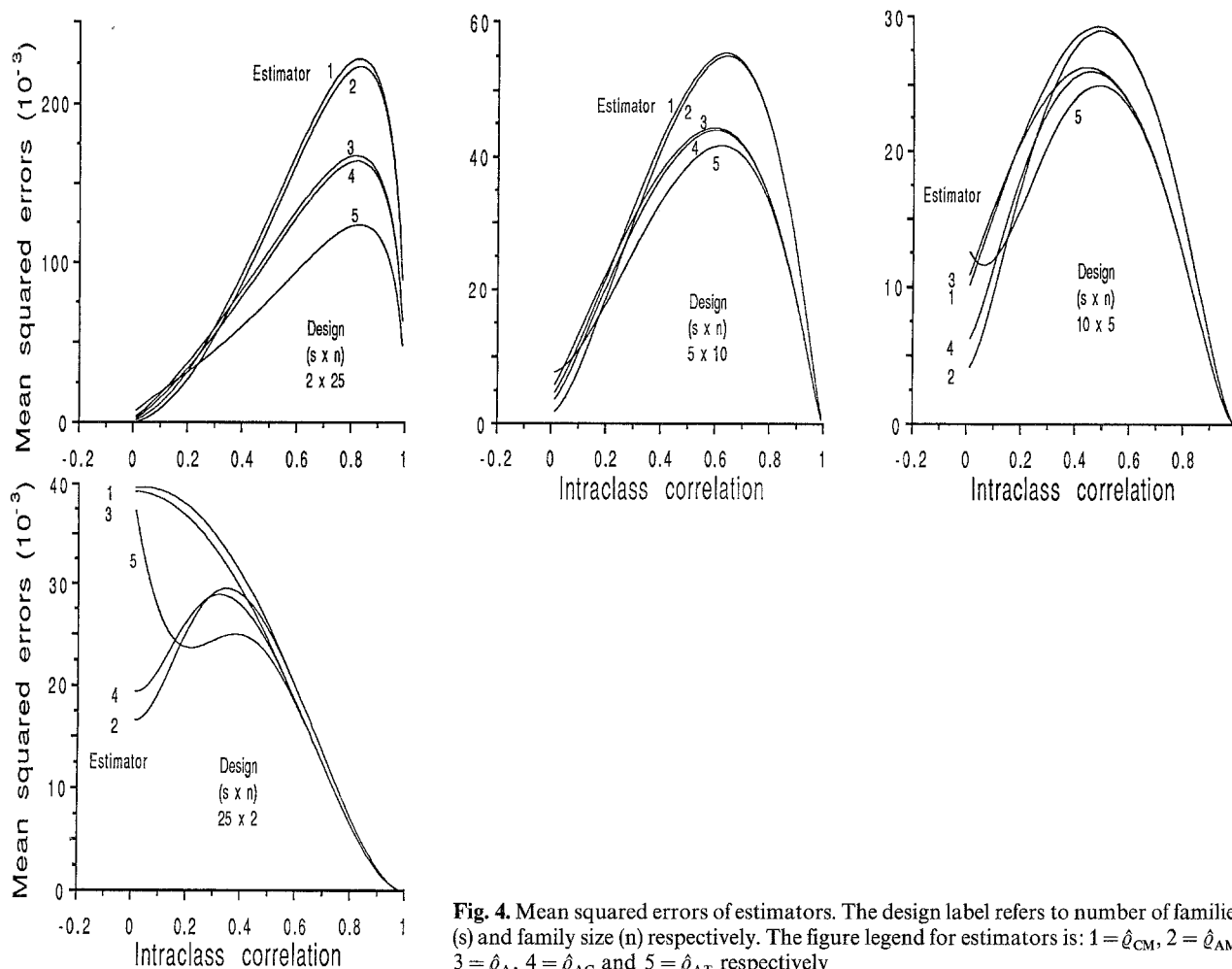


Fig. 4. Mean squared errors of estimators. The design label refers to number of families (s) and family size (n) respectively. The figure legend for estimators is: 1 = $\hat{\varrho}_{CM}$, 2 = $\hat{\varrho}_{AM}$, 3 = $\hat{\varrho}_A$, 4 = $\hat{\varrho}_{AC}$ and 5 = $\hat{\varrho}_{AT}$ respectively

Notice the closeness between the results of balanced data and that of unbalanced data for each of the five estimators. In general, the former had less bias, less variance and less MSE than the latter, as expected. The differences, in terms of variance and MSE, between balanced and unbalanced layouts were usually no greater than $\pm 10\%$ of relative error though occasionally was up to $\pm 15\%$. This finding was true for other layouts not presented in the paper, including the results from the random Poisson n -patterns.

A comparison of $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AM}$ with true REML and ML estimators in the unbalanced case

Results from the simulation for the unbalanced case are summarized in Tables 4 and 5, along with the appropriate exact results computed from the exact sampling densities for the balanced case. The exact results of the balanced case served as a reference basis for comparisons. Table 4 gives relative biases of estimators. The values of RB in Table 4 exceeding $\pm 10\%$ are italicized.

The dominant feature of Table 4 is that, in general, the RB of REML (ML) estimators and $\hat{\varrho}_{AC}$ ($\hat{\varrho}_{AM}$) in the unbalanced case were of the same magnitude, and their RB were similar to that of REML (ML) estimators in the balanced case. Some differences existed in case of $\varrho = 0.025$; however, these differences did not exhibit any particular pattern. In other words, the true REML (ML) estimator was no better than $\hat{\varrho}_{AC}$ ($\hat{\varrho}_{AM}$), in terms of bias, for the layouts considered in this study.

Mean squared errors (Table 5) between $\hat{\varrho}_{AC}$ ($\hat{\varrho}_{AM}$) and REML (ML) estimators in the unbalanced case were very close, generally less than $\pm 10\%$ of relative error. The largest relative error between $\hat{\varrho}_{AC}$ ($\hat{\varrho}_{AM}$) and REML (ML) estimators was 31.97% (31.39%), when $\varrho = 0.025$ with a design of $s = 5$ and a p_5 n -pattern. It is also clear that the MSE of the above mentioned estimators in the unbalanced case very similar to that of their counterparts in the balanced case.

The results obtained above were essentially the same as those of Swallow and Monahan (1984) in comparing the performances of ML, REML and ANOVA estimators

Table 3. A comparison of five estimators between balanced and unbalanced designs

| | $n_i(n)$ ($s = 20$) | $\hat{\varrho}^a$ | Parametric values of ϱ | | | | |
|----------------------|---------------------------|-------------------|--------------------------------|--------|--------|--------|--------|
| | | | 0.025 | 0.100 | 0.200 | 0.500 | |
| RB (%) | P_1^b | 1 | -41.55 | -16.68 | -11.18 | -7.28 | |
| | | 2 | 49.44 | -9.33 | -10.54 | -7.28 | |
| | | 3 | 1.62 | -4.18 | -4.07 | -4.25 | |
| | | 4 | 78.25 | 1.67 | -3.60 | -4.25 | |
| | | 5 | 188.25 | 19.75 | -0.59 | -4.25 | |
| | 5(Equal) ^c | 1 | -45.44 | -15.20 | -10.12 | -6.14 | |
| | | 2 | 46.52 | -8.60 | -9.62 | -6.14 | |
| | | 3 | -3.06 | -2.80 | -3.04 | -3.12 | |
| | | 4 | 74.23 | 2.42 | -2.67 | -3.12 | |
| | | 5 | 190.34 | 18.89 | -0.26 | -3.12 | |
| | Variance (10^{-3}) | P_1 | 1 | 5.571 | 7.858 | 10.799 | 13.892 |
| | | | 2 | 2.598 | 6.107 | 10.271 | 13.892 |
| | | | 3 | 5.890 | 8.215 | 11.120 | 13.709 |
| | | | 4 | 3.084 | 6.698 | 10.704 | 13.709 |
| | | | 5 | 3.006 | 5.724 | 9.843 | 13.709 |
| 5(Equal) | | 1 | 5.511 | 7.605 | 10.033 | 12.051 | |
| | | 2 | 2.608 | 6.023 | 9.618 | 12.050 | |
| | | 3 | 5.829 | 7.953 | 10.339 | 11.890 | |
| | | 4 | 3.095 | 6.587 | 10.010 | 11.889 | |
| | | 5 | 3.051 | 5.688 | 9.295 | 11.882 | |
| MSE (10^{-3}) | | P_1 | 1 | 5.679 | 8.136 | 11.299 | 15.217 |
| | | | 2 | 2.751 | 6.194 | 10.715 | 15.217 |
| | | | 3 | 5.890 | 8.232 | 11.186 | 14.161 |
| | | | 4 | 3.467 | 6.701 | 10.756 | 14.161 |
| | | | 5 | 5.221 | 6.114 | 9.844 | 14.161 |
| | 5(Equal) | 1 | 5.640 | 7.836 | 10.443 | 12.994 | |
| | | 2 | 2.743 | 6.097 | 9.988 | 12.993 | |
| | | 3 | 5.829 | 7.961 | 10.376 | 12.133 | |
| | | 4 | 3.440 | 6.593 | 10.039 | 12.132 | |
| | | 5 | 5.316 | 6.045 | 9.295 | 12.125 | |

^a $1 = \hat{\varrho}_{CM}$, $2 = \hat{\varrho}_{AM}$, $3 = \hat{\varrho}_A$, $4 = \hat{\varrho}_{AC}$ and $5 = \hat{\varrho}_{AT}$
^b Based on 1000 simulation runs
^c Exact results from numerical integration

of variance components for the one-way random model with unbalanced data.

Conclusions

(1) We compared the performance of five estimators of intraclass correlation in small and moderate samples. All estimators were biased. The ANOVA estimator had downward bias. The ML estimator from the CCM had further downward bias than the ANOVA estimator, particularly for small s . The resulting biases of the concentrated and truncated estimators from the ANOVA estimator were large for small ϱ . Hence, the ANOVA estimator was preferred from the bias point of view, because the bias of the ANOVA estimator was usually not greater than 10% of RB, even for very small samples ($N = 50$).

Table 4. Relative biases of estimators of ϱ under both balanced and unbalanced layouts (%)

| Design | Estimator | Parametric values of ϱ | | | | | |
|-----------|----------------------------|---|------------------------------|---------------|--------------|--------|--------|
| | | 0.025 | 0.100 | 0.200 | 0.500 | | |
| s | $n_i(n)$ | | | | | | |
| 20 | 5(equal) | REML($\hat{\varrho}_{AC}$) ^a | 74.23 | 2.42 | -2.67 | -3.11 | |
| | | P_1 REML ^b | 49.33 | -4.25 | -8.07 | -6.19 | |
| | | P_1 $\hat{\varrho}_{AC}$ ^c | 78.25 | 1.67 | -3.60 | -4.25 | |
| | 5(equal) | ML($\hat{\varrho}_{AM}$) ^a | 46.52 | -8.60 | -9.62 | -6.14 | |
| | | P_1 ML ^b | 20.88 | -16.25 | -15.73 | -9.51 | |
| | | P_1 $\hat{\varrho}_{AM}$ ^c | 49.44 | -9.33 | -10.54 | -7.28 | |
| | 10 | 10(equal) | REML($\hat{\varrho}_{AC}$) | 43.60 | -1.11 | -4.59 | -5.85 |
| | | | REML | 46.32 | 0.91 | -3.26 | -3.28 |
| | | P_2 | $\hat{\varrho}_{AC}$ | 50.08 | -3.75 | -7.42 | -7.22 |
| | | | ML($\hat{\varrho}_{AM}$) | 9.51 | -16.65 | -15.27 | -11.35 |
| | | P_2 | ML | 9.20 | -15.21 | -14.27 | -8.84 |
| | | | $\hat{\varrho}_{AM}$ | 15.14 | -18.85 | -17.98 | -12.72 |
| | 10 | 5(equal) | REML($\hat{\varrho}_{AC}$) | 120.62 | 8.76 | -3.83 | -6.44 |
| | | | P_4 REML | 115.98 | 15.80 | -2.93 | -6.30 |
| | | | P_4 $\hat{\varrho}_{AC}$ | 129.66 | 12.78 | -0.39 | -6.53 |
| 5(equal) | | ML($\hat{\varrho}_{AM}$) | 68.23 | -11.02 | -16.75 | -12.49 | |
| | | P_4 ML | 61.25 | -5.70 | -16.37 | -12.64 | |
| | | P_4 $\hat{\varrho}_{AM}$ | 76.00 | -7.49 | -13.43 | -12.56 | |
| 5 | | 10(equal) | REML($\hat{\varrho}_{AC}$) | 76.10 | 1.16 | -8.27 | -12.27 |
| | | | REML | 111.63 | 8.17 | -4.71 | -7.98 |
| | | P_5 $\hat{\varrho}_{AC}$ | 80.13 | 0.29 | -6.60 | -12.06 | |
| 10(equal) | ML($\hat{\varrho}_{AM}$) | 11.76 | -26.88 | -27.98 | -23.03 | | |
| | P_5 ML | 34.54 | -22.98 | -26.41 | -19.03 | | |
| | P_5 $\hat{\varrho}_{AM}$ | 13.67 | -28.07 | -26.47 | -22.81 | | |

^a REML and ML estimators in the balanced case, exact results from numerical integration
^b REML and ML estimators in the unbalanced case, based on 300 simulation runs
^c $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AM}$ as defined in the text in the unbalanced case, based on 1000 simulation runs

(2) Three concentrated and truncated estimators ($\hat{\varrho}_{AM}$, $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AT}$) had smaller variances than their counterparts ($\hat{\varrho}_{CM}$ and $\hat{\varrho}_A$), because the concentration and the truncation reduced the variance. The differences were largest when n was small, where the truncation and the concentration were the severest.

(3) Three ANOVA estimators ($\hat{\varrho}_A$, $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AT}$) generally had advantages over the ML estimators ($\hat{\varrho}_{CM}$ and $\hat{\varrho}_{AM}$) in terms of MSE. However, for the case of small ϱ combined with small n , the situation was similar to the variance comparison, i.e., the concentrated estimators had less MSE than their counterparts.

(4) The estimators from the unbalanced designs performed comparably to those of the balanced ones, given the settings in this study. The difference between $\hat{\varrho}_{AC}$ ($\hat{\varrho}_{AM}$) and REML (ML) was small in the unbalanced case.

(5) The choice of estimator of the intraclass correlation is largely based on personal preference and judgement. All

Table 5. Mean squared errors (MSE) of estimators of ϱ under both balanced and unbalanced layouts (10^{-3})

| Design | Estimator | Parametric values of ϱ | | | | |
|--------|-----------|---|-------|--------|--------|--------|
| | | 0.025 | 0.100 | 0.200 | 0.500 | |
| s | $n_i(n)$ | | | | | |
| 20 | 5(equal) | REML($\hat{\varrho}_{AC}$) ^a | 3.440 | 6.593 | 10.039 | 12.132 |
| | p_1 | REML ^b | 2.563 | 5.953 | 9.428 | 14.170 |
| | p_1 | $\hat{\varrho}_{AC}$ ^c | 3.467 | 6.701 | 10.756 | 14.161 |
| | 5(equal) | ML($\hat{\varrho}_{AM}$) ^a | 2.743 | 6.097 | 9.988 | 12.993 |
| | p_1 | ML ^b | 1.919 | 5.561 | 9.695 | 15.709 |
| | p_1 | $\hat{\varrho}_{AM}$ ^c | 2.751 | 6.194 | 10.715 | 15.217 |
| 10 | 10(equal) | REML($\hat{\varrho}_{AC}$) | 2.215 | 5.958 | 10.997 | 19.235 |
| | p_2 | REML | 2.485 | 5.546 | 11.802 | 18.930 |
| | p_2 | $\hat{\varrho}_{AC}$ | 2.317 | 6.219 | 11.095 | 20.157 |
| | 10(equal) | ML($\hat{\varrho}_{AM}$) | 1.565 | 5.313 | 10.774 | 21.528 |
| | p_2 | ML | 1.732 | 5.754 | 11.354 | 20.657 |
| | p_2 | $\hat{\varrho}_{AM}$ | 1.614 | 5.549 | 11.080 | 22.795 |
| 10 | 5(equal) | REML($\hat{\varrho}_{AC}$) | 6.868 | 11.565 | 18.240 | 25.776 |
| | p_4 | REML | 7.525 | 12.782 | 18.799 | 24.358 |
| | p_4 | $\hat{\varrho}_{AC}$ | 7.524 | 12.191 | 19.069 | 27.346 |
| | 5(equal) | ML($\hat{\varrho}_{AM}$) | 4.785 | 9.672 | 17.493 | 28.947 |
| | p_4 | ML | 5.261 | 10.485 | 17.765 | 27.846 |
| | p_4 | $\hat{\varrho}_{AM}$ | 5.302 | 10.130 | 17.903 | 30.528 |
| 5 | 10(equal) | REML($\hat{\varrho}_{AC}$) | 4.603 | 10.891 | 20.547 | 42.094 |
| | p_5 | REML | 7.069 | 12.128 | 20.052 | 43.017 |
| | p_5 | $\hat{\varrho}_{AC}$ | 4.850 | 11.164 | 21.859 | 43.252 |
| | 10(equal) | ML($\hat{\varrho}_{AM}$) | 2.499 | 8.311 | 19.016 | 49.675 |
| | p_5 | ML | 3.888 | 8.967 | 18.354 | 49.811 |
| | p_5 | $\hat{\varrho}_{AM}$ | 2.645 | 8.652 | 19.848 | 50.719 |

^a REML and ML estimators in the balanced case, exact results from numerical integration

^b REML and ML estimators in the unbalanced case, based on 300 simulation runs

^c $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AM}$ as defined in the text in the unbalanced case, based on 1000 simulation runs

estimators are inadmissible in the sense that no one estimator is dominant over others in terms of MSE. There is always a tradeoff between bias, variance and MSE. Due to the dependency of the estimators on design, there are no simple answers. It is quite striking that the ML estimator from the common correlation model generally had deficiencies in every aspect. If one prefers MSE efficiency of the estimators, then the concentrated and truncated estimators were superior to their counterparts, especially for small n and small ϱ . However, this MSE efficiency was the by-product of the bias deficiency of the estimator. The larger the MSE efficiency, the larger the bias deficiency. Recall that the biases of $\hat{\varrho}_{AC}$ and $\hat{\varrho}_{AT}$ were usually not less than 20% when $\varrho \leq 0.10$. On the other hand, the bias of estimators, $\hat{\varrho}_A$ is preferred for its smaller bias, usually not greater than 10%.

(6) One of the major aims of quantitative geneticists is to estimate genetic parameters of a population; for example, the intraclass correlation in this study. Given limited resources, it is usually difficult to obtain an accurate estimate from a single experiment. Thus some geneticists tend to gather information from literature and try to get

a pooled estimate which is hopefully better than a single estimate. The common practices of treating a negative estimate as zero or not reporting a negative estimate could introduce sizable biases in this pooling process, especially for small ϱ . In this regard, we recommend the reporting negative estimates of $\hat{\varrho}_A$ to avoid a large potential bias being introduced.

Acknowledgements. We thank the college of Agriculture and Life Sciences, University of Wisconsin-Madison, for supporting this work. Computing was made on SUN 3 Computer of the Department of Statistics, and the CALS computer of Agriculture and Life Sciences, University of Wisconsin-Madison. Consultations provided by Sook-Fwe Yap on several occasions are highly appreciated. We are indebted to the editor and a reference for comments and suggestions which improved the paper substantially.

References

- Becker RH, Chambers JM, Wilks AR (1988) The new S language: a programming environment for data analysis and graphics. Wadsworth and Brooks/Cole, Pacific Grove, California
- Donner A, Koval JJ (1980) The estimation of intraclass correlation in the analysis of family data. *Biometrics* 36:19–25
- Donner A, Koval JJ (1983) A note on the accuracy of Fisher's approximation to the large sample variance of an intraclass correlation. *Commun Statist – Simula Computa* 12:443–449
- Falconer DS (1989) Introduction to quantitative genetics, 3rd edn. Longman, London New York
- Fisher RA (1918) The correlation between relatives on the supposition of mendelian inheritance. *Trans Roy Soc Edin* 52:399–433
- Fisher RA (1925) Statistical methods for research workers. Oliver and Boyd, Edinburgh London
- Gill JL, Jensen EL (1968) Probability of obtaining negative estimates of heritability. *Biometrics* 24:517–526
- Ginsburg EH (1973) On the planning of the experiment on estimation of intraclass correlation. *Biom Z* 15:47–52
- Henderson CR (1953) Estimation of variance and covariance components. *Biometrics* 9:226–252
- Herbach LH (1959) Properties of Model II type analysis of variance tests. *Ann Math Statist* 30:939–959
- IMSL, INC. (1989) IMSL®, MATH/LIBRARY: Fortran Subroutines for Mathematical Applications. Houston, Texas
- Kempthorne O (1957) An introduction to genetic statistics. Wiley, New York
- Leone FC, Nelson LS (1966) Sampling distributions of variance components. I. Empirical studies of balanced nested designs. *Technometrics* 8:457–468
- O'Neill ME, Prasetyo LH, Kirby AC, Nicholas FW (1987) A comparison of parent-offspring correlation estimators in terms of large sample mean squared error and bias. *Biometrics* 43:37–44
- Osborne R, Paterson WSB (1952) On the sampling variance of heritability estimates derived from variance analyses. *Proc Roy Soc Edin* B64:456–461
- Ponzoni RW, James JW (1978) Possible biases in heritability estimates from intraclass correlation. *Theor Appl Genet* 53:25–27
- Rosner B, Donner A, Hennekens CH (1977) Estimation of intraclass correlation from familial data. *Appl Stat* 26:179–187

- SAS Institute Inc. (1985) SAS® User's Guide: Statistics, 5th edn. Cary, North Carolina
- Searle SR (1971) Topics in variance component estimation. *Biometrics* 27:1–76
- Searle SR (1988) Mixed models and unbalanced data: wherefrom, whereat and whereto? *Commun Statist – Theor Meth* 17:935–968
- Snedecor GW, Cochran WG (1980) *Statistical methods*, 17th edn. The Iowa State University Press, Ames
- Smith CAB (1956) On the estimation of intraclass correlation. *Ann Hum Genet* 21:363–373
- Swallow WH, Monahan JF (1984) Monte Carlo comparison of ANOVA, MIVQUE, REML, and ML estimators of variance components. *Technometrics* 26:47–55
- Swiger LA, Harvey WR, Everson DO, Gregory KE (1964) The variance of intraclass correlation involving groups with one observation. *Biometrics* 20:818–826
- Wang CS, Yandell BS, Rutledge JJ (1991) Bias of maximum likelihood estimator of intraclass correlation. *Theor Appl Genet* 82:421–424