

Kansas State University Libraries

New Prairie Press

Conference on Applied Statistics in Agriculture

2001 - 13th Annual Conference Proceedings

ASSESSING THE PERFORMANCE OF CLOSED-FORM APPROXIMATIONS TO THE REML ESTIMATOR OF HERITABILITY

Brent D. Burch

Ian R. Harris

Follow this and additional works at: <https://newprairiepress.org/agstatconference>



Part of the [Agriculture Commons](#), and the [Applied Statistics Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](#).

Recommended Citation

Burch, Brent D. and Harris, Ian R. (2001). "ASSESSING THE PERFORMANCE OF CLOSED-FORM APPROXIMATIONS TO THE REML ESTIMATOR OF HERITABILITY," *Conference on Applied Statistics in Agriculture*. <https://doi.org/10.4148/2475-7772.1218>

This is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact cads@k-state.edu.

ASSESSING THE PERFORMANCE OF CLOSED-FORM APPROXIMATIONS TO THE REML ESTIMATOR OF HERITABILITY

Brent D. Burch¹ and Ian R. Harris²

¹Department of Mathematics and Statistics
Northern Arizona University, Flagstaff, Arizona 86011, U.S.A.

²Department of Statistical Science
Southern Methodist University, Dallas, Texas 75275, U.S.A.

ABSTRACT

For many researchers the restricted maximum likelihood (REML) method of estimation is the procedure of choice for estimating heritability. In most applications the REML estimate can only be obtained via an iterative method. In some cases the algorithm used to compute the REML estimate may be slow or fail to converge. These predicaments have provided the motivation to develop closed-form approximations to the REML estimator of heritability in mixed linear models having two variance components. These estimators are compared to the REML estimator by considering their large and small sample properties. We provide guidance on how to select the closed-form estimator that provides the best approximation to the REML estimator. A simple one-way random effects model and an animal breeding model with correlated genetic effects are presented.

1 Introduction

Heritability, the proportion of total variation in the phenotypic values attributable to additive genetic effects, is an important parameter in plant and animal breeding studies. If σ_1^2 is the variation in the phenotypic values due to additive genetic effects and σ_2^2 is the variation in the phenotypic values due to other effects, then $\sigma_1^2 + \sigma_2^2$ is the total phenotypic variance. It follows that heritability, denoted by ρ , is equal to $\sigma_1^2 / (\sigma_1^2 + \sigma_2^2)$ and is the ratio of the additive genetic variance to the phenotypic variance.

The restricted maximum likelihood (REML) method is a commonly used procedure to estimate variance components and hence heritability. The REML method of estimation is a maximum likelihood procedure based on that part of the likelihood function which is free of fixed effects. See Anderson and Bancroft (1952), Russell and Bradley (1958), and W. A. Thompson (1962) for a description of the REML procedure. Harville (1974, 1977) suggests that REML estimators can be obtained by maximizing a likelihood function based on error contrasts. In our paper we illustrate the use of quadratic forms of the error contrasts, which may also be viewed as quadratic forms of linear combinations of the data, to estimate ρ .

In most applications the REML estimate is not available in closed-form and thus iterative procedures are required to find its value. Searle, Casella, and McCulloch (1992)

suggest that there is no consensus on the best iterative procedure since the commonly used techniques do not guarantee convergence to a global maximum from an arbitrary starting value. Even in relatively simple models such as unbalanced one-way random effects models, Swallow and Monahan (1984) found that the REML estimation technique failed to converge in 20 iterations.

These results provide the motivation to obtain closed-form approximations to the REML estimator of ρ . These estimators are compared to the REML estimator by evaluating their asymptotic variances in large-sample applications and by evaluating their mean squared errors in small-sample applications. Small-sample comparisons also take into account the fact that estimators are truncated if computed values fall outside the parameter space. These large-sample as well as small-sample comparisons suggest there exists a closed-form estimator that is a competitor to and may even outperform the REML estimator.

2 Mixed Linear Models with Two Variance Components and the REML Estimator of ρ

The mixed linear model under consideration is

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}, \tag{1}$$

where \mathbf{Y} is a $n \times 1$ vector of observable random variables, $\boldsymbol{\beta}$ is a $p \times 1$ vector of location parameters, and \mathbf{u} and \mathbf{e} are vectors of unobservable random variables of size $m \times 1$ and $n \times 1$, respectively. The matrices \mathbf{X} and \mathbf{Z} are known and without loss of generality, $\text{rank}(\mathbf{X}) = p$. It is assumed that $\mathbf{u} \sim N(\mathbf{0}, \sigma_1^2 \mathbf{A})$ and independently $\mathbf{e} \sim N(\mathbf{0}, \sigma_2^2 \mathbf{I}_n)$. It follows that $\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma_2^2 \mathbf{I}_n + \sigma_1^2 \mathbf{Z}\mathbf{A}\mathbf{Z}')$. The known matrix \mathbf{A} is referred to as the relationship matrix since it describes the degree to which the elements of \mathbf{u} are related. In the usual manner, we take $\sigma_1^2 \geq 0, \sigma_2^2 > 0$ so that $0 \leq \rho < 1$.

To find the REML estimators of the variance components and hence ρ , one can maximize the restricted likelihood function based on a set of independently distributed quadratic forms. The quadratic forms, denoted by (Q_1, \dots, Q_d) , are a set of minimal sufficient statistics associated with the reduced linear model void of the fixed effects. The quadratic forms may be obtained by diagonalizing the variance-covariance matrix of a linear transformation of the observations. The linear transformation of \mathbf{Y} is $\mathbf{H}'\mathbf{Y} \sim N(\mathbf{0}, \sigma_2^2 \mathbf{I}_{n-p} + \sigma_1^2 \mathbf{H}'\mathbf{Z}\mathbf{A}\mathbf{Z}'\mathbf{H})$, which is a $n - p$ dimensional vector whose distribution does not depend on the $\boldsymbol{\beta}$. \mathbf{H} is a $n \times (n - p)$ matrix whose columns span the space orthogonal to the space spanned by the columns of \mathbf{X} and satisfies $\mathbf{H}'\mathbf{H} = \mathbf{I}_{n-p}$. It can be shown that the quadratic forms, $Q_i \sim \sigma_2^2(1 + \Delta_i \rho / (1 - \rho)) \chi_{r_i}^2, i = 1, \dots, d$, are independently distributed where $0 \leq \Delta_1 < \dots < \Delta_d$ are the distinct eigenvalues of $\mathbf{H}'\mathbf{Z}\mathbf{A}\mathbf{Z}'\mathbf{H}$ having multiplicities r_1, \dots, r_d , respectively. See Burch and Harris (2000) for additional details.

The simplest model under the umbrella of models given by (1) is the one-way random effects model. In industrial applications, this model serves as a tool to highlight how a

specific part of the manufacturing process influences the variability in the finished product. In animal breeding studies, the one-way random effects model is referred to as the sire model since it quantifies how the genetic material the sire passes to its progeny influences the variability in the attributes of the progeny. The one-way random effects model is

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij} \tag{2}$$

where $i = 1, \dots, a$ and $j = 1, \dots, b_i$. It is assumed that α_i and ϵ_{ij} are independently distributed where $\alpha_i \stackrel{iid}{\sim} N(0, \sigma_1^2)$ and $\epsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma_2^2)$. In this model ρ is called the intraclass correlation coefficient since it is the correlation between measurements in the same class (or level) of the random effect α . The ANOVA table associated with the one-way random effects model is given in Table 1.

It is interesting to note that for the balanced one-way random effects model $d = 2$ so that the total variation in the measurements is the sum of the two quadratic forms Q_1 and Q_2 . In addition, the eigenvalues Δ_i and their replications r_i are simply $\Delta_1 = 0$, $r_1 = a(b-1)$, $\Delta_2 = b$, and $r_2 = a-1$ where $b = b_i$ for all i . The smallest eigenvalue takes on the value of zero since there are replications within a class (or level) of the random effect α . r_1 corresponds to the degrees of freedom within classes and it follows that $Q_1 \sim \sigma_2^2 \chi_{a(b-1)}^2$. The second eigenvalue in this case is the number of measurements per class, r_2 is the degrees of freedom between classes and $Q_2 \sim \sigma_2^2(1 + b\rho/(1 - \rho)) \chi_{a-1}^2$. It can be shown that the REML estimator of ρ in balanced one-way random effects models is

$$\hat{\rho} = \frac{r_1 Q_2 - r_2 Q_1}{(\Delta_2 - 1)r_2 Q_1 - (\Delta_1 - 1)r_1 Q_2} \tag{3}$$

which must be truncated if it is outside of the parameter space.

In general, the REML estimator of ρ may be obtained by maximizing the likelihood function of ρ and σ_2^2 based on Q_1, \dots, Q_d . It can be shown that the REML estimator of ρ is

$$\hat{\rho} = \frac{\sum_{i=1}^d u_i(\rho, \Delta_i, r_i) Q_i}{\sum_{i=1}^d v_i(\rho, \Delta_i, r_i) Q_i} \tag{4}$$

where u_i and v_i are coefficients of Q_i . In essence, the REML estimator of ρ is a ratio of linear combinations of the quadratic forms Q_1, \dots, Q_d whose coefficients depend in part on the unknown parameter. In this manner the value of $\hat{\rho}$ must be obtained iteratively by selecting a starting value and relying on the convergence of the procedure. Only for case when $d = 2$ are the coefficients of Q_i free of the parameter value and hence a closed-form expression for $\hat{\rho}$ available. We use this fact to build closed-form approximations to the REML estimator of ρ .

3 Closed-form Approximations to the REML Estimator of ρ

As noted in Section 2, a closed-form expression for the REML estimator is not available if $d > 2$, where d is the dimension of (Q_1, \dots, Q_d) . To obtain a closed-form estimator of ρ when $d > 2$, we start by compressing (Q_1, \dots, Q_d) into two non-overlapping sums $\sum_i Q_i$ and $\sum_j Q_j$. Since $i \neq j$, the two sums are independent. The next step is to determine which Q 's should be used in the first sum and which Q 's should be used in the second sum.

For example, in the one-way random effects model the total variation in the measurements may be written as

$$\begin{aligned} \sum_{i=1}^a \sum_{j=1}^{b_i} (Y_{ij} - \bar{Y})^2 &= Q_1 + \dots + Q_d \\ &= (Q_1 + \dots + Q_k) + (Q_{k+1} + \dots + Q_d) \\ &= Q_1^* + Q_2^*. \end{aligned} \tag{5}$$

In this manner the sum of d quadratic forms has been compressed into the sum of two quadratic forms. Note that $Q_1^* = \sum_{i=1}^k Q_i$ and $Q_2^* = \sum_{i=k+1}^d Q_i$ are such that the quadratic forms associated with the smaller eigenvalues are in the first sum and the quadratic forms associated with the larger eigenvalues are in the second sum. The rationale for grouping the Q_i 's in this manner is in part that the ratio of the sums is an approximate pivotal quantity for ρ which can be inverted to produce confidence intervals. See Burch and Iyer (1997) for more details on the relationship between pivotal quantities and confidence intervals for ρ . Placing Q_1 by itself and Q_2, \dots, Q_d in a sum for the one-way random effects model corresponds to dividing the total variation into the familiar within and between sums of squares. See LaMotte (1976) for additional details.

Although the concept of partitioning and compressing Q_1, \dots, Q_d was illustrated using the one-way random effects model, it can be applied to any model of the form (1). It follows that closed-form approximations to the REML estimator of ρ may be written as

$$\hat{\rho}_k = \frac{r_1^* Q_2^* - r_2^* Q_1^*}{(\Delta_2^* - 1)r_2^* Q_1^* - (\Delta_1^* - 1)r_1^* Q_2^*} \tag{6}$$

confined to the parameter space where

$$\Delta_1^* = \frac{\sum_{i=1}^k r_i \Delta_i}{\sum_{i=1}^k r_i}, \quad \Delta_2^* = \frac{\sum_{i=k+1}^d r_i \Delta_i}{\sum_{i=k+1}^d r_i} \tag{7}$$

and

$$r_1^* = \sum_{i=1}^k r_i, \quad r_2^* = \sum_{i=k+1}^d r_i. \tag{8}$$

Using the results of Satterthwaite (1946),

$$Q_1^* \underset{\text{approx}}{\sim} \frac{\sigma_2^2}{1 - \rho} (1 + \rho(\Delta_1^* - 1)) \chi_{r_1^*}^2 \quad (9)$$

$$Q_2^* \underset{\text{approx}}{\sim} \frac{\sigma_2^2}{1 - \rho} (1 + \rho(\Delta_2^* - 1)) \chi_{r_2^*}^2. \quad (10)$$

The Satterthwaite approximation is improved if the Q_i 's for each sum are selected according to the similarity of their eigenvalues, which provides further motivation for the groupings suggested here. One can show that the closed-form approximations to the REML estimator of ρ are obtained by maximizing the approximate loglikelihood function based on (9) and (10). Note the similarities between $\hat{\rho}_k$ given in (6) and $\hat{\rho}$ given in (3). The closed-form estimators, indexed by k , have the same structure as the REML estimator of ρ when $d = 2$.

Conceptually, $\hat{\rho}_k$ is the true REML estimator of ρ if one considers a mixed linear model having $d = 2$, where the first eigenvalue is Δ_1^* having replication r_1^* and the second eigenvalue is Δ_2^* having replication r_2^* . Of course, in most cases such a model does not really exist and the conception of one is just a convenient vehicle for understanding the partitioning and compressing of information in order to obtain analytic expressions for point estimators of ρ . The choice of k and quantifying the information lost in order to obtain a closed-form estimator of ρ will be examined later in this paper.

4 Comparing the Closed-form Estimators to the REML Estimator of ρ

The asymptotic properties of $\hat{\rho}$ (the REML estimator of ρ) and $\hat{\rho}_k$ (closed-form approximations to the REML estimator of ρ) can be determined using regularity conditions. From Burch and Harris (2001) it can be shown that

$$\hat{\rho} \underset{\text{asympt}}{\sim} N(\rho, Var(\hat{\rho})) \quad (11)$$

and

$$\hat{\rho}_k \underset{\text{asympt}}{\sim} N(\rho, Var(\hat{\rho}_k)). \quad (12)$$

The forms of $Var(\hat{\rho})$ and $Var(\hat{\rho}_k)$ are given in Burch and Harris (2000).

We compare $\hat{\rho}_k$ to $\hat{\rho}$ by examining their large sample properties. The asymptotic relative efficiency of $\hat{\rho}_k$ with respect to $\hat{\rho}$ is defined as

$$\frac{Var(\hat{\rho})}{Var(\hat{\rho}_k)}. \quad (13)$$

Note that $0 \leq Var(\hat{\rho})/Var(\hat{\rho}_k) \leq 1$ and the asymptotic relative efficiency depends on ρ since the asymptotic variances depend on ρ .

As an example, consider the data in Table 2 which represents a scenario in which an unbalanced one-way random effects design was employed. There of 3 groups having 3, 5, and 7 observations, respectively. The number of distinct eigenvalues in this example is $d = 3$. Table 3 summarizes the values associated with Q_i , Δ_i , and r_i for $i = 1, 2, 3$. Maximizing the restricted loglikelihood function with respect to ρ and σ_2^2 gives a REML estimator for ρ of 0.80. As $d = 3$, there are two ways in which to compress the three quadratic forms into Q_1^* and Q_2^* . If $k = 1$, then $Q_1^* = Q_1$ and $Q_2^* = Q_2 + Q_3$ are the resulting statistics. The closed form estimator is then $\hat{\rho}_1 = 0.81$. If $k = 2$, the closed-form estimator is built from the statistics $Q_1^* = Q_1 + Q_2$ and $Q_2^* = Q_3$. In this case $\hat{\rho}_2 = 0.95$.

The question that arises is which estimator is best, $\hat{\rho}_1$ or $\hat{\rho}_2$? In other words, which closed-form estimator best approximates the REML estimator of ρ ? It is clear that both estimators are asymptotically unbiased so we turn our attention to their asymptotic variances. In Figure 1, the asymptotic relative efficiencies of $\hat{\rho}_k$ as compared to $\hat{\rho}$ are presented for $k = 1$ and $k = 2$. Figure 1 suggests that the estimator corresponding to $k = 1$ is vastly superior to the estimator corresponding to $k = 2$. In addition, the asymptotic properties of $\hat{\rho}_1$ rival those of the REML estimator of ρ .

Comparisons of the estimators based on small-samples are performed via simulation. Using scaled chi-squared distributions, 10000 realizations of the quadratic forms Q_1, Q_2, Q_3 were obtained for values of ρ ranging from 0.0 to 1.0 in increments of 0.05. The mean-squared error of the estimators and thus $MSE(\hat{\rho})/MSE(\hat{\rho}_k)$ were computed for these values of ρ . In the simulation study an estimate of ρ was assigned the value of zero (or one) if the computed value fell outside the parameter space. Figure 2 is a display of the relative mean-squared errors for $k = 1$ and $k = 2$. As in the large-sample comparisons presented in Figure 1, it is clear that $\hat{\rho}_1$ outperforms $\hat{\rho}_2$.

Comparisons among the estimators are also made in terms of their truncation percentages. The estimators $\hat{\rho}_1$ and $\hat{\rho}_2$ are truncated if their computed values fall outside the parameter space $[0, 1)$. By definition, the REML estimator $\hat{\rho}$ is bound by the endpoints of the parameter space and thus remains in the unit interval. For a given value of ρ we define "truncation" as the proportion of estimates that are less than or equal to 0.001 plus the proportion of estimates that are greater than or equal to 0.999 using the 10000 realizations of the quadratic forms. In this manner we compare $\hat{\rho}_1$, $\hat{\rho}_2$, and $\hat{\rho}$ in terms of their propensity to hover at or near the endpoints of the parameter space. The results are displayed in Figure 3. One may infer that $\hat{\rho}_1$ and $\hat{\rho}$ are only truncated at the lower end of the parameter space whereas $\hat{\rho}_2$ experiences truncation at both ends of the parameter space. We conclude that $\hat{\rho}_1$ mimics the performance of $\hat{\rho}$ in terms of truncation percentage.

A second example uses 171 yearling bulls from a Red Angus seed stock herd in Montana (Evans et al. (1995)). A trait of interest was the loineye (i.e., ribeye) muscle area measured in square inches. Ultrasound techniques were used to procure these measurements which were located on the dorso-ventral line between the 12th and 13th ribs on the left side of each animal. The fixed effect was age of dam and in the analysis five categories were used: 2 years, 3 years, 4 years, 5-9 years, and 10 or more years. The random effects in the model are the animal's additive genetic effect (\mathbf{u}) and the error (\mathbf{e}).

In the mixed linear model (1), \mathbf{Y} is a 171×1 vector of observable random variables, \mathbf{X} is a 171×5 incidence matrix, $\boldsymbol{\beta}$ is a 5×1 vector of unknown parameters for the 5 fixed effect categories, $\mathbf{Z} = \mathbf{I}_{171}$, and \mathbf{u} and \mathbf{e} are vectors of unobservable random variables of size 171×1 . The relationship matrix \mathbf{A} was determined using a recursive method given in Henderson (1976). It uses knowledge of the animal's sire, dam, and grandparents. Note that some animals are inbred so that it is possible that $Var(u_i) > \sigma_1^2$. For instance, it turns out that $Var(u_1) = 1.03125\sigma_1^2$.

The number of distinct eigenvalues is $d = 165$. Eigenvalues range in magnitude from $\Delta_1 = 0.56569$ to $\Delta_{165} = 8.65925$. Except for $\Delta_{61} = 0.67188$ having $r_{61} = 2$, all eigenvalues have a multiplicity of one. Maximizing the restricted loglikelihood function with respect to ρ and σ_2^2 gives a REML estimator for ρ of 0.10. Since $d = 165$, there are 164 possible values of k . In other words, there are 164 ways to partition the information $Q_1, \dots, Q_k, Q_{k+1}, \dots, Q_{165}$ into two pieces in order to obtain a closed-form approximation to the REML estimator of ρ given in (6).

In Figure 4, the asymptotic relative efficiencies of $\hat{\rho}_k$ as compared to $\hat{\rho}$ are presented for selected values of k . As in the first example, one can see that some information about ρ was surrendered in order to obtain the closed-form estimators. Figure 4 suggests that if k is too small, the asymptotic relative efficiency of $\hat{\rho}_k$ is poor when ρ is small. Likewise, if k is too large, the asymptotic relative efficiency of $\hat{\rho}_k$ is poor when ρ is large.

Small-sample comparisons of the estimators are made via simulation. Using scaled chi-squared distributions, 10000 realizations of the quadratic forms Q_1, \dots, Q_{165} were obtained for values of ρ ranging from 0.0 to 1.0 in increments of 0.05. The ratio of mean-squared error of the estimators $MSE(\hat{\rho})/MSE(\hat{\rho}_k)$ was computed for these values of ρ . Figure 5 is a display of the ratio for selected values of k . It is interesting to see that $\hat{\rho}_{150}$ and $\hat{\rho}_{160}$ (as well as other estimators not pictured) outperform $\hat{\rho}$ for a portion of the parameter space in this simulation.

Figure 6 displays "truncation percentages" for selected $\hat{\rho}_k$'s and $\hat{\rho}$. It appears that $\hat{\rho}_{150}$ has results that are similar to those of $\hat{\rho}$. The asymptotic and simulation comparisons yield similar selections of closed-form estimators and suggest that there is a closed-form estimator, $\hat{\rho}_{150}$, that is a viable contender to $\hat{\rho}$. It is interesting to note that $\hat{\rho}_{150} = 0.08$ compared to $\hat{\rho} = 0.10$.

5 Discussion and Conclusions

This paper presents a set of analytic expressions which approximate the REML estimator of heritability, or the intraclass correlation coefficient, in a mixed linear model having two variance components. The model takes into account the possibility that the elements of the random vector may be correlated with one another. The resulting estimators, which can be obtained using a non-iterative procedure, are relatively easy to compute.

We offer the technique of partitioning and compressing information contained in a sample as a way to obtain analytically expressed estimators which are viable alternatives to the

iteratively generated REML estimator. The estimators are built using quadratic forms of the data as well as eigenvalues of a variance-covariance matrix. The method developed in this paper may also be used to obtain approximate confidence intervals for ρ which possess analytic endpoints. In this manner the approximation to the REML estimator of ρ may be associated with a specific interval estimate of ρ .

By examining the asymptotic variances of the estimators in large-sample scenarios and the mean-squared error as well as truncation percentages of the the estimators in small sample scenarios, one can compare the closed-form estimators with the iteratively generated REML estimator. The examples presented in this paper suggest that there is a closed-form estimator that contends with the REML estimator of ρ . Readers are encouraged to see Burch and Harris (2001) for a more detailed discussion of this topic. Future work will explore the possibility of extending the techniques discussed in this paper to mixed linear models having more than two variance components.

REFERENCES

- Anderson, R. L. and Bancroft, T. A. (1952). *Statistical Theory in Research*. New York: McGraw-Hill.
- Burch, B. D. and Harris (2001). Closed-form approximations to the REML estimator of a variance ratio (or heritability) in a mixed linear model. *Biometrics* **57**, 43–51.
- Burch, B. D. and Harris, I. R. (2000). Point estimators of heritability based in confidence intervals: A closed-form approximation to the REML estimator. *Proceedings of the Kansas State University Conference on Applied Statistics in Agriculture*, 95–102.
- Burch, B. D. and Iyer, H. K. (1997). Exact confidence intervals for a variance ratio (or heritability) in a mixed linear model. *Biometrics* **53**, 1318–1333.
- Evans, J. L., Golden, B. L., Bailey, D. R. C., Gilbert, R. P., and Green, R. D. (1995). Genetic parameter estimates of ultrasound measures of backfat thickness, loin eye muscle area, and gray shading score in red angus cattle. *Proceedings, Western Section, American Society of Animal Science* **46**, 202–204.
- Harville, D. A. (1974). Bayesian inference for variance of components using only error contrasts. *Biometrika* **61**, 383–385.
- Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association* **72**, 320–340.
- Henderson, C. R. (1976). A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. *Biometrics* **32**, 69–83.

- LaMotte, L. R. (1976). Invariant Quadratic estimators in the random, one-way ANOVA model. *Biometrics* **32**, 793–804.
- Russell, T. S. and Bradley, R. A. (1958). One-way variances in a two-way classification. *Biometrika* **45**, 111–129.
- Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin* **2**, 110–114.
- Searle, S. R., Casella, G., and McCulloch, C. E. (1992). *Variance Components*. New York: Wiley.
- Swallow, W. H. and Monahan, J. F. (1984). Monte Carlo comparison of ANOVA, MIVQUE, REML, and ML estimators of variance components. *Technometrics* **26**, 47–57.
- Thompson, W. A., Jr. (1962). The problem of negative estimates of variance components. *Annals of Mathematical Statistics* **33**, 273–289.

Table 1: ANOVA Table for the One-way Random Effects Model

| Source | df | SS |
|---------|------------------------|--|
| Between | $a - 1$ | $Q_2 + \dots + Q_d$ |
| Within | $\sum_{i=1}^a b_i - a$ | Q_1 |
| Total | $\sum_{i=1}^a b_i - 1$ | $\sum_{i=1}^a \sum_{j=1}^{b_i} (Y_{ij} - \bar{Y})^2$ |

Table 2: Unbalanced One-way Random Effects Data Set

| Group 1 | Group 2 | Group 3 |
|---------|---------|---------|
| 2 | 12 | 1 |
| 4 | 12 | 2 |
| 8 | 16 | 3 |
| | 16 | 3 |
| | 19 | 5 |
| | | 7 |
| | | 9 |

Table 3: Q_i , Δ_i , and r_i Values for the One-way Random Effects Data Set

| Q_i | Δ_i | r_i |
|----------------|-------------------|------------|
| $Q_3 = 281.17$ | $\Delta_3 = 5.92$ | $r_3 = 1$ |
| $Q_2 = 93.67$ | $\Delta_2 = 3.55$ | $r_2 = 1$ |
| $Q_1 = 104.10$ | $\Delta_1 = 0.00$ | $r_1 = 12$ |

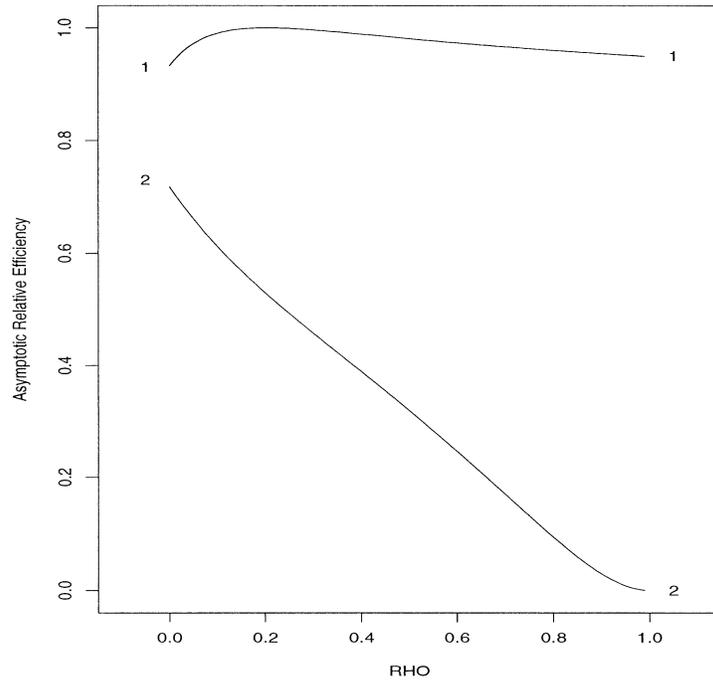


Figure 1: Asymptotic relative efficiency of $\hat{\rho}_k$ for the simple one-way random effects data.

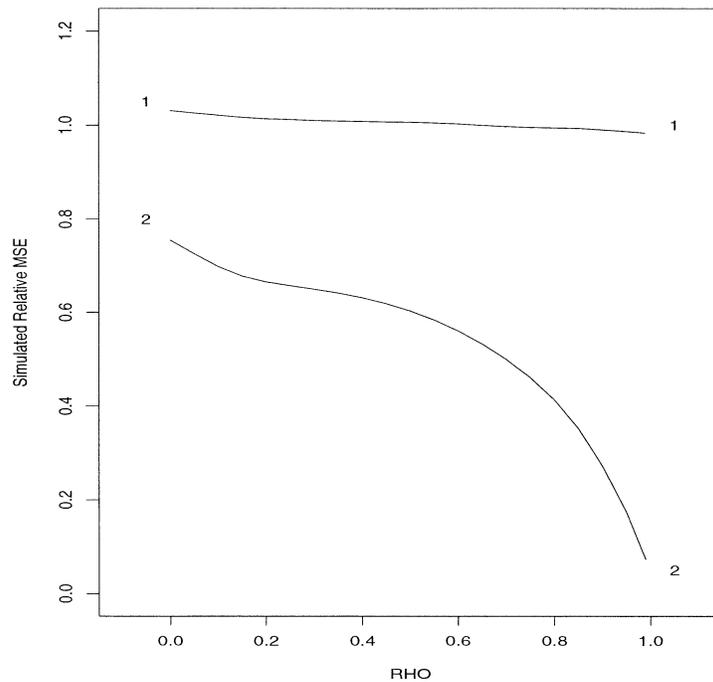


Figure 2: Simulated relative MSE of $\hat{\rho}_k$ for the simple one-way random effects data.

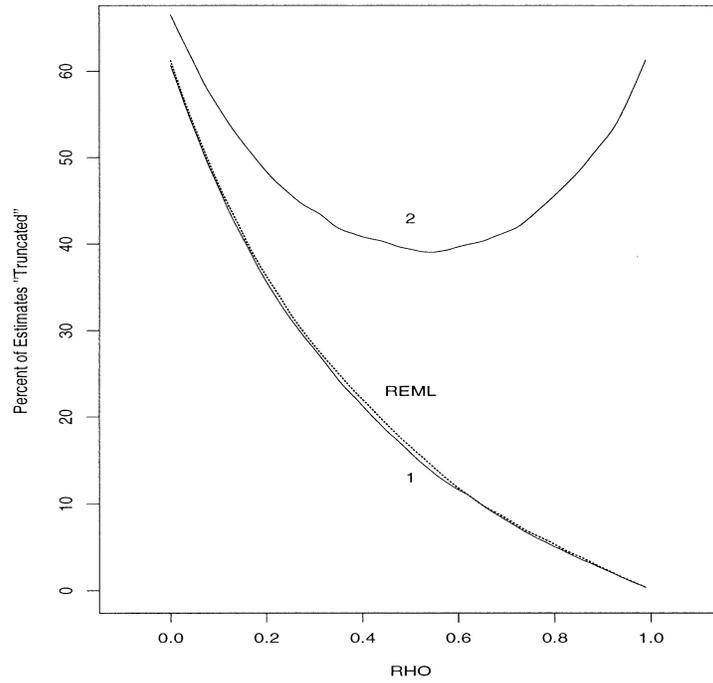


Figure 3: Percentage of estimates affected by bounds on parameter space for simple data.

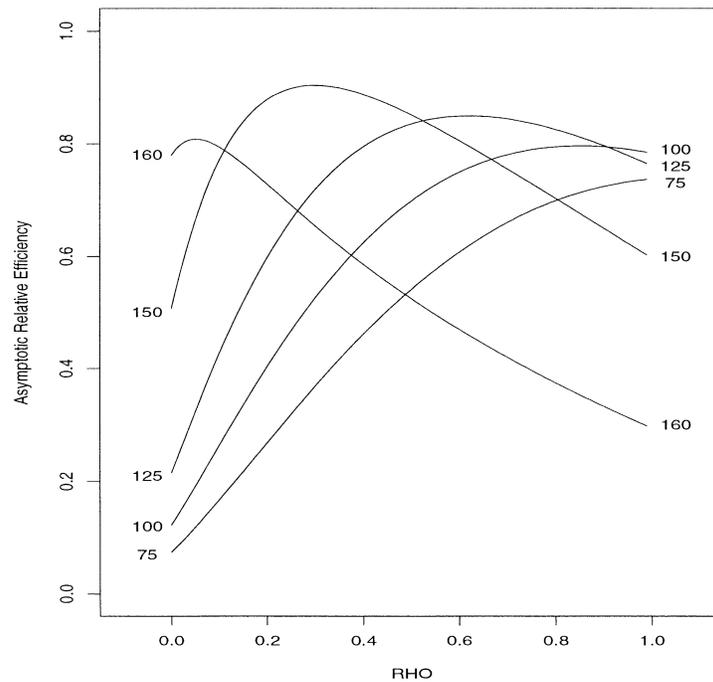


Figure 4: Asymptotic relative efficiency of $\hat{\rho}_k$ for selected values of k using the ribeye data.

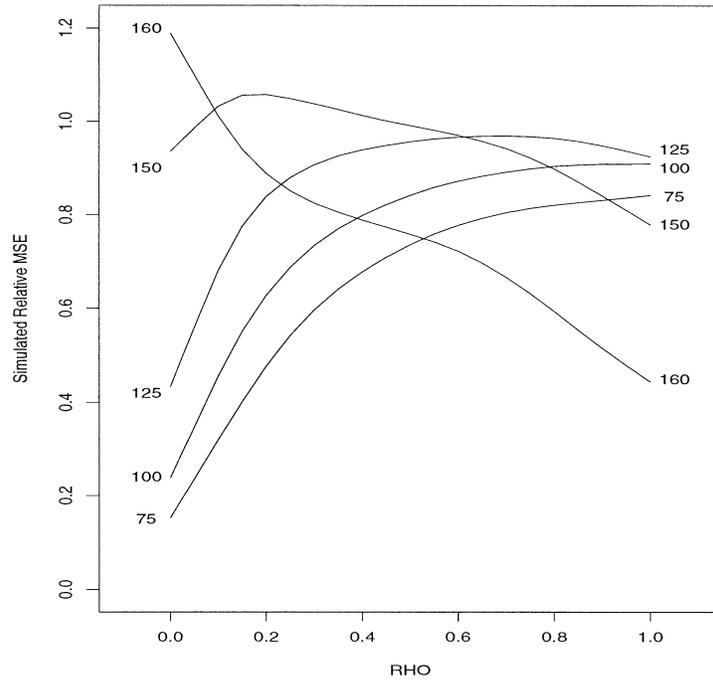


Figure 5: Simulated relative MSE of $\hat{\rho}_k$ for selected values of k using the ribeye data.

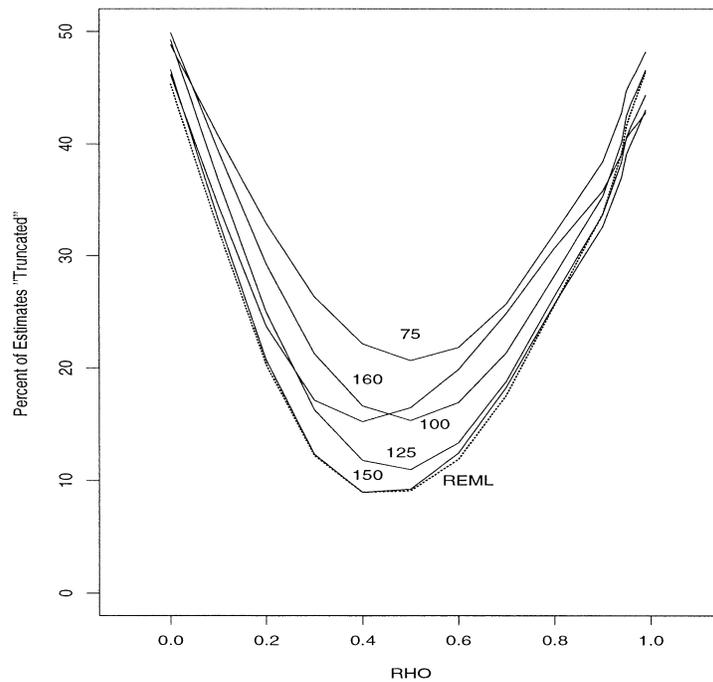


Figure 6: Percentage of estimates affected by bounds on parameter space for ribeye data.