Department of Statistical Science
Southern Methodist University
Technical report SMU-TR-307
July 2nd 2003

The probability of preponderancy, an alternative to the Intraclass Correlation

Ian R. Harris

Department of Statistical Science
Southern Methodist University
Dallas, Texas 75275
(iharris@mail.smu.edu)

Brent D. Burch

Department of Mathematics and Statistics
Northern Arizona University
Flagstaff, Arizona 86011
(brent.burch@nau.edu)

The probability of preponderancy, an alternative to the intraclass correlation

By IAN R. HARRIS

Department of Statistical Science, Southern Methodist University
Dallas, TX 75275, U.S.A.
iharris@mail.smu.edu

and BRENT D. BURCH

Department of Mathematics and Statistics, Northern Arizona University
Flagstaff, AZ 86011-5717, U.S.A.
Brent.Burch@nau.edu

Summary

We propose a new parameter for measuring the influence of a random effect in a mixed linear model. This is the probability of preponderance of the random effect under study over the other random effects. In a one-way random effects model, this is simply the probability the group random effect is larger in absolute size than the individual random effect (or error). We discuss the meaning of the parameter and relate it to the more familiar intraclass correlation coefficient. The new parameter is invariant under a broad set of transformations of the error distributions, and thus is not tied to a particular parametric family. This is in contrast to the intraclass correlation, which has it's origins in normally distributed random effects. Furthermore, the new parameter directly measures the random effect's impact on the observations whereas the intraclass correlation relies on the variances (second moments) of the random effects. We suggest parametric and nonparametric estimators of the parameter, and demonstrate the applicability of the results using real data. We also indicate how to extend the ideas presented in this paper to models having more than two sources of variation.

Some key words: Bootstrap confidence intervals; Nature versus nurture; One-way random effects model; U-statistic: Variance components

1. Introduction

Using variance to measure variation and conducting inferences based on normality assumptions are prevalent throughout the random and mixed effects model literature. Fisher (1918), defined variance as the square of the standard deviation, and proposed using proportions of total variance to describe the contribution of a particular effect. Later, (Fisher 1925) he pointed out that this ratio is the intraclass correlation, "...the correlation merely

measures the relative importance of two groups of factors causing variation". Fisher (1918) specifically stated that variance is to be used to measure variation as the data are usually to be taken to be normally distributed. This point is a primary motivation for the present paper. If the random effects are normally distributed then variance is clearly the correct measure of variation. However, if the random effects follow, for example, a Laplace (double exponential) distribution then one might argue that absolute deviations are more appropriate. The fact that variance may not be the best measure of variation is also a motivation behind the work of Cox and Hall (2002), who examine additive random effects and consider estimating cumulants of the distributions of the random effects. In practice, one might not be sure of the correct error distribution, and so desire a parameter which has a consistent interpretation across a reasonably large family of errors. To meet this goal, we introduce a new parameter, which we term the probability of preponderancy, and denote it by θ.

The probability of preponderancy is the probability that the random effect under study is larger in absolute size than the other random effects in the model. Consider the one-way random effects model where the observation on the j^{th} individual in the i^{th} group, denoted by Y_{ij}, is

$$Y_{ij} \quad = \quad \mu + A_i + \epsilon_{ij}, \tag{1}$$

μ is the overall population mean, A_i is the group random effect, and ϵ_{ij} is the random error. The parameter under study is

$$\theta = P(|A_i| > |\epsilon_{ij}|). \tag{2}$$

The new parameter has an important role to play when measuring the influence of random effects on the observations. By definition, θ is the proportion of individuals for which the random effect outweighs the error. In essence, θ quantifies how important the random effect is to the observations without using variance components. In Section 3 we show that the preponderance probability is invariant under transformations between symmetric error distributions, and thus is not tied to a particular parametric family. The intraclass correlation coefficient is not invariant under the same transformations.

The remainder of the paper is organized as follows. Section 2 briefly reviews the history of variance components and the intraclass correlation coefficient and also robustness of common estimators to nonnormality. Section 3 introduces the preponderance probability θ, explores basic properties and considers its relationship to ρ. Section 4 discusses classical inference about θ under normality assumptions. In Section 5 we introduce a nonparametric estimator based on a U-statistic. We recommend using bootstrap methods to construct confidence intervals based on a nonparametric estimator, and discuss this in the section. Examples are given in Section 6 which illustrate the usefulness of the parameter. There are some concluding comments in Section 7 and a brief discussion that extends the method to models with more than two sources of variation.

2. Background

According to Scheffé (1956) and Searle, Casella, and McCulloch (1992), the first explicit use of the one-way random effects model was made by Airy (1861). Shortly after,

Chauvenet (1863a, 1863b) employed a one-way random effects model to study how many repeated observations of the right ascension of a star are required to obtain accurate distance estimates. Fisher (1918, 1925) used a two variance component model in quantitative genetics by considering additive genetic effects as one random effect and lumping together non-additive genetic effects with environmental effects to form a second random effect. See Bennett (1983) for details.

The intraclass correlation coefficient, first named by Fisher (1925) is generally defined in the terms of variance components. Usually it is assumed that errors have normal distributions. Even when this is not an explicit assumption, it can be argued that the use of variance to measure variation is an implicit assumption of normality, (Fisher, 1918, Cox and Hall, 2002). For a thorough overview of the intraclass correlation coefficient in the one-way random effects model, see Donner (1986). Despite the clear dominance in the literature of the assumptions of normality and using variance to measure variation, there are some papers where the intraclass correlation coefficient is not based on these assumptions. Rothery (1979) defines a nonparametric measure of intraclass correlation as a probability of certain types of concordances among the observations. The estimator is a function of the ranks of the observations. Shirahata (1982) compares the estimator derived by Rothery (1979) to two additional nonparametric estimators. Commenges and Jacqmin (1994) provide a definition of the intraclass correlation coefficient which depends on variances but not normal theory. The statistic used to test the hypothesis of null intraclass correlation is related to the pairwise correlation coefficient, which predates the ANOVA estimator of ρ. See Karlin, Cameron, and Williams (1981) for a general class of weighted pairwise correlation coefficients. Bansal and Bhandary (1994) discuss the properties of M-estimators of the intraclass correlation cofficient. Müller and Büttner (1994) note that selecting the appropriate estimator of ρ depends on the underlying sampling theory. Vogler, Wette, McGue, and Rao (1995) compare estimators of ρ under a variety of sampling conditions. Cox and Hall (2002) propose a linear model, but do not assume the random effects are normally distributed.

It is our intention to use an underlying linear model structure, that is, consider scenarios in which one can identify sources that produce variation in the observations, and where the sources combine additively. Furthermore, the parameter of interest should not be tied to a particular parametric family.

3. The Probability of Preponderancy

Although θ can be defined in more complicated mixed linear models, it is conceptually beneficial to consider the one-way random effects model and relate the probability of preponderancy to the intraclass correlation coefficient. Using (1), suppose $i = 1, ..., a$, $j = 1, ..., b_i$, and $\sum_{i=1}^{a} b_i = n$. The a groups in the model are assumed to be randomly selected from some large population of groups. Furthermore, a random sample of size b_i has been obtained from the i^{th} group. We will assume that $A_i \overset{iid}{\sim} S(0, \sigma_1^2)$, $\epsilon_{ij} \overset{iid}{\sim} S(0, \sigma_2^2)$ (where S denotes a symmetric distribution), and that A_i and ϵ_{ij} are mutually independent, with $\sigma_1^2 \geq 0$ and $\sigma_2^2 > 0$. In Section 4 of this paper we will assume the random effects are

normally distributed, but in general this is not the case.

Observations within the same group are correlated ($Cov(Y_{ij}, Y_{ij'}) = \sigma_1^2$) and observations from different groups are uncorrelated. In addition, $Var(Y_{ij}) = Var(A_i) + Var(\epsilon_{ij}) = \sigma_1^2 + \sigma_2^2$. The intraclass correlation coefficient is $\rho = \sigma_1^2/(\sigma_1^2 + \sigma_2^2)$ where $0 \leq \rho < 1$. ρ may be interpreted as the proportion of the variation in the Y_{ij}'s attributed to factor A. Note that ρ is a function of the variances of the random effects; and is not a direct comparison of random effects themselves. The preponderance probability, $\theta = P(|A_i| > |\epsilon_{ij}|)$, has a direct interpretation in terms of random effects as it compares the influence of A and ϵ on Y . For example, in genetic applications it can be interpreted as the probability that additive genetic effects are more important than "other" effects for the phenotype under investigation.

Consider the simple scenario where A and ϵ are independent and normally distributed with $\sigma_1^2 = \sigma_2^2 = 1$. In this case $\theta = 0.5$ since the two random effects contribute equally to the formation of Y. Figure 1 provides a picture of θ, the volume under the bivariate normal probability density function contrained by the domain $|A| > |\epsilon|$. Of course, if $\sigma_1^2 \neq \sigma_2^2$, then one of the random effects outweighs the other and $\theta \neq 0.5$.

The probability of preponderancy is invariant under a set of transformations described as follows. Suppose A, ϵ are both continuous and symmetrically distributed about 0, with densities from the same scale family f. We will assume possibly different scale parameters. Consider a model transformation of the members of f to a different symmetric about 0 scale family g, (an example would be from a normal family to a Laplace family). A simple way to define such a transformation is to define $h : A \to h(A)$ so that h maps quantiles of A to quantiles of $h(A)$. Examples of such transformations are: the scale transformations ($A \to kA$); signed-power transformations ($A \to$ sign $(A)|A|^k$); and the normal to Laplace transformation. Note that this transformation must be monotonic increasing on $|A|$, due to the preservation of symmetry about 0.

Now for theoretical consistency we will apply the same transformation to ϵ, that is $h : \epsilon \to h(\epsilon)$. Note that applying the same transformation to A and ϵ is critical. The whole point of these models is to be able to make comparisons of the relative size or variability of the random effects. If we were to allow or consider separate transformations for A and ϵ we would have to allow transformations that dramatically and "unequally" scaled them, for example $A \to 100A$ paired with $\epsilon \to \epsilon/100$. Since h is monotonic increasing on absolute values, this implies $P(|h(A)| > |h(\epsilon)|) = P(|A| > |\epsilon|)$. That is, θ is invariant under simultaneous transformation of A and ϵ by h. Note also that in general the ratio of second moments of A and ϵ will not be preserved under these transformations, only the scale transformation will preserve this ratio. Thus in general, ρ is not invariant under transformation. So in this sense, θ is not tied to any particular parametric form, and "means the same", whether the errors are Laplace, normal or some other symmetric scale family distribution, in contrast to ρ.

Closed-form expressions relating θ to ρ can be obtained for specific distributions. If the random effects are normally distributed, it follows that

$$\theta \;=\; P(A_i^2 > \epsilon_{ij}^2)$$

$$
\begin{aligned}
&= P\left(\frac{\epsilon_{ij}^2/\sigma_2^2}{A_i^2/\sigma_1^2} < \sigma_1^2/\sigma_2^2\right) \\
&= F_{1,1}(\sigma_1^2/\sigma_2^2) \qquad (3)
\end{aligned}
$$

where $F_{1,1}(.)$ is the cumulative distribution function of an F-distributed variate having numerator and denominator degrees of freedom equal to one. By definition, $0 \leq \theta < 1$. In this case the parameter can be written as

$$
\begin{aligned}
\theta &= F_{1,1}(\rho/(1-\rho)) \\
&= \frac{2}{\pi} tan^{-1}\sqrt{\frac{\rho}{1-\rho}} \\
&= \frac{2}{\pi} sin^{-1}\sqrt{\rho} \qquad (4)
\end{aligned}
$$

and thus

$$
\rho = sin^2(\frac{\pi}{2}\theta). \qquad (5)
$$

Note that when $\rho = 0$, $\theta = 0$, when $\rho = 1/2$, $\theta = 1/2$, and as ρ approaches one, θ approaches one.

One alternative distribution for the random effects is Laplace (double exponential). Suppose that A_i has a Laplace distribution with mean 0 and variance σ_1^2, ϵ_{ij} is Laplace with mean 0 and variance σ_2^2, and A_i and ϵ_{ij} are mutually independent. In this scenario it is known that $|A_i|$ and $|\epsilon_{ij}|$ have exponential distributions with means equal to $\sigma_1/\sqrt{2}$ and $\sigma_2/\sqrt{2}$, respectively. It follows that

$$
\begin{aligned}
\theta &= P(|A_i| > |\epsilon_{ij}|) \\
&= P\left(\frac{|\epsilon_{ij}|/\sigma_2}{|A_i|/\sigma_1} < \sigma_1/\sigma_2\right) \\
&= F_{2,2}(\sigma_1/\sigma_2) \qquad (6)
\end{aligned}
$$

where $F_{2,2}(.)$ is the cumulative distribution function of an F-distributed random variable having numerator and denominator degrees of freedom equal to two. The parameter can also be written as

$$
\begin{aligned}
\theta &= F_{2,2}(\rho^{1/2}/(1-\rho)^{1/2}) \\
&= 1 - \left(1 + \frac{\rho^{1/2}}{(1-\rho)^{1/2}}\right)^{-1} \\
&= \frac{\rho^{1/2}}{(1-\rho)^{1/2} + \rho^{1/2}}. \qquad (7)
\end{aligned}
$$

Alternatively,

$$
\rho = \frac{\theta^2}{(1-\theta)^2 + \theta^2}. \qquad (8)
$$

Another alternative distribution for the random effects is uniform (rectangular). Suppose that A_i has a uniform distribution with mean 0 and variance σ_1^2, ϵ_{ij} has a uniform distribution with mean 0 and variance σ_2^2, and A_i and ϵ_{ij} are mutually independent. Then

$$\theta = \begin{cases} \frac{1}{2}\sqrt{\frac{\rho}{1-\rho}}, & \text{if } \rho \leq \frac{1}{2} \\ 1 - \frac{1}{2}\sqrt{\frac{1-\rho}{\rho}}, & \text{if } \rho > \frac{1}{2} \end{cases}$$

or

$$\rho = \begin{cases} \frac{4\theta^2}{1+4\theta^2}, & \text{if } \theta \leq \frac{1}{2} \\ \frac{1}{1+4(1-\theta)^2}, & \text{if } \theta > \frac{1}{2} \end{cases}.$$

As in the normal case, $0 \leq \theta < 1$. Furthermore, when $\rho = 0$, $\theta = 0$, when $\rho = 1/2$, $\theta = 1/2$, and as ρ approaches one, θ approaches one. For other values of ρ, the relationship between ρ and θ is somewhat different depending on which of these distributional forms we adopt. Figure 2 displays the values of θ as a function of $0 \leq \rho < 1$ for the normal, Laplace, and uniform distributional assumptions.

One can make some general statements about the relationship between θ and ρ, at least in the case of continuous distributions with finite second moments. Note that $\rho = 0$ iff $\sigma_1^2 = 0$ and $\sigma_2^2 > 0$. Of course, $\sigma_1^2 = 0$ iff $A = 0$ almost everywhere. Since ϵ is continuous with positive variance, we then have that these conditions are satisfied iff $P(|A| > |\epsilon|) = 0$, i.e., $\theta = 0$. Similar reasoning also establishes that ρ approaches one iff θ approaches one. It seems reasonable that $\rho > 1/2$ iff $\theta > 1/2$, but we as yet have no proof of this fact. If the variables in question are discrete then it is possible for other relationships to occur, for example, ρ can be very small, but θ be quite large.

It is important to note that θ and ρ measure somewhat different phenomena. θ relates to individuals, telling us for what proportion of individuals the group random effect is "more important" than the individual random effect. ρ tells us what proportion of the variance of a trait in a population is due to the group effect. Using variance to measure variation is a concept tied to normal theory. It follows that using ρ to measure the relative contribution of the random effect to the variation in the observations may be misleading, if the errors are not really normal. For example, one might think that $\rho = 0.01$ would imply that group effects are negligible, but under the assumption of Laplace distributions, in over 9% of individuals, the group effect outweighs the individual effect. Again, $\rho = 0.1$ would usually be taken to be a small group effect, but under Laplace distributions, $\theta = 0.25$, so that for 25% of individuals, the group effect dominates. We offer θ as an alternative to ρ and suggest its appropriateness is not dependent on a specific distribution. In the next two sections we take up the problem of inference from data, first under parametric, and then under nonparametric assumptions.

4. Inference under normal distribution assumptions

Although the main intent of this paper is to introduce a parameter and associated estimator that do not assume normality, it is still useful to look at inference about the

parameter when normality does hold. Assume that A_i is $N(0, \sigma_1^2)$ and ϵ_{ij} is $N(0, \sigma_2^2)$, with the usual independence assumptions. A commonly choosen estimator of ρ is the restricted maximum likelihood (REML) estimator, which we denote by $\widehat{\rho}$. See Searle, Casella, and McCulloch (1992, pages 90ff, 159ff, 249ff) for a general description of REML estimators of variance components. For alternatives to REML for estimating ρ, see Vogler et al. (1995). Due to the invariance of maximum likelihood estimators, the REML estimator of θ is

$$\widehat{\theta} \;=\; \frac{2}{\pi} sin^{-1}\sqrt{\widehat{\rho}}. \tag{9}$$

The analysis of variance table for the balanced one-way random effects model is given in Table 1. For balanced data, the REML estimator of ρ is the same as the ANOVA estimator

Table 1: ANOVA Table

Source	df	Sum of Squares
Between Groups	$a-1$	$Q_2 = \sum_{i=1}^{a}\sum_{j=1}^{b}\left(\overline{Y}_{i.} - \overline{Y}_{..}\right)^2$
Within Groups	$a(b-1)$	$Q_1 = \sum_{i=1}^{a}\sum_{j=1}^{b}\left(Y_{ij} - \overline{Y}_{i.}\right)^2$
Total	$ab-1$	$\sum_{i=1}^{a}\sum_{j=1}^{b}\left(Y_{ij} - \overline{Y}_{..}\right)^2$

of ρ bound to the parameter space. In these scenarios an exact confidence interval for θ is readily available. It is well known that Q_1 (the between group sum of squares) and Q_2 (the within group or error sum of squares) are independently distributed where

$$\frac{Q_1}{\sigma_2^2} \;\sim\; \chi^2(a(b-1)) \tag{10}$$

and

$$\frac{Q_2}{\sigma_2^2 + b\sigma_1^2} \;\sim\; \chi^2(a-1). \tag{11}$$

It follows that

$$\left(1 + b\frac{\sigma_1^2}{\sigma_2^2}\right)\frac{(a-1)Q_1}{a(b-1)Q_2} \sim F(a(b-1), a-1). \tag{12}$$

Let $F_{\alpha/2}$ and $F_{1-\alpha/2}$ be the $\alpha/2$ and $1-\alpha/2$ percentiles of the F distribution having numerator and denominator degrees of freedom equal to $a(b-1)$ and $a-1$, respectively. A 100(1-α)% equal-tailed confidence interval for θ is obtained by recognizing that

$$1-\alpha = P\left(F_{\alpha/2} < \left(1 + b\frac{\sigma_1^2}{\sigma_2^2}\right)\frac{(a-1)Q_1}{a(b-1)Q_2} < F_{1-\alpha/2}\right)$$

$$= P\left(\frac{1}{b}\left(F_{\alpha/2}\frac{a(b-1)Q_2}{(a-1)Q_1}-1\right) < \frac{\sigma_1^2}{\sigma_2^2} < \frac{1}{b}\left(F_{1-\alpha/2}\frac{a(b-1)Q_2}{(a-1)Q_1}-1\right)\right)$$
$$= P\left(F_{1,1}\left[\frac{1}{b}\left(F_{\alpha/2}\frac{a(b-1)Q_2}{(a-1)Q_1}-1\right)\right] < \theta < F_{1,1}\left[\frac{1}{b}\left(F_{1-\alpha/2}\frac{a(b-1)Q_2}{(a-1)Q_1}-1\right)\right]\right) \quad (13)$$

since $F_{1,1}(.)$ is a monotone increasing function. Simplifying, a 100(1-α)% equal-tailed confidence interval for θ is

$$\left(\frac{2}{\pi}tan^{-1}\sqrt{\frac{1}{b}\left(F_{\alpha/2}\frac{a(b-1)Q_2}{(a-1)Q_1}-1\right)}, \frac{2}{\pi}tan^{-1}\sqrt{\frac{1}{b}\left(F_{1-\alpha/2}\frac{a(b-1)Q_2}{(a-1)Q_1}-1\right)}\right). \quad (14)$$

The asymptotic properties of the REML estimator of θ follow from the asymptotic properties of $\hat{\rho}$. Burch and Harris (2001) show that

$$\hat{\rho} \overset{asymp}{\sim} N\Big(\rho, V(\hat{\rho})\Big) \quad (15)$$

where $V(\hat{\rho})$ is the variance of the asymptotic distribution of $\hat{\rho}$. While asymptotic normality does not depend on underlying distributional assumptions, the form of $V(\hat{\rho})$ does. Assuming normality, $Var(\hat{\rho})$ is given by

$$V(\hat{\rho}) = \frac{2(n-1)(1-\rho)^2}{(n-a)(a-1)}\frac{\Big(1+\rho(b-1)\Big)^2}{b^2}. \quad (16)$$

This formula was first derived by Fisher (1925). Using a Taylor expansion approach, it is well known that if a function $g(.)$ is differentiable at ρ and $V(\hat{\rho})$ goes to zero as the sample size increases, then

$$\hat{\theta} \overset{asymp}{\sim} N(g(\rho), (g'(\rho))^2 V(\hat{\rho})). \quad (17)$$

In our case, $g(\rho) = 2/\pi sin^{-1}\sqrt{\rho}$ and the asymptotic distribution $\hat{\theta}$ can we expressed of terms of ρ as

$$\hat{\theta} \overset{asymp}{\sim} N\left(\frac{2}{\pi}sin^{-1}\sqrt{\rho}, \frac{2(n-1)}{\pi^2(n-a)(a-1)}\frac{(1-\rho)\Big(1+\rho(b-1)\Big)^2}{b^2\rho}\right) \quad (18)$$

and in terms of θ as

$$\hat{\theta} \overset{asymp}{\sim} N\left(\theta, \frac{2(n-1)}{\pi^2(n-a)(a-1)}\frac{\Big(1+sin^2(\frac{\pi}{2}\theta)(b-1)\Big)^2}{b^2tan^2(\frac{\pi}{2}\theta)}\right). \quad (19)$$

5. Nonparametric Inference

For maximum utility, inference about θ should be robust to choice of the parametric family. One way to achieve this is to use nonparametric methods. In this section it is

assumed that $A_i \overset{iid}{\sim} S(0, \sigma_1^2)$, $\epsilon_{ij} \overset{iid}{\sim} S(0, \sigma_2^2)$, and that A_i and ϵ_{ij} are mutually independent. Along with a nonparametric point estimator, we construct a nonparametric confidence interval for θ using a bootstrap sampling procedure for hierarchical data.

We begin by recognizing that if A and ϵ were directly observed, each of the a levels of A could be compared with each of the ab values of ϵ to form a U-statistic. Specifically,

$$\tilde{\theta} = \frac{1}{a^2 b} \sum_{k=1}^{a} \sum_{i=1}^{a} \sum_{j=1}^{b} I(|A_k| > |\epsilon_{ij}|)$$

where $I(.)$ is an indicator function that takes the value one if the condition is true and zero if not. Clearly, $\tilde{\theta}$ can never be calculated as it depends on the actual values of the random effects. Nevertheless, it is instructive to explore the properties of $\tilde{\theta}$.

THEOREM 1; *The quantity $\tilde{\theta}$ is consistent for θ and has an asymptotic normal distribution.*

Proof:

The quantity $\tilde{\theta}$ is a U-statistic. It can be written as

$$\tilde{\theta} = \frac{1}{a^2 b} \sum_{i,j,k} \phi(A_i, \epsilon_{jk}),$$

where ϕ is 1 if $|A| > |\epsilon|$, and 0 otherwise.

Clearly $\tilde{\theta}$ is unbiased for $E(\phi)$, which is $\theta = P(|A| > |\epsilon|)$. Using standard theory of U-statistics (see for example, Lehmann (1999, sec 6.1), one can show that

$$VAR(\tilde{\theta}) = \frac{(ab-1)\sigma_{10}^2 + (a-1)\sigma_{01}^2 + \sigma_{11}^2}{a^2 b}.$$

Where

$$\sigma_{11}^2 = var(\phi) = \theta(1-\theta),$$

$$\sigma_{10}^2 = cov(\phi_{ijk}, \phi_{ilm}) = \phi_1 - \theta^2,$$

$$\phi_1 = P(|A| > max(|\epsilon_1|, |\epsilon_2|)),$$

and

$$\sigma_{01}^2 = cov(\phi_{ijk}, \phi_{ljk}) = \phi_2 - \theta^2,$$

with

$$\phi_2 = P(|\epsilon| < min(|A_1|, |A_2|)).$$

As $a \to \infty$, it is easy to show that

$$Var(\sqrt{a}\tilde{\theta}) \to \sigma_{10}^2 + \sigma_{01}^2/b.$$

Or anotherwords, for large a, we have that $Var(\tilde{\theta})$ is approximately

$$\frac{\sigma_{10}^2 + \sigma_{01}^2/b}{a}.$$

If in addition $b \to \infty$, then the last term vanishes, and we have that $Var(\tilde{\theta})$ is about σ_{10}^2/a. Thus $\tilde{\theta}$ is consistent. To see that it has an asymptotic normal distribution we apply a standard result for U-statistics, which again is illustrated in Lehmann (1999). Define

$$\phi_{10}(a) = E_\epsilon(\phi(a, \epsilon)).$$

Then clearly $E(\phi_{10}(A)) = \theta$. In addition we can define

$$\phi_{01}(e) = E_A(\phi(A, e)).$$

Again, ϕ_{01} is unbiased for θ. Now it is possible to show that

$$var(\phi_{10}(A)) = \sigma_{10}^2,$$

and that

$$var(\phi_{01}(\epsilon)) = \sigma_{01}^2.$$

For completeness, we will illustrate the first of these. Note that

$$var(\phi_{10}(A)) = Var_A(E_\epsilon \phi(A, \epsilon)),$$

where the subscript notation means taking expectation with respect to the variable in question. This latter quantity is then

$$E_A E_\epsilon^2 \phi(A, \epsilon) - \theta^2.$$

Now for this to be equal to

$$\sigma_{10}^2 = cov(\phi(A, \epsilon_1), \phi(A, \epsilon_2)),$$

we need

$$E_A E_\epsilon^2 \phi(A, \epsilon) = E_{A, \epsilon_1, \epsilon_2}(\phi(A, \epsilon_1)\phi(A, \epsilon_2)).$$

This is true if

$$E_\epsilon^2 \phi(A, \epsilon) = E_{\epsilon_1, \epsilon_2}(\phi(A, \epsilon_1)\phi(A, \epsilon_2)).$$

That this is true follows from the fact that the right hand side of the latter can be factored as

$$E_{\epsilon_1}(\phi(A, \epsilon_1))E_{\epsilon_2}(\phi(A, \epsilon_2)).$$

By a similar argument, the other result follows. Now consider the quantity

$$T_a = \sqrt{a}(\tilde{\theta} - \theta).$$

We want to show that this quantity has a limiting normal distribution. We already know its variance. The limit here will be as $a \to \infty$. To prove this we require the lemma

stated in Lehmann (1999, page 378), that T_a and T_a^* have the same limiting distribution if $E(T_a^* - T_a)^2 \to 0$ as $a \to \infty$. Adapting standard results (see Lehmann (1999)), we construct our T_a^* to be

$$T_a^* = \sqrt{a}\{1/a(\sum_i \phi_{10}(A_i) - \theta) + 1/ab(\sum_{j,k} \phi_{01}(\epsilon_{jk}) - \theta)\}.$$

Clearly, $E(T^*) = 0$, and careful examination shows that the two sums in T^* are sums of independent quantities, and by the central limit theorem, have limiting normal distributions. In fact

$$var(T_a^*) \to \sigma_{10}^2 + 1/b\sigma_{01}^2.$$

This is the same limiting variance as T_a. Since they both have zero expectation, and the same limiting variance, the squared difference goes to zero if $cov(T_a, T_a^*)$ has the same limiting expression as $Var(T_a)$ also. In fact,

$$cov(T_a, T_a^*) = cov(\sqrt{a}\{1/a^2 b \sum \phi_{ijk}\}, \sqrt{a}\{1/a \sum \phi_{10}(A_i) + 1/ab \sum \phi_{01}(\epsilon_{jk})\}).$$

Which is

$$\frac{1}{a^2 b} cov(\sum_{ijk} \phi_{ijk}, \sum_i \phi_{10}(A_i)) + \frac{1}{a^2 b^2} cov(\sum_{ijk} \phi_{ijk}, \sum_{jk} \phi_{01}(\epsilon_{jk})).$$

Each of these sums can be simplified. For example, the first becomes

$$\sum_{ijk} cov(\phi_{ijk}, \phi_{10}(A_i)),$$

as all other terms have zero covariance. There are thus $a^2 b$ terms, each of which is

$$cov(\phi_{ijk}, \phi_{10}(A_i)).$$

These terms individually are

$$E_{A,\epsilon}(\phi(A, \epsilon)\phi_{10}(A)) - \theta^2$$

which is

$$E_A(E_\epsilon(\phi(A, \epsilon)\phi_{10}(A))) - \theta^2,$$

or

$$E_A \phi_{10}^2(A) - \theta^2 = \sigma_{10}^2,$$

from before. A similar result holds for the second set of covariances, and thus the covariance has the required form, which means that T_a has a limiting normal distrbution with the specified variance.

Of course, we cannot observe $\tilde{\theta}$, and instead must work with estimated values of A and ϵ. This introduces considerable complication to the theory. Using simple, or naive estimators such as

$$\widehat{A}_i^{naive} = \overline{Y}_{i.} - \overline{Y}_{..} \quad (20)$$

$$\widehat{\epsilon}_{ij}^{naive} = Y_{ij} - \overline{Y}_{i.} \quad (21)$$

we have

$$\widehat{A}_i^{naive} = A_i - \bar{A}_\cdot + \bar{\epsilon}_{i\cdot} - \bar{\epsilon}_{\cdot\cdot},$$

$$\widehat{\epsilon}_{ij}^{naive} = \epsilon_{ij} - \bar{\epsilon}_{i\cdot},$$

and

$$\widehat{\theta}^{naive} = \frac{1}{a^2 b} \sum_{k=1}^{a} \sum_{i=1}^{a} \sum_{j=1}^{b} I(|\widehat{A}_k^{naive}| > |\widehat{\epsilon}_{ij}^{naive}|).$$

The presence of the extra terms induces dependence among the $\widehat{A}^{naive}$'s, and the distribution of $\widehat{A}_i^{naive}$ depends on the distribution of ϵ. These dependences only disappear if we have $b \to \infty$ as well as $a \to \infty$ as outlined above. From a practical viewpoint this might seem problematic, as b is usually small. However, simulation results outlined later are encouraging for $b \geq 4$.

The estimators of A and ϵ actually employed depend on one's desire to match the properties of $\widehat{A}$ and $\widehat{\epsilon}$ to those of A and ϵ, respectively. While it is true that $\widehat{A}_i^{naive}$ and $\widehat{\epsilon}_{ij}^{naive}$ play the role of A_i and ϵ_{ij} in that $E(\widehat{A}_i^{naive}) = E(A_i)$ and $E(\widehat{\epsilon}_{ij}^{naive}) = E(\epsilon_{ij})$, the variances of $\widehat{A}_i^{naive}$ and $\widehat{\epsilon}_{ij}^{naive}$ do not equal the variances of A_i and ϵ_{ij}. In this manner comparing the magnitudes of $\widehat{A}_i^{naive}$ and $\widehat{\epsilon}_{ij}^{naive}$ in $\widehat{\theta}_{naive}$ may lead to erroneous results. It is readily known, however, that

$$Var\left(\sqrt{\frac{a}{a-1}\frac{b\sigma_1^2}{\sigma_2^2 + b\sigma_1^2}}(\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot})\right) = Var(A_i) \tag{22}$$

$$Var\left(\sqrt{\frac{b}{b-1}}(Y_{ij} - \overline{Y}_{i\cdot})\right) = Var(\epsilon_{ij}) \tag{23}$$

and under normal theory

$$E\left[\frac{a}{a-1}\left(1 - \frac{(a-3)Q_1}{a(b-1)Q_2}\right)\right] = \frac{a}{a-1}\frac{b\sigma_1^2}{\sigma_2^2 + b\sigma_1^2}. \tag{24}$$

Simulation results indicate that estimating $a/(a-1)b\sigma_1^2/(\sigma_2^2 + b\sigma_1^2)$ jointly using normal assumptions is better than estimating the individual variance components and then computing an estimate of the scalar. We consider

$$\widehat{A}_i^{scalar} = \sqrt{max\left\{0, \frac{a}{a-1}\left(1 - \frac{(a-3)Q_1}{a(b-1)Q_2}\right)\right\}}(\overline{Y}_{i\cdot} - \overline{Y}_{\cdot\cdot}) \tag{25}$$

$$\widehat{\epsilon}_{ij}^{scalar} = \sqrt{\frac{b}{b-1}}(Y_{ij} - \overline{Y}_{i\cdot}) \tag{26}$$

where $\widehat{A}_i^{scalar}$ is an estimator A and $\widehat{\epsilon}_{ij}^{scalar}$ is an estimator ϵ. The corresponding estimator of θ is denoted by $\widehat{\theta}_{scalar}$. Note that the computations of $E(\widehat{A}_i^{scalar})$ and $Var(\widehat{A}_i^{scalar})$

involve two correlated quantities (the argument of the square root and $\overline{Y}_{i.} - \overline{Y}_{..}$) as well as truncation and thus will not identically match $E(A_i)$ and $Var(A_i)$.

We also consider an estimator of A based on jackknife versions of Q_1 and Q_2. That is, determine $Q_{1(-i)}$ and $Q_{2(-i)}$, where $(-i)$ denotes that Q_1 and Q_2 are computed by excluding the observations in the i^{th} group. It follows that

$$Var\left(\sqrt{\frac{a}{a-1}\left(1-\frac{(a-4)Q_{1(-i)}}{(a-1)(b-1)Q_{2(-i)}}\right)}(\overline{Y}_{i.}-\overline{Y}_{..})\right) \;=\; \sigma_1^2 \qquad (27)$$

so an alternative estimator of A to consider is

$$\widehat{A}_i^{jackknife} \;=\; \sqrt{max\left\{0, \frac{a}{a-1}\left(1-\frac{(a-4)Q_{1(-i)}}{(a-1)(b-1)Q_{2(-i)}}\right)\right\}}(\overline{Y}_{i.}-\overline{Y}_{..}). \qquad (28)$$

Since the argument of the square root and $\overline{Y}_{i.}-\overline{Y}_{..}$ are uncorrelated, $E(\widehat{A}_i^{jackknife}) = E(A_i)$ and $Var(\widehat{A}_i^{jackknife})$ should be close to $Var(A_i)$. We use $\widehat{\theta}_{jackknife}$ to denote the estimator of θ based on $\widehat{A}_i^{jackknife}$ and $\widehat{\epsilon}_{ij}^{scalar}$.

We now evaluate the practical use of the estimators of θ. Note that while the A's are independent of one another and the ϵ's are independent of one another in $\tilde{\theta}$, such is not the case for the $\widehat{A}$'s and $\widehat{\epsilon}$'s in the various forms of $\widehat{\theta}$. It is also clear that estimating A and ϵ to compute $\widehat{\theta}$ will contribute to the variation in the resulting estimator. To accompany the point estimators of θ, we consider bootstrap confidence intervals of θ based on the various forms of $\widehat{\theta}$.

Efron and Tibshirani (1998) discuss nonparametric bootstrap estimation techniques when resampling is based on nonhierarchical data. The natural layering or nested feature of the data in the problems we consider present a complication when compared to the usual bootstrap resampling methods. Davison and Hinkley (1997, p.100-102) provide an outline of the resampling procedure for hierarchical data having two stages of sampling. The strategy recommended by Davison and Hinkley (1997) attempts to match the resampling variations of the statistics to the variational properties of the data. As mentioned by Davison and Hinkley (1997), we note that the resampling procedure works well when $a \geq 10$. Using model (1), the resampling strategy employed involves obtaining resampled values of $\widehat{A}_k$ and $\widehat{\epsilon}_{ij}$, denoted by $\widehat{A}_k^*$ and $\widehat{\epsilon}_{ij}^*$, respectively, as follows:

1. Compute $\overline{Y}_{K_i.}$, where $K_1, ..., K_a$ are randomly sampled with replacement from $\{1, ..., a\}$.
2. Compute $\overline{Y}_{K_i.} - \overline{Y}_{..}$, where $\overline{Y}_{..} = \sum_{K_i} \overline{Y}_{K_i.}/a$.
3. Compute $Y_{I_iJ_j} - \overline{Y}_{I_i.}$, where $I_1, ..., I_a$ are randomly sampled with replacement from $\{1, ..., a\}$ and $J_1, ..., J_b$ are randomly sampled with replacement from $\{1, ..., b\}$.
4. Compute $\widehat{A}_k^* = f(\overline{Y}_{K_i.} - \overline{Y}_{..})$ and $\widehat{\epsilon}_{ij}^* = g(Y_{I_iJ_j} - \overline{Y}_{I_i.})$, where the functions f and g depend on the naive, estimated scalar, or jackknife forms of the estimators (and thus Q_1^*, Q_2^*, $Q_{1(-i)}^*$, $Q_{2(-i)}^*$).

The bootstrap estimator of θ, denoted by θ^*, is computed using the appropriate forms of $\widehat{A}_k^*$ and $\widehat{\epsilon}_{ij}^*$. The algorithm listed above can be repeated to obtain θ_1^*, ..., θ_B^*, where B is the

number of bootstrap replications. Bootstrap confidence intervals for θ are then constructed from the empirical distribution of θ^*. The bootstrap confidence interval procedure we use are bias-corrected (BC) but are not accelerated. The performance of the bias-corrected and accelerated (BC_a) bootstrap confidence interval depends on the estimated acceleration constant. In nonparametric problems such as ours, an accurate estimator of the acceleration constant is not easily obtained and resulting BC_a intervals may perform poorly. See Shao and Tu (1995) for additional details.

A simulation study was conducted to evaluate the performance of the nonparametric BC bootstrap confidence intervals using the naive, scalar, and jackknife forms of the estimators. Performance was judged by the simulated coverage probability of 90% confidence intervals. For various combinations of a and b using $B = 2000$, 10000 nonparametric BC bootstrap intervals were built from normal, Laplace, and uniform distributed data for $\theta = 0.1$, 0.5, and 0.9. Simulation coverage probabilities (CP) and expected lengths (EL) of the intervals are displayed in Tables 2 and 3. When $\theta = 0.1$, the simulated coverage probabilities using the naive and scalar versions of the estimator fall far short of the nomimal 0.90 level. In general, the coverage probabilities associated with the jackknife estimator are more apt to be close to the nominal level. We recommend using $\widehat{\theta}_{jackknife}$ over the estimators considered for studies where $a \geq 10$, $b \geq 4$.

The usefulness of the bootstrap method to estimate θ is illustrated by a comparison of the confidence intervals associated with $\widehat{\theta}_{jackknife}$ to the confidence intervals based on the asymptotic properties of $\widehat{\theta}$. From (19), an approximate 90% confidence interval for θ (assuming normality) is

$$\widehat{\theta} \;\pm\; 1.645\sqrt{\frac{2(n-1)}{(n-a)(a-1)}}\frac{1+sin^2(\frac{\pi}{2}\widehat{\theta})(b-1)}{\pi b tan(\frac{\pi}{2}\widehat{\theta})}. \tag{29}$$

For small values of θ, $\widehat{\theta}$ may be small or zero and the expected length of confidence intervals using asymptotic results can be very large no matter what sample size is considered. Furthermore, for intermediate to large values of θ, (29) may not produce acceptable results. Consider scenarios where $a = 10, 20$ and $b = 4, 5, 10$ for Laplace distributed data. For $\theta = 0.9$, the simulated coverage probabilities using (29) range from 0.74 to 0.76; far short of the nominal 0.90 level. Using $\widehat{\theta}_{jackknife}$ and its accompanying BC bootstrap confidence interval provide reliable results for estimating θ based on intermediate sized samples having a variety distributions.

6. Examples

Gibbons and Bhaumik (2001) compared the results of copper concentrations as determined by seven laboratories. Water samples, prepared by an independent source, were analyzed by the laboratories in a blind interlaboratory study. For brevity reasons, we consider that part of the dataset consisting of five replications per laboratory based on a copper concentration of zero μg/L. The copper concentrations as measured by the laboratories are given in Table 4. Negative values are possible since the copper concentrations are based on a linear calibration function. Using a one-way random effects model, one

Table 2: $a = 10$. Comparison of different estimation methods. Nominal value 90 % Tabulated value of CP is a %, EL is expected length.

			Distribution					
			Normal		Laplace		Uniform	
b	θ	Estimator	CP	EL	CP	EL	CP	EL
4	0.1	Naive	49	0.36	22	0.38	76	0.34
		Scalar	69	0.31	68	0.33	71	0.29
		Jackknife	81	0.34	82	0.38	80	0.31
	0.5	Naive	94	0.44	95	0.45	93	0.45
		Scalar	91	0.55	92	0.57	89	0.54
		Jackknife	92	0.56	93	0.59	89	0.55
	0.9	Naive	88	0.24	85	0.33	89	0.35
		Scalar	87	0.35	83	0.34	88	0.36
		Jackknife	87	0.35	83	0.34	88	0.36
5	0.1	Naive	42	0.32	12	0.35	76	0.30
		Scalar	69	0.29	68	0.32	73	0.28
		Jackknife	80	0.32	82	0.36	81	0.29
	0.5	Naive	94	0.42	94	0.43	93	0.43
		Scalar	92	0.50	94	0.53	90	0.49
		Jackknife	93	0.51	94	0.55	90	0.51
	0.9	Naive	89	0.34	87	0.33	89	0.34
		Scalar	88	0.34	86	0.33	89	0.35
		Jackknife	88	0.34	86	0.33	89	0.35
10	0.1	Naive	43	0.23	06	0.26	83	0.21
		Scalar	73	0.24	63	0.25	82	0.22
		Jackknife	82	0.25	76	0.28	86	0.23
	0.5	Naive	92	0.37	93	0.39	92	0.38
		Scalar	92	0.40	93	0.43	91	0.40
		Jackknife	92	0.41	93	0.44	91	0.41
	0.9	Naive	88	0.32	87	0.31	89	0.33
		Scalar	88	0.32	87	0.31	89	0.34
		Jackknife	88	0.32	87	0.31	89	0.34

Table 3: $a = 20$. Comparison of estimation methods. Nominal value 90 %, tabulated values are a % for CP, and EP is expected length.

			Distribution					
			Normal		Laplace		Uniform	
b	θ	Estimator	CP	EL	CP	EL	CP	EL
4	0.1	Naive	10	0.24	1	0.26	36	0.22
		Scalar	69	0.22	62	0.22	73	0.21
		Jackknife	79	0.24	76	0.26	78	0.22
	0.5	Naive	93	0.30	94	0.31	92	0.31
		Scalar	84	0.39	87	0.43	81	0.37
		Jackknife	84	0.40	87	0.44	81	0.38
	0.9	Naive	86	0.21	80	0.21	87	0.21
		Scalar	85	0.22	77	0.22	88	0.22
		Jackknife	85	0.22	77	0.22	88	0.22
5	0.1	Naive	12	0.22	1	0.24	52	0.21
		Scalar	70	0.21	63	0.22	76	0.20
		Jackknife	79	0.23	76	0.25	81	0.21
	0.5	Naive	93	0.29	94	0.29	93	0.30
		Scalar	89	0.34	91	0.37	85	0.34
		Jackknife	89	0.34	91	0.38	85	0.35
	0.9	Naive	87	0.21	81	0.21	88	0.21
		Scalar	87	0.22	79	0.22	89	0.22
		Jackknife	87	0.22	79	0.22	89	0.22
10	0.1	Naive	13	0.16	1	0.18	73	0.14
		Scalar	77	0.18	62	0.18	85	0.17
		Jackknife	84	0.19	75	0.20	85	0.17
	0.5	Naive	92	0.26	93	0.27	92	0.27
		Scalar	92	0.27	94	0.29	90	0.28
		Jackknife	92	0.27	94	0.29	90	0.28
	0.9	Naive	87	0.21	83	0.21	88	0.21
		Scalar	87	0.21	82	0.21	88	0.21
		Jackknife	87	0.21	82	0.21	88	0.21

Table 4: Copper Concentrations (μg/L)

Lab 1	Lab 2	Lab 3	Lab 4	Lab 5	Lab 6	Lab 7
3.000	2.100	0.800	1.661	0.090	7.226	0.018
2.000	0.300	-0.185	1.996	-2.510	-1.000	-3.000
-1.000	2.000	0.990	0.000	7.270	0.000	0.000
1.000	1.300	0.905	2.993	7.140	10.244	-2.000
-1.000	2.000	0.365	2.042	0.280	-2.177	-2.000

can determine how the variability of the laboratories contributes to the overall variability of the measurements. If σ_1^2 represents the variance of the laboratories and σ_2^2 represents the variance of the measurements within laboratories (or error), the intraclass correlation coefficient is $\rho = \sigma_1^2/(\sigma_1^2 + \sigma_2^2)$. For this data, $Q_1 = 230.32$, $Q_2 = 60.08$, and $\widehat{\rho} = 0.04$. That is, the percentage of variance in copper concentrations associated with the variance in laboratories is 4%.

Recall that the probability of preponderancy (θ) and ρ measure different things. θ relates to individual measurements, indicating the proportion of individuals for which the laboratory effect is "more important" than the error effect. Under normal distribution assumptions, $\widehat{\theta} = 0.13$ and a 95% confidence interval for θ is (0.00, 0.51). The value of 0.13 should be interpreted as follows; "in 13% of determinations, the laboratory effect is greater than individual measurement error". In this application we find that the intraclass correlation coefficient understates the direct impact laboratories have on the copper concentration measurements. The nonparametric estimators are $\widehat{\theta}_{naive} = 0.43$, $\widehat{\theta}_{scalar} = 0.31$, and $\widehat{\theta}_{jackknife} = 0.24$.

Consider the example presented by Vangel (1992) in which tensile-strength measurements were made on five consecutive batches of composite material used to make aircraft components. Five measurements per batch were obtained for a total sample size of 25. The coded strength measurements for this application are displayed in Table 5. For these

Table 5: Tensile-strength Measurements

Batch 1	Batch 2	Batch 3	Batch 4	Batch 5
379	363	401	402	415
357	367	402	387	405
390	382	407	392	396
376	381	402	395	390
376	359	396	394	395

data, $Q_1 = 1578.40$, $Q_2 = 4163.36$, and $\widehat{\rho} = 0.71$. That is, the percentage of variability in tensile strength associated with the variability in batches is 71%. It may also be of

interest to determine how the batches directly influence the tensile strength measurements, a matter not addressed by the intraclass correlation coefficient. The probability the batch effect outweighs the individual measurement (or error) effect is estimated by $\widehat{\theta}$. Under normal distribution assumptions, $\widehat{\theta} = 0.64$ and a 95% confidence interval for θ is (0.41, 0.87). Again the .64 should be interpreted as "in 64 % of measurements, the batch effect is larger than the measurement error." (Is it measurement error, or variation within batches?). The nonparametric estimators are $\widehat{\theta}_{naive} = 0.84$, $\widehat{\theta}_{scalar} = 0.82$, and $\widehat{\theta}_{jackknife} = 0.82$. We do not recommend computing the nonparametric BC bootstrap confidence interval for either of these examples since $a < 10$.

7. Discussion

We have presented a new parameter, which we call the probability of preponderancy. In the one-way random effects model, this probability is the proportion of individuals for whom the group random effect is larger than the individual random effect. In genetic applications this could be interpreted as the proportion of individuals for whom the genetic effect dominates environmental effects. This parameter is an appealing complement to the more familiar intraclass correlation coefficient since its relevancy does not hinge on a particular distribution. We have presented both parametric and nonparametric estimators. Bias-corrected bootstrap confidence intervals associated with the nonparametric estimators may employed when $a \geq 10, b \geq 4$. The estimator incorporating a jackknife approximation to a scalar exhibits the most consistent coverage probabilities. Actual confidence levels depend on the underlying distributions.

It is possible to extend the ideas presented in this paper to models having more than two sources of variation. For example, suppose $Y = \mu + A + B + \epsilon$. Without worrying too much about indicies, or whether we have crossing or nesting, there are at least two possible extensions to θ. The first is

$$\theta_A \quad = \quad P(|A| > max\{|B|, |\epsilon|\}), \tag{30}$$

the second is

$$\lambda_A = P(|A| > |B + \epsilon|). \tag{31}$$

The first attempts to find for what proportion of individuals is A the greatest of the three effects. The second finds for what proportion of individuals is A greater than all other effects combined. Similar expressions could be formed for B. Under normal theory, with A having variance σ_1^2, B having variance σ_2^2, and ϵ with variance σ_3^2, then one can show that

$$\lambda_A = F_{1,1}\left(\frac{\sigma_1^2}{\sigma_2^2 + \sigma_3^2}\right). \tag{32}$$

The properties of such parameters and their estimators is a subject of current research by the authors of this paper.

Acknowledgements

The first author acknowledges Bill Schucany for helpful conversations. The second author would like to thank Robert Gibbons for providing additional information about the interlaboratory study.

References

Airy, G. B. (1861). *On the Algebraical and Numerical Theory of Errors of Observations and the Combinations of Observations.* London: MacMillan.

Bansal, N. K. & Bhandary, M. (1994). Robust M-estimation of the intraclass correlation coefficient. *Australian Journal of Statistics* **36**, 287–301.

Burch, B. D. & Harris, I. R. (2001). Closed-form approximations to the REML estimator of a variance ratio (or heritability) in a mixed linear model. *Biometrics* **57**, 1148–1156.

Bennett, J. H. (1983). *Natural selection, heredity, and eugenics: including selected correspondence of R.A. Fisher with Leonard Darwin and others.* Oxford: Clarendon Press.

Chauvenet, W. (1863a). *A Manual of Spherical and Practical Astronomy, 1: Spherical Astronomy.* Philadelphia: Lippincott, pp. 363–367.

Chauvenet, W. (1863b). *A Manual of Spherical and Practical Astronomy, 2: Theory and Use of Astronomical Instruments.* Philadelphia: Lippincott, pp. 194-197.

Commenges, D. & Jacqmin, H. (1994). The intraclass correlation coefficient: distribution-free definition and test. *Biometrics* **50**, 517–526.

Cox, D. R. & Hall, P. (2002). Estimation in a simple random effects model with non-normal distributions. *Biometrika* **89**, 831–840.

Davison, A. C. & Hinkley, D. V. (1997). *Bootstrap Methods and their Applications.* Cambridge: Cambridge University Press.

Donner, A. (1986). A review of inference procedures for the intraclass correlation coefficient in the one-way random effects model. *International Statistical Review* **54**, 67–82.

Efron, B. & Tibshirani, R. J. (1998). *An Intoduction to the Bootstrap.* Boca Raton, FL: CRC Press LLC.

Fisher, R. A. (1918). The correlation between relatives on the supposition of Mendelian inheritance. *Transactions of the Royal Society of Edinburgh* **52**, 399–433.

Fisher, R. A. (1925). *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd.

Gibbons, R. D. & Bhaumik, D. K. (2001). Weighted random-effects regression models with application to interlaboratory calibration. *Technometrics* **43**, 192–198.

Karlin, S., Cameron, E. C. & Williams, P. T. (1981). Sibling and parent-offspring correlation estimation with variable family size. *Proceedings of the National Academy of Sciences, U.S.A.* **78**, 2664–2668.

Lehmann, E. L. (1999). *Elements of Large-Sample Theory.* New York: Springer-Verlag.

Müller, R. & Büttner, P. (1994). A critical discussion of intraclass correlation coefficients. *Statistics in Medicine* **13**, 2465–2476.

Rothery, P. (1979). A nonparametric measure of intraclass correlation. *Biometrika* **66**, 629–639.

Searle, S. R., Casella, G. & McCulloch, C. E. (1992). *Variance Components.* New York: John Wiley.

Scheffé, H. (1956). Alternative models for the analysis of variance. *The Annals of Mathematical Statistics* **27**, 251–271.

Shao, J. & Tu, D. (1995). *The Jackknife and Bootstrap.* New York: Springer-Verlag.

Shirahata, S. (1982). Nonparametric measures of intraclass correlation. *Communications in Statistics - Theory and Methods* **11**, 1707–1721.

Vangel, M. G. (1992). New methods for one-sided tolerance limits for a one-way balanced random-effects ANOVA model. *Technometrics* **34**, 176–185.

Vogler, G. P., Wette, R., McGue, M. K. & Rao, D. C. (1995). Properties of alternative estimators of familial correlations under variable sibship size. *Biometrics* **51**, 276–283.

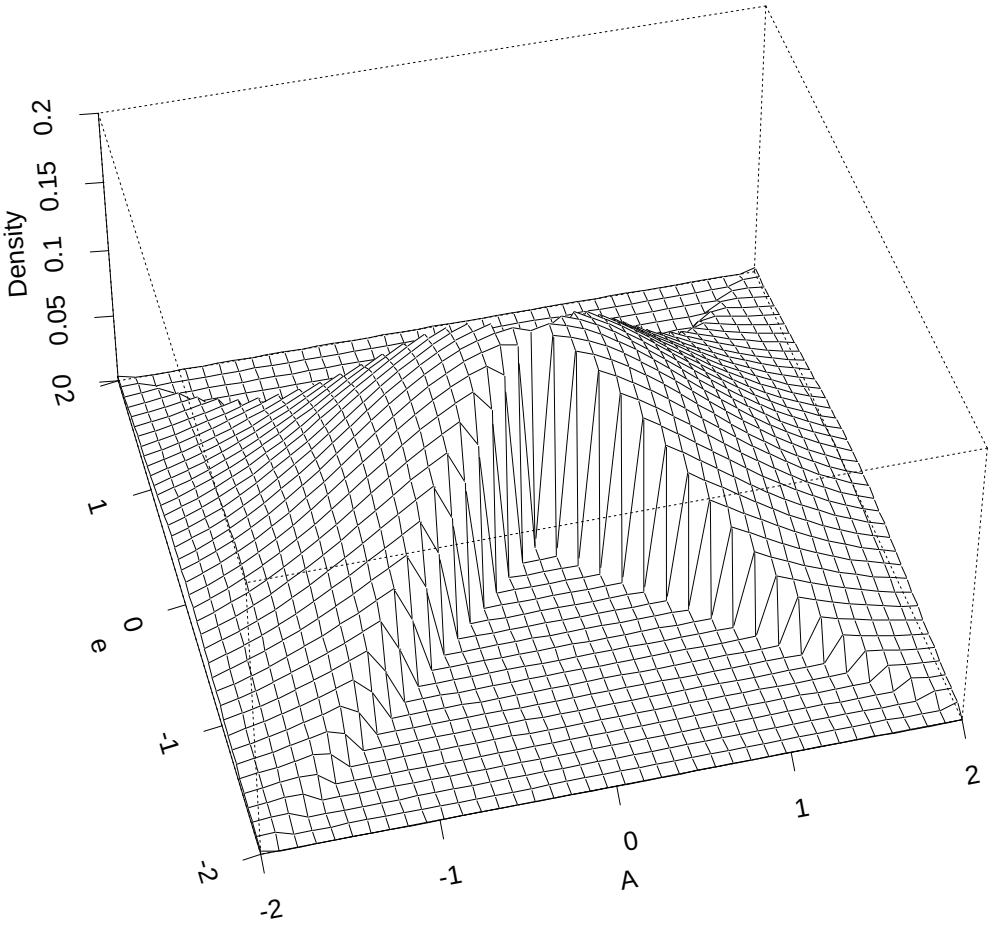

Figure 1: θ based on the bivariate normal distribution with $\sigma_1^2 = \sigma_2^2 = 1$

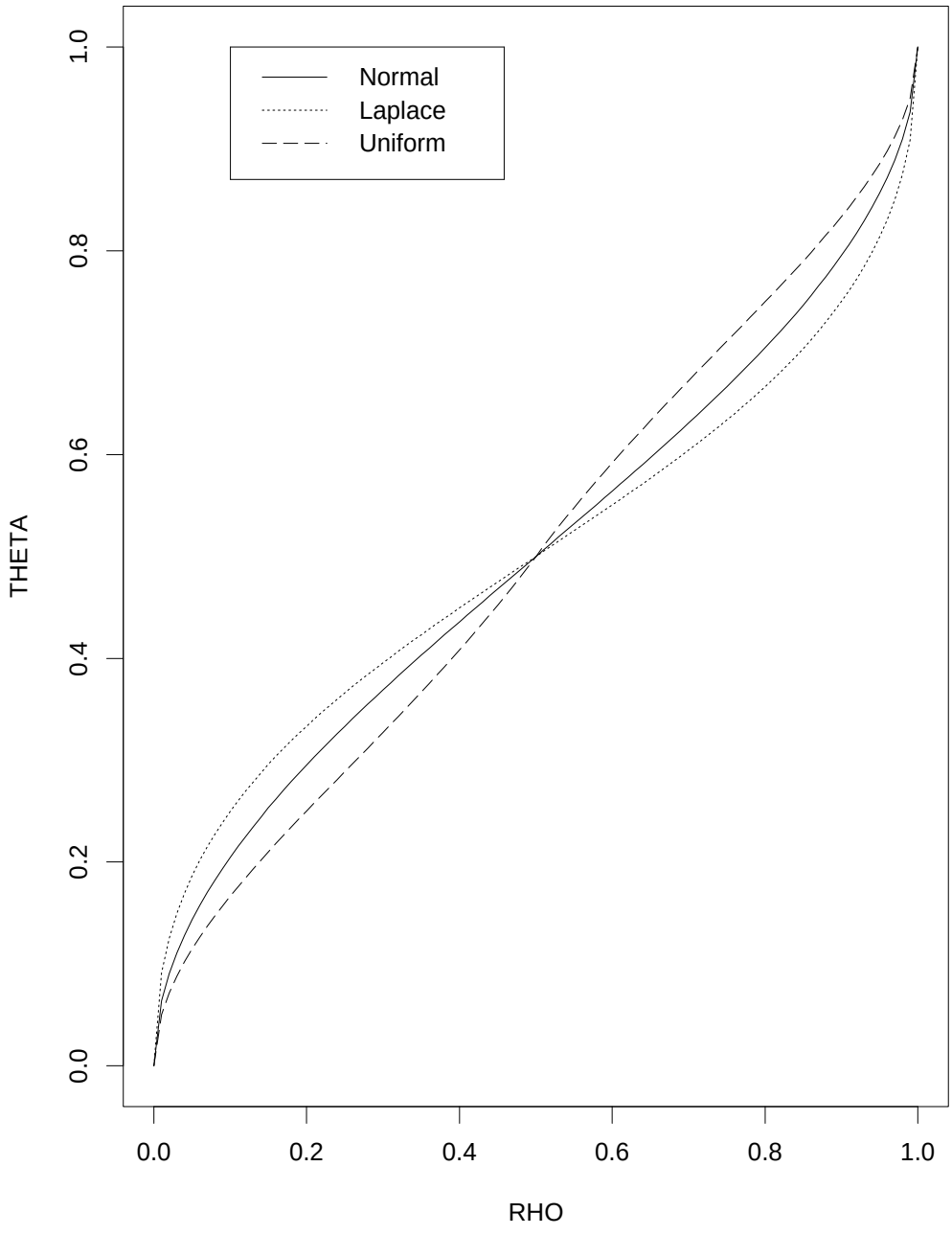

Figure 2: Relationship between ρ and θ