Department of Statistical Science

Southern Methodist University

Technical report SMU-TR-329

Dec 10th 2004

Measuring relative importance of sources

of variation without using variance.

Ian R. HARRIS

Department of Statistical Science, Southern Methodist University

Dallas, TX 75275

iharris@smu.edu

Brent D. BURCH

Department of Mathematics and Statistics, Northern Arizona University

Flagstaff, AZ 86011

Brent.Burch@nau.edu

Authors' footnote

Ian R. Harris is Associate Professor, Department of Statistical Science, Southern Methodist University, Dallas TX, 75214 (e-mail: iharris@smu.edu). Brent D. Burch is Associate Professor, Department of Mathematics and Statistics, Northern Arizona University, Flagstaff, AZ, 86011-5717 (e-mail: brent.burch@nau.edu).

Abstract

We propose a new parameter, the group dominance probability, to measure the relative importance of random effects in one-way random effects models. This parameter is the probability the group random effect is larger in absolute size than the individual (or error) random effect. The new parameter compares the middle-part of the distributions of the two sources of variation, and pays little attention to the tails of the distributions. This is in contrast to the traditional approach of comparing the variances of the random effects, which can be heavily influenced by the tails of the distributions. We suggest parametric and nonparametric estimators of the group dominance probability, and demonstrate the applicability of the ideas using data on blood pressure measurements.

key words: Bootstrap confidence interval; One-way random effects model; Group dominance probability; U-statistic; Variance component

1. INTRODUCTION

Using variance to measure variation is a long-standing tradition in statistics, particularly in the random and mixed effects model literature. According to Scheffé (1956) and Searle, Casella, and McCulloch (1992, page 23), the first explicit use of a one-way random effects model was made by Airy (1861, part IV). Fisher (1918) 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. In addition, Fisher (1918) defined variance as the square of the standard deviation, and used

$$\rho = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \tag{1}$$

to describe the proportion of total variance due to the first effect, where σ_1^2 and σ_2^2 are the variances of two independent sources of variation. Fisher (1918) specifically stated his belief that use of standard deviation (and hence variance) to measure variation followed from the assumption that the data are 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 distributions are not normal or are unknown, one might consider measuring the relative contribution of a source of variation without using variance. One can show that the parameter ρ is very sensitive to the behavior of the tails of the distributions. We introduce a parameter which compares the bulk or middle-part of the random effects' distributions. This new parameter, which we term the group dominance probability and denote by θ, is a useful alternative and complement to ρ in judging relative sizes of random effects.

Note that ρ is also the intraclass correlation coefficient, and was first described in this

fashion by Fisher in 1925 (see Fisher (1973)). For a thorough overview of the intraclass correlation coefficient in the one-way random effects model, see Donner (1986). More recently, Bansal and Bhandary (1994), Commenges and Jacqmin (1994), Müller and Büttner (1994), Vogler et. al. (1995), Giraudeau et. al. (1996), Gervini and Yohai (1998), and Lohr (1999) provide additional results about the intraclass correlation coefficient. We do not propose that θ be used as an alternative to ρ if the intent is to look at the correlation between members of the same group.

The group dominance probability is the probability that the random effect under study is larger in absolute size than the other random effects in the model. Consider a one-way random effects model, where the jth observation from group i is given by

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

$i = 1, ..., a$ and $j = 1, ..., b_i$. μ is the overall population mean, A_i the group random effect, and ϵ_{ij} is the individual (or error) random effect. We assume that the observations are measured on a ratio scale. Furthermore, the A_i's are a random sample from a distribution with mean 0 and variance σ_1^2, ϵ_{ij}'s are a random sample from a distribution with mean 0 and variance σ_2^2, and A_i and ϵ_{ij} are mutually independent, with $\sigma_1^2 \geq 0$ and $\sigma_2^2 > 0$. The group dominance probability is

$$\theta = P(|A_i| > |\epsilon_{ij}|). \tag{3}$$

The remainder of the paper is organized as follows. An application involving blood pressure measurements is given in Section 2 which illustrates potential problems in using variance to compare the random effects. Section 3 explores the basic properties of θ and considers its relationship to ρ. The example presented in this section shows that θ is not overly influenced by the tail behavior of the distributions being compared. Section 4 discusses

3

inference about θ under normality assumptions. In Section 5 we introduce nonparametric estimators of θ based on U-statistics. We recommend using bootstrap methods to construct confidence intervals based on nonparametric estimators, and briefly discuss this in the section. There are some concluding comments in Section 6.

2. EXAMPLE

Cox and Solomon (2003, pages 24-27) describe the analysis of blood pressure measurements taken from a large scale study of patients with hypertension. In the particular data they use, the focus is on 25 men whose systolic and diastolic blood pressures are measured twice at each of 16 occasions. For simplicity, we average the two measurements taken at each occasion to result in 16 measurements of systolic and 16 measurements of diastolic blood pressure for each of the 25 men. In this example, we consider the systolic and diastolic blood pressure measurements separately and ignore possible serial correlation or trends. A simple model for one set of blood pressure values, say the systolic values Y_{ij}, is the one-way random effects model (2) where A_i is the random effect associated with man i (the subject effect) having variance σ_1^2, and ϵ_{ij} is the random effect associated with man i's visit on the jth occasion (the occasion effect) having variance σ_2^2. A similar equation holds for the diastolic values. Thus each set of data (systolic and diastolic) are examples of the balanced one-way random effects model with $a = 25$ and $b = 16$. The original data values are available at the web site http://www.maths.adelaide.edu.au/people/psolomon.

Taking the analysis of systolic values first, we find that the ANOVA estimator of ρ is $\hat{\rho} = 0.52$. The estimate $\hat{\rho} = 0.52$ should be interpreted as "52% of the variance in systolic blood pressure measurements is due to variance in subject effect". Note the use of variance in this interpretation, instead of the phrase variation. For the diastolic measurements,

4

$\hat{\rho} = 0.34$. The lower value of $\hat{\rho}$ for the diastolic measurements would suggest that the subject effect is relatively less important for diastolic than systolic measurements.

One can compare the subject (A_i) and the day-to-day (ϵ_{ij}) effects directly by looking at simple estimators,

$$\widehat{A}_i = \overline{Y}_{i.} - \overline{Y}_{..} \tag{4}$$

$$\widehat{\epsilon}_{ij} = Y_{ij} - \overline{Y}_{i.} \tag{5}$$

where $\overline{Y}_{i.}$ is the ith subject sample mean and $\overline{Y}_{..}$ the overall sample mean. Using the ANOVA estimate of ρ is equivalent to comparing the variances of $\widehat{A}_i$ and $\widehat{\epsilon}_{ij}$, as the sample variance of the former is an unbiased estimator of $(\sigma_1^2 + \sigma_2^2/b)$, and the sample variance of the latter is an unbiased estimator of $a(b-1)/(ab-1)\sigma_2^2$. If $\sigma_1^2 = \sigma_2^2$, then one would expect the sample variances of the estimated effects to be very close, indeed for this design their ratio would be about 1.13. For the diastolic measurements the actual ratio is 0.61, and for the systolic measurements it is 1.21, reflecting the conclusion based on ρ that the subject effect is relatively less important for diastolic than systolic measurements. An alternative analysis shown below indicates that for most observations, the subject effect is of equal importance in diastolic and systolic measurements, and that it is only for relatively few observations that the previous conclusion holds.

Histograms of the $\widehat{A}_i$ and the $\widehat{\epsilon}_{ij}$ are shown in figure 1, for both systolic and diastolic measurements. The plots suggest that there may be some differences in tail behaviors of the effects, both between subject and day-to-day effects, and diastolic versus systolic measurements. Examination of interquartile ranges (IQR) of the effect estimates sheds some light. For the diastolic measurements, the subject IQR is 9.85, the day-to-day IQR is 8.26, for a ratio of 1.19. This suggests that for the middle 50% of observations, the variability in

subject effect is similar to the variability of day-to-day effect. This is a somewhat different message from the comparison of variance performed by ρ, in which the variance of the subject effect is noticeably less than the variance of the day-to-day effect. For the systolic measurements the IQR's are 15.58 for subject effect, and 15.26 for day-to-day effect, for a ratio of 1.02, again suggesting that for the middle 50% of observations, the variability of subject effect is comparable to the variability of day-to-day effect. Together the IQR values suggest that for "typical" observations, subject effect is equally important for diastolic and systolic measurements. The different conclusion reached by comparing sample variances, or looking at ρ, seems to be a result of differences in the tails of the subject and day-to-day distributions.

Focusing on only the middle 50% of data is restrictive so one might look at say, the 15th to 85th percentiles. Any such choice is however ad-hoc, and it would be preferable to use some procedure which incorporates all the data. One approach is to use robust estimators of ρ. Our paper deals with a completely different approach, altering the fundamental scale of comparison of the random effects by comparing them directly through the group dominance probability θ, rather than through variance. For the systolic values, the nonparametric jackknife estimator (section 5) of θ is $\hat{\theta}_{jackknife} = 0.52$. This value should be interpreted as "in 52% of systolic blood pressure measurements, the subject effect is of greater magnitude than the day-to-day effect." For the diastolic measurements, $\hat{\theta}_{jackknife} = 0.49$. The diastolic value is very consistent with the value for the systolic measurements, with essentially the same conclusion as obtained by looking at interquartile range. Ninety percent bootstrap confidence intervals (section 5)for θ using the systolic and diastolic blood pressure

measurements are $(0.38, 0.66)$ and $(0.39, 0.60)$, respectively.

3. THE GROUP DOMINANCE PROBABILITY

An example is used to illustrate the comparative behavior of ρ and θ in relation to the tail behavior of the distributions of the two random effects in model (2). Suppose that ϵ has the probability density function

$$
\begin{aligned}
f_\epsilon(t) &= 1/10(k-1), \quad -k < t < -1 \\
&= 4/10, \qquad\qquad -1 < t < 1 \\
&= 1/10(k-1), \quad 1 < t < k
\end{aligned}
\tag{6}
$$

and that A has the probability density function

$$
\begin{aligned}
f_A(t) &= 1/10(c-1), \quad -c < t < -1 \\
&= 4/10, \qquad\qquad -1 < t < 1 \\
&= 1/10(c-1), \quad 1 < t < c.
\end{aligned}
\tag{7}
$$

Note that these distributions both have 80% of the probability uniformly distributed between -1 and 1, with uniform tails whose lengths are controlled by the sizes of k and c. In this example $k, c > 1$. These distributions for ϵ and A have been chosen for mathematical simplicity, and to illustrate the effect of tail behavior on ρ and θ, with no intent to say that in an actual application the distributions of ϵ and A would follow these forms. The relative contributions of the tails can vary dramatically, and can be summarized by looking at the kurtosis compared to the normal distribution. If k (or c) is 1.25, then the distribution in question is uniform, with kurtosis -1.2, indicating very short tails. In general, the kurtosis

7

is

$$\frac{0.9(0.4 + \frac{k^5 - 1}{10(k-1)})}{(0.4 + \frac{k^3 - 1}{10(k-1)})^2} - 3, \tag{8}$$

for the distribution of ϵ, with the same expression for the distribution of A with c replacing k. As k (or c) goes to infinity, the kurtosis increases monotonically to 6, indicating a heavier and heavier tail effect.

One can show that the proportion of total variance due to the random effect A is

$$\rho = \frac{5 + c(c+1)}{10 + c(c+1) + k(k+1)}. \tag{9}$$

Thus for $k \approx c$, ρ is close to $1/2$, but by varying the relative sizes of c and k, and hence the tails, ρ may take any value in $(0, 1)$. Clearly ρ is very influenced by tail behavior. In contrast, the group dominance probability is

$$\theta = 0.5 + 0.02 \frac{c - k}{max(c, k) - 1}. \tag{10}$$

One can see that θ lies in the interval $(0.48, 0.52)$, regardless of the lengths of the tails. The group dominance probability reflects the fact that most of the time, both the random effects are uniformly distributed between $(-1, 1)$, and so are equally likely to be larger than each other.

It is possible to illustrate the same behavior of the two parameters in other situations using, say, tails of only 5% or 1% each. Any percentage in the tails, no matter how small, can have the same effect on ρ, if the tail is stretched sufficiently. In the robustness literature, this concept is essentially the breakdown point of the estimator, see Huber (1981). More specifically, ρ has a breakdown of zero, in that no percentage of the distribution can be arbitrarily large without affecting ρ dramatically, and causing ρ to become "arbitrarily bad". In robustness studies breakdown is typically applied to estimators, not parameters,

but the essence of the idea is the same. For θ the breakdown point is in fact 1, or 100%. Although large breakdowns (above 50%) are relatively meaningless, the high value shows that in real life problems, where contamination of 1% to 10% of the data is fairly common (Hampel et. al, page 28), θ will be very resistant to this contamination. Contamination here means addition of a few outliers or large values to the data.

Note that ρ and θ measure different things, and so in general one should not say one is better than another, but that which one to use will depend on the question of interest. The group dominance probability θ relates to individual observations, telling us for what proportion of the observations is the group random effect "more important" than the individual random effect. The proportion of variance ρ tells us what proportion of the variance of a response is due to the group effect. The previous discussion might also suggest that using θ is related to robustness and outliers. The real issue is one of tail behavior, although that is usually where outliers are found. If the intent is to compare the bulk of the distributions of the random effects, θ will be useful regardless of tail behavior. If one uses ρ, there is an understanding that the tails of the distributions are very influential.

Closed-form expressions relating θ to ρ can be obtained for specific distributions. Suppose that A_i has a normal distribution with mean 0 and variance σ_1^2, ϵ_{ij} is normal with mean 0 and variance σ_2^2, and A_i and ϵ_{ij} are mutually independent. It follows that

$$
\begin{aligned}
\theta &= P(A_i^2 > \epsilon_{ij}^2) \\
&= P\left(\frac{\epsilon_{ij}^2/\sigma_2^2}{A_i^2/\sigma_1^2} < \sigma_1^2/\sigma_2^2\right) \\
&= F_{1,1}(\sigma_1^2/\sigma_2^2)
\end{aligned}
\tag{11}
$$

where $F_{1,1}(.)$ is the cumulative distribution function of an F-distributed variate having numerator and denominator degrees of freedom equal to one. In this case the parameter

9

can be written as

$$\theta = F_{1,1}(\rho/(1-\rho))$$

$$= \frac{2}{\pi}sin^{-1}\sqrt{\rho} \tag{12}$$

and thus

$$\rho = sin^2(\frac{\pi}{2}\theta). \tag{13}$$

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

Figure 2 displays the values of θ as a function of ρ for normally distributed random effects. Also displayed are the relationships between ρ and θ using Laplace (double-exponential) and uniform distributional assumptions. The kurtosis corresponding to each of these distributions is 0, 3, and -1.2, respectively. These three cases are representative of distributions having normal, long, and short tails.

Because ρ has a long history and is widely used, conventions have been developed to define the meaning of small, medium, and large effects. For example, Zyzanski et. al. (2004) state that in medicine, $\rho = 0.05$ is small, $\rho = 0.10$ medium and $\rho = 0.15$ is large. Using (12) applied to these values, with a little rounding we can similarly define $\theta = 0.15, 0.20, 0.25$ as reasonable cutoffs for small, medium and large effects, respectively. It is important to note that these suggestions are very much context and area specific, and we recommend in general using (12) to convert standard conventions for ρ into similar conventions for θ.

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

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 chosen estimator of ρ is the restricted maximum likelihood estimator, which we denote by $\widehat{\rho}$. See Searle et. al. (1992, pages 90ff, 159ff, 249ff) for a general description of restricted maximum likelihood estimators of variance components. For alternatives to restricted maximum likelihood for estimating ρ, see Vogler et al. (1995). Due to the invariance of maximum likelihood estimators, the restricted maximum likelihood estimator of θ is

$$\widehat{\theta} = \frac{2}{\pi} sin^{-1}\sqrt{\widehat{\rho}}. \tag{14}$$

Since $\widehat{\theta}$ and $\widehat{\rho}$ are related by the simple one-to-one relationship (14), inferences for ρ in model (2) based on normality assumptions can easily be carried over to θ. For example, confidence intervals and hypothesis tests for θ are readily available. We illustrate this by forming a confidence interval for θ in a balanced one-way random effects model. Let Q_1 be the between group sum of squares and Q_2 be the within group or error sum of squares associated with model (2) having $b_i = b$ for all i. Then Q_1 and Q_2 are independently distributed and

$$\left(1 + btan^2(\frac{\pi}{2}\theta)\right)\frac{(a-1)Q_1}{a(b-1)Q_2} \sim F(a(b-1), a-1), \tag{15}$$

where b is the common group size. 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 given by

$$\left(\frac{2}{\pi}tan^{-1}\sqrt{\frac{1}{b}\left(F_{\alpha/2}\frac{a(b-1)Q_2}{(a-1)Q_1} - 1\right)}, \frac{2}{\pi}tan^{-1}\sqrt{\frac{1}{b}\left(F_{1-\alpha/2}\frac{a(b-1)Q_2}{(a-1)Q_1} - 1\right)}\right). \tag{16}$$

Note that this is an exact interval under normality assumptions.

5. NONPARAMETRIC INFERENCE

We have seen that θ is insensitive to tail behavior, hence it is desirable to estimate θ in a way that is also insensitive to tail behavior as well as normality assumptions. This suggests the use of nonparametric estimators. 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, one could calculate $\tilde{\theta}$,

$$\tilde{\theta} \;=\; \frac{1}{a^2 b}\sum_{k=1}^{a}\sum_{i=1}^{a}\sum_{j=1}^{b} I(|A_k| > |\epsilon_{ij}|) \tag{17}$$

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:

See the Southern Methodist University Department of Statistical Science technical report SMU-TR-307, by the present authors, available at http://www.smu.edu/statistics/TechReports/tech-rpts.htm, hereafter referred to as "the technical report".

Of course, we cannot observe $\tilde{\theta}$, and instead must work with estimated values of A and ϵ. Using the simple estimators of A and ϵ given in (4) and (5), we can form

$$\widehat{\theta}_{simple} \;=\; \frac{1}{a^2 b}\sum_{k=1}^{a}\sum_{i=1}^{a}\sum_{j=1}^{b} I(|\widehat{A}_k| > |\widehat{\epsilon}_{ij}|). \tag{18}$$

12

The large-sample properties of $\widehat{\theta}_{simple}$ are difficult to obtain, as there is dependence among the $\widehat{A}$'s, and the distribution of $\widehat{A}$ depends on the distribution of ϵ. These dependences disappear when $a, b \rightarrow \infty$, and under these conditions we expect the asymptotic behavior of $\widehat{\theta}_{simple}$ to be comparable to that of $\tilde{\theta}$. Of course, the usefulness of an asymptotic result depends on how "quickly" the result applies to finite sample sizes. Simulation results outlined later are encouraging for $a \geq 10$ and $b \geq 4$ using a modified estimator of θ.

While it is true that $E(\widehat{A}_i) = E(A_i)$ and $E(\widehat{\epsilon}_{ij}) = E(\epsilon_{ij})$, the variances of $\widehat{A}_i$ and $\widehat{\epsilon}_{ij}$ do not equal the variances of A_i and ϵ_{ij}, respectively. Thus comparing the magnitudes of $\widehat{A}_i$ and $\widehat{\epsilon}_{ij}$ in $\widehat{\theta}_{simple}$ may lead to erroneous results. To remedy this problem we 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 ith group. It follows that

$$var\left(\sqrt{\left(\frac{a}{a-1}\left(1 - \frac{(a-4)Q_{1(-i)}}{(a-1)(b-1)Q_{2(-i)}}\right)\right)}(\overline{Y}_{i.} - \overline{Y}_{..})\right) = \sigma_1^2 \tag{19}$$

so an alternative estimator of A is

$$\widehat{A}_i^* = \sqrt{\left(max\left\{0, \frac{a}{a-1}\left(1 - \frac{(a-4)Q_{1(-i)}}{(a-1)(b-1)Q_{2(-i)}}\right)\right\}\right)}(\overline{Y}_{i.} - \overline{Y}_{..}). \tag{20}$$

Since the argument of the square root and $\overline{Y}_{i.} - \overline{Y}_{..}$ are uncorrelated, $E(\widehat{A}_i^*) = E(A_i)$ and $var(\widehat{A}_i^*)$ should be close to $var(A_i)$. The jackknife estimator of θ is

$$\widehat{\theta}_{jackknife} = \frac{1}{a^2 b}\sum_{k=1}^{a}\sum_{i=1}^{a}\sum_{j=1}^{b} I(|\widehat{A}_k^*| > |\widehat{\epsilon}_{ij}^*|) \tag{21}$$

where $\widehat{\epsilon}_{ij}^* = \sqrt{b}/\sqrt{(b-1)}(Y_{ij} - \overline{Y}_{i.})$. Note that $var(\widehat{\epsilon}_{ij}^*) = var(\epsilon) = \sigma_2^2$. Of course one could use other estimators of the random effects, such as empirical Bayes estimates, as outlined in say Raudenbush and Bryk (2002). The properties of such estimators based on these estimates are a potential subject of future research.

13

We now evaluate the practical use of the estimators of θ. To accompany the point estimators of θ, we consider bootstrap confidence intervals of θ based on $\widehat{\theta}_{simple}$ and $\widehat{\theta}_{jackknife}$. Davison and Hinkley (1997, p.100-102) provide an outline of the resampling procedure for hierarchical data having two stages of sampling. As mentioned by Davison and Hinkley (1997), the resampling procedure works well when $a \geq 10$. The bootstrap confidence interval procedures we use are bias-corrected but are not accelerated since in nonparametric problems such as ours, an accurate estimator of the acceleration constant is not easily obtained and the resulting intervals may perform poorly. See Shao and Tu (1995, p.168 and p.176) for additional details on acceleration. See the technical report for more details on the bootstrap procedure.

A simulation study was conducted to evaluate the performance of the nonparametric bias-corrected (BC) bootstrap confidence intervals using the simple 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 $B = 2000$ bootstrap replications, 10000 nonparametric BC bootstrap intervals were built from normal, Laplace, and uniform distributed data for $\theta = 0.1$, 0.5, and 0.9. For brevity we present only the results for $a = 10$, the minimum number of groups suggested by Davison and Hinkley (1997), using $b = 4$ and 10. See the technical report for additional combinations of a and b. Simulation coverage probabilities (CP) and expected lengths (EL) of the intervals are displayed in Table 1. To investigate how the normal-based estimator performs when the random effects are not normally distributed, we also include results for the confidence interval given by (16).

Table 1 about here

When $\theta = 0.1$, the simulated coverage probabilities using the simple version of the es-

14

timator fall far short of the nominal 0.90 level. When $\theta = 0.5$ or 0.9, the normal-based approach produces intervals that are too short and thus have inadequate coverage probabilities assuming Laplace distributed random effects. Overall, the coverage probabilities associated with the jackknife estimator are more apt to be close to the nominal level. If the distributions of the random effects are unknown, we recommend using $\widehat{\theta}_{jackknife}$ to estimate θ for studies where $a \geq 10$, $b \geq 4$.

6. DISCUSSION

We have presented a new parameter, which we call the group dominance probability, that measures the relative importance of sources of variation without using variance. In the one-way random effects model, this probability is the proportion of individual observations for which the group random effect is larger than the observation random effect. The group dominance probability is relatively insensitive to tail behavior of the random effects' distributions. Essentially, the group dominance probability compares the bulk of the two distributions.

We have discussed both parametric and nonparametric estimators of the new parameter. Bias-corrected bootstrap confidence intervals associated with the nonparametric estimators may be employed when the number of groups is at least ten, and the number of individuals per group is at least four. The estimator incorporating a jackknife approximation to a scalar exhibits the most consistent coverage probabilities. Actual confidence levels depend on the underlying distributions.

The main results presented in this paper assume a one-way random effects models. It is possible, however, to extend the concepts to mixed effects models with multiple random effects. It is also possible to consider models with multiple fixed effects. Parametric estimators

under normality theory can easily be obtained by transformation of maximum likelihood estimators, but nonparametric estimators may be harder to find, as one needs good estimates of the random effects. The properties of such parameters and their estimators is beyond the scope of this paper, and is a subject of future research.

ACKNOWLEDGMENTS

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.

Commenges, D. & Jacqmin, H. (1994), The Intraclass Correlation Coefficient: Distribution-Free Definition and Test. *Biometrics*, 50, 517–526.

Cox, D. R. & Solomon, P. J. (2003), *Components of Variance*, Chapman and Hall.

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.

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. (1973), *Statistical Methods for Research Workers*, 14th Edition. Edinburgh: Oliver and Boyd.

Gervini, D. and Yohai, V. J. (1998), Robust estimation of variance components, *The Canadian Journal of Statistics*, 26, 419–430.

Giradeau, B., Mallet, A., and Chastang, C. (1996). Case influence on the intraclass correlation coefficient estimate. *Biometrics*, 52, 1492–1497.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. and Stahel, W. A. (1986), *Robust Statistics: The Approach Based on Influence Functions*, New York: Wiley.

Huber, P. J. (1981), *Robust Statistics*, New York: Wiley.

Lohr, S. L. (1999), *Sampling: Design and Analysis*, New York: Duxbury Press.

Müller, R. & Büttner, P. (1994), A Critical Discussion of Intraclass Correlation Coefficients, *Statistics in Medicine*, 13, 2465–2476.

Raudenbush, S. W. and Bryk, A. S. (2002), *Hierarchical Linear Models, Applications and Data Analysis Methods, 2nd Ed*, Thousand Oaks: Sage.

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.

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.

Zyzanski, S. J., Flocke, S. A. & Dickinson, L. M. (2004), On the Nature and Analysis of Clustered Data, *Annals of Family Medicine*, 2, 199–200.
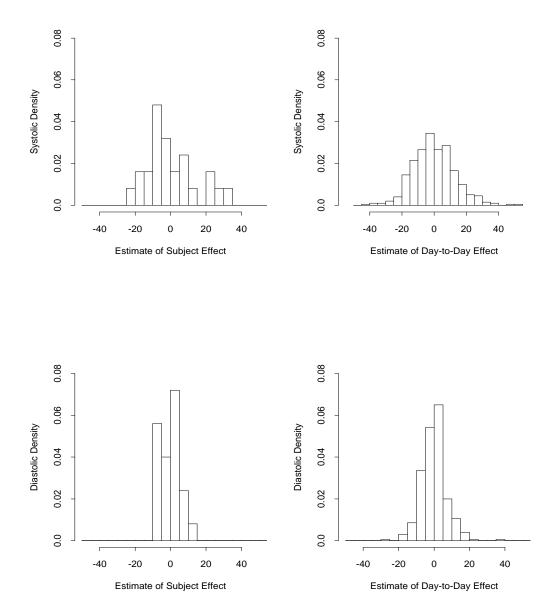
Figure 1: Estimates of Subject and Day-to-Day Effects for Systolic and Diastolic Blood Pressures. For systolic blood pressures, the IQR's as well as the sample variances are about the same for the estimated subject and day-to-day effects. For diastolic blood pressures, the IQR's are about the same but the sample variances for the estimated subject and day-to-day effects are quite different.
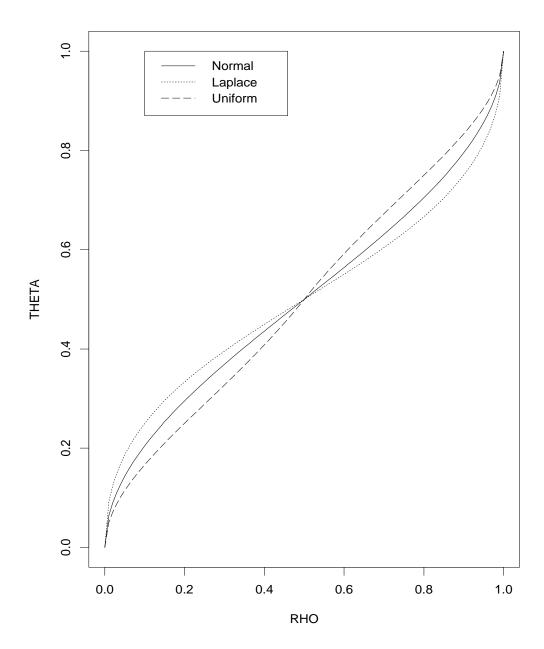
Figure 2: Relationship between ρ and θ. The parameters ρ and θ have approximately a linear relationship for the short-tailed uniform distribution. As the length of the tails of the distributions increase, the relationship between ρ and θ is increasingly nonlinear.

Table 1: $a = 10$. Comparison of the nonparametric estimation methods. Nominal coverage value is 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	Simple	49	0.36	22	0.38	76	0.34
		Jackknife	81	0.34	82	0.38	80	0.31
		Normal	90	0.36	91	0.35	90	0.37
	0.5	Simple	94	0.44	95	0.45	93	0.45
		Jackknife	92	0.56	93	0.59	89	0.55
		Normal	90	0.37	81	0.37	95	0.37
	0.9	Simple	88	0.24	85	0.33	89	0.35
		Jackknife	87	0.35	83	0.34	88	0.36
		Normal	90	0.10	59	0.07	91	0.11
10	0.1	Simple	43	0.23	06	0.26	83	0.21
		Jackknife	82	0.25	76	0.28	86	0.23
		Normal	90	0.26	90	0.25	89	0.28
	0.5	Simple	92	0.37	93	0.39	92	0.38
		Jackknife	92	0.41	93	0.44	91	0.41
		Normal	90	0.29	78	0.28	97	0.29
	0.9	Simple	88	0.32	87	0.31	89	0.33
		Jackknife	88	0.32	87	0.31	89	0.34
		Normal	90	0.09	56	0.07	90	0.10