JOURNAL OF FORENSIC SCIENCES

J Forensic Sci, November 2018, Vol. 63, No. 6
doi: 10.1111/1556-4029.13748
Available online at: onlinelibrary.wiley.com

TECHNICAL NOTE

ANTHROPOLOGY

Amanda B. Lee,[1] *B.A.; and Lyle W. Konigsberg*,[1] *Ph.D.*

Univariate and Linear Composite Asymmetry Statistics for the "Pair-Matching" of Bone Antimeres

ABSTRACT: This paper examines the distributional properties of univariate and linear composite measures of long bone asymmetry. The goal of this paper is to examine models that best fit the distribution of asymmetries with implications for the improvement of forensic pair-matching techniques. We use the software R to model reference data ($N = 2343$) and test data ($N = 71$) as normal distributions, an exponential power distribution, and a skew exponential power distribution—the latter two include the normal as a special case. Our results indicate that the data best fit the latter two distributions because the data are nonnormal. We also show how asymmetry statistics that use absolute values of side differences can be fit as folded distributions. This obviates the need for empirical distributions or for transformations that attempt to convert nonnormal distributions to normal distributions. The results of this study lay the framework for improving pair-matching methods that use comparative reference data.

KEYWORDS: forensic science, bone asymmetry, osteometric sorting, pair-matching, commingling, folded distributions

The matching of antimeres (in this case left and right femora) from commingled remains has been suggested as a basis for estimating the most likely number of individuals in an archeological or forensic assemblage (1–5). In contrast to the minimum number of individuals estimate, which only sets a lower bound, the most likely number of individuals gives a more realistic assessment of the scale for a mass disaster in an "open population" (6) setting. This matching of antimeres may be performed by visual comparisons of left with right bones, but the development of algorithms that use one or more measurements is generally viewed as desirable from the standpoint of method standardization. To make matches on the basis of measurements, it is necessary to first characterize the directional and fluctuating asymmetry for paired elements. Previous methods that rely on multiple measurements have generally used a linear composite, such as the sum of differences between left and right side measurements (7–10). After finding that this linear composite did not follow a normal distribution, Vickers et al. (10) suggested using the sum of absolute differences between sides. However, they did not note that such a measure would follow a "folded" distribution. Similarly, Thomas et al. (11) produced tables for "pair-matching" on single measurements that could have been obviated as the statistic they used should follow a folded normal distribution if the signed differences were normally distributed.

[1]Department of Anthropology, University of Illinois, 109 Davenport Hall, 607 South Matthews Avenue, Urbana, IL 61801.

Received 1 Sept. 2017; and in revised form 28 Dec. 2017; accepted 10 Jan. 2018.

Osteometric pair-matching algorithms like those by Thomas et al. (11) and Byrd and LeGarde (9) that rely on *t*-tests, necessarily rest on assumptions of asymptotic normality. Although technically these methods rely on *t*-distributions, generally the reference sample sizes are large enough that the normal distribution should be a good approximation. Demonstrating the lack of normality invalidates these methods and points to a gap in the field that needs to be addressed. In correctly characterizing the distributional models, it is possible to estimate the levels of asymmetry that should be present in a case, thereby increasing the accuracy of pair-matching methods.

The first part of this paper examines univariate measures of asymmetry, or in other words asymmetry based on one measurement from each side. This presentation runs counter to Thomas et al.'s (11) argument for using empirical distributions. As Vickers et al. have called into question the distributional assumption underlying the linear composite approach, the second part of this paper examines the distribution of this composite on a large reference sample. This is performed by estimating an exponential power distribution (12) for the linear composite. The "absolute value of sum difference" suggested by Vickers and co-workers (10:103) can then be shown to follow a folded exponential power distribution. For the "summed absolute value of differences" as described by Lynch and co-workers (13:2), it was necessary to fit a skew exponential power distribution (14). Rather than only assessing a single quantile, which in Vickers et al.'s case was at the 0.10 probability value, this paper uses complete quantile-quantile plots to check the fit of the composite score to theoretical distributions. More importantly, this paper also uses an independent test sample and quantile-quantile plots to compare the distribution of univariate and composite scores in the test sample to distributions derived from the reference sample.

Materials and Methods

The Samples and Data

The reference sample used in this study is taken from two large databases. The first is the online Goldman data set (15,16) available at http://web.utk.edu/~auerbach/GOLD.htm from which the maximum femoral length, mid-shaft mediolateral diameter, and mid-shaft anteroposterior diameter measured on the right and left femora were extracted. Complete data were available for 1326 individuals from this collection. The second source was the forensic database from which complete data on these six measurements were available for 1017 individuals. The total reference sample size was consequently 2343 individuals. The test data consists of the same measurements as taken on the reference sample for 59 individuals in the collections of the American Museum of Natural History and twelve individuals from the Office of the Chief Medical Examiner, New York City for a total of 71 individuals.

Distribution of a Univariate Statistic: M *Statistic*

Thomas et al. (11) have defined what they refer to as an M statistic, which is:

$$M = |R - L|/((L + R)/2), \tag{1}$$

where L and R represent the same measurement taken on left and right bones. Thomas et al. wrote M using the absolute value of L-R instead of R-L used here. While the absolute values are the same, there will be an occasion to use the signed difference later in this paper. Asymmetry is usually calculated using right minus left measurements (17). Thomas et al. noted the similarity of M to a statistic used by Van Valen (18). In fact, Thomas et al.'s M statistic is identical to what Palmer and Strobeck (17) referred to as "Index 2" of fluctuating asymmetry (their Table 1), a commonly used measure in studies of fluctuating asymmetry. Palmer (19:339) notes that measures based on the absolute value "will be very biased if either DA or antisymmetry is present" (emphasis in the original). "DA" refers to directional asymmetry, the tendency for the measurement to be larger on a given side. As there is known directional asymmetry for long bones (15), it is difficult to argue in favor of using Thomas et al.'s M statistic. With that said, the absence of a statistical hypothesis testing framework for M is also problematic. Thomas et al. provided a table of the empirical 90th, 95th, and 100th percentile values from their study, but did not give summary statistics that might be used to calculate probability values for future observed pairings of antimeric bones. If the signed version of Thomas et al.'s M statistic is normally distributed, then the absolute value converts the distribution to that of a folded normal (20) which makes statistical hypothesis testing readily available. If the folded normal distribution provides an adequate fit, then reference sample data can be compared to the fitted distribution from the test sample using a quantile-quantile plot. Oldford (21) describes "self-calibrating quantile-quantile plots" that can be fit using the R package "qqtest." These plots simulate from the hypothetical (fitted) distribution to produce confidence envelopes around the empirical quantile-quantile plot for the test data.

If the signed version is not normally distributed, then one can consider more general distributions, of which the exponential power distribution is probably sufficient. Writing x for the signed version of Thomas et al.' M statistic (i.e., without the absolute value), the density function for the exponential power distribution is:

$$f_{\mathrm{EP}}(x|\mu, \sigma, \alpha) = \frac{1}{2\alpha^{(1/\alpha)}\,\Gamma(1+1/\alpha)\sigma}\exp\left[\frac{-|x-\mu|^{\alpha}}{\alpha\sigma^{\alpha}}\right], \tag{2}$$

where μ is a location parameter (mean),σ is a scale parameter (the standard deviation when $\alpha = 2$), and α is a shape parameter (14). Equation 2 is identical with Equation 4 in Mineo (12) and with equation 4.1 in Azzalini and Capitanio (22) assuming $\mu = 0$ and $\sigma = 1$. When $\alpha = 2$ the denominator in the first term is equal to $\sqrt{2\pi}\sigma$ and the distribution is normal. For α equal 1.0 the distribution is a Laplace (double exponential) while between 1.0 and 2.0 the distribution is similar to a normal but with greater weight in the tails. As α increases above 2.0, the distribution approaches a uniform distribution. The density in Eq. 2 can be fit to data using the function "estimatep" in the R package normalp. If this density gives an adequate fit, then it can be converted to the folded distribution using the "folded" scripts in the R package "gendist." Again, self-calibrating quantile-quantile plots can be used with the reference and the test data.

Distribution of the Linear Composite: Byrd's D *Statistic*

The linear composite is as defined in Byrd and co-worker's publications (7–9), which is:

$$D = \sum_{j=1}^{p}\left(R_j - L_j\right), \tag{3}$$

where there are p measurements on both the right and left sides within an individual. To make the notation clearer, particularly with reference to the possibility of taking absolute values, we rewrite Eq. 3 in the equivalent form:

$$D = \sum_{j=1}^{p} R_j - \sum_{j=1}^{p} L_j. \tag{4}$$

Unfortunately, Byrd (7) in his Table 10.2 and Byrd and Legarde (9) in their Table 8.2 either reversed the subtraction to left measurements minus right measurements, or they reversed the labeling of columns for "Left" and "Right." The sum of the measurements in the column labeled "Right" is 643 while the sum from the column labeled "Left" is 698. The value of D should consequently be −55, but in both the 2008 and the 2014 Tables the listed value is 55. This led Vickers et al. (10:103) to suggest that in both Tables what was actually used was the "*absolute value* of sum difference," or:

$$D = \left|\sum_{j=1}^{p} R_j - \sum_{j=1}^{p} L_j\right|. \tag{5}$$

Vickers et al. then noted a poor fit, yet did not mention that the D value as defined in Eq. 5 should follow a folded form. We look at the distributional form for the D statistic in Eqs 4 and 5 much as we did for the simple univariate statistic. Note that the "absolute value of sum difference" as defined by Vickers et al. is generally not the same value as the sum of the absolute values of the individual differences. More formally, we have:

$$\left|\sum_{j=1}^{p} R_j - \sum_{j=1}^{p} L_j\right| \leq \sum_{j=1}^{p} |R_j - L_j| \tag{6}$$

Lynch et al. (13:2) refer to the value on the right-hand side as the "summed absolute value of differences" and suggest that this statistic should have a half-normal distribution. The half normal is a folded normal distribution where the "folding" at zero coincides with a mean of zero. In point of fact, if *R-L* for a given measurement has a normal distribution with a mean of zero, then $|R - L|$ will have a half-normal distribution, the sum of two such variables will have a skewed normal distribution, and as the number of absolute values in the sum increases the distribution will approach a normal distribution under the central limit theorem.

In order to fit the "summed absolute value of differences," we had to use a skew exponential power distribution. The density function for this is (14):

$$f_{\mathrm{SEP}}(x|\mu,\sigma,\lambda,\alpha) = (2/\alpha)\,\Phi(w) f_{\mathrm{EP}}((x-\mu)/\sigma|\mu=0,\sigma=1,\alpha),$$
$$w = \mathrm{sgn}(x-\mu)\frac{|x-\mu|^{\alpha/2}}{\sigma^{\alpha/2}}\lambda\sqrt{\frac{2}{\alpha}}, \tag{7}$$

where $\Phi(w)$ is the standard normal integral up to w while f_{EP} is the density function from Eq. 2.

Results

Univariate Statistics

Figures 1–3 show the folded normal distributions for Thomas et al.'s statistic for maximum femoral length, mid-shaft medial-lateral diameter, and mid-shaft anterior-posterior diameter. The Figures are drawn as "self-calibrating quantile-quantile plots" where the sample quantiles are from the 71 test individuals and the hypothetical quantiles are from the 2343 reference individuals. These Figures also show the mean right minus left (signed) measurements divided by the individual averages, the signed version of Thomas et al.'s *M* statistic, and the standard deviations of the signed *M* in the upper right corner. The variable names are shown in the bottom right corners. It is clear that while the femoral length and the medial-lateral diameter fit folded normal distributions for Thomas et al.'s *M* statistic, such is not the case for the anterior-posterior diameter. Figure 4 shows the folded exponential power distribution, which does fit. Note that the fitted exponential power of 1.149 is considerably less than 2.0, the power for a normal distribution. Indeed, the fitted distribution is closer to a Laplace distribution (with a power of 1.0) than it is to a normal distribution.

Linear Composite

Figures 5 and 6 show the self-calibrating quantile-quantile plots for Byrd's *D* statistic for the composite scores. Figure 5 shows that the data do not fit a normal distribution, however, they do fit the exponential power distribution as shown in Fig. 6. As the signed differences of sums follow an exponential power distribution, then the absolute values of the differences of sums should follow a folded distribution. This is shown in Fig. 7, which compares the parametric model (a folded exponential power distribution) to the boundary kernel density plot for the reference data. Here the boundary is at zero and the kernel density was drawn using the R library bde with Vitale's (23) method. Figure 8 is a self-calibrating quantile-quantile plot that shows that the absolute value of the

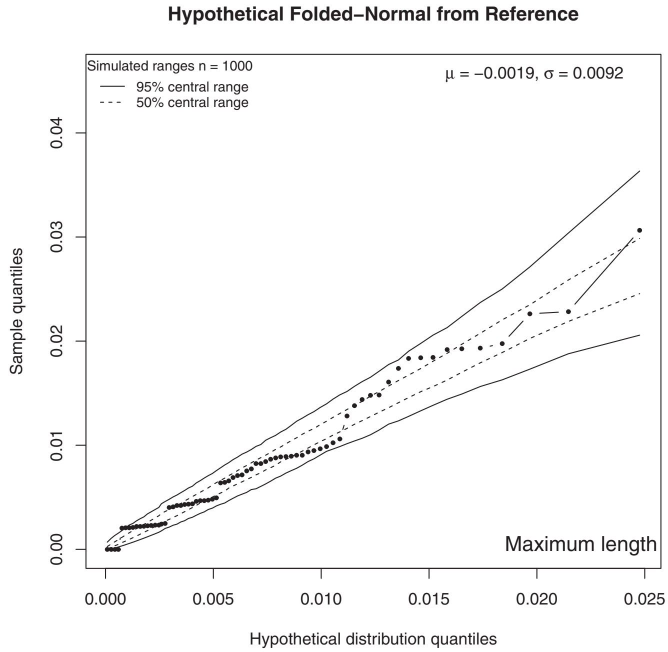

FIG. 1—*Self-calibrating quantile-quantile plot (21) using a hypothetical folded normal from the reference sample for maximum femoral length and the 71 test cases. The signed parameters from the reference sample are shown in the upper right corner. The hypothetical and sample quantiles are from Thomas and co-workers (11)* M *statistic shown as Eq. 1.*

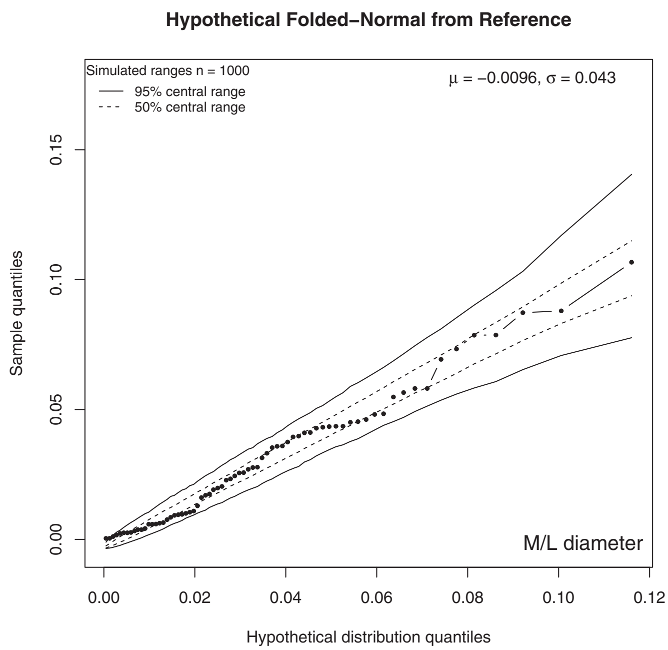

FIG. 2—*As in Fig. 1 but for the medial-lateral mid-shaft diameter from the femur.*

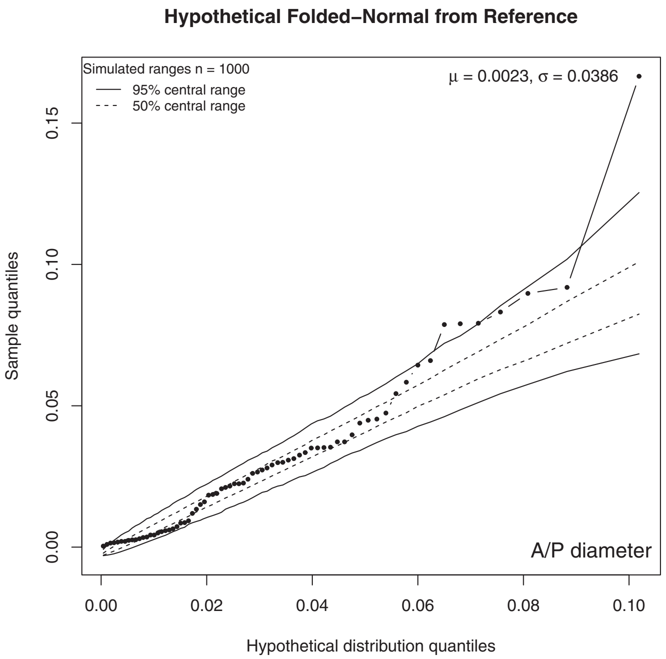

FIG. 3—*As in Fig. 1 but for the anterior-posterior mid-shaft diameter from the femur. Note that the 65th–67th and 71st sorted points for the test sample fall outside of the 95% range.*

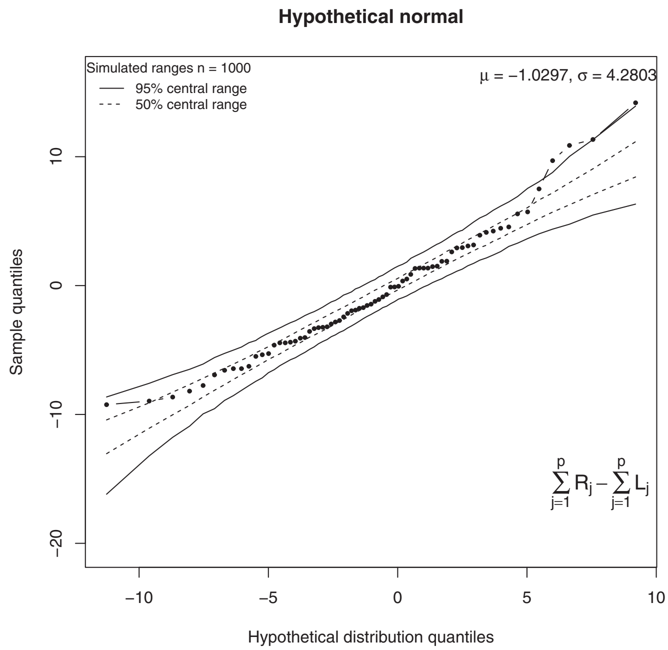

FIG. 5—*Self-calibrating quantile-quantile plot using the hypothetical normal of the difference of sums for the three measurements from the reference sample. Note that the 68th and 69th ordered points from the test sample fall outside of the 95% central range.*

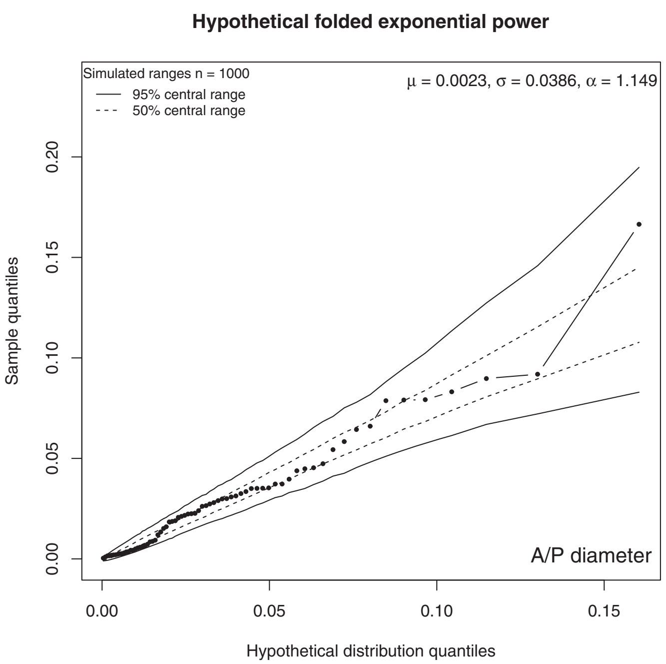

FIG. 4—*As in Fig. 3 but using a hypothetical folded exponential power distribution which brings all of the test points within the 95% central range. The three parameters in the upper right corner are the mean, scale, and exponential power from the signed distribution.*

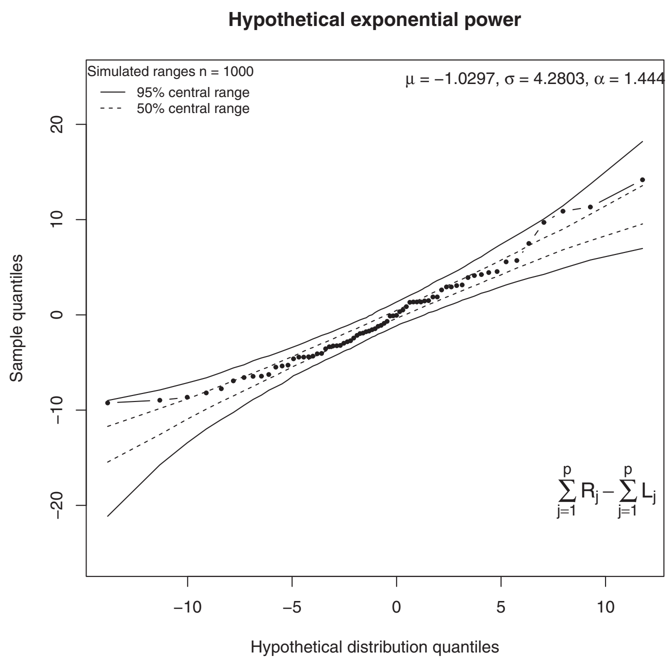

FIG. 6—*As in Fig. 5 but using an exponential power distribution from the reference and test data.*

difference of sums for the test data does indeed follow a folded exponential power distribution.

Figure 9 shows a comparison of the parametric model (a skew exponential power distribution) for the test data using the sum of absolute values of differences to the boundary kernel density plot. Again, the boundary is at zero, but this time we used an ordinary kernel density estimator with a Gaussian kernel and the default bandwidth. It was possible to use an ordinary kernel density estimator without a boundary because the density at the boundary is so low. Figure 10 shows the self-calibrating quantile-quantile plot for the test data against the hypothetical skew exponential power distribution from the reference data. This shows, as expected, that the sum of absolute differences should follow some form of skewed normal distribution, and not the half normal claimed in Lynch et al. (13).

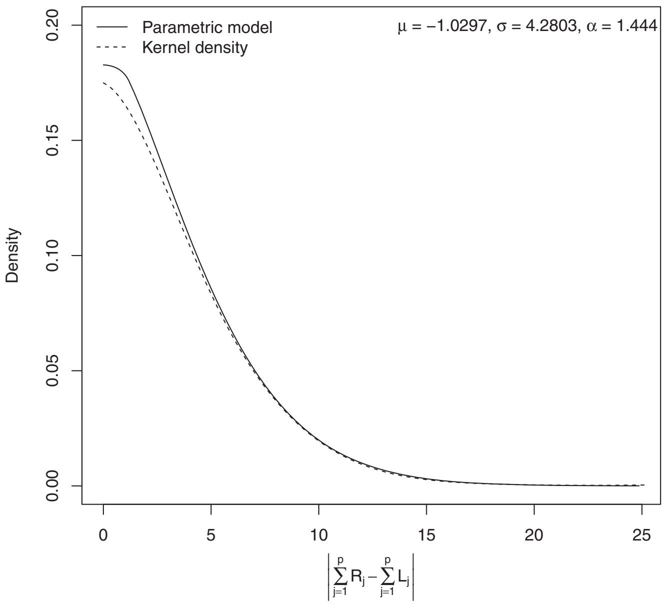

FIG. 7—*Comparison of the folded exponential power distribution from Fig. 6 (parametric model) to the kernel density fit for the absolute values of difference of sums in the reference samples. Note that a folded distribution is appropriate.*

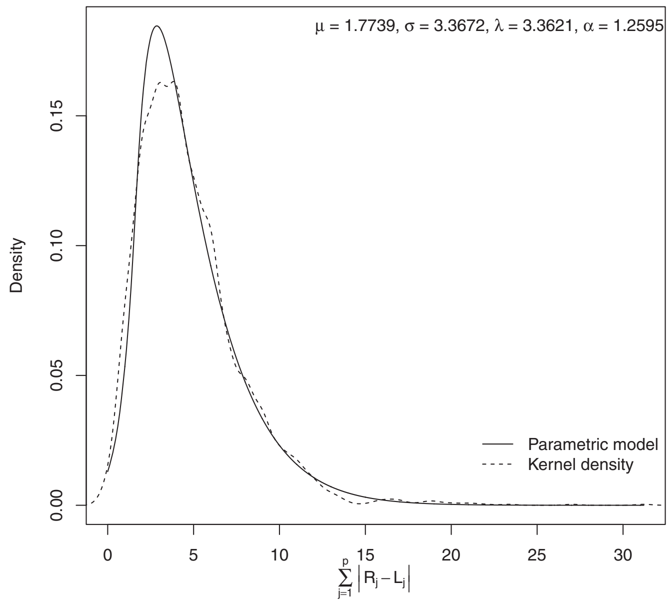

FIG. 9—*Comparison of a skew exponential power distribution (14) for the sum of absolute values of differences in the reference sample (parametric model) to the kernel density fit. Note that the parametric model and kernel density fit are comparable. Parameters in the upper right corner are location, scale, and shape parameters related to skewness and kurtosis.*

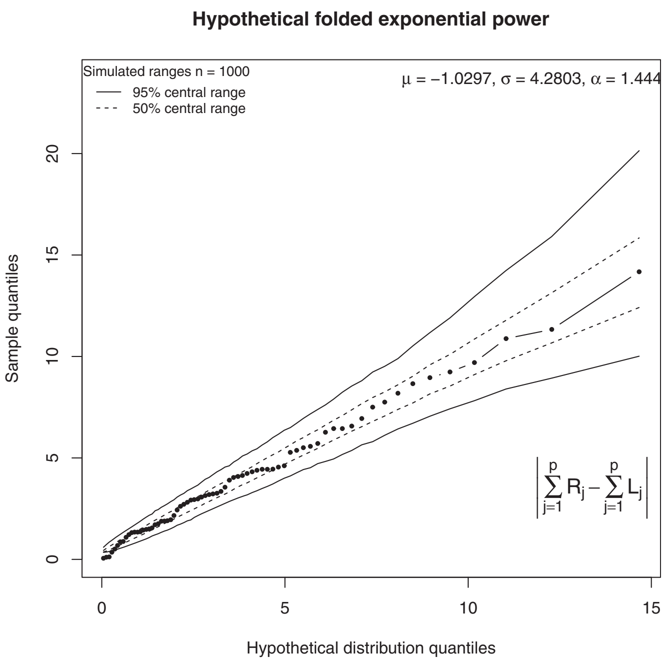

FIG. 8—*Self-calibrating quantile-quantile plot for the hypothetical folded exponential power distribution (from the reference sample) for the absolute value of the difference of sums.*

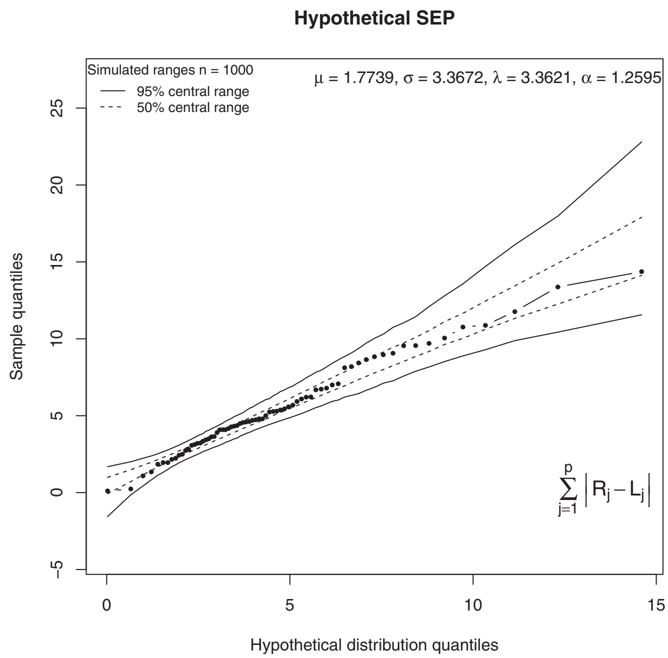

FIG. 10—*Self-calibrating quantile-quantile plot for the hypothetical skew exponential power distribution for the sum of the absolute values of differences.*

Discussion

The presence of asymmetry in long bones introduces complications when estimating the number of individuals in a commingled assemblage. Previous studies have found significant levels of long bone directional asymmetry in the human population, suggesting that pair-matching methods predicated on an assumption of zero asymmetry would prove problematic. Precise distributional models are important in the creation of automated pair-matching methods, a concept that has been a recent subject of interest in the literature (13,24–26) and ostensibly represents the future in forensic analysis of commingled assemblages.

While we are not the first to suggest that the distribution of side differences violates the normality assumption, this study investigates the best fit model for the true distribution. The results show that an exponential power distribution for the signed composite score and the signed univariate anterior-posterior diameter data and best fit these nonnormal data. The folded distributions are an accurate representation of the unsigned (absolute value) differences and lend themselves well to the exponential power models put forth in this paper. The exponential powers used in the models were closer to 1.0 than to 2.0 ("normal"), which makes sense considering that the model needed to be adjusted to fit a heavy-tailed, kurtotic distribution. These results not only confirmed that the data violated assumptions of normality but also characterized the nonnormality.

Our work also points out some methodological issues with recent analyzes. Lynch (25) and Lynch and co-workers (13) have suggested that the "summed absolute value of differences" should follow a half-normal distribution, in other words, a folded normal with a mean of zero on the signed scale. We have argued instead that the expected distributional form is a skewed normal if the number of paired variables is small and a (symmetric) normal distribution if the number of paired variables is large. Rather than referencing their statistics to a half or a folded normal distribution, these researchers have applied a Box-Cox (27) transformation for normality. Additionally, rather than estimating the parameters for the Box-Cox transformation, these researchers used values of $\lambda_1 = 0.33$ and $\lambda_2 = 5 \times 10^{-5}$. Swaddle and co-workers (28:987) wrote for univariate (one paired variable) that "values of λ_1 around 0.3 and λ_2 set to be somewhat smaller than the smallest nonzero asymmetry work well," while Graham and co-workers (29:474), citing Swaddle and co-workers, wrote that "One can normalize $|d_i|$ with a power transform: $(|d_i| + 0.00005)^{0.33}$." While in the past the Box-Cox transformation was a useful tool to transform nonnormal data to normal data, we believe that it is more useful now to find the distributional form of potentially nonnormal data. Further, if the Box-Cox transformation is used it is more appropriate when it is estimated on the relevant data rather than based on independent values from the literature.

While our work has focused on the statistical approach to univariate and composite measures of asymmetry, we believe that a more useful approach to asymmetry will be the true multivariate characterization (5,29,30). For those who choose not to take a multivariate approach, this study represents the first step toward characterizing univariate and composite measures of asymmetry as used in the forensic pair-matching method. Our results better model the distributional properties of asymmetry, particularly when absolute values are used, and define the relationship between the sample and reference distributions. These results lay the framework for future applications of these models in improving and automating current standard practices. Further work also needs to be focused on the sequential testing setting, particularly when the number of right and left bones may not be equal because of the loss of elements.

References

1. Adams BJ, Konigsberg LW. Estimation of the most likely number of individuals from commingled human skeletal remains. Am J Phys Anthropol 2004;125(2):138–51.
2. Adams BJ, Konigsberg LW. How many people? Determining the number of individuals represented by commingled human remains. In: Adams BJ, Byrd JE, editors. Recovery, analysis, and identification of commingled human remains. Totowa, NJ: Springer, 2008;241–55.
3. Nikita E. Estimation of the original number of individuals using multiple skeletal elements. Int J Osteoarchaeol 2014;24(5):660–4.
4. Nikita E, Lahr MM. Simple algorithms for the estimation of the initial number of individuals in commingled skeletal remains. Am J Phys Anthropol 2011;146(4):629–36.
5. Konigsberg LW, Adams BJ. Estimating the number of individuals represented by commingled human remains: a critical evaluation of methods. In: Adams BJ, Byrd JE, editors. Commingled human remains: methods in recovery, analysis, and identification. San Diego, CA: Elsevier, 2014;193–220.
6. Christensen AM, Passalacqua NV, Bartelink EJ. Forensic anthropology: current methods and practice. San Diego, CA: Elsevier, 2014.
7. Byrd JE. Models and methods for osteometric sorting. In: Adams BJ, Byrd JE, editors. Recovery, analysis, and identification of commingled human remains. Totowa, NJ: Springer, 2008;199–220.
8. Byrd JE, Adams BJ. Osteometric sorting of commingled human remains. J Forensic Sci 2003;48(4):717–24.
9. Byrd JE, LeGarde CB. Osteometric sorting. In: Adams BJ, Byrd JE, editors. Commingled human remains: methods in recovery, analysis, and identification. San Diego, CA: Elsevier, 2014;167–91.
10. Vickers S, Lubinski PM, Henebry DeLeon L, Bowen JT. Proposed method for predicting pair matching of skeletal elements allows too many false rejections. J Forensic Sci 2015;60(1):102–6.
11. Thomas RM, Ubelaker DH, Byrd JE. Tables for the metric evaluation of pair-matching of human skeletal elements. J Forensic Sci 2013;58 (4):952–6.
12. Mineo AM. On the estimation of the structure parameter of a normal distribution of order p. Statistica 2007;63(1):109–22.
13. Lynch JJ, Byrd J, LeGarde CB. The power of exclusion using automated osteometric sorting: pair-matching. J Forensic Sci 2017. https://doi.org/10.1111/1556-4029.13560. Epub: 2017 May 26.
14. DiCiccio TJ, Monti AC. Inferential aspects of the skew exponential power distribution. J Am Stat Assoc 2004;99(466):439–50.
15. Auerbach BM, Ruff CB. Limb bone bilateral asymmetry: variability and commonality among modern humans. J Hum Evol 2006;50(2):203–18.
16. Auerbach BM, Ruff CB. Human body mass estimation: a comparison of "morphometric" and "mechanical" methods. Am J Phys Anthropol 2004;125(4):331–42.
17. Palmer AR, Strobeck C. Fluctuating asymmetry: measurement, analysis, patterns. Annu Rev Ecol Syst 1986;17:391–421.
18. Van Valen L. A study of fluctuating asymmetry. Evolution 1962;16 (2):125–42.
19. Palmer AR. Fluctuating asymmetry analyses: a primer. In: Markow TA, editor. Developmental instability: its origins and evolutionary implications. Totowa, NJ: Springer, 1994;335–64.
20. Leone F, Nelson L, Nottingham R. The folded normal distribution. Technometrics 1961;3(4):543–50.
21. Oldford RW. Self-calibrating quantile–quantile plots. Am Stat 2016;70 (1):74–90.
22. Azzalini A, Capitanio A. The skew-normal and related families. Institute of Mathematical Statistics Monographs. Cambridge, U.K.: Cambridge University Press, 2014.
23. Vitale RA. A Bernstein polynomial approach to density function estimation. In: Puri ML, editor. Statistical inference and related topics. New York, NY: Academic Press, 1975;87–99.
24. Lynch JJ. An analysis on the choice of alpha level in the osteometric pair-matching of the os coxa, scapula, and clavicle. J Forensic Sci 2017. https://doi.org/10.1111/1556-4029.13599. Epub: 2017 Jul 18.
25. Lynch JJ. The automation of regression modeling in osteometric sorting: an ordination approach. J Forensic Sci 2017. https://doi.org/10.1111/1556-4029.13597. Epub: 2017 July 21.
26. Box GEP, Cox DR. An analysis of transformations. J R Stat Soc Series B Stat Methodol 1964;26:211–43.
27. Swaddle JP, Witter MS, Cuthill IC. The analysis of fluctuating asymmetry. Animal Behav 1994;48(4):986–9.
28. Graham JH, Raz S, Hel-Or H, Nevo E. Fluctuating asymmetry: methods, theory, and applications. Symmetry 2010;2(2):466–540.
29. Livshits G, Smouse PE. Multivariate fluctuating asymmetry in Israeli adults. Hum Biol 1993;65(4):547–78.
30. O'Brien M, Storlie CB. An alternative bilateral refitting model for zooarchaeological assemblages. J Taphonomy 2011;9(4):245–68.

Additional information and reprint requests:
Amanda B. Lee, B.A.
Department of Anthropology
University of Illinois
109 Davenport Hall
607 South Matthews Avenue
Urbana
IL 61801
E-mail: alee103@illinois.edu