

# On the Asymmetry of Biological Frequency Distributions

Nicholas J. Schork, Alan B. Weder, and M. Anthony Schork

*Division of Hypertension, Department of Medicine (N.J.S., A.B.W.) and Departments of Statistics (N.J.S.) and Biostatistics (M.A.S.), University of Michigan, Ann Arbor*

The long-standing problem of determining whether the skewness in a sample frequency distribution is the manifestation of the intermixing of disparate groups characterizable by a normal mixture distribution or the manifestation of non-mixture, skew-producing determinants is discussed. Biometrical tools for modeling and quantifying the significance of the skewness in a trait of interest that invite interpretations other than those formed in mixtures or “subgroups” are elaborated. Statistical methods for testing whether a normal mixture distribution better characterizes a set of data than the proposed (or any other) skewed, single-population-oriented models are offered. The power of these tests is examined through Monte Carlo experimentation. A brief application in hypertension research demonstrates some of the problems and methods discussed in the paper.

**Key words:** normal mixture distributions, commingling analysis, bootstrap tests, skewness, blood pressure

## I. INTRODUCTION

The use of normal mixture distributions in quantitative genetics has a history that spans more than 100 years, dating back to the time of Karl Pearson [1893]. Despite this long history, one crucial problem with the application of normal mixture distribution arguments in genetics remains: determining whether the asymmetry in the potentially mixed distribution of the trait under scrutiny—the asymmetry which often initially suggested the possibility of a mixture—can actually be characterized by an asymmetrical distributional model that does *not* invite the mixture interpretation. In the words of Karl Pearson [1895], the problem concerns

Received for publication May 7, 1990; revision accepted August 2, 1990.

Address reprint requests to Nicholas Schork, Division of Hypertension, Department of Medicine, University of Michigan, R6592 Kresge I, Ann Arbor, MI 48109-0500.

. . . how [we are] to discriminate between a true curve of skew type and a compound [i.e., mixture] curve, supposing we have no reason to suspect our statistics a priori of mixture. I have at present been unable to find any general condition among the moments, which would be impossible for a skew curve and possible for a compound, and so indicate compoundness. I do not, however, despair of one being found [Pearson, 1895, p. 394].

Though criteria for distinguishing distributions of a “skewed” type from normal mixtures based on moments have not been developed (see Murphy [1979] for a related discussion), general methods designed for this purpose have been proposed [MacLean et al., 1976; Ott, 1979b; and Schork and Schork, 1988]. Unfortunately, these methods have been misunderstood, improperly implemented, inadequate to the task, or, most commonly, ignored altogether. Many modern applications of mixture distributions arguments compare mixed normal hypotheses *only* against single normal (or lesser number of component normal mixture distribution) hypotheses and not against *skewed* (i.e., *non-normal*) *single-population-oriented* alternatives, despite the argument presented by Schork and Schork [1988] and, less directly, by Preston [1953], that, given a certain level of skewing in a data set, one will almost always conclude that a mixture distribution fits the data if a single normal distribution is taken as the alternative against which it is compared.

The purpose of this paper is to elaborate both models of skewed data *and* tests for determining the significance of normal mixtures when given skewed, single-population alternatives. As such our work can be viewed as an extension of Pearson’s system of curves (i.e., types I–V) which he developed as homogenous, single-population-oriented alternatives to the mixture distribution (i.e., subgroup) explanation of skewed data [Pearson, 1895]. The format of this paper is as follows. Section II will describe mathematical models for skewed data that are intuitive and heuristically appealing biologically. Section III will describe tests one can use to discriminate between normal mixture and skew alternative hypotheses, and offers results of Monte Carlo studies investigating the power of these tests and aspects of the robustness of some of the modeling devices (e.g., covariate adjustment) that go into them. Section IV provides a brief application of the methods described in sections II and III by investigating the distribution of human blood pressure, since issues relating to mixture distributions and blood pressure generated a very heated debate about the pathophysiology of hypertension in the 1950s and 1960s—the famous “Pickering/Platt” debate [Swales, 1985]—whose impact remains with medical and genetics researchers today. Part V offers a discussion and some brief summary remarks.

## II. MODELING ASYMMETRICAL DISTRIBUTIONS

There is no question that the mere fitting of distributions to data in an effort to determine a “best-fitting” model is both dangerous and potentially counterproductive. The motivation behind model fitting practices in the biological and medical sciences typically is to provide interpretable models. In this light, normal mixture distributions provide an excellent model for the intermixing of homogenous subpopulations, as described in the following.

Writing  $\phi(x)$  as the normal density function:

$$\phi(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \exp[-(x - \mu)^2/2\sigma^2], \quad (1)$$

we consider the existence of  $G$  different subpopulations or subtypes occurring in the proportions  $p_g$ ;  $g = 1, \dots, G$ , such that  $\sum_{g=1}^G p_g = 1$ , each characterized by a normal distribution with a unique mean,  $\mu_g$ , and variance,  $\sigma_g^2$ . The probability density at an observed trait value,  $x$ , taken from a sample comprised of these various subgroups is then:

$$\phi G(x|p_1, \dots, p_G; \mu_1, \dots, \mu_G; \sigma_1^2, \dots, \sigma_G^2) = \sum_{g=1}^G p_g \cdot \phi(x|\mu_g, \sigma_g^2). \quad (2)$$

Often one can make the assumption of equal variances (i.e.,  $\sigma_1^2 = \sigma_2^2 = \dots = \sigma_G^2$ ) among the groups.

Though mixture distribution analyses have wide applicability, they have been particularly useful in genetics. MacLean et al. [1976] considered the case of a quantitative trait determined by a genetic locus with 2 alleles  $A$  and  $a$  whose genotypes ( $AA$ ,  $Aa$ , and  $aa$ ) produce phenotype values that are normally distributed around a mean value unique to each genotype. Modeling the phenotypic expression of the locus as a mixture of three normal distributions, they considered the theoretically appealing restriction of the proportion parameters or "mixing weights" to values dictated by Hardy-Weinberg equilibrium:  $p_{AA} = q^2$ ;  $p_{Aa} = 2q(1 - q)$ ; and  $p_{aa} = (1 - q)^2$ , where  $q$  is the frequency of the "A" allele. For data,  $\hat{x}$ , collected on a quantitative trait with unknown genetic determinants, MacLean et al. considered comparing the likelihood of a homoscedastic (i.e., equal variance) mixture of three normal distributions,

$$L_G = {}_3(q; \mu_1, \mu_2, \mu_3; \sigma^2 | \hat{x}) = \prod_{n=1}^N \phi_3(x_n | q; \mu_1, \mu_2, \mu_3; \sigma^2), \quad (3)$$

(where  $N$  is the number of observations in the sample,  $\hat{x}$  denotes a vector, and the subscripts 1, 2, and 3 index the genotypes  $AA$ ,  $Aa$ , and  $aa$ ) to the likelihood of a single normal distribution (i.e.,  $G = 1$ ) to test hypotheses about the monogenetic control of the trait. MacLean et al. further suggested that by restricting the mixture to two-component normals, one could test for possible dominance or recessivity (i.e.,  $\mu_{AA} = \mu_{Aa}$  or  $\mu_{aa} = \mu_{AA}$ ). Such work with normal mixture models for loci controlling quantitative traits manifested itself in the development of segregation analysis for quantitative traits [Elston and Stewart, 1971; Morton and MacLean, 1974; Ott, 1979a]. This analytic technique essentially uses normal mixture distributions to model the genetic transmission of a quantitative trait from generation to generation. This paper will not delve into segregation analysis; instead we confine our attention entirely to the use of normal mixture distributions with cross-sectional, as opposed to generational or familial, data.

There has been a great deal of research on the statistical comparison of normal mixture distributions with  $G$  components to a single normal ( $G = 1$ ) or lesser number

of component normal mixtures ( $G' < G$ ) [see the book by Titterington et al., 1985, for an extensive list of references]. In addition to their interpretability, normal mixture distributions have the favorable quality of being able to take on a wide variety of shapes, as depicted in Figure 1. However, as intimated in the opening section, this flexibility also poses some problems. These problems arise from the fact that normal mixture distributions that show skewness (Fig. 1, top right) as opposed to multimodality (Fig. 2, bottom right) generally look and behave mathematically like a number of other distributions whose parameters invite interpretations wholly unlike those of normal mixture distributions. Pearson, of course, recognized this, as have others in the more recent past [Murphy, 1964, 1979]. In what follows we develop mathematical models that can rival the normal mixture model of *skewed* data. Each model has characteristics which make it heuristically appealing from a biological standpoint. Later, we offer methods for discriminating among these models.

### Log-Normal Distributions

It is well known that the distribution of phenomena determined by numerous independent and additive forces can generally be modeled as a normal distribution. When determining forces are not additive but rather interactive or "multiplicative," the resulting phenomena may be log-normally distributed [see Aitchison and Brown, 1957 for a discussion]. The likelihood function for the mean,  $\mu$ , and variance,  $\sigma^2$ , given data,  $\hat{x}$ , which are assumed to follow a log-normal distribution, is

$$L_{log}(\mu, \sigma^2 | \hat{x}) = \prod_{n=1}^N \frac{1}{\sqrt{2\pi\sigma}x_n} \exp[-(\log(x_n) - \mu)^2 / 2\sigma^2], \quad (4a)$$

where  $x > 0$ . The relationship of the log-normal distribution to the normal distribution is obvious: taking the logs of data following a log-normal distribution will produce normally distributed data. The log-normal distribution is extremely flexible, can exhibit an extremely large range of skewness-levels (see Fig. 2), and characterizes a number of biological phenomena [Aitchison and Brown, 1957].

In addition to non-additivity, a quantitative trait may possess an intrinsic "lower bound" or threshold; e.g., "0" is an absolute lower bound for blood pressure and heart rate. Pearson recognized this fact and therefore felt compelled to work with discrete distributions (such as the binomial) whose forms more easily integrate threshold phenomena, since discrete distributions typically have well-defined, finite ranges, unlike many continuous distributions which can be defined on  $-\infty < x < \infty$  [Pearson, 1895, pp. 388–389]. In modeling phenomena that exhibit non-additivity and a threshold, one can use the three-parameter log normal distribution with threshold parameter  $\tau$ , mean  $\mu$ , and variance  $\sigma^2$ . The likelihood function for these parameters given data,  $\hat{x}$ , is

$$L_{3log}(\tau, \mu, \sigma | \hat{x}) = \prod_{n=1}^N \frac{1}{\sqrt{2\pi\sigma}(x_n - \tau)} \exp[-(\log(x_n - \tau) - \mu)^2 / 2\sigma^2], \quad (4b)$$

where  $\tau < x < \infty$ . Estimation of the parameters of the three-parameter log-normal distribution is discussed in Cohen [1988].

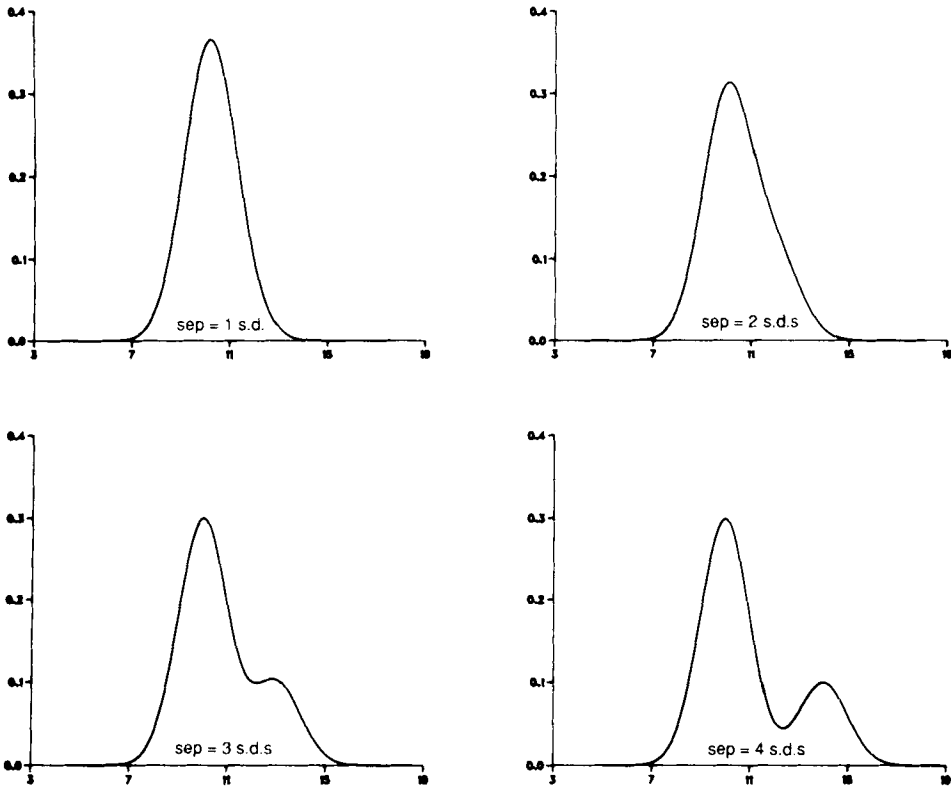


Fig. 1. Homoscedastic mixtures of two normal distributions with mixing weight 0.75 and varying separations (in standard deviation units) between the means. Note that as the mean separation gets larger, the mixture moves from a *skewed* appearance (**upper right**) to a *bimodal* appearance (**lower right**).

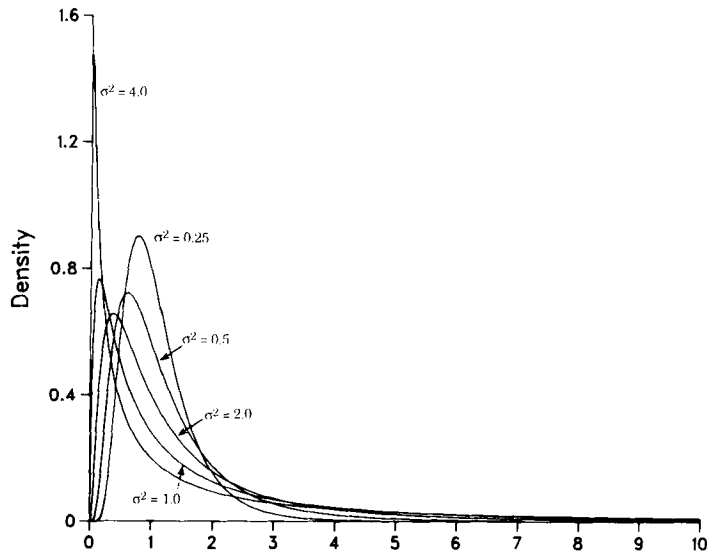


Fig. 2. Log-normal distributions with mean 0 and different variances.

### Truncated Distributions

As a direct analog to the notion of a “threshold,” one can think in terms of the “truncation points” of certain phenomena. As an example, consider a quantitative trait that would be normally distributed except that low values of the trait, e.g., extremely low blood pressures, are lethal. The lower tail of the distribution would then be missing. Such a trait would not manifest itself as a patently obvious normal distribution, but would show a degree of skewing toward higher trait values that would be roughly proportional to the size of that part of the distribution that has been lost. If we let  $\kappa$  stand for the truncation point, then the likelihood of the parameters for a truncated *normal* distribution given data  $\hat{x}$  is

$$L_t(\kappa, \mu, \sigma^2 | \hat{x}) = \prod_{n=1}^N \frac{\phi(x_n | \mu, \sigma^2)}{1 - \Phi(\kappa | \mu, \sigma^2)}, \quad (5)$$

where  $x > \kappa$  and  $\Phi(x | \mu, \sigma^2)$  is the cumulative normal distribution function with mean  $\mu$  and variance  $\sigma^2$  evaluated at the point  $x$ .

Of course, one can model truncation events with distributions other than the normal. Aitchison and Brown [1957] discuss the truncated log-normal and related distributions, as well as parameter estimation techniques for them.

### Convolutions of Distributions

It may be the case that a variable is determined by the sum of independent factors whose individual distributions are known. For the sake of argument, let us say a variable,  $Z$ , is determined by the sum of a factor,  $X$ , with distribution function  $f_x(x)$ , and a factor,  $Y$ , with distribution function  $f_y(y)$ . The density function,  $f_z(z)$ , of  $Z = X + Y$  can be evaluated through the convolution integral:

$$f_z(z) = \int_{-\infty}^{+\infty} f_x(x) f_y(z - x) dx. \quad (6)$$

Convolutions can not only produce a wide variety of distributional shapes, but also admit a wide range of interpretations. For example, the convolution of normal and exponential distributions can produce distributions that are skewed and unimodal. Again using blood pressure as an example, the normal/exponential convolution might invite the following interpretation: blood pressure is determined by the sum of a basic variable that, because of the additivity of its own determinants, is normally distributed, and a component that increases the pressure exponentially (say, because it reflects the aging process and blood pressure’s tendency to increase through time or with repeated exposures to certain environmental stimuli). Those with a high “basic” pressure component (i.e., a pressure value in the upper tail of the normal distribution characterizing this basic component) would be most likely to suffer the greatest effect of the age or environmental component. Such activity would create skewing toward higher values.

Estimation of parameters implicated in a convolution is difficult. Davis and Kutner [1976] discuss parameter estimation in the case of the normal-exponential(s) convolution. Considerable analytic and computational advantages for the estimation of convolution parameters can be gained through the use of characteristic functions and/or Fourier transforms [Lukacs, 1970].

### Covariate-Dependent Distributions

Many traits exhibit regular associations or dependencies with other traits. Accordingly one may want to “remove” the effect of the associated trait in order to better quantify certain other properties of the distribution of the trait of interest. Certainly weight, age, and sex are all-too-obvious sources of variation in need of removal for the study of the distribution of blood pressure. Makuch et al. [1979] and Murphy [1979] argue further that age-mediated effects may impact on the symmetry of biological distributions.

If we assume a variable,  $Y$ , is linearly related to another variable,  $X$ , such that the distribution of the adjusted variate  $\hat{y} = y - bx$ , (where  $b$  is the regression coefficient) is  $f_{\hat{y}}(\hat{y})$ , and if we assume  $f_{\hat{y}}(\hat{y})$  (i.e., the distribution of the residuals or adjusted variates  $\hat{y}$ ) has parameter vector,  $\theta$ , then the resulting likelihood function is

$$L_{adj}(\theta, b|y, x) = \prod_{n=1}^N f_{\hat{y}}(\hat{y}_n). \quad (7)$$

Extensions to multiple covariates and non-linear associations are straightforward. Note that for estimation purposes equation 7 assumes that  $b$  is estimated simultaneously with  $\theta$ . Problems may arise if one adjusts the variate *prior* to estimating the parameters,  $\theta$ , as will be discussed in the next section. Of course, the form of the distribution function,  $f$ , should reflect biological properties appropriate to the adjusted variate. Thus, one may assume log-normality, mixed-normality, etc. for the distribution of the adjusted variate  $\hat{y}$ .

### Arbitrarily Skewed Distributions

Skewness in a distribution can arise for reasons or through disturbances not easily quantifiable such as biased sampling, measurement artifact, or the effect of unknown factors. If one assumes that apart from these disturbances, the relevant trait values follow a distribution,  $f$ , then one may wish to adjust for skewness induced by these disturbances while estimating the parameters associated with  $f$ . This adjustment can be performed through the use of a flexible transformation—flexibility is the key since the ultimate form of the disturbance is unknown—such as the power transform

$$y = \begin{cases} (x^\lambda - 1)/\lambda + \lambda, & \text{if } \lambda \neq 0; \\ \log(x), & \text{if } \lambda = 0, \end{cases} \quad (8a)$$

where  $x > 0$  is the variate under scrutiny. Equation 8a is variant of a transformation discussed in Box and Cox [1964], which has the advantage that with  $\lambda = 1$  the data remain untransformed [see Ott, 1979b]. Using the Jacobian of the appropriate transformation one can then compute the likelihood of the parameters,  $\theta$ , assumed in the distribution function,  $f$ . For example, the likelihood equation assuming the transformation given in equation 8a is

$$L_{pow}(\theta, \lambda|\hat{x}) = \prod_{n=1}^N f(y_n) \cdot x_n^{\lambda-1}. \quad (8b)$$

Here  $f$  can be the normal distribution, the log-normal distribution, a mixture of normal distributions, etc., MacLean et al. [1976], Ott [1979b], and Schork and Schork [1988] all consider use of the power transformation in settings where  $f$  is a mixture of normal distributions. Two problems arise in the use of the power transformation in equation 8b: its implementation and its interpretation. As with concomitant variate adjustment, some investigators transform their data *prior* to fitting the mixture whereas others estimate it simultaneously, as implied in equation 8b. In the next section we show how one can use both the prior and simultaneous estimation of the power transformation parameter,  $\lambda$ , in omnibus tests of the hypothesis that a mixture of normal distributions better explains the variation of a trait than a single, non-mixed, arbitrarily skewed distribution. Questions about the interpretability of the transformation will be addressed in the final section.

### III. TESTING ASYMMETRICAL DISTRIBUTIONS

In this section we first describe procedures used to test hypotheses about whether skewing in data with unknown determinants results from a mixture or from causative agents whose expression is characterizable by a log-normal or some other inherently skewed, single-population-oriented distribution. We then describe problems that may arise in covariate adjustment strategies, since, as mentioned earlier, covariate relationships can impact on the distribution of a trait of interest.

For expository purposes, we consider two distinct kinds of tests: those in which the alternative to the mixture hypothesis is specified (e.g., a log-normal alternative, a truncated normal alternative, etc.) and more omnibus tests where the alternative is simply that the skewness does not arise from a mixture. The same testing strategies are used in both cases and are flexible enough to be used in virtually any parametric, data-dependent hypothesis-testing situations.

To aid the discussion of the first testing scenario, we describe tests involving mixtures of normal and log-normal alternative distributional hypotheses, though any of the models described in section II (or any other model, for that matter) can be used in place of the log-normal alternative. We consider the direct comparison of the likelihoods of the mixture model (equation 2 or 3) and the log-normal model (equation 4a or 4b). Akaike [1974] suggests transforming the likelihood of each model by first multiplying it by  $-2$  and then adding two times the number of free parameters in the model to this product. Akaike argues that the model with the lowest value should be considered the best model. Though not a statistical test in the Neymann-Pearson sense (i.e., there are no probabilities associated with the outcome), the "Akaike Information Criterion" (AIC), as it is known, has been used widely. However, at least two problems should be mentioned. First, the term "2 times the number of free parameters in the model," which is based on entropy and information-theoretic devices, is one among many asymptotic criteria, some of which may have superior analytic and asymptotic justifications [see Linhart and Volkers, 1984; Linhart and Zucchini, 1984, for example]. Second, it has been argued that for large samples, the AIC tends to favor models with more parameters [McDonald, 1989]. As an alternative to the AIC, we consider testing (i.e., using  $P$ -values) the log-ratio of likelihoods computed under the mixture and log-normal hypotheses. Since the distribution of the resulting log-ratio is unknown, we simulate it and estimate critical values and significance levels from the simulated dis-



tribution. Schork and Schork [1989] describe two tests for segregation analysis settings, the parametric bootstrap (PB) and SOS tests, that are easily extended to the situations outlined herein. Though we forego an indepth elaboration of these tests since they are fully described in Schork and Schork [1989], the following provides the basic motivation behind them. One first computes a test statistic,  $t_N$ , (e.g., likelihood ratio), whose distribution is unknown, over some data,  $x_1, \dots, x_N$ . One then draws  $r$  random samples,  $x_{r1}, \dots, x_{rN}$ , from the distribution assumed in the null hypothesis (e.g., using random number generators) and computes, for each of these  $r$  samples, the statistic,  $t_N^*$ . The number of  $t_N^*$  exceeding  $t_N$  divided by  $r$  gives the  $P$ -value for the observed  $t_N$ . Such test constructions are not entirely new and have received sporadic attention in the statistics literature [Birnbaum, 1974; Hope, 1968].

In order to examine the usefulness of the PB and SOS tests in discriminating between normal mixture and log-normal hypotheses, we performed a few Monte Carlo experiments. Since it is generally acknowledged that it is better to err on the side of non-mixture (e.g., non-major locus) hypotheses, we consider the case where the log-normal distribution is taken as the null hypothesis. In a first experiment 250 deviates following log-normal distributions with the variances given in Figure 2 were generated to which normal mixture and log-normal distributions were fit. For each variance level, this experiment was replicated 100 times with each replicate's resulting log-ratio being tested with the PB and SOS tests of the hypothesis of log-normality. Nineteen bootstrap replications were used to conduct each test [Schork and Schork, 1989]. Results are given in Table I and suggest that, with the exception of the small variance case for the SOS test, the PB and SOS tests are at or near the expected significance levels. We also note that the AIC (correctly) accepted the log-normal hypothesis in each experimental setting.

We next assessed the power of the PB, SOS, and AIC tests to reject the log-normality hypothesis when the data actually follow a two-component normal mixture. We again considered samples of size 250 and generated mixtures with a mixing weight of 0.75 and distances between the two means ranging from 1 to 4 standard deviations (see Fig. 1)—the idea being that as bimodality became more pronounced (i.e., there is greater separation between the means) the log-normal hypothesis would be more easily rejected. The results are depicted in Figure 3 and indeed suggest that all three tests work better as the separation between the means increases, but that the AIC and SOS

**TABLE I. Estimated Significance Levels for the Parametric Bootstrap and SOS Test of the Hypothesis That A Mixture of Two Homoscedastic Normal Distributions Fit Log Normally Distributed Data With Different Variances Better Than a Two-Parameter Log-Normal Distribution\***

Variance	PB			SOS			AIC
	0.05	0.10	0.20	0.05	0.10	0.20	
0.25	0.02	0.10	0.24	0.19	0.32	0.56	0.00
0.50	0.02	0.04	0.13	0.11	0.23	0.44	0.00
1.00	0.00	0.01	0.08	0.02	0.09	0.29	0.00
2.00	0.01	0.06	0.13	0.08	0.15	0.31	0.00
4.00	0.00	0.00	0.10	0.06	0.15	0.26	0.00

\*PB = parametric bootstrap test; SOS = SOS criterion test; AIC = proportion log normal hypothesis rejected based on the Akaike information criterion; .05, .10, and .20 = theoretical significance levels; variance = variance used with mean 0 to generate the log normal data.

tests have better power at lower separations. We repeated this experiment with a 3-parameter log-normal hypothesis as the null hypothesis. The results are depicted in Figure 4, and suggest that all these tests behave roughly the same, though the AIC and SOS tests have somewhat better power at lower mean separation levels.

To carry out omnibus tests of the hypothesis that a mixture of normal distributions better characterizes a set of data than an arbitrarily skewed, non-mixture, distribution, we consider testing the likelihoods of a normal mixture and a power-transformed single normal distribution (i.e., equation 8b with  $f = \phi$ ) using the 3 methods discussed in the foregoing remarks on "specified alternative" tests. We again employed Monte Carlo experiments to assess the usefulness of the tests in this situation. Table II shows the estimated significance levels for 250 data points generated with different  $\lambda$  values (from equation 8) based on 100 replications; it suggests that the SOS and AIC tests too easily reject the null hypothesis of power-transform-induced normality when it is true. Figure 5 shows the power the tests have in rejecting the power-transform-induced normality hypothesis when data actually come from a two-component normal mixture with varying mean separations. The power of the SOS and AIC tests should be interpreted with caution since they do not appear to behave like true level  $\alpha$  tests (i.e., they too easily reject true null hypotheses—see Table II). Figure 5 also suggests that discriminating mixtures from non mixture, skewed distributions is difficult when the mean components are separated by less than 2 standard deviations—the "critical" point for bimodality for certain parameter settings [Behboodian, 1970]. This is intuitive because, as has been emphasized throughout this paper, normal mixtures without multimodality generally look like simply skewed distributions.

It would have been impossible to assess the properties of the AIC, SOS, and PB tests in all situations (i.e., mixing weights of 0.5, sample sizes of 100, etc.). Although

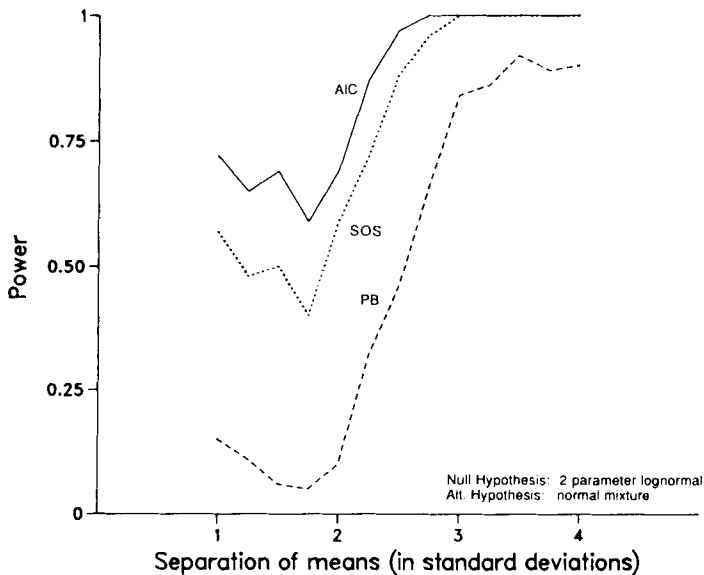


Fig. 3. Power of the AIC, PB, and SOS tests to reject the log-normal hypothesis for samples of size 250 as the mean separation in a two-component normal mixture with mixing weight 0.75 gets larger.

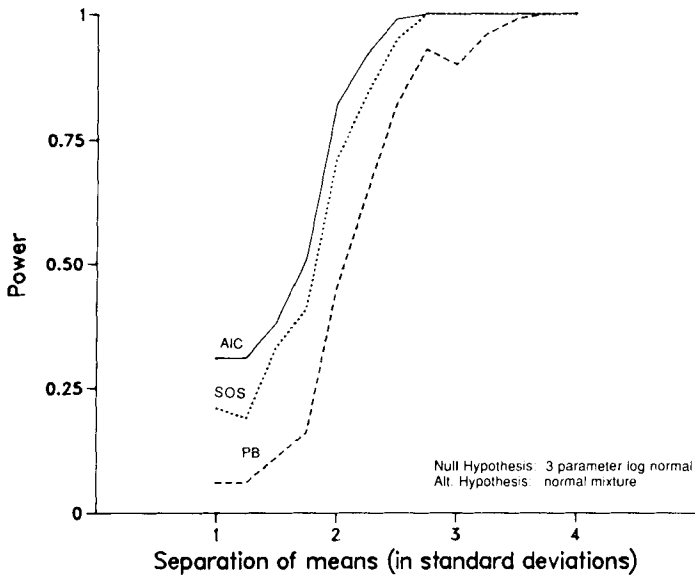


Fig. 4. Power of the AIC, PB, and SOS tests to reject the three-parameter log-normal hypothesis for samples of size 250 as the mean separation in a two-component normal mixture with mixing weight 0.75 gets larger.

**TABLE II. Estimated Significance Levels for the Parametric Bootstrap and SOS Tests of the Hypothesis That A Mixture of Two Homoscedastic Normal Distributions Fit Normal Distributed Power-Transformed Data Better Than a Power-Transformed Normal Distribution\***

Transformation	PB			SOS			AIC
	0.05	0.10	0.20	0.05	0.10	0.20	
0.10	0.00	0.04	0.10	0.16	0.37	0.50	0.01
0.30	0.06	0.15	0.28	0.36	0.57	0.77	0.26
0.50	0.03	0.08	0.20	0.34	0.60	0.82	0.25
0.70	0.04	0.08	0.21	0.36	0.62	0.88	0.21
0.90	0.01	0.04	0.15	0.40	0.65	0.88	0.16

\*PB, SOS, AIC, .05, .10, .20 (see Table III); transformation = the value of the power transformation (see text) whose inverse was used along with a mean of 2.0 and variance 0.25 to generate skewed data.

Schork [1990] investigates the power of the tests as a function of sample size, further experiments are called for. Despite this, our results suggest these tests have promise, if nothing else.

Mention should be made of the omnibus testing strategy originally developed by MacLean et al. [1976], which was later elaborated by Ott [1979b] and Schork and Schork [1988]. This strategy involves estimating a power transformation parameter (e.g., equation 8a) simultaneously with the normal mixture parameters and using  $\chi^2$ -based likelihood ratio tests to determine the number of components in the mixture and the significance of the skewing in the data. This testing procedure has a number of problems. Scaling problems inherent in the simultaneous estimation process make it difficult and often unreliable. However, these scaling problems can be addressed through the clever use of Taylor series approximations (the ‘‘delta method’’) as suggested by

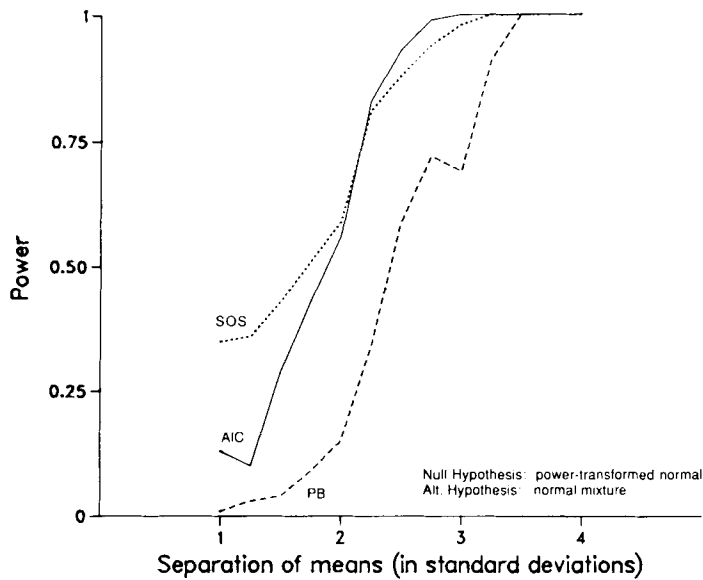


Fig. 5. Power of the AIC, PB, and SOS tests to reject the power-transform-induced normal distribution hypothesis (see text) for samples of size 250 as the mean separation in a two-component normal mixture with mixing weight 0.75 gets larger.

Ott [1979b]. In addition, it is known that mixture distributions violate the regularity conditions assumed for the log-likelihood ratio statistic to have a  $\chi^2$  distribution, as mixture models are not identifiable in the statistical sense [Ghosh and Sen, 1985; Hartigan, 1985] and often produce test statistics whose distributions are not independent of their parameters [Wolfe, 1971; Everitt, 1980]. Thus, use of  $\chi^2$  tests in mixture model settings is invalid. Schork and Schork [1988] also showed that a variant of the MacLean et al. test has low power for rejecting incorrect models. It is our belief that the “specified alternative” and omnibus testing procedures based on the AIC, PB, and SOS tests outlined above provide intuitive, reliable, flexible, and easily implemented alternatives to the MacLean et al. [1976] testing strategy.

As mentioned earlier, and as will become clear in the following, covariate dependence may impact on the distribution of the trait of interest and therefore influence testing strategies, like those mentioned above, which are meant to assess the significance of the mixture hypothesis. Current covariate adjustment techniques for a trait whose distribution is to be tested as a mixture of normal distributions typically involve performing the adjustment *prior* to fitting the mixture. That is, one first regresses the variable of interest on known concomitants or other factors, saves the residuals from the resulting regression equation, and then subjects these residuals to a mixture analysis (say, by using maximum likelihood techniques and equations 2 or 3). However, when the prior regression analysis is carried out, one may not want to assume, say, a log-normal distribution for the error terms, simply because it is this error distribution that is being tested in the later mixture analysis. For instance, if one assumes a log-normal distribution for the error terms in the prior regression and then finds evidence for a normal mixture using the residuals saved from this prior regression analysis, then

this would really only suggest that the regression model used initially (along with its assumption of  $\log(\hat{y}) \sim \phi(x)$ ) was incorrect! Thus, the estimates of the regression coefficients would be invalidated. To avoid this, one can use least squares estimates (or estimates derived assuming  $\hat{y} \sim \phi(x)$  since the relevant likelihood equations produce coefficient estimates equivalent to the least squares estimates). By the Gauss-Markov theorem it is known that least squares estimates have the smallest possible variance of any linear unbiased estimators. Since this holds irrespective of the true distribution of the error terms, it seems that use of prior least squares regression adjustment of covariates in mixture analysis settings is ideal.

However, two problems with this strategy arise. First, Gauss-Markov results hold only if the error terms have a common variance, thus invalidating the optimality of the prior least squares regression approach in settings in which the components in the normal mixture characterizing the adjusted variate possess unique variances. In addition, in using prior covariate adjustment through least squares techniques and then estimating the mixture parameters from residual values, one is relying on two different estimation techniques (e.g., least squares and maximum likelihood) for parameters meant to characterize a single phenomenon. It is not clear what properties such "mixed" estimation procedures have for finite samples. Thus, in certain situations, the simultaneous estimation of "b" and "θ" parameters, as suggested in equation 7, may seem appropriate.

In order to investigate the aforementioned problems, two small Monte Carlo studies were performed. Both involved the comparison of results gleaned from the simultaneous estimation of regression and mixture parameters with results gleaned from the two-step approach of least-squares regression adjustment followed by a maximum likelihood based (i.e., equation 2 with  $G = 2$ ) mixture analysis of residual values obtained from the regression.

The first experiment involved the generation of deviates known to conform to a homoscedastic (i.e., equal variance) covariate-dependent normal mixture distribution. One thousand samples of size 25, 50, 100, 200, and 400 were generated from which mixture and regression coefficient parameter estimates were obtained using the simultaneous and two-step least squares/maximum likelihood estimation approaches. The mean squared error (MSE) and bias of the mixing weight, mean separation, and regression coefficient were computed from the 1,000 replicates at each sample size. The results are described in Table III. The two models mentioned in Table III assumed different parameter settings to generate the data. Model 1 assumed a mixing weight of 0.75, means of 0.00 and 2.00, a regression coefficient of 1.00, and a residual variance of 1.00. Model 2 assumed a mixing weight of 0.75, means of 0.00 and 2.00, a regression coefficient of 0.35, and a residual variance of 1.5. Table III suggests that the simultaneous and two step procedures behave similarly, and show similar rates of decline in MSE and bias for increasing sample sizes.

The second experiment was similar to the first except that data following a heteroscedastic covariate-dependent normal mixture were generated. The results are described in Table IV. Model 1 assumed a mixing weight of 0.50, means of 0.00 and 2.00 and variances of 1.00 and 1.50 for the two components, and a regression coefficient of 1.00. Model 2 assumed a mixing weight of 0.50, means of 0.00 and 2.00 and variances of 1.00 and 2.00 for the two components, and a regression coefficient of 1.00. Table IV suggests that the two-step procedure does a little better in estimating the regression coefficient, but much worse in estimating the separation between the means.

**TABLE III. Comparison of the MSE and Bias (in Parentheses) of Parameter Estimates for a Homoscedastic Covariate-Dependent Normal Mixture Using the Simultaneous and Two-Step Regression Coefficient Estimation Procedures**

Sample size	Simultaneous			Two-step		
	$P$	$\mu_2 - \mu_1$	Slope	$P$	$\mu_2 - \mu_1$	Slope
<b>Model 1</b>						
25	.035 (0.051)	7.250 (1.93)	0.109 (0.027)	.060 (0.107)	4.920 (1.590)	0.077 (0.011)
50	.025 (0.030)	1.880 (0.829)	0.040 (0.014)	.042 (0.062)	1.370 (0.749)	0.038 (0.003)
100	.018 (0.008)	0.676 (0.367)	0.018 (-0.000)	.025 (0.024)	0.549 (0.338)	0.017 (-0.000)
200	.010 (0.047)	0.320 (0.228)	0.018 (-0.000)	.011 (0.008)	0.279 (0.204)	0.009 (0.003)
400	.047 (-0.000)	0.130 (0.081)	0.071 (0.008)	.041 (-0.000)	0.126 (0.073)	0.004 (0.002)
<b>Model 2</b>						
25	.097 (0.147)	2.630 (0.592)	0.178 (0.030)	.100 (0.181)	1.630 (0.536)	0.135 (0.010)
50	.089 (0.124)	0.843 (-0.240)	0.074 (-0.010)	.091 (0.149)	0.469 (-0.150)	0.064 (0.003)
100	.047 (0.056)	0.690 (-0.590)	0.032 (-0.010)	.079 (0.123)	0.512 (-0.530)	0.030 (-0.010)
200	.034 (0.008)	0.842 (-0.780)	0.015 (-0.000)	.063 (0.080)	0.702 (-0.730)	0.014 (-0.000)
400	.021 (0.004)	0.979 (-0.910)	0.008 (-0.000)	.039 (0.047)	0.879 (-0.880)	0.008 (-0.000)

**TABLE IV. Comparison of the MSE and Bias (in Parentheses) for Estimates of the Parameters of a Heteroscedastic Covariate-Dependent Normal Mixture Using the Simultaneous and Two-Step Regression Coefficient Estimation Procedures**

Sample size	Simultaneous			Two-step		
	<i>P</i>	$\mu_2 - \mu_1$	Slope	<i>P</i>	$\mu_2 - \mu_1$	Slope
<b>Model 1</b>						
25	.035 (-0.103)	0.742 (0.081)	0.303 (0.408)	.037 (-0.092)	36.61 (3.871)	0.081 (0.010)
50	.028 (-0.102)	1.200 (0.183)	0.237 (0.372)	.029 (-0.110)	23.93 (3.392)	0.041 (-0.010)
100	.020 (-0.101)	1.040 (0.146)	0.175 (0.328)	.021 (-0.103)	19.80 (3.410)	0.019 (0.003)
200	.016 (-0.104)	1.540 (0.257)	0.164 (0.009)	.016 (-0.100)	18.42 (3.623)	0.009 (0.003)
400	.013 (-0.100)	1.910 (0.329)	0.163 (0.362)	.014 (-0.101)	20.11 (4.044)	0.005 (-0.019)
<b>Model 2</b>						
25	.048 (-0.160)	0.059 (-0.023)	1.780 (-0.161)	.055 (-0.171)	27.80 (2.530)	0.183 (-0.012)
50	.048 (-0.180)	0.399 (-0.071)	1.201 (-0.262)	.055 (-0.202)	16.21 (1.882)	0.092 (-0.001)
100	.047 (-0.190)	0.340 (-0.102)	0.880 (-0.344)	.054 (-0.223)	8.802 (1.310)	0.046 (-0.000)
200	.049 (-0.210)	0.099 (-0.191)	0.349 (-0.281)	.052 (-0.229)	4.811 (0.794)	0.023 (-0.000)
400	.050 (-0.220)	0.150 (-0.196)	0.223 (-0.280)	.054 (-0.233)	1.812 (0.226)	0.011 (-0.001)

Also of note is the fact that for each replication we used the Akaike information criterion [Akaike, 1974] to determine if a single normal distribution better fit the data than the resulting mixture. In the first experiment the two estimation procedures provided similar powers for rejecting the single normal hypothesis. In the second experiment it was found that the two-step procedure had 10–15% less power to reject the single normal hypothesis for small samples, but virtually equal power for larger samples. These studies suggest some caution and/or discretion should be used when employing the two-step estimation procedure, since typically one does not know, a priori, if the equal variance assumption should be made.

#### IV. PICKERING, PLATT, AND THE BLOOD PRESSURE DISTRIBUTION

In this section we apply some of the methods outlined in sections II and III to the distribution of systolic and diastolic blood pressure. Our treatment of the blood pressure distribution is not meant to be exhaustive but rather expository, since the normal mixture distribution argument was a keystone in one of the most widely publicized medical disputes in this century: the famous Pickering/Platt debate on the nature of hypertension (for a description the Pickering/Platt debate we refer the reader to the excellent book by Swales [1985] on the subject and all the references contained therein).

Pickering and Platt were two noted English internists with differing views on the etiology of essential hypertension. Platt claimed hypertension was a “disease” with discernible characteristics, whose underlying pathophysiologic determinants (genetic or otherwise) admitted almost “qualitative” interpretations: one simply either had them or did not. Platt placed a great deal of emphasis on evidence derived from his personal observations that the distribution of blood pressure values has a skewness that may be the manifestation of the effects of a Mendelian dominant gene (i.e., the blood pressure distribution admits a mixture). Pickering staunchly opposed Platt’s interpretation, arguing that the designation “hypertension” was entirely arbitrary and that in sufficiently large samples no mixture was evident. For Pickering, the determinants of blood pressure were numerous and of small effect individually. As such, hypertension, Pickering maintained, is not a qualitative disease but merely a label assigned arbitrarily to those with pressure readings in the upper tail of the distribution. Pickering held that any evidence for a mixture in the blood pressure distribution was most likely due to measurement artifact (“digit preference”) and possibly to non-additivity.

Many researchers have tried to settle the dispute by fitting normal mixture distributions to large samples of blood pressure values [Cicchinelli, 1962; Clark et al., 1968; and McManus, 1983], but results have been inconclusive. An overriding criticism of these studies has been aimed at, not surprisingly given the context of this paper, their testing strategy: the alternative hypothesis to the mixture distribution hypothesis posited in each has been a single normal distribution and not a single skewed, non-mixed distribution.

We applied the techniques outlined in sections II and III to a large population-based sample of blood pressure values. The alternatives to the normal mixture distribution hypothesis considered were the two-parameter log-normal (equation 4a) and the power-transform-induced normal distribution (equations 8a and 8b). It should be understood that the log-normal distribution is a special case of the power-transformed normal as defined in equation 8a. Our data consisted of systolic and diastolic blood



pressures collected on 941 white male subjects participating in a random, statewide blood pressure screening in Michigan [Schork, et al., 1990].

In order to adjust the systolic and diastolic pressures for the effects of age, height, and weight (see Table V) while avoiding the problems discussed in section III, the following adjustment procedure was employed. Age was stratified by decade (the range of the ages of the participants in the sample was 8–83), and weight was broken into quintiles. Blood pressure values were then standardized within each age-decade  $\times$  weight quintile category. The original sample means of systolic and diastolic pressure were added back to the standardized values to preserve positive observational units [Schork et al., 1977]. No regression procedure was employed. This standardization procedure effectively reduced the contribution of age, height, and weight to blood pressure variability, as indicated in Table V.

The distributions of the standardized values are plotted in Figure 6, using an optimal width for display as described in Schork and Schork [1990]. The best fitting homoscedastic two-component normal mixture distribution, log-normal distribution, and power-transformed-induced normal distribution were then determined for each distribution by maximizing likelihoods for each. The log-likelihoods were, for the systolic pressure distribution,  $-1,291.68$ ,  $-1,315.21$ , and  $-1,311.62$ ; and for the diastolic pressure distribution,  $-1,310.02$ ,  $-1,314.48$ , and  $-1,312.23$ . In each case, the null hypothesis was either the log-normal or power-induced normal. By the AIC, the mixture distribution better characterized each pressure curve than either skewed distribution. However, as remarked earlier, the AIC may not be an optimal criterion and does not appear to be overly effective in comparing mixed normal and power-induced normal distributions (see Table II). We therefore used the PB test to determine the most likely distribution. The difference in the log-likelihoods for systolic pressure were 23.53 for a log-normal null-hypothesis-based test and 19.94 for a power-transformed normal null-hypothesis-based test. Critical values determined from the PB test at the 5% level for these situations were, respectively, 2.57 and 2.83. Thus in each case the skewed distribution null hypothesis was rejected. For diastolic pressure, the differences in log-likelihoods were 4.46 for the log-normal case and 2.18 for the power-transformed normal case. Critical values obtained from the PB test in these settings at the 5% level were 2.75 and 3.02. Thus, though the log-normal hypothesis was rejected, the power-induced normal hypothesis was not. This may not have been surprising since Figure 6 suggests that greater skewing—and hence greater “potential” for a normal mixture—exists in the systolic distribution. Also, there appears to be a greater discrepancy between the best fitting normal mixture and the best fitting log-normal distribution for systolic pressure rather than diastolic pressure, as also evidenced in Figure 6.

**TABLE V. Pearson Product-Moment Correlation Coefficients and *P*-Values (in Parentheses) Between Anthropometric Measures and Systolic and Diastolic Blood Pressure, Before and After Covariate Adjustment**

Measure	Unadjusted		Adjusted	
	Systolic	Diastolic	Systolic	Diastolic
Age	0.328 (<.0001)	0.163 (<.0001)	0.007 (.83)	0.002 (.95)
Body mass index	0.210 (<.0001)	0.246 (<.0001)	0.034 (.25)	0.021 (.50)
Height	-0.018 (.15)	0.047 (.76)	-0.050 (.12)	-0.025 (.44)
Weight	0.185 (<.0001)	0.249 (<.0001)	0.010 (.76)	0.001 (.77)

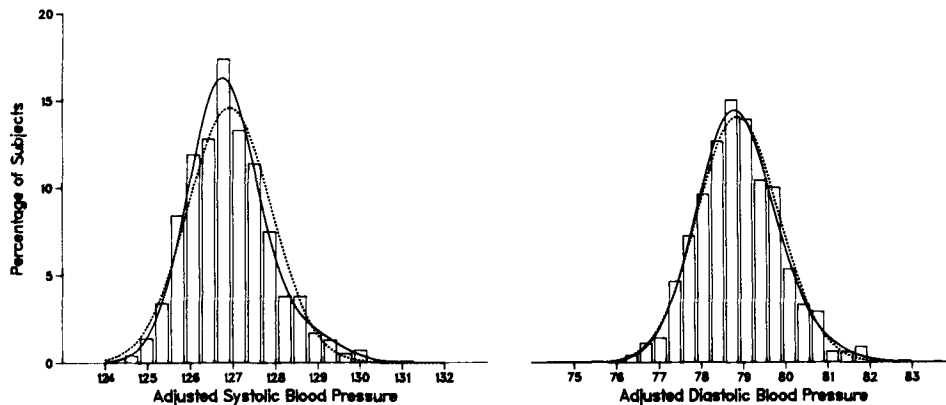


Fig. 6. Histograms showing the distributions for weight- and age-adjusted (see text) systolic and diastolic blood pressures taken from a sample of 941 white men. The superimposed solid-line curve is the best fitting two-component, homoscedastic, normal mixture distribution and the superimposed dotted-line curve is the best fitting log-normal distribution.

As emphasized earlier, this analysis of the blood pressure distribution is not meant to exhaust the issues raised in the Pickering/Platt debate. Consideration should be given to heteroscedastic mixture distributions, the mean blood pressure distribution [Weder and Schork, 1989], multivariate situations [Weder and Schork, 1989; Schork et al., 1990], other alternatives to the normal mixture, and, most importantly, the *interpretability* of the test outcomes.

## V. SUMMARY AND CONCLUSIONS

The problem of distinguishing skewness resulting from normal mixture distributions and skewness inherent in distributions such as the log-normal and others is not a new problem for the statistical geneticist. We have suggested that this problem was actually confronted as early as 1895 by the highly influential mathematician Karl Pearson. Despite this lengthy history, the recent genetics literature tends to describe the problem without actually combatting it. Most workers have followed the practice of transforming genetic data to normality, working entirely with the transformed data and then interpreting the results as though the transformations used were arbitrary and non-meaningful biologically, or have simply used single normal distributions as an alternative hypothesis to mixture hypotheses. What is counterintuitive in these practices is that most geneticists recognize that skewness may be an integral part of a biological trait, and may, in fact, have a biological meaning. Certainly for quantitative traits with purely polymorphic, monogenic determinants the skewness in data sampled for this trait "means" a mixture exists. As such, the practice of indiscriminant transformation to normality should be avoided since by *inducing normality (via transformation) one has, in effect, chopped off the very branch (i.e., skewness) upon which the more interesting genetic hypotheses rest: namely, monogenic determination, non-additivity, and the like*. In addition, by positing single normal distributions (or lesser component mixed normals) as alternatives to the mixture distribution hypothesis, one is providing a less than optimal and very unconservative testing environment for the mixture hypothesis.

Two points concerning skewness should be mentioned at this point. First, often-times the use of a specific transformation is called for on purely a priori grounds. For instance, if it is known, say, that the variability in a measure increases with increases in the measure itself—thereby inducing skewness in the distribution of the measure—then one may want to perform a transformation to induce homoscedasticity. Often, however, such a priori knowledge is absent and forces one to *look* for explanations of the skewing in a reliable and flexible way. The chapter on scaling in Falconer [1981] contains an excellent discussion on this issue and does much to convince the reader that transformations can actually effect the revelation of certain genetic phenomena (e.g., dominance effects can be masked by log transformations). Second, though the focus of this paper is on *normal* mixture distribution models, there is no reason why the effects of various genotypes cannot produce distributions that are themselves skewed. In such situations one can model the phenotype distribution as, say, a mixture of log-normal or Weibull distributions. The testing procedures outlined earlier would be entirely applicable in such settings.

In an erudite and somewhat rhetorical fashion Murphy [1964] explored issues relating to the use of normal mixture distributions. His message, that one should be aware of the dangers that come with the indiscriminant use of mixture analyses, was well taken. It is our hope that the message and methods outlined in this paper will pick up where Murphy left off by both cautioning the use of, and promoting further research into, the use of mixture distribution arguments in genetics research.

## REFERENCES

- Aitchison J, Brown JAC (1957): "The Log-Normal Distribution." Cambridge: Cambridge University Press.
- Akaike H (1974): A new look at the statistical model identification. *IEEE Trans Automatic Control* 19:719–723.
- Behboodian J (1970): On the modes of a mixture of two normal distributions. *Technometrics* 12:131–139.
- Birnbaum ZW (1974): Computers and unconventional test statistics. In Proschan F, Serfling RJ (eds): "Reliability and Biometry." Philadelphia: SIAM, pp 441–458.
- Box GEP, Cox DR (1964): An analysis of transformations. *J R Stat Soc* 26B:211–252.
- Cicchinelli AL (1962): The composite of two gaussian distributions as a model for blood pressure in man. Ph.D. thesis, the University of Michigan, Ann Arbor.
- Clark VA, Chapman JM, Coulson AH, Hasselblad V (1968): Dividing the blood pressures from the Los Angeles Heart Study into two normal distributions. *Johns Hopkins Med J* 122:77–83.
- Cohen AC (1988): Three-parameter estimation. In "Lognormal Distributions: Theory and Practice." Crow EL, Shimuzu K (eds): New York: Marcel Dekker, pp 113–137.
- Davis GC, Kutner MH (1976): The lagged normal family of probability density functions applied to indicator-dilution curves. *Biometrics* 32:669–676.
- Elston RC, Stewart J (1971): A general model for the genetic analysis of pedigree data. *Hum Hered* 21:523–542.
- Everitt BS (1980): A Monte Carlo investigation of the likelihood ratio for the number of components in a mixture of normal distributions. *Multivariate Behav Res* 16:171–180.
- Falconer DS (1981): "Introduction to Quantitative Genetics," 2nd edition. United Kingdom: Longman.
- Ghosh JK, Sen PK (1985): On the asymptotic performance of the log likelihood ratio statistic for the mixture model and related issues. In LeCam LM, Olshen RA (eds): "The Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer," Vol 2. Monterey: Wadsworth, pp 789–806.
- Hartigan JA (1985): A failure of likelihood asymptotics for normal mixtures. In LeCam LM, Olshen RA (eds): "Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer," Vol 2. Monterey: Wadsworth, pp 807–810.
- Hope AC (1968): A simplified Monte Carlo significance test procedure. *J R Stat Soc B* 30:582–598.
- Linhart H, Volkens P (1984): Asymptotic criteria for model selection. *O R Spektrum* 6:161–165.

- Linhart H, Zucchini W (1984): Model selection by the discrepancy of Cramer and von Mises. *Statistics Decisions* 57:297–306.
- Lukacs E (1970): "Characteristic Functions," 2nd edition. New York: Hafner.
- MacLean CJ, Morton NE, Elston RC, Yee S (1976): Skewness in commingled distributions. *Biometrics* 32:695–699.
- Makuch RW, Freeman DH, Johnson MF (1979): Justification for the lognormal distribution as a model for blood pressure. *J Chronic Dis* 32:245–250.
- McDonald RP (1989): An index of goodness-of-fit based on noncentrality. *J Classification* 6:97–103.
- McManus IC (1983): Bimodality of blood pressure levels. *Statistics Med* 2:253–258.
- Morton NE (1966): The detection of major genes under additive continuous variation. *Am J Hum Genet* 19:23–34.
- Morton NE, MacLean CJ (1974): Analysis of family resemblance. III. Complex segregation of quantitative traits. *Am J Hum Genet* 26:489–503.
- Murphy EA (1964): One cause? Many causes? The argument from the bimodal distribution. *J Chronic Dis* 17:301–324.
- Murphy EA (1979): Segregation of noisy mendelian traits and the effect of age-dependence: A prolegomenon. *Am J Med Genet* 4:173–190.
- Ott J (1979a): Maximum likelihood estimation by counting methods under polygenic and mixed models in human pedigrees. *Am J Hum Genet* 31:161–175.
- Ott J (1979b): Detection of rare major genes in lipid levels. *Hum Genet* 51:79–91.
- Pearson K (1893): Contributions to the mathematical theory of evolution. *Philos Trans A* 185:71–110.
- Pearson K (1895): Contributions to the mathematical theory of evolution. V. Skew variation in homogeneous material. *Philos Trans A* 186:342–414.
- Preston EJ (1953): A graphical method for the analysis of statistical distributions into two normal components. *Biometrika* 40:460–464.
- Schork MA, Schull WJ, Harburg E, Roepers P, Chape C (1977): Heredity, stress, and blood pressure, a family set method. IV. Blood Pressure adjustment techniques. *J Chronic Dis* 30:671–682.
- Schork NJ (1990): Bootstrapping likelihood ratios in quantitative genetics. To appear in *The Proceedings of the 214th meeting of the Institute of Mathematical Statistics: Special Topics Session on the Bootstrap*. John Wiley.
- Schork NJ, Schork MA (1988): Skewness and mixtures of normal distributions. *Communications Statistics Theory Methods* 17:3951–3969.
- Schork NJ, Schork MA (1989): Testing separate families of segregation hypotheses: Bootstrap methods. *Am J Hum Genet* 45:803–813.
- Schork NJ, Schork MA (1990): Histograms: Multimodal or poorly constructed? *Am J Hum Genet* (letter) 46:397,398.
- Schork NJ, Weder AB, Schork MA, Bassett DR, Julius S (1990): Disease entities, mixed multinormal distributions, and the role of the hyperkinetic state in the pathophysiology of hypertension. *Statistics Med* 9:301–314.
- Swales JD (ed) (1985): "Platt Vs. Pickering: An Episode in Recent Medical History." Cambridge: The Keynes Press.
- Titterton DM, Smith AFM, Makov UE (1985): "Statistical Analysis of Finite Mixture Distributions." London: John Wiley and Sons.
- Weder AB, Schork NJ (1989): Mixture analysis of erythrocyte lithium-sodium countertransport and blood pressure. *Hypertension* 13:145–150.
- Wolfe JH (1971): A Monte Carlo study of the sampling distributions of the likelihood ratio for mixtures of multinormal distributions. *Naval Personnel and Training Research Laboratory, Technical Bulletin STB* 72–2, San Diego.

**Edited by D.C. Rao**