ESTIMATION OF VARIANCE FUNCTIONS IN ASSAYS WITH POSSIBLY UNEQUAL REPLICATION AND NONNORMAL DATA

M. Davidian
Department of Statistics
Box 8203
North Carolina State University
Raleigh, North Carolina 27695-8203 USA

January, 1989

ABSTRACT

Estimation of parametric variance functions for assays relies on transformation of standard deviations based on replication at each concentration. The quality of such estimates has been shown to have direct impact on the quality of inference based on the fitted calibration curve. The theory of Davidian and Carroll (1987) is used to demonstrate that ignoring unequal replication can lead to bias and inefficiency in estimation and that weighting the common log-linear estimator to account for it can improve efficiency. A comparison of efficiency of estimation for different transformations for nonnormal distributions is given. An alternate method to account for such bias is investigated that may be useful for assays with large numbers of replicates and several outliers, and leads to a comparison of Gini's mean difference to sample standard deviation. A method for computing all of these estimators using standard software is described.

Key words: Efficiency, Gini's mean difference, Heteroscedasticity, Minimum detectable concentration, Nonnormality, Variance estimation

1. INTRODUCTION

Assay data in biological and clinical science exhibit systematic heterogeneity of variance. A nonlinear heteroscedastic regression model is fitted assuming independent responses y_{ij} at concentrations x_i, for $i = 1, \ldots, N, j = 1, \ldots, m_i$:

$$E(y_{ij}) = \mu_i = f(x_i, \beta); \ \mathrm{Var}(y_{ij}) = \sigma_i^2 = \left\{\sigma \, g(x_i, \beta, \theta)\right\}^2 \tag{1.1}$$

for regression parameter β and $q \times 1$ variance parameter θ. The variance function g often depends on β through μ_i, as for the power variance function μ_i^θ; write $g(\mu_i, \theta)$ in general. The fit $\hat{\beta}$ is by generalized least squares with weights based on estimates $\hat{\theta}$ and $\hat{\mu}_i$, and the goal is prediction, calibration and estimation of detection limits, the quality of which depends crucially on how well one estimates θ so that good estimates are essential. See Finney (1978), Carroll and Ruppert (1988), and Davidian, Carroll and Smith (1988), and Davidian (1989).

Assay responses may be "standards" with known x_i or "unknowns" with x_i unknown. Standards only may contribute to $\hat{\beta}$, but unknowns, which can comprise a large proportion of available data, also contain information about variability. Thus, estimation of θ for assays does not depend on f and uses a "regression" based on means $\bar{y}_i$ and transformations of standard deviations s_i at each x_i, as in the log-linear estimator of Rodbard and Frazier (1975) and that of Sadler and Smith (1985). Methods that use f in absolute residuals and predicted values can not use unknowns but do not require and are unaffected by replication. Often, $m_i = 2$, 3 or 4, see Raab (1981), so Davidian, et al. (1988) suggested combining both methods by weighted average since residual-based estimators can be more efficient for known x_i and small m_i. For assays with more replicates the difference may not be as profound.

Davidian and Carroll (1987) gave a general account of variance function estimation but did not explicitly indicate considerations for unequal $\{m_i\}$. Raab (1981) stated that it is common to take more replicates at higher $\{x_i\}$ where variability is greater and noted that the log-linear estimator should be biased for unequal $\{m_i\}$. These articles compared properties of different replication-based estimators for normal data only. Since good replication-based estimates are essential for assays, we investigate the effects of nonnormality and unequal $\{m_i\}$ for the general class of estimators for θ based on transformations of $\{s_i\}$. We show how not accounting for unequal $\{m_i\}$ can lead to inconsistency and inefficiency for the general class and that it is worthwhile to consider transformations of $\{s_i\}$ besides the common logs and squares if even slight deviations from normality are suspected. Correction for inconsistency would usually depend on distributional assumptions, so we investigate one way to modify the general class to account for bias without distributional knowledge that has potential to reduce bias and produce modest gains in efficiency for some problems.

In Section 2 we describe the effects of unequal replication on the general class and in Section 3 describe the modification. In Section 4 we compare different transformations of $\{s_i\}$ and the modified estimator for efficiency at nonnormal distributions. In Section 5 we report Monte Carlo results, and describe a general computational method in Section 6.

2. THE GENERAL CLASS AND UNEQUAL REPLICATION

We use the "small σ" asymptotic theory in Davidian and Carroll (1987), letting $N \to \infty$ and $\sigma \to 0$ simultaneously. "Small" σ is a good approximation to reality for most assays, see Davidian, et al. (1988). Define errors $\epsilon_{ij} = (y_{ij} - \mu_i)/\{\sigma\, g(\mu_i, \theta)\}$, $E(\epsilon_{ij}^2) = 1$, and

$$q_i^2 = (m_i - 1)^{-1} \sum_{j=1}^{m_i} (\varepsilon_{ij} - \bar{\varepsilon}_i)^2,$$

where $\bar{\varepsilon}_i$ is the mean at x_i.

Following Davidian and Carroll (1987), the general class of estimators for θ based on transformations of $\{s_i\}$ is defined as follows. Let T, $M_i(\eta,\theta,\mu_i)$, $V_i(\eta,\theta,\mu_i)$, and $H_i(\eta,\theta,\mu_i)$ be smooth functions, and let η be a general scale parameter. Then solve in η and θ

$$N^{-1/2}\sum_{i=1}^{N} H_i(\eta,\theta,\bar{y}_i)\,\{T(s_i) - M_i(\eta,\theta,\bar{y}_i)\}\,V_i^{-1}(\eta,\theta,\bar{y}_i) = 0. \qquad (2.1)$$

H_i is for us the partial derivative of M_i with respect to (η,θ), so (2.1) may be regarded as "normal equations" with "mean" M_i, "weights" V_i^{-1} and "design" H_i. First consider two common special cases of (2.1). Let β, η, and θ denote the true parameter values.

The log-linear estimator $\hat{\theta}_{LL}$ popularized in this context by Rodbard and Frazier (1975) is obtained by unweighted regression of $\log s_i$ on $\log\{\sigma\, g(\bar{y}_i, \theta)\}$, since $\log s_i = \log\sigma + \log g(\mu_i, \theta) + \frac{1}{2}\log q_i^2$, so $T(x) = \log x$, and $T(q_i) = \frac{1}{2}\log q_i^2$. For equal $\{m_i\}$ and i.i.d. $\{\varepsilon_{ij}\}$, the mean and variance of $T(q_i)$ are constant, and we expect an unbiased estimator for θ. For unequal $\{m_i\}$, both mean and variance of $T(q_i)$ will be different for different m_i, even for i.i.d. $\{\varepsilon_{ij}\}$, e.g., for normal $\{\varepsilon_{ij}\}$, $E\{T(q_i)\} = \frac{1}{2}[\psi\{(m_i - 1)/2\} - \log\{(m_i - 1)/2\}]$, $\mathrm{Var}T\{(q_i)\} = \frac{1}{4}\psi'\{(m_i - 1)/2\}$, where ψ and ψ' are the digamma and trigamma functions. Raab (1981) noted that bias should result and suggested basing the regression on $(\log s_i - b_i^*)$, where $b_i^* = E\{T(q_i)\}$ assuming normal $\{\varepsilon_{ij}\}$. In (2.1), with $\eta = \log\sigma$, $M_i = \eta + \log g(\mu_i,\theta) + b_i^*$ and $V_i \equiv 1$. For normal data, under regularity conditions as $N \to \infty$, $\sigma \to 0$, (2.1) is an unbiased estimating equation for (η,θ), but if the data are not normal and $E\{T(q_i)\} = b_i \neq b_i^*$, (2.1) is not unbiased, so there is potential for some bias. This is now a "heteroscedastic" regression problem with known variances so instead of the usual choice $V_i \equiv 1$, we ought choose $V_i = v_i^*$ for some $\{v_i^*\}$ for a weighted regression; a likely choice is $v_i^* = \mathrm{Var}\{T(q_i)\}$ for normality.

Raab (1981) proposed to estimate θ by maximizing in σ, θ, and $\{\mu_i\}$ a normal likelihood "modified" to correct bias. Sadler and Smith (1985) replaced μ_i by $\bar{y}_i$ to obtain $\hat{\theta}_{SS}$ solving (2.1) with $T(x) = x^2$, $\eta = \log \sigma$, $M_i = e^{2\eta} g^2(\mu_i,\theta)$, and $V_i = 2g^4(\mu_i,\theta)/(m_i-1)$, so $T(q_i) = q_i^2$. $E(s_i^2) = e^{\eta} g^2(\mu_i,\theta)\ E\{T(q_i)\}$, and $E\{T(q_i)\} = 1$ regardless of the distribution of $\{\epsilon_{ij}\}$, so this choice of M_i is always exact for (2.1) to be unbiased estimating equations for (η,θ), let $N \to \infty$, $\sigma \to 0$. $\mathrm{Var}(s_i^2) = e^{4\eta} g^4(\mu_i,\theta)\ \mathrm{Var}\{T(q_i)\}$, $\mathrm{Var}\{T(q_i)\} = (2+\kappa)/(m_i-1) + 2/\{m_i(m_i-1)\}$ for i.i.d. $\{\epsilon_{ij}\}$ with kurtosis κ, so for normal $\{\epsilon_{ij}\}$ ($\kappa = 0$) the "weights" V_i^{-1} are those required to account for "heteroscedasticity" induced by unequal $\{m_i\}$ in the spirit of generalized least squares. We show later that we prefer an estimator based on another transformation if $\kappa > 0$.

The examples illustrate the potential under unequal replication for inconsistency and inefficiency for the class (2.1). Define $T_\alpha(x) = x^\alpha$, $\alpha \neq 0$; $= \log x$, $\alpha = 0$. The theory of Davidian and Carroll (1987) defines $M_i = E\{T(s_i)\}$, so that as $N \to \infty$, $\sigma \to 0$ (2.1) are unbiased estimating equations for (η,θ), informally legitimizing the assumption of consistency. If we let $\eta = \log \sigma$, then $T_\alpha(s_i) = e^{\alpha\eta} g^\alpha(\mu_i,\theta)\ T_\alpha(q_i)$, $\alpha \neq 0$; $= \eta + \log g(\mu_i,\theta) + T_\alpha(q_i)$, $\alpha = 0$. Unless $\alpha = 2$, $E\{T(q_i)\}$ may depend on m_i, so for unequal $\{m_i\}$ choose $M_i = b_i^* e^{\alpha\eta} g^\alpha(\mu_i,\theta)$, $\alpha \neq 0$; $= \eta + \log g(\mu_i,\theta) + b_i^*$, $\alpha = 0$ for some $\{b_i^*\}$; the necessary definition requires $b_i^* = b_i = E\{T_\alpha(q_i)\}$. Raab's suggestion for $\hat{\theta}_{LL}$ is the choice assuming normality for $\alpha = 0$; for $\alpha \neq 0$ and i.i.d. normal $\{\epsilon_{ij}\}$, $E\{T_\alpha(q_i)\} = \Gamma\{(m_i-1+\alpha)/2\}\ 2^{\alpha/2}[\Gamma\{(m_i-1)/2\}(m_i-1)^{\alpha/2}]^{-1}$. For equal $\{m_i\}$ and i.i.d. $\{\epsilon_{ij}\}$, the common b is absorbed into a redefined scale parameter $e^\eta = b\ \sigma^\alpha$ for $\alpha \neq 0$ and replaced by 0 for $\alpha = 0$. For estimators based on residuals, a similar scale parameter is defined for i.i.d. $\{\epsilon_{ij}\}$, see Davidian and Carroll (1987). Thus, for estimation based on (2.1) with equal replication or absolute residuals, consistency ought to obtain, but ignoring the need to specify $\{b_i^*\}$ because it is not required for the former cases or incorrectly specifying $\{b_i^*\}$ for unequal $\{m_i\}$ can lead to estimation inconsistent for (η,θ).

Simulation evidence suggests that for small σ, if we choose M_i as above for some $\{b_i^*\}$, where $b_i^* \equiv 0$ for $\alpha \neq 0$ or 1 for $\alpha = 0$ correspond to ignoring unequal $\{m_i\}$, instead of the true values $\{b_i\}$, $(\hat{\eta}^*,\hat{\theta}^*)$ solving (2.1) converge in probability to some (η^*,θ^*). The role of $\sigma \to 0$ is essentially to replace $\bar{y}_i$ by μ_i, so that as $N \to \infty$, $\sigma \to 0$, we have approximately that (2.1) are unbiased estimating equations for (η^*,θ^*) if $N^{-1}\Sigma\ E\ [H_i(\eta^*,\theta^*,\mu_i)\{T(s_i)-M_i(\eta^*,\theta^*\mu_i)\}\ V_i^{-1}(\eta^*,\theta^*,\mu_i)\ ] \to 0$. For definiteness consider the power variance function. Let $\nu_i = \log \mu_i$, $A_r^\alpha = \lim N^{-1}\Sigma\ (b_i b_i^* \mu_i^{\alpha(\theta-\theta^*)} \nu_i^r/v_i^*)$, $\alpha \neq 0$; $= \lim N^{-1}\Sigma\ \{(b_i^*-b_i)\ \nu_i^r/v_i^*\}$, $\alpha = 0$, and $C_r^\alpha = \lim N^{-1}\ \Sigma\ (\ b_i^{*2}\nu_i^r/v_i^*)$, $\alpha \neq 0$; $= \lim N^{-1}\Sigma\ (\nu_i^r/v_i^*)$, $\alpha = 0$, where all limits exist.

Result 1. Let $g(\mu_i,\theta) = \mu_i^\theta$, $\eta = \log \sigma$, let M_i be as above, and let $V_i = v_i^* e^{2\alpha\eta} g^{2\alpha}(\mu_i,\theta)$, $\alpha \neq 0$; $= v_i^*$, $\alpha = 0$ for some $\{v_i^*\}$. Suppose for (2.1) with T_α, $\{b_i^*\}$, and $\{v_i^*\}$ that the solution $(\hat{\eta}^*,\hat{\theta}^*) \to (\eta^*,\theta^*)$ in probability, where $\eta^* = \eta$, $\theta^* = \theta$ if $b_i^* = b_i$. Then $N^{-1}\Sigma$ E [

$H_i(\eta^*,\theta^*,\mu_i)\{T(s_i)-M_i(\eta^*,\theta^*\mu_i)\}\ V_i^{-1}(\eta^*,\theta^*,\mu_i)\ ] \to 0$ iff θ^* satisfies

$$A_0^\alpha C_1^\alpha - A_1^\alpha C_0^\alpha = 0,\ \alpha \neq 0 \tag{2.2}$$

$$\theta^* = \theta + [\ A_1^\alpha C_0^\alpha - A_0^\alpha C_1^\alpha\ ]\ [\ C_2^\alpha\ C_0^\alpha - (C_1^\alpha)^2\ ]^{-1},\ \alpha = 0. \tag{2.3}$$

If $\alpha = 2$, the logical choice is always $b_i^* = b_i \equiv 1$, as for $\hat{\theta}_{SS}$. (2.3) may be obtained by direct evaluation of the limit in probability of $\hat{\theta}^*$ as $N \to \infty$, $\sigma \to 0$ and is a formal statement of Raab's caution of bias for $\hat{\theta}_{LL}$ when $b_i^* \equiv 0$. If the small σ assumption is valid and a unique θ^* exists, (2.2) and (2.3) may be used to assess the severity of misspecifying $\{b_i^*\}$ on point estimation. For numerical illustration, consider the data of Table 1 of Davidian, et al. (1988), for which $\theta \approx 0.472$. Ignore $x_i = 0$, and consider the common situation with m' replicates at the first $N/2$ x_i and m'' at the remainder. For unweighted log-linear estimation, $v_i^* \equiv 1$, if we use $b_i^* \equiv 0$, ignoring unequal replication, $m' = 2$ and $m'' = 3$, and the data are normal, (2.6) implies $\theta^* = \theta + 0.187$, a bias 38% of the "true" value; for $m'' = 4$, the bias is 0.232, a factor of 49%. Choosing $\{b_i^*\}$ based on normality may be reasonable even if the data are not normal; for true distribution 5% contaminated normal with contamination standard deviation 3 times that of the remaining data, $m' = 2$, $m'' = 4$, the bias is less than approximately 1%. If the true distribution is 10% contaminated normal with contamination standard deviation 5 times that of the remaining data, $m' = 2$, $m'' = 4$, the bias is roughly 5%. Davidian (1989) has cautioned that upwardly biased estimates of θ can yield estimates of minimum detectable concentration that are too optimistic in that asymptotically they can be both biased down toward 0 and less variable than estimates based on the true θ. A reasonable correction for unequal $\{m_i\}$ is thus essential. For many problems, that based on the normal distribution may be reasonable. Effect for nonnormal distributions is shown in an example in Section 5.

When (2.2) or (2.3) hold it is possible to evaluate the asymptotic normal distribution of $\hat{\theta}^*$ in Result 1. Unless the difference between θ and θ^* is small, it is unlikely that we would be satisfied with $\hat{\theta}^*$. Thus, to evaluate the choice of $\{v_i^*\}$ in V_i defined in Result 1, which arises because Var $\{T_\alpha(s_i)\} = v_i e^{2\alpha\eta} g^{2\alpha}(\mu_i,\theta)$, $\alpha \neq 0$; $= v_i$, $\alpha = 0$, $v_i = \mathrm{Var}\{T_\alpha(q_i)\}$, consider the "ideal" case in which we may essentially regard $\{b_i^*\}$ as equal to the true $\{b_i\}$, so that $\theta^* = \theta$. Let $\nu_i = \partial/\partial\theta \log(\mu_i,\theta)$ for general variance function, $\nu_i = \log \mu_i$ for power variance, and define $\tau_i = (1, \alpha^2\nu_i^T)^T$, $\alpha \neq 0$, $= (1, \nu_i^T)^T$, $\alpha = 0$. The following is immediate from Theorem 4.3 of Davidian and Carroll (1987) under the conditions stated there.

Result 2. Suppose for (2.1) with T_α, M_i with $b_i^* = b_i$ and V_i for some v_i^* as defined above, and $v_i = \mathrm{Var}\ \{T_\alpha(q_i)\}$, the solution $(\hat{\eta},\hat{\theta})$ is $N^{1/2}$-consistent as $N \to \infty$, $\sigma \to 0$. Then if $N^{1/2}\sigma \to$

λ, $0 \leq \lambda < \infty$, $\hat{\theta}$ is asumptotically normally distributed with mean $\theta + N^{-1/2}D_N$ and covariance matrix N^{-1} times the lower right $q \times q$ submatrix of

$$\left\{N^{-1}\sum_{i=1}^{N}\tau_i\tau_i^T\frac{d_i^2}{v_i^*}\right\}^{-1}\left\{N^{-1}\sum_{i=1}^{N}\tau_i\tau_i^T\frac{d_i^2 v_i}{v_i^{*2}}\right\}\left\{N^{-1}\sum_{i=1}^{N}\tau_i\tau_i^T\frac{d_i^2}{v_i^*}\right\}^{-1} \tag{2.4}$$

where $d_i = b_i$, $\alpha \neq 0$ and $d_i \equiv 1$, $\alpha = 0$ and $D_N \to 0$ in probablility if $\{\epsilon_{ij}\}$ are symmetric.

Consider symmetric errors, since our goal is simple insight. The choice of $\{v_i^*\}$ is optimized in the sense of Gauss-Markoff when $v_i^* = v_i$, e.g., $v_i = 1 - \{2/(m_i-1)\}\,[\Gamma(m_i/2)\,/\,\Gamma\{(m_i-1)/2\}]^2$ if $\alpha = 1$ and the $\{\epsilon_{ij}\}$ are i.i.d. normal. If $\{m_i\}$ are all equal, $v_i^* \equiv 1$ will suffice and comparison among estimators of the class depends only on the common true v and b, as in Davidian and Carroll (1987). For unequal $\{m_i\}$, comparisons depend on the design and $\{v_i^*\}$, but ordering of efficiency for different T will be the same as for equal $\{m_i\}$. (2.4) shows that gains may be possible by not ignoring "heteroscedasticity" in the problem induced by unequal $\{m_i\}$. In the numerical example for log-linear estimation with $m' = 2$, $m'' = 4$ and normal data, if $\{b_i^*\}$ and $\{v_i^*\}$ are correct, the asymptotic relative efficiency (ARE) of the usual log-linear estimator with $v_i^* \equiv 1$ relative to a weighted estimator with $v_i^* = v_i$ is 88%. If $\{v_i^*\}$ are based on normality but the data have the 5% and 10% contaminated normal distributions above, the AREs are approximately 90% and 96%, respectively, so if we slightly misspecify the weights, we can still gain improvement which diminishes as the deviation from normality increases. Gains of 10% to 12% may seem modest, but virtually no extra effort was required.

3. A MODIFIED ESTIMATOR

For unequal $\{m_i\}$, estimators of form (2.1) depend on making an assumption about the distribution. In many examples we have observed that the choice based on normality behaves reasonably under moderate deviations from normality, but for some designs or sufficiently nonnormal distributions there is potential for more serious inconsistency, so we consider an alternative. Suppose the $\{m_i\}$ are not all equal, and let $m^* = \min\{m_i, i = 1, \dots, N\}$. For given T, replace $T(s_i)$ in (2.1) by $T_i^U = T^U(y_{i1}, \dots, y_{im_i})$,

$$T^U(z_1, \dots, z_{m_i}) = \frac{1}{\binom{m_i}{m^*}} \sum_{\beta \epsilon B} h(z_{\beta_1}, \dots, z_{\beta_{m^*}}),$$

$$h(z_1, \dots, z_{m^*}) = T[\,\{(m^*-1)^{-1}\sum_{j=1}^{m^*}(z_j - \bar{z})^2\}^{1/2}\,],$$

where $B = \{\beta : \beta$ is one of the nonordered subsets of m^* integers chosen without replacement from $(1, \ldots, m_i)\}$. Thus, $T_i^U = T(s_i)$ if $m_i = m^*$. If $m_i \neq m^*$, T_i^U is the average over all possible subsets of $\{y_{i1}, \ldots, y_{im_i}\}$ of size m^* of T on the subset, e.g., if $T(x) = x$, T_i^U is the average of the standard deviations of each possible subset. If the $\{\epsilon_{ij}\}$ are i.i.d., $E\{T^U(\epsilon_{i1}, \ldots, \epsilon_{im_i})\} = a$, a constant for all i, so for T_α, (2.1) with $T(s_i)$ replaced by T_i^U and $M_i = e^\eta g^\alpha(\mu_i,\theta)$, $e^\eta = \sigma^\alpha a$, $\alpha \neq 0$; $= \eta + \log g(\mu_i,\theta)$, $\eta = \log \sigma$, $\alpha = 0$, should be unbiased as $N \to \infty$, $\sigma \to 0$, and the estimator should be consistent under regularity conditions. Consistency is obtained directly for power variance and $\alpha = 0$. Since T^U is a U-statistic with symmetric kernel h, see Serfling (1980, Ch. 5), call this the U-statistic transformation estimator based on T. By an argument similar to that of Theorem 4.3 of Davidian and Carroll (1987), we have

Result 3. For i.i.d. $\{\epsilon_{ij}\}$, suppose $(\hat\eta^U, \hat\theta^U)$ solve (2.1) with $T(s_i)$ replaced by T_i^U and are $N^{1/2}$-consistent as $N \to \infty$, $\sigma \to 0$. Then if $N^{1/2}\sigma \to \lambda$, $0 \leq \lambda < \infty$, with M_i as above,

$$B_N N^{1/2}\begin{pmatrix}\hat\eta^U - \eta \\ \hat\theta^U - \theta\end{pmatrix} = N^{-1/2}\sum_{i=1}^{N} C_i^U + \lambda D_N + o_p(1), \tag{3.1}$$

$B_N = N^{-1}\Sigma H_i H_i^T V_i^{-1}$, $C_i^U = (T_i^U - M_i) H_i V_i^{-1}$, and $D_N \to 0$ in probablility for symmetric $\{\epsilon_{ij}\}$.

From (3.1), for symmetry, if $V_i = u_i^* g^{2\alpha}(\mu_i,\theta)$, $\alpha \neq 0$; $= u_i^*$, $\alpha = 0$, for some $\{u_i^*\}$ and $u_i = \mathrm{Var}\{T^U(\epsilon_{i1}, \ldots, \epsilon_{im_i})\}$, $\hat\theta^U$ is asymptotically normal with mean θ and covariance matrix (2.4) with v_i and v_i^* replaced by u_i and u_i^* and $d_i \equiv \alpha$, $\alpha \neq 0$; $d_i \equiv 1$, $\alpha = 0$. To "optimize" choice of $\{u_i^*\}$ requires knowledge of the error distribution, but inconsistency is not a problem. From Serfling (1980, p. 183), with $m^* = 2$, $u_i = 2\{2(m_i-2)\zeta_1 + \zeta_2\}/\{m_i(m_i-1)\}$, $\zeta_c = \mathrm{Cov}\{ h(\epsilon_{i\beta_1}, \ldots, \epsilon_{i\beta_{m^*}}), h(\epsilon_{i\delta_1}, \ldots, \epsilon_{i\delta_{m^*}}) \}$, $\beta, \delta \in B$ with c common integers, for evaluation of u_i for given T and distribution, e.g., for normality and $\alpha = 1$, $\zeta_1 = 0.0814$, $\zeta_2 = 0.3634$.

4. NONNORMALITY

Result 1 gives a way to evaluate seriousness of misspecification of $\{b_i^*\}$ in (2.1) based on $T(s_i)$ for specific design and error distribution. We compare theoretical efficiencies of estimators for nonnormal distributions using (2.4) for the ideal case with $\{b_i^*\}$, $\{v_i^*\}$, and $\{u_i^*\}$ correctly specified for the true distribution for various m_i. In practice, these would be chosen based on normality, so comparisons could be optimistic. For unequal $\{m_i\}$, actual efficiencies also depend on the design. For nonnormal distributions, $\{b_i\}$ and $\{v_i\}$ were evaluated by averaging 20 Monte Carlo experiments based on 5000 replications each, while evaluation of these values for normality and $\{u_i\}$ and a for all distributions in Table 2 was exact. Thus, values given are approximate to the extent of the error in the Monte Carlo values; the largest

Monte Carlo coefficient of variation was 4.5%, with average of about 2%.

Table 1 lists AREs for estimators based on (2.1) with $T(x) = \log x$ and x relative to using $T(x) = x^2$ for contaminated normal distributions, so is the approximate analogue of Table 2 of Davidian and Carroll (1987) for estimators based on transformations of $\{s_i\}$. As m_i increases and the proportion of "bad" data increases, the superiority of these transformations is evident. The identity transformation seems preferable to log for slight deviations from normality, and conversely for more profound deviations. Thus, with or without unequal replication, for even slight deviations from normality, these transformations may be preferred to the squares of Sadler and Smith (1985). The identity transformation does not do too badly at normality, particularly for higher numbers of replicates.

To compare the U-statistic transformation estimator to that based on $T(s_i)$, let $m^* = 2$ and $\alpha = 1$; one could prepare a similar table for other α and m^*. Then $h(z_1,z_2) = | z_1 - z_2 | / \sqrt{2}$, so T_i^U is like replacing s_i by Gini's mean difference in (2.1), see Johnson and Kotz (1970, p. 67), which is noted for high efficiency relative to sample standard deviation as an estimator of population standard deviation at normality. Table 2 lists AREs for various m_i for using s_i in (2.1) relative to T_i^U with $\alpha = 1$ and so is of independent interest as a comparison of Gini's mean difference to sample standard deviation for nonnormality; $\gamma = 0.00$ corresponds to the second line of Table 7 of Johnson and Kotz (1970, p. 70). For small numbers of replicates and nonnormality, the advantage in efficiency of T_i^U relative to T is negligible. Gains of roughly 20% are possible for large m_i and moderate deviation from normality, suggesting that usefulness of the estimator may be limited to such cases. Gini's mean difference as an estimator of standard deviation may only be preferred for fairly large data sets with several outliers, despite promising asymptotics ($m_i \to \infty$). A table for $\alpha = 0$ shows conditions favoring T_i^U must be even more extreme although this may perform better than $\alpha = 1$ for some cases. Properties for $m^* > 2$ are more difficult but could be investigated by simulation.

5. SIMULATION RESULTS

We ran several simulations based on the data in Table 1 of Davidian, et al. (1988) with f as in their equation (1.2) and power variance, $\beta_1 = 29.52744$, $\beta_2 = 1.88638$, $\beta_3 = 1.57933$, $\beta_4 = 1.00218$, $\sigma = 0.08718$, and $\theta = 0.7$. Estimates of θ were constrained to lie in [0, 1.5], and each study used 500 data sets. There were three error distributions: contaminated normal with variance standardized to 1 with contamination proportion γ and contamination standard deviation b with $\gamma = 0.05$ and b=3 and $\gamma = 0.10$ and $b = 5$ to represent moderate and more substantial nonnormality, and standard normal, $\gamma = 0.00$. In Table 3, LL, AB, and SS correspond to T_α with $\alpha = 0$, 1, and 2, respectively, with "nc" denoting ignoring unequal $\{m_i\}$

in choice of $\{b_i^*\}$, "U" the U-statistic estimator with $m^* = 2$ and $\{u_i^*\}$ based on normality, and "b" correction for bias and "bv" correction for bias and "heteroscedasticity" based on normality. The first situation had 12 design points, starting with 0.000 and then every other x in Table 1 of Davidian et al. (1988), with $m_i = 2$ for the first 6 $\{x_i\}$, $m_i = 3$ or 4 for the remainder. The second 3 columns had these $\{x_i\}$ with $m_i = m \equiv 4$, and the final column $m_i = m \equiv 8$ with 7 $\{x_i\}$ including 0.000 for a situation with a large number of replicates.

The top part of the Table gives Monte Carlo bias. Ignoring unequal $\{m_i\}$ yields bias on the order suggested in Result 1. The U-statistic is less biased when we would expect, although the significance of numerical values as these depends on context. The lower portion gives efficiencies based on Monte Carlo mean squared error (MSE). The bottom two rows show qualitatively the behavior predicted by Table 1. Note how $\alpha = 1$ produces negligible efficiency loss at normality vs. $\alpha = 2$. The first row shows a roughly 10% gain for simply weighting the log-linear estimator. The third row compares favorably with behavior predicted by Table 2. The second and third rows taken with the bias results show that the extra effort required to compute the U-statistic transformation might be justified only for situations with fairly substantial nonnormality or large numbers of replicates, but it will not do worse.

6. A COMPUTATIONAL METHOD

A method to compute the general classes of estimators (2.1) with $T(s_i)$ or T_i^U can be based on a procedure proposed by Giltinan and Ruppert (1988) in a similar context and also described by Carroll and Ruppert (1988). The principle may be used for general T and similarly to compute estimates based on general transformations of absolute residuals.

Let $T_i = T(s_i)$ or T_i^U based on T_α for given $\alpha \neq 0$. Computation of T_i^U involves more work but need only be done once. Johnson and Kotz (1970, p. 67) give a simpler formula for $\alpha = 1$, $m^* = 2$. Let $M_i(\eta,\theta,\bar{y}_i) = b_i^* e^{\eta\alpha} g^\alpha(\bar{y}_i,\theta)$ and $V_i(\eta,\theta,\bar{y}_i) = v_i^* g^{2\alpha}(\bar{y}_i,\theta)$ for some $\{b_i^*\}$ and $\{v_i^*\}$. For $T(s_i)$, $e^\eta = \sigma$; for T_i^U, always set $b_i^* \equiv 1$ and then $e^\eta = \sigma\, b^{1/\alpha}$, $b = E\{T_\alpha(q_{m^*})\}$. Let

$$L = \prod_{i=1}^{N} \{ e^{\eta\alpha} g^\alpha(\bar{y}_i,\theta)\sigma\}^{-b_i^{*2}/(\alpha v_i^*)} \exp\left\{ - \frac{T_i\, b_i^*}{\alpha\, v_i^*\, e^{\eta\alpha}\, g^\alpha(\bar{y}_i,\theta)} \right\}.$$

Setting partial derivatives of log L with respect to η and θ equal to 0 yields (2.1) with $T(s_i)$ replaced by T_i and $H_i = \partial M_i/\partial [\eta\ \theta^T]^T$, so maximizing log L is like solving the appropriate equations. With $d_i = b_i^{*2}/v_i^*$ and $\dot{g}_\alpha = \{ \Pi\, g(\bar{y}_i,\theta)^{\alpha d_i/\Sigma d_j}\}^\theta$, log L maximized over η is, ignoring constants, $-(1/\alpha)\,(\Sigma\, d_i) \log [\, \Sigma\, T_i\, b_i^* \dot{g}_\alpha/\{ v_i^*\, g^\alpha(\bar{y}_i,\theta)\}]$. Thus, to maximize L, minimize

$$\sum_{i=1}^{N} \left(\left\{ \frac{T_i\, b_i^*\, \dot{g}_\alpha}{v_i^*\, g^\alpha(\bar{y}_i,\theta)} \right\}^{1/2} \right)^2.$$

One can use a nonlinear regression program, regressing a dummy variable identically 0 on $[\, T_i\, b_i\, \dot{g}_\alpha / \{ v_i\, g^\alpha(\bar{y}_i, \theta) \}]^{1/2}$. This must allow multiple passes through the data to compute general $\dot{g}_\alpha$; for $g(\mu_i, \theta) = \mu_i^\theta$, this can be done external to the program, since then $\dot{g}_\alpha = (\Pi\, \bar{y}_i^{\alpha d_i / \Sigma d_j})^\theta$.

7. DISCUSSION

For estimation of variance functions in assays by methods based on transformations of within-concentration standard deviations, care must be taken to account for unequal replication. Basing this on the normal distribution will often suffice. Modifying log-linear estimation to be a weighted regression with known weights for unequal replication produces some gain in efficiency with minimal effort. An alternate way to account for unequal replication without distributional assumptions met with limited success in cases we considered. It may be useful for problems with large numbers of replicates or several "unusual" observations in which cases it will do no worse and has potential for moderate gains in efficiency and bias reduction. Efficiency of estimation for equal or unequal replication is sensitive to choice of transformation when the data are even slightly nonnormal. The Sadler and Smith (1985) estimator is favored for strict normality, but other transformations may outperform it at slight deviations from normality, depending on numbers of replicates. The identity transformation is particularly promising. In practice, one may consider our results as a rough guide to choosing a method when taken with faith in the normal assumption. For data with several "unusual" observations, the identity transformation may be safest.

In some instances the assumption of independent responses may not be appropriate, as when "replicates" at a concentration are really subsamples from a "batch." A components-of-variance model with possibly both components heteroscedastic would be more suitable. Recent personal communications with statisticians in the pharmaceutical industry suggest that some scientists have been successfully encouraged to abandon the "batch" approach when possible and conduct the assay by design to avoid problems of intra-batch correlation.

REFERENCES

Carroll, R.J. and Ruppert, D. (1988). *Transformation and Weighting in Regression.* New York: Chapman and Hall.

Davidian, M. (1989). A note on estimating minimum detectable concentration. Preprint.

Davidian, M. and Carroll, R.J. (1987). Variance function estimation. *J. Am. Statist. Assoc.* 82, 1079-91.

Davidian, M., Carroll, R.J., and Smith, W. (1988). Variance functions and the minimum detectable concentration in assays. *Biometrika* 75, 549-56.

Finney, D.J. (1978). *Statistical Method in Biological Assay,* 3rd ed. High Wycombe, Bucks.: Griffin.

Giltinan, D.M. and Ruppert, D. (1988). Fitting heteroscedastic regression models to individual pharmacokinetic data using standard statistical software. *Journal of Pharmacokinetics and Biopharmaceutics*, to appear.

Johnson, N.L. and Kotz, S. (1970). *Continuous Univariate Distributions, Vol. I.* New York: John Wiley and Sons.

Raab, G.M. (1981). Estimation of a variance function, with application to radioimmunoassay. *Appl. Statist.* 30, 32-40.

Rodbard, D. and Frazier, G.R. (1975). Statistical analysis of radioligand assay data. *Meth. Enzy.* 37, 3-22.

Sadler, W.A. and Smith, M.H. (1985). Estimation of the response-error relationship in immunoassay. *Clin. Chem.* 31/11, 1802-5.

Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics.* New York: John Wiley and Sons.

Table 1.

Approximate asymptotic relative efficiencies for using T(x) in (2.1) relative to x^2 for contaminated normal distributions with contamination proportion γ and standard deviation 3 for a function T of numbers of replicates m_i.

	T(x)									
	log x					x				
m_i	2	3	4	9	10	2	3	4	9	10
γ										
0.000	0.40	0.61	0.71	0.88	0.89	0.88	0.92	0.94	0.97	0.98
0.002	0.44	0.68	0.81	0.99	1.01	0.93	1.00	1.03	1.07	1.07
0.010	0.56	0.92	1.09	1.32	1.34	1.14	1.25	1.28	1.28	1.27
0.050	0.85	1.37	1.60	1.67	1.64	1.45	1.58	1.53	1.41	1.38
0.100	0.87	1.40	1.55	1.47	1.46	1.42	1.45	1.45	1.26	1.26

Table 2.

Approximate asymptotic relative efficiencies for contaminated normal distributions with contamination proportion γ and standard deviation 3 for $T(x) = x$ of standard deviation relative to Gini's mean difference. The values for $\gamma = 0.000$ and $m_i \to \infty$ are exact.

γ	m_i						
	2	3	4	5	9	10	∞
0.000	1.000	1.008	1.013	1.015	1.019	1.019	1.023
0.002	1.000	1.002	1.009	1.006	0.998	0.996	0.900
0.010	1.000	0.997	0.977	0.969	0.919	0.904	0.681
0.050	1.000	0.995	0.934	0.933	0.814	0.795	0.558
0.100	1.000	0.972	0.952	0.941	0.798	0.797	0.620

Table 3

Monte Carlo Results

	$m'=2$, $m''=3/4$, N=12						m=4, N=12			m=8 N=7
	$\gamma=0.00$		$\gamma=0.05$ b=3		$\gamma=0.10$ b=5		$\gamma=0.00$	$\gamma=0.05$ b=3	$\gamma=0.10$ b=5	$\gamma=0.05$ b=3
Monte Carlo bias as % of true θ										
sign of bias	+	+	+	+	+	+	+	+	+	−
LL_{nc}	25.3	32.6	26.1	32.1	26.6	38.2	*n/a*	*n/a*	*n/a*	*n/a*
AB_{nc}	10.4	16.1	11.3	15.3	11.8	22.3	*n/a*	*n/a*	*n/a*	*n/a*
LL_{bv}	1.0	0.7	2.2	0.5	3.0	7.0	2.1	1.7	3.5	0.9
LL_u	1.5	1.3	1.8	0.6	0.9	3.6	2.3	2.0	3.9	0.1
AB_{bv}	3.9	4.4	4.9	3.0	5.3	9.5	2.2	2.2	3.9	0.9
AB_u	1.9	4.5	2.7	2.7	2.7	8.1	2.3	2.3	4.0	0.6
SS	3.0	6.3	3.2	4.7	4.2	10.5	2.4	2.5	4.1	1.0
Efficiency based on Monte Carlo MSE (%)										
LL_b rel. LL_{bv}	91.5	88.9	98.0	88.5	96.6	100.5	*n/a*	*n/a*	*n/a*	*n/a*
LL_{bv} rel. LL_u	105.8	106.5	104.3	104.9	99.1	95.0	133.5	122.9	90.8	95.6
AB_{bv} rel. AB_u	99.2	100.2	98.1	99.2	97.4	95.0	102.1	97.0	90.4	88.2
SS rel. LL_{bv}	164.5	144.1	140.9	121.2	74.7	80.4	128.5	82.1	56.2	83.4
SS rel. AB_{bv}	103.1	101.3	94.5	90.8	79.8	75.3	104.9	79.9	70.7	88.2

n/a: not applicable