Published in final edited form as:

Biometrics. 2011 December ; 67(4): 1461–1470. doi:10.1111/j.1541-0420.2011.01569.x.

A Moment Adjusted Imputation Method for Measurement Error Models

Laine Thomas^{1,*}, Leonard Stefanski^{2,**}, and Marie Davidian^{2,***}

¹Department of Biostatistics and Bioinformatics, Duke University, Durham, North Carolina 27705, U.S.A.

²Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695-8203, U.S.A.

Summary

Studies of clinical characteristics frequently measure covariates with a single observation. This may be a mis-measured version of the “true” phenomenon due to sources of variability like biological fluctuations and device error. Descriptive analyses and outcome models that are based on mis-measured data generally will not reflect the corresponding analyses based on the “true” covariate. Many statistical methods are available to adjust for measurement error. Imputation methods like regression calibration and moment reconstruction are easily implemented but are not always adequate. Sophisticated methods have been proposed for specific applications like density estimation, logistic regression, and survival analysis. However, it is frequently infeasible for an analyst to adjust each analysis separately, especially in preliminary studies where resources are limited. We propose an imputation approach called Moment Adjusted Imputation (MAI) that is flexible and relatively automatic. Like other imputation methods, it can be used to adjust a variety of analyses quickly, and it performs well under a broad range of circumstances. We illustrate the method via simulation and apply it to a study of systolic blood pressure and health outcomes in patients hospitalized with acute heart failure.

Keywords

Conditional score; Measurement error; Non-linear models; Regression calibration

1. Introduction

In clinical studies, biological covariates are often measured only at baseline, and this measurement includes noise due to natural fluctuations or other sources. The quantity of interest may be the average over fluctuations. For example, using data from the Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF) registry, Gheorgiade et al. (2006) studied the relationship between mortality and systolic blood pressure at hospital admission in patients with acute heart failure, in-hospital by fitting logistic regression models and post-discharge via Cox proportional hazard models. Blood pressure was determined by a single, in-hospital measurement; however, many studies have demonstrated large fluctuations in systolic blood

*laine.thomas@duke.edu

**stefanski@stat.ncsu.edu

***davidian@ncsu.edu

Supplementary Materials: Web Appendices A-G and R code implementing the methods are available under the Paper Information link at the *Biometrics* website <http://www.biometrics.tibs.org>.

pressure, and the average of many longitudinal measurements is more strongly correlated with outcomes (Brueren et al., 1997; Pickering et al., 2005; Marshall, 2008). In other words, outcomes are more directly related to an underlying blood pressure level, averaged over fluctuations. Similarly, in descriptive analysis, an unobserved “true” value may be more relevant than a noisy baseline measure.

Measurement error in covariates can also be attributed to device error, assay error, and instrumentation. Complications arise when an imperfect measurement, W , is observed in place of a latent variable, X , and desired inferences involve X . When interest focuses on estimation of the density of X , this presents an obvious problem because the mis-measured data, W , are over-dispersed relative to the distribution of X (Eddington, 1940; Tukey, 1974). When X is a covariate in a regression model for an outcome Y , estimators obtained when W is substituted for X may be substantially biased (Liu et al., 1978) and statistical power may be compromised (Freudenheim and Marshall, 1988; Carroll et al., 2006).

Many strategies to adjust for measurement error depend on the specific modeling or regression context. Correction for measurement error in linear model covariates is commonly achieved by regression calibration (RC) (Carroll and Stefanski, 1990; Gleser, 1990), which substitutes an estimate of the conditional mean $E(X|W)$ for the unknown X . The resulting linear regression estimates the underlying parameters of interest. When the distribution of X is of interest, $E(X|W)$ is over-corrected in terms of having reduced spread (Eddington, 1940; Tukey, 1974). Regression calibration is also implemented in non-linear models because of its simplicity but is typically most effective for general linear models when the measurement error is not large (Rosner, Spiegelman, and Willett, 1989; Carroll et al., 2006). Other approaches for non-linear models are described by Carroll et al. (2006) and include structural models, which regard X as a random variable, and functional models in which X is treated as a fixed parameter. Structural methods like maximum likelihood yield efficient estimation but require that the density of X , $f_X(x)$, be known or well approximated. Conditional score methods use estimating equations derived from the distribution of the observed data conditional on sufficient statistics for the unobserved X and include estimators that are efficient among functional methods. These have been developed for generalized linear models (Stefanski and Carroll, 1985, 1987), survival analysis (Tsiatis and Davidian, 2001), and joint models for longitudinal data and a primary endpoint (Li, Zhang, and Davidian, 2004).

The preceding methods target estimation of parameters in a specific regression context; a different method must be implemented for every type of regression model in which X is used. This would be burdensome in the OPTIMIZE-HF study, where the mis-measured variable is used in multiple analyses. An alternative approach is to focus on re-creating the true X from the observed W , at least approximately, as the primary quantity of interest or as a means to improving parameter estimation. This has been explored from a Bayesian perspective (Tukey, 1974; Louis, 1984; Shen and Louis, 1998; Freedman et al., 2004). Freedman et al. (2004) aim to replace the mis-measured data W with estimators that have asymptotically the same joint distribution with Y as does X . They implement a more practical approximation to this idea by focusing on only the first two moments of the joint distribution. Their moment reconstruction (MR) method is based on adjusted data, X_{MR} , defined so that $E(X_{MR})$, $\text{Var}(X_{MR})$ and $\text{Cov}(X_{MR}, Y)$ are equal to the corresponding moments of X . They impute estimates, $\hat{X}_{MR}$, in a range of applications and demonstrate that this approach yields good results for normally distributed X (Freedman et al., 2004, 2008).

When X has a normal distribution, it may suffice to match two moments; bias in linear and logistic regression parameter estimators will be eliminated. More generally, this is not adequate, and an extension of MR to higher-order moments and cross-products is suggested.

To the best of our knowledge, the idea of computing higher-order, moment-adjusted estimates of the true X originated with the unpublished dissertation research of Bay (1997). We expand on this work, calling the method Moment Adjusted Imputation (MAI). Our method retains the convenience of other imputation methods, in that, once the adjusted values are obtained, it can be used across a variety of analyses on the same data set using standard software.

In this paper, we demonstrate the benefit of MAI, particularly for density estimation and logistic regression, where X is non-normal. In Section 2, we define the MAI algorithm and relate it to other imputation methods. In Sections 3 and 4, we compare adjustment procedures for estimating kernel density and non-linear regression coefficients by simulation studies, respectively. We adjust the previous OPTIMIZE-HF analysis to account for measurement error and obtain estimates that describe the features of “true” blood pressure in Section 5.

2. The Moment Adjusted Imputation Method

We describe the method of Bay (1997). Consider mis-measured observations W_i , $i = 1, \dots, n$, assumed independent across i , where $W_i = X_i + U_i$ for $U_i \sim N(0, \sigma_{ui}^2)$, $\mathbf{U} = (U_1, \dots, U_n)^T$ independent of $\mathbf{X} = (X_1, \dots, X_n)^T$, and U_i are mutually independent. The W_i may be mis-measured versions of a scalar covariate X_i in a regression model or subject-specific estimates of scalar random effects X_i in a mixed effects model with estimation uncertainty represented by U_i . No assumptions are made about the unobservable latent variables $X_1, \dots, X_n$; they could be an independent, identically distributed (iid) random sample from some unknown distribution, as in a structural model, or fixed constants. We focus on the iid case where notation is simpler. Assume that σ_{ui} are known, as is common in measurement error models.

The objective is to construct adjusted versions of the W_i , $\hat{X}_i$, say, where the first M sample moments of $\hat{X}_i$ unbiasedly estimate the corresponding moments of X_i ; i.e.,

$E(n^{-1} \sum_{i=1}^n \hat{X}_i^r) = E(X^r)$, $r = 1, \dots, M$. The distribution of $\hat{X}_i$ thus approximates that of X_i up to M moments. We can also match cross-product moments between model variables. Consider the simple linear regression model, $E(Y_i|X_i) = \beta_0 + \beta_X X_i$. The naive estimator of $\beta = (\beta_0, \beta_X)^T$, based on W_i , is $\hat{\beta}^N = A^{-1}(\sum_i Y_i, \sum_i W_i Y_i)^T$, where $\text{vech}(A) = (n, \sum_i W_i, \sum_i W_i^2)^T$.

Inspection of $\hat{\beta}^N$ suggests matching the first two moments of X_i and the cross-product with Y_i so that $M = 2$ and $E(n^{-1} \sum_{i=1}^n \hat{X}_i Y_i) = E(XY)$. Freedman et al. (2004) match these same moments. It is straightforward to show that $\hat{\beta} = A^{-1}(\sum_i Y_i, \sum_i \hat{X}_i Y_i)^T$, where

$\text{vech}(A) = (n, \sum_i \hat{X}_i, \sum_i \hat{X}_i^2)^T$, is consistent for β . Similarly, in the multiple linear regression model $E(Y_i|X_i, Z_i) = \beta_0 + \beta_X X_i + \beta_Z Z_i$, with error-free covariate Z_i , both β_X and β_Z are consistently estimated if $\hat{X}_i$ also satisfy $E(n^{-1} \sum_{i=1}^n \hat{X}_i Z_i) = E(XZ)$. In non-linear models, where parameter estimators depend on higher-order moments and cross-products, we propose to match these as well.

In general, we wish to find $\hat{X}_i$ with $E(n^{-1} \sum_{i=1}^n \hat{X}_i^r V_{ik}) = E(X^r V_k)$ for $r = 1, 2, \dots, M_k$, where V_{ik} is the (i, k) element of $\mathbf{V} = (\mathbf{1}, \mathbf{Y}, \mathbf{Z})$, $\mathbf{1}$ is a $n \times 1$ vector of ones, $\mathbf{Y} = (Y_1, \dots, Y_n)^T$, and $\mathbf{Z}$ is a $n \times (K - 2)$ matrix whose columns are the values of $K - 2$ error-free covariates for $i = 1, \dots, n$ and $k = 1, \dots, K$. Because $\mathbf{V}$ includes a vector of ones, matching cross-products with the columns of $\mathbf{V}$ includes matching moments. We make the common *surrogacy* assumption that W_i is conditionally independent of V_{ik} given X_i (Carroll et al. 2006, Section 2.5) and use this in the following implementation. Note that $\mathbf{V}$ could be defined to include higher powers

of the components Y and Z , thus matching moments of the form $E(X^r V_k^r)$. We do not take this approach here, as we have found it unnecessary to achieve good results.

2.1 Implementation

The first step is to find estimators $\hat{m}_{rk}$ so that $E(\hat{m}_{rk}) = E(X^r V_k)$, $k = 1, \dots, K$. Based on the normality of U_i , unbiased estimators for the moments of X_i can be found as follows. Define Hermite polynomials by the recursive formulæ $H_0(z) = 1$, $H_1(z) = z$, $H_r(z) = zH_{r-1}(z) - (r-1)H_{r-2}(z)$ for $r \geq 2$ (Cramer, 1957). Stulajter (1978) proved that, if $W \sim N(\mu, \sigma^2)$, then $E\{\sigma^r H_r(W/\sigma)\} = \mu^r$ (Stefanski, 1989; Cheng and Van Ness, 1999). Letting

$P_{ri}(w) = \sigma_{ui}^r H_r(w/\sigma_{ui})$, we have $E\{P_{ri}(W_i)|X_i\} = X_i^r$. The estimators $\hat{m}_{rk} = n^{-1} \sum_{i=1}^n P_{ri}(W_i) V_{ik}$ are unbiased under the surrogacy assumption because

$$E\{E(\hat{m}_{rk}|\mathbf{X}, \mathbf{V})\} = E[n^{-1} \sum_{i=1}^n E\{P_{ri}(W_i)|X_i\} V_{ik}] = n^{-1} \sum_{i=1}^n E(X_i^r V_{ik}) = E(X^r V_k).$$

The adjusted $\hat{X}_i$ are obtained by minimizing $\sum_{i=1}^n (W_i - X_i)^2$ subject to constraints on the moments and cross-products. For the k^{th} column of $\mathbf{V}$, M_k constraints are imposed. For a particular matrix $\mathbf{V}$, the vector $\mathbf{M} = (M_1, \dots, M_K)$ describes the number of cross-products matched with each of its columns. Using Lagrange multipliers $(\lambda_{11}, \dots, \lambda_{M_k K}) = \mathbf{\Lambda}$, the objective function is

$$Q_{MK}(X_1, \dots, X_n, \mathbf{\Lambda}) = n^{-1} \sum_{i=1}^n \frac{1}{2} (W_i - X_i)^2 + \sum_{k=1}^K \sum_{r=1}^{M_k} \frac{\lambda_{rk}}{r} (n^{-1} \sum_{i=1}^n X_i^r V_{ik} - \hat{m}_{rk}). \quad (1)$$

We take the derivative of Q_{MK} with respect to $(X_1, \dots, X_n, \mathbf{\Lambda})$, equate this to 0, and solve for $(\hat{X}_1, \dots, \hat{X}_n, \hat{\mathbf{\Lambda}})$ by Newton-Raphson (see Web Appendix A). The resulting adjusted data are defined implicitly as $\hat{X}_i = h(W_i, \mathbf{V}_i, \hat{\mathbf{\Lambda}})$. The solution $\hat{X}_i$ is then substituted for X_i in the standard methods of estimation that would be performed if X_i were observed.

In a simple case it is possible to obtain an analytical solution that minimizes objective function (1). For a scalar X in the absence of additional covariates, the estimator that matches two moments is $\hat{X}_i = W_i \hat{\alpha} + \bar{W}(1 - \hat{\alpha})$, where $\bar{W} = n^{-1} \sum_{i=1}^n W_i$, $\hat{\alpha} = (\hat{\sigma}_x^2 / \hat{\sigma}_w^2)^{1/2}$, $\hat{\sigma}_w^2 = n^{-1} \sum_{i=1}^n (W_i - \bar{W})^2$, $\hat{\sigma}_x^2 = \hat{\sigma}_w^2 - \bar{\sigma}_u^2$ and $\bar{\sigma}_u^2 = n^{-1} \sum_{i=1}^n \sigma_{ui}^2$. It is easy to see that $\hat{X}_i$ are not independent because they depend on estimated moments. In the general case, $\hat{X}_i = h(W_i, \mathbf{V}_i, \hat{\mathbf{\Lambda}})$ are dependent for the same reason. In applications where $\hat{X}_i$ is substituted for X_i , the usual standard errors, which assume independent data, should not be used. We recommend that standard errors for analyses involving $\hat{X}_i$ be obtained by bootstrapping. In specific applications, below, we derive modifications of the usual standard errors.

2.2 Implications

The MAI estimator, $\hat{X}_i = W_i \hat{\alpha} + \bar{W}(1 - \hat{\alpha})$, is familiar. A very similar estimator was developed via a multi-stage loss function by Louis (1984), with the intention of matching only two

moments. Except for the exponent in $\hat{\alpha} = (\hat{\sigma}_x^2 / \hat{\sigma}_w^2)^{1/2}$, $\hat{X}_i$ resembles the empirical Bayes estimator, which is known to have variance smaller than the posterior expected variance (Louis, 1984). MAI puts more weight on W_i than does empirical Bayes and thus provides an alternative to empirical Bayes when one is concerned about the problem of over-shrinkage.

MAI maintains the desirable properties of RC and MR with additional benefits. In fact, in simple linear regression we can replicate the RC and MR parameter estimates by matching two moments and a cross-product with the response, so $\mathbf{M} = (M_1, M_2) = (2, 1)$ for $\mathbf{V} = (\mathbf{1}, \mathbf{Y})$.

The adjusted data are $\widehat{X}_i = W_i \widehat{c} + \overline{W}(1 - \widehat{c}) + (Y_i - \overline{Y})(\widehat{\sigma}_{wy}/\widehat{\sigma}_y^2)(1 - \widehat{c})$, where $\overline{Y} = n^{-1} \sum_{i=1}^n Y_i$, $\widehat{\sigma}_y^2 = n^{-1} \sum_{i=1}^n (Y_i - \overline{Y})^2$, $\widehat{c} = \{(\widehat{\sigma}_x^2 \widehat{\sigma}_y^2 - \widehat{\sigma}_{wy}^2)/(\widehat{\sigma}_w^2 \widehat{\sigma}_y^2 - \widehat{\sigma}_{wy}^2)\}^{1/2}$, and $\widehat{\sigma}_{wy} = n^{-1} \sum_{i=1}^n (W_i - \overline{W})(Y_i - \overline{Y})$. The estimator $\widehat{\beta}_x = \sum_{i=1}^n (\widehat{X}_i - \overline{\widehat{X}})(Y_i - \overline{Y}) / \sum_{i=1}^n (\widehat{X}_i - \overline{\widehat{X}})^2$ is equivalent to the RC estimator, $(\widehat{\sigma}_w^2 / \widehat{\sigma}_x^2) \widehat{\beta}_x^N$, which is also identical to MR (Bay, 1997; Freedman et al., 2004). In addition, the $\widehat{X}_i$ converge in probability to

$X_i^* = W_i c + E(X)W_i(1 - c) + \{Y_i - E(Y)\}(\sigma_{wy}/\sigma_y^2)(1 - c)$, where $c = \{(\sigma_x^2 \sigma_y^2 - \sigma_{wy}^2)/(\sigma_w^2 \sigma_y^2 - \sigma_{wy}^2)\}^{1/2}$. (X^*, Y) has the same distribution as (X, Y) when this is multivariate normal. This allows for consistent estimation of the regression error variance (Freedman et al., 2004). RC and MR are not identical in logistic regression, where only the later may be consistent under normality (Freedman et al., 2004). When Y is binary and $(X|Y)$ is normal, letting $\mathbf{M} = (2, 2)$ produces $\widehat{X}$ that converge to some X^{**} , and (X^{**}, Y) has the same distribution as (X, Y) . As with MR, this leads to consistency of parameter estimation in linear logistic regression, but also in quadratic logistic regression (see Web Appendix B). We do not expect parameter estimators based on imputation methods to be consistent outside of specific cases in linear and logistic regression. However, substantial reduction in bias may be achieved with reasonable convenience.

The proposed method replicates the adjusted estimator of Cheng and Schneeweiss (1998) for polynomial regression, which is obtained without adjusting data. In polynomial regression, there is a closed form solution for the coefficients that depends only on sample moments and cross-products, so the unbiased estimators $\widehat{m}_{rk}$ can be substituted directly into this solution. These authors use the same Hermite polynomials as a method to obtain $\widehat{m}_{rk}$ for normally distributed measurement error; because MAI creates adjusted data with these unbiased moments, it replicates theirs. For consistent estimation of a quadratic polynomial regression in X it is necessary to substitute unbiased estimators for four moments and second order cross-products, i.e. $\mathbf{M} = (4, 2)$ for $\mathbf{V} = (\mathbf{1}, \mathbf{Y})$. Although closed form solutions are not available for many non-linear models, non-linearity is often well approximated by a lower-order polynomial. This suggests that these estimators are largely determined by lower-order moments and cross-products, so MAI could result in negligible bias.

2.3 Practical considerations

In practice, the $\widehat{m}_{rk}$ may not be a valid moment sequence. If our aim is only to match moments, but not cross-products ($k = 1$), it is well known (Shohat and Tamarkin, 1943) that a sequence of $2q + 1$ moments is valid if

$$\left| \begin{array}{cc} 1 & \widehat{m}_{11} \\ \widehat{m}_{11} & \widehat{m}_{21} \end{array} \right| \geq 0, \dots, \left| \begin{array}{cccc} 1 & \widehat{m}_{11} & \dots & \widehat{m}_{q1} \\ \vdots & \vdots & \ddots & \vdots \\ \widehat{m}_{q1} & \widehat{m}_{(q+1)1} & \dots & \widehat{m}_{(2q)1} \end{array} \right| \geq 0.$$

Checking these determinants will identify the number of valid moments for a given data set. We address how many of these valid moments should be used in the sequel.

For the purpose of matching an arbitrary number of moments and cross-products, it is less clear how to identify a valid collection. For this discussion, let there be a single error-free covariate Z . Consider a limited set of $\mathbf{M} = (M_1, M_2 = M_1/2, M_3 = M_1/2)$, where the number of moments, M_1 , is even and the order of cross-products is $(M_1/2 + 1)$. This corresponds to matching the variance-covariance matrix of $(1, X, \dots, X^{M_1/2}, Y, Z)$, and therefore has a nice interpretation. It is also the set of moments that must be matched to achieve consistent parameter estimation in polynomial regression of order $M_1/2$. In addition, simulations indicate that letting M_1 be odd can lead to “outlying” $\hat{X}_i$ if the true distribution of X_i is extremely skewed (Bay, 1997). When $M_1 = 4$, we obtain $\hat{m}_{11}, \hat{m}_{21}, \hat{m}_{31}, \hat{m}_{41}, \hat{m}_{12}, \hat{m}_{22}, \hat{m}_{13}$, and $\hat{m}_{23}$, with expectations $E(X), E(X^2), E(X^3), E(X^4), E(XY), E(X^2Y), E(XZ)$, and $E(X^2Z)$, respectively. This set of moments defines the estimated variance-covariance matrix for $(1, X, \dots, X^{M_1/2}, Y, Z)$ given by

$$\begin{bmatrix} 1 & \hat{m}_{11} & \hat{m}_{21} & n^{-1} \sum Y_i & n^{-1} \sum Z_i \\ \hat{m}_{11} & \hat{m}_{21} & \hat{m}_{31} & \hat{m}_{12} & \hat{m}_{13} \\ \hat{m}_{21} & \hat{m}_{31} & \hat{m}_{41} & \hat{m}_{22} & \hat{m}_{23} \\ n^{-1} \sum Y_i & \hat{m}_{12} & \hat{m}_{22} & n^{-1} \sum Y_i^2 & n^{-1} \sum Y_i Z_i \\ n^{-1} \sum Z_i & \hat{m}_{13} & \hat{m}_{23} & n^{-1} \sum Y_i Z_i & n^{-1} \sum Y_i^2 Z_i \end{bmatrix}.$$

We propose checking that this matrix is positive definite to verify a valid set of moment estimates. The MAI data are denoted $\hat{X}_M$. For example, adjusted data based on matching two moments are $\hat{X}_2$ and $\hat{X}_{4,2}$ are derived by matching four moments and two cross-products.

Occasionally, numerical problems may arise when these moments can not be matched with a data set of size n . Even if the moments form a valid sequence, there may not be an empirical distribution function that takes jumps of size $1/n$ at each of n points having these moments. When this occurs, the Newton-Raphson algorithm will not converge to a solution. It is rare to encounter this problem when only moments, but not cross-products, are involved. When this does occur, we propose matching fewer moments.

3. Histogram and Kernel Density Estimation

In this and the next section, we demonstrate the utility of MAI in several representative analysis contexts. A simple and useful application is to adjust mis-measured data to approximate an underlying error-free variable when interest focuses on its distribution. A histogram or kernel density estimate (KDE) based on mis-measured data is too flat and dispersed. We illustrate for three distributions of X : $N(0,1)$; chi-square with 4 degrees of freedom; and a bimodal mixture of normals, $0.30N(0,1) + 0.70N(5,1)$, the latter two standardized to have mean 0 and variance 1. Measurement error with variance $\sigma_u^2=1$ is added to X , corresponding to large measurement error, with reliability ratio (RR) $\text{Var}(X)/\text{Var}(W) = 0.5$. Using the default bandwidth in the R `density()` package, KDEs for a simulated data set for each distribution are displayed in Figure 1. Those based on $\hat{X}_4$ have features more like the density estimates that would be obtained from the true X . Because the normal distribution is completely defined by its first two moments, there is no benefit to matching additional moments. However, there is great improvement from matching four moments when X is chi-square or bimodal, as the KDE based on $\hat{X}_4$ is substantially closer to that based on X .

We conducted a Monte Carlo simulation to investigate more generally the extent of improvement possible and to identify the best number of moments to match. Several situations were considered, including the three distributions for X above; two levels of

measurement error with RRs of 0.75 and 0.5; and three sample sizes, $n = 300, 1000$, and 2000 , typical for measurement error applications (Stefanski and Carroll, 1985; Freedman et al., 2004).

We compare KDE based on MAI to that based on alternative methods of obtaining adjusted data. The first is regression calibration, where $E(X_i|W_i)$ is estimated by the best linear unbiased predictor $\hat{X}_{RC,i} = \hat{\mu}_x + (\hat{\sigma}_{wx}/\hat{\sigma}_w^2)(W_i - \hat{\mu}_w)$; and $\hat{\mu}_x$, $\hat{\sigma}_{wx}$, and $\hat{\sigma}_w^2$ are estimates obtained by method of moments. This is the empirical Bayes estimator for X_i when X is normally distributed; however, the same linear estimator is used regardless of the actual distribution of X . To account for non-normality, we also consider a different estimator of $E(X_i|W_i)$ obtained by assuming X_i has density represented by the flexible family of semi-nonparametric (SNP) densities $f_X(x|\mu, \sigma, \alpha)$ (Gallant and Nychka, 1987; Davidian, 1992, 1993; Zhang, 2001; Carroll et al., 2006), which involve parameters μ , σ , and α and can approximate many potential latent variable distributions. The family has a convenient form so that the corresponding density $f_W(w|\mu, \sigma, \alpha)$ of W can be obtained by integration over x , and the parameters (μ, σ, α) can be estimated from the observed data by maximum likelihood. This approach naturally provides density estimation in the form of $f_X(x|\hat{\mu}, \hat{\sigma}, \hat{\alpha})$. We took the extra step of estimating $E(X_i|W_i)$ as $\hat{X}_{SNP,i} = \int x f_{X,W}(x, W_i|\hat{\mu}, \hat{\sigma}, \hat{\alpha}) dx / f_W(W_i|\hat{\mu}, \hat{\sigma}, \hat{\alpha})$.

The versions of $\hat{X}_i$ are evaluated according to their closeness to the underlying X_i as

measured by $MSE(\hat{X}) = B^{-1} \sum_{b=1}^B n^{-1} \sum_{i=1}^n (\hat{X}_{i,b} - X_{i,b})^2$ for B simulated data sets, and by computing the integrated squared error between the empirical distribution functions, given by $ISE(G_{\hat{X}}) = B^{-1} \sum_{b=1}^B \int \{G_{\hat{X},b}(t) - G_{X,b}(t)\}^2 dt$, where $G_X(t) = n^{-1} \sum_{i=1}^n I_{(X_i \leq t)}$, $-\infty < t < \infty$. When X is assumed to have the SNP density, we can estimate the cumulative distribution function, cdf_{SNP} , directly from $f_X(x|\hat{\mu}, \hat{\sigma}, \hat{\alpha})$. Density estimation using the SNP family is well established, so we calculate $ISE(cdf_{SNP}) = B^{-1} \sum_{b=1}^B \int \{cdf_{SNP,b}(t) - G_{X,b}(t)\}^2 dt$ as a gold standard. We report the MSE ratio $MSE(W)/MSE(\hat{X})$ and ISE ratio $ISE(G_W)/ISE(G_{\hat{X}})$, so that larger ratios indicate a greater reduction in error. Standard errors for these ratios are obtained by the delta method and are reported as a “coefficient of variation,” which is the ratio standard error divided by the ratio itself.

Results for $n = 1000$ are displayed in Table 1; those for $n = 300$ and $n = 2000$ are similar (see Web Appendix C). For normally distributed X , there is very little difference in the MSE ratio based on matching two or four moments. However, there is no additional benefit from matching six moments. When it is not important to get each $\hat{X}_i$ close to the original X_i , but instead we want an ensemble of $\hat{X}_1, \dots, \hat{X}_n$ with distribution similar to that of $X_1, \dots, X_n$, ISE ratios indicate that it is better to match only two moments. However, when $f_X(x)$ is chi-square or bimodal normal mixture, both ratios indicate that it is better to match four moments rather than two or six. Moments greater than four may not be as essential in describing distributions, and their estimators are likely to be highly variable. A general recommendation is to match four moments.

Comparison of MAI, $\hat{X}_4$, to $\hat{X}_{RC}$ and $\hat{X}_{SNP}$ shows that the latter two methods do a better job getting each $\hat{X}_i$ close to X_i and have slightly larger MSE ratios than those for $\hat{X}_4$, regardless of the distribution of X . It is not surprising that conditional expectations yield good estimation of the individual $\hat{X}_i$. However, these expectations are known to be less variable than the original data, so the distributions of $\hat{X}_{RC}$ and $\hat{X}_{SNP}$ may not resemble that of X (Eddington (1940), Tukey (1974)). In fact, the ISE ratios for these methods are close to 1 or even smaller than 1, indicating that the empirical distributions based on these $\hat{X}$ are no better

than that based on W . Thus, when interest focuses on density estimation, imputation of $\hat{X}_{RC}$ or $\hat{X}_{SNP}$ is inadequate. ISE ratios for $\hat{X}_4$ are at least as large as those of the gold standard, cdf_{SNP} , particularly for chi-square X . The SNP density estimator used here is based on the normal distribution and is not ideal for estimating skewed densities. ISE ratios for $\hat{X}_4$ reflect large improvement in density estimation relative to W , confirming the impression suggested by Figure 1. When the ensemble of measurements is of interest, rather than the individual specific observations, MAI provides an attractive alternative.

4. Regression Models

We evaluate various methods that adjust for covariate measurement error in common nonlinear regression models over a range of conditions, including distribution of X ($N(0,1)$, standardized chi-square $df=4$, standardized bimodal mixture of normals; see Section 3); sample size ($n = 300, 1000, 2000$); measurement error variance (moderate, $\sigma_u^2=0.33$, $RR=0.75$; large, $\sigma_u^2=1.0$, $RR=0.50$); and underlying model parameters (levels depend on the specific model). An additional error-free covariate Z was generated from the same distribution as X such that $\text{Corr}(Z, X)=0.4$. The underlying model parameters control the extent to which the model deviates from linearity. To get a general understanding of performance in different circumstances, we varied the model parameters and strength of the covariate effects.

We consider three imputation methods as well as the conditional score method (Stefanski and Carroll, 1987; Tsiatis and Davidian, 2001). For RC, $\hat{X}_{RC}$ is the best estimated linear unbiased estimator of $E(X|W, Z)$. We use the modification of MR proposed by Freedman et al. (2004) that involves conditioning on the error-free covariate so that $\hat{X}_{MR,i} = \hat{E}(W|Y_i, Z_i) (1 - \hat{G}) + W_i \hat{G}$ where $\hat{G} = \sqrt{\text{Var}(X|Y_i, Z_i)}^{1/2} / \sqrt{\text{Var}(W|Y_i, Z_i)}^{1/2}$. For MAI, we use $\hat{X}_{4,2,2}$ for which $E(X)$, $E(X^2)$, $E(X^3)$, $E(X^4)$, $E(XY)$, $E(X^2Y)$, $E(XZ)$, $E(X^2Z)$ are matched. Some alternatives are discussed below in the context of specific models.

When $\hat{X}_{4,2,2}$ are used in a regression model, the usual standard errors for regression parameter estimates are not correct. As described previously, standard errors can be obtained by bootstrapping. When the estimators are M-estimators, standard errors may be obtained by the empirical sandwich approach. The equations that determine $\hat{\Lambda}$ can be stacked with the usual equations in which the unknown X_i are replaced with $h(W_i, V_i, \hat{\Lambda})$ (Bay, 1997; Stefanski and Boos, 2002; Carroll et al., 2006); see Web Appendix D.

4.1 Logistic Regression

The model for the outcome is $P(Y = 1|X, Z) = F(\beta_0 + \beta_X X + \beta_Z Z)$ where $F(v) = \{1 + \exp(-v)\}^{-1}$. We simulated data from two parameter settings, $(\beta_0, \beta_X, \beta_Z) = (-1.5, 1, 1)$ and $(\beta_0, \beta_X, \beta_Z) = (-0.6, 0.3, 0.3)$. The first is similar to Freedman et al. (2004) and corresponds to substantial non-linearity, strong covariate effects, and event rate $P(Y = 1) \approx 0.30$. For the second, $P(Y = 1|X, Z)$ is nearly linear in the range of X , the effect of X is moderate, and the event rate $P(Y = 1) \approx 0.36$. The observed data are Y_i , W_i , and Z_i , for $i = 1, \dots, n$.

Boxplots of the estimated coefficients $\hat{\beta}_X$ from $B = 500$ simulations are displayed in Figure 2 for the case where $(\beta_0, \beta_X, \beta_Z) = (-1.5, 1, 1)$, $n = 2000$, and $\sigma_u^2=1.0$. When X is normally distributed, the RC estimator for β_X shows slight bias, but has the least variability. The other methods are unbiased and have similar variability. RC and MR are expected to perform well when X is normally distributed, and there is nothing to be gained from information about higher-order moments. However, it is reassuring to see that the increase in variability from matching additional moments is not substantial. When the latent variable distribution is

either chi-square or bimodal, the RC and MR estimators are biased. Only MAI and CS appear unbiased, and these have similar variability.

Other results are presented in Web Appendix E. When the underlying coefficients are $(\beta_0, \beta_X, \beta_Z) = (-1.5, 1, 1)$, results are similar to those in Figure 2. Due to the near linearity of the model, when the underlying coefficients are $(\beta_0, \beta_X, \beta_Z) = (-0.6, 0.3, 0.3)$, all adjustment methods are nearly identical in terms of estimator bias and variability. The measurement error in X also affects the estimation of the error-free covariate effect, β_Z . The naive estimator for β_Z is biased. This can be corrected by adjusting for measurement error in X . However, in our simulations the regression calibration estimator for β_Z is biased, even when the latent variable is normally distributed. The other adjustment procedures perform similarly and demonstrate negligible bias in β_Z (Web Appendix E).

Moment matching is a good alternative in logistic regression when the underlying latent variable distribution is unknown. In logistic regression, we recommend that four moments and two cross-products with important covariates be matched. This level of matching is necessary to render negligible bias in our simulations (see Web Appendix E). In Web Appendix E, Tables 10 and 11, we compare the sandwich and bootstrap variance estimators to the Monte Carlo variance of parameter estimates. The sandwich estimator appears reasonable across a variety of settings (Table 10) and the bootstrap variance is similar except in the case of large measurement error and small sample sizes (Table 11).

4.2 Cox Proportional Hazard Model

Another common non-linear model is the Cox proportional hazard model for a time to event outcome. For subject $i = 1, \dots, n$, let T_i denote failure time and C_i denote censoring time. The failure time T_i is not available for all subjects, but instead $Y_i = \min(T_i, C_i)$ and $\delta_i = I(T_i \leq C_i)$ are observed. The hazard of failure $\lambda(t|X, Z)$ is related to the covariates by

$$\lambda(t|X, Z) = \lim_{h \rightarrow 0^+} \left\{ h^{-1} P(t \leq T \leq t+h | T \geq t, X, Z) \right\} = \lambda_0(t) \exp(\beta_X X + \beta_Z Z),$$

where $\lambda_0(t)$ is an underlying baseline hazard function.

We consider two scenarios. The first is similar to that of Wang (2006), where failure times occur according to the hazard function $\lambda(t|X, Z)$ with $(\lambda_0, \beta_X, \beta_Z) = (0.2, 0.7, 0.7)$, and 50% of subjects are censored uniformly. This implies a very strong covariate effect with a hazard ratio of 2 for each unit change in X and hazard ratio of 66 for the largest value of X compared to the smallest [$\exp(0.7) \approx 2$ and $\exp\{0.7 \text{ range}(X)\} \approx 66$]. As a moderate alternative, we generated failure times from $\lambda(t|X, Z)$ with $(\lambda_0, \beta_X, \beta_Z) = (1.0, 0.3, 0.3)$ and 40% censoring, corresponding to a hazard ratio of 1.4 for each unit change in X and a hazard ratio of 6 overall [$\exp(0.3) \approx 1.4$ and $\exp\{0.3 \text{ range}(X)\} \approx 6$].

In logistic regression, the available data consist of (Y, W, Z) , whereas for time to event data we have (Y, δ, W, Z) . In this case, we extend the moment matching to include δ and target the joint distribution of (Y, δ, X, Z) . We considered several approaches. The simplest approach would match the variance-covariance matrix of $(1, X, \dots, X^{M1/2}, Y, \delta, Z)$. We could also match the variance-covariance matrix of $(1, X, \dots, X^{M1/2}, Y, \delta, Z)$ within each level of δ . Alternatively, we could match on risk sets, which is to re-match $(1, X, \dots, X^{M1/2}, Y, \delta, Z)$ at different points in time for those subjects who are still at risk. We tried all of these and saw little difference in the results. We therefore recommend the first and simplest method. The adjusted data are $\hat{X}_{4,2,2,2}$, for which $E(X), E(X^2), E(X^3), E(X^4), E(XY), E(X^2Y), E(X\delta), E(X^2\delta), E(XZ)$, and $E(X^2Z)$ are matched. We compared this to a lesser adjustment for which

only $E(X)$, $E(X^2)$, $E(XY)$, $E(X\delta)$, $E(XZ)$ are matched and observed more bias and similar variability in the estimators from the lesser adjustment. Accordingly, we use $\hat{X}_{4,2,2,2}$ and results are presented for this version of MAI only.

Boxplots of the estimated coefficients $\hat{\beta}_X$ from $B = 500$ simulations, where $n = 2000$ and $\sigma_u^2 = 1.0$, are shown in Figure 3. When the true parameter values are $\lambda_0 = 0.2$, $\beta_X = 0.7$, and $\beta_Z = 0.7$, only the CS estimator shows no evidence of bias, regardless of the distribution of X . Bias in RC and MAI estimators is evident, though relatively small. The variability in these estimators is similar, though the RC estimator is somewhat less variable. When the true parameter values are $\lambda_0 = 1$, $\beta_X = 0.3$ and $\beta_Z = 0.3$, and X has a normal distribution, all of the methods have similar variability and no detectable bias. However, for a chi-square or bimodal latent variable, the RC estimator is biased. Both the MAI and CS estimators appear unbiased and have similar variability. Results for other sample sizes and levels of measurement error are similar (Web Appendix F).

As in logistic regression, the measurement error in X impacts estimation of β_Z , and the naive estimator is biased. In Web Appendix F, we see that the RC estimator for β_Z is over-corrected, particularly for the larger underlying parameter values. MAI and CS estimators are nearly unbiased and have similar variability to RC.

In our simulations, the CS approach is preferable. However, this method may be excessively time consuming or infeasible for complicated Cox models. Imputation approaches, although imperfect, offer a practical solution. Both RC and MAI are easy to implement and yield great improvement over the naive method. For estimation of β_X , neither can be recommended over the other based on our simulations. However, when β_Z is also of interest, MAI is preferred.

5. Application to OPTIMIZE-HF

We carry out the OPTIMIZE-HF analyses performed by Gheorgiade et al. (2006), accounting for measurement error. The data set includes information on $n = 48,612$ subjects, aged 18 or older, with heart failure. There are two outcomes of interest, in-hospital mortality and post-discharge mortality. We use the models reported by Gheorgiade et al. (2006), which include baseline systolic blood pressure, many baseline covariates, and linear splines and truncation that account for non-linearity in continuous covariates. Their model for in-hospital mortality is

$$P(Y=1|Z, X) = \{1 - \exp(\beta_0 + \beta_Z^T Z + \beta_1 S)\}^{-1},$$

where Z includes error-free covariates listed in Web Appendix G, Table 17; S is a truncated version of blood pressure, i.e. $S = -\{XI(X < 160) + 160I(X \geq 160)\}$; and X represents true systolic blood pressure in 10-mm Hg units. A pre-specified subset of patients ($n = 5,791$) was followed for 60 to 90 days after discharge. In this group, post-discharge mortality is described by the Cox proportional hazard model

$$\lambda(t|X, Z) = \lambda_0(t) \exp(\beta_Z^T Z + \beta_1 S_1 + \beta_2 S_2)$$

where Z includes error free covariates listed in Web Appendix G, Table 18, and S_1 and S_2 fit a linear spline to blood pressure, i.e. $S_1 = -\{140I(X \geq 140)\}$ and $S_2 = -\{0I(X < 140) + XI(X$

≥ 140). Gheorgiade et al. (2006) fit these models using observed baseline systolic blood pressure, W , in place of X . We adjust the mis-measured W , by matching four moments and two cross-product moments with response, and impute $\hat{X}$ in place of X .

Adjustment procedures assume that the measurement error variance, σ_u^2 is known. In practice, σ_u^2 is usually replaced by a good estimate. It would be best to estimate the measurement error variance from replicate measures of systolic blood pressure, taken over a period of time. Replicate measures were not available in the OPTIMIZE-HF data set; however, variability in blood pressure has been extensively studied. One source is the Framingham data set (Carroll et al., 2006), which includes four measurements of blood pressure, two taken at the first exam and two taken at a second exam. The average standard deviation in four measurements is 9 mm Hg, which corresponds to a reliability ratio of about 0.75. Based on the information from other external studies, the measurement error may actually be larger (Marshall, 2008). For the purpose of illustration, we use a reliability ratio of 0.75 for adjustment. In practice, it is critical to obtain replicate data for this purpose.

The estimated density of baseline systolic blood pressure is altered by adjustment (Figure 4). The adjusted version shows a higher peak and smaller tails and conveys the impression that patients' blood pressures are more similar to each other.

In Table 2, we compare the MAI odds ratios and hazard ratios to those obtained by Gheorgiade et al. (2006). We report odds ratios per 10-mm Hg change in S and hazard ratios for S_1 , and S_2 per 10-mm Hg change. Wald-type 95% confidence intervals for the odds ratios and hazard ratios are based on standard errors from 1000 bootstrap samples. The adjusted estimates indicate a stronger effect of systolic blood pressure. The RC estimates move in the same direction as MAI, but are closer to the naive estimates. The impact of adjustment is not substantial in this case; however, we have assumed relatively moderate measurement error. Many studies have reported higher variability in replicate blood pressure measurements, and adjustment could be more important in estimating effect size.

Further adjustment for measurement error could be implemented by matching cross-products with all, or some, of the 18 covariates listed in Web Appendix G. We did not use these for two reasons. First, most of these covariates are only weakly correlated with systolic blood pressure, and we expect little gain from matching moments with such covariates. The only covariate strongly correlated with systolic blood pressure is diastolic blood pressure, which, realistically, is also measured with error; moreover, the measurement error due to biological fluctuations in both variables is likely correlated. Appropriate adjustment would involve a multivariate version of MAI. Extension to this case is possible; the implementation involves non-trivial considerations, and we will report this work elsewhere.

6. Discussion

We have introduced MAI as a means for adjusting mis-measured data to reflect the latent variable distribution and improve parameter estimation in non-linear regression models. The method does not require any assumptions on the latent variable distribution. Under the general recommendation of matching four moments, it performs well for a variety of distributions. For density estimation, MAI is typically superior to the simple alternatives we considered. In simulations of logistic regression, the method is similar to MR when the latent variable is normally distributed, but is a superior imputation method when the latent variable is non-normal. In the Cox proportional hazards model, RC and MAI provide substantial improvement over the naive approach, but do not eliminate bias. Of the functional approaches that we considered, the conditional score is the only method that eliminates bias in Cox model parameter estimators. The performance of MAI may differ for

other latent variable distributions, such as heavy-tailed distributions. Investigation of tail effects would require larger samples than those considered here.

The OPTIMIZE-HF study of systolic blood pressure of Gheorgiade et al. (2006) is illustrative of a typical data analysis. The mis-measured variable is included in descriptive analyses and in multiple, complex models involving splines to account for non-linearity. In practice, models could include splines, squared terms or interactions with the latent variable of interest. These are easily accommodated by imputation methods, and other approaches such as conditional score may be difficult or impossible to implement. In these circumstances, an imputation approach may be desirable.

We have developed MAI for the case of normally distributed measurement error. The method depends on correct specification of the measurement error distribution. Analysts should take care to verify that normality of measurement error is a reasonable assumption. The MAI method can be applied for other types of measurement error, as long as the distribution is known. Work on such extensions is reported elsewhere.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH grants R01CA085848, T32HL079896, R37AI031789, and P01CA142538 and NSF grants DMS 0906421 and DMS 0504283

References

- Bay, J. Unpublished dissertation thesis. 1997.
- Brueren MM, van Limpt P, Schouten HJA, de Leeuw PW, van Ree JW. Is a series of blood pressure measurements by the general practitioner or the patient a reliable alternative to ambulatory blood pressure measurement? A study in general practice with reference to short-term and long-term between-visit variability. *American Journal of Hypertension*. 2008; 101:879–885.
- Carroll, RJ.; Ruppert, D.; Stefanski, LA.; Crainiceanu, CM. *Measurement Error in Nonlinear Models: A Modern Perspective*. 2nd. Boca Raton, Florida: Chapman and Hall; 2006.
- Carroll RJ, Stefanski LA. Approximate quasi-likelihood estimation in models with surrogate predictors. *Journal of the American Statistical Association*. 1990; 85:652–663.
- Cheng CL, Schneeweiss H. Polynomial regression with errors in the variables. *Journal of the Royal Statistical Society, Series B*. 1998; 60:189–199.
- Cheng, CL.; Van Ness, J. *Statistical Regression with Measurement Error*. London: Arnold Publishers; 1999.
- Cramer, H. *Mathematical Methods of Statistics*. Princeton: Princeton University Press; 1957.
- Davidian M, Gallant AR. Smooth nonparametric maximum likelihood estimation for population pharmacokinetics, with application to quinidine. *Journal of Pharmacokinetics and Biopharmaceutics*. 1992; 20:529–556. [PubMed: 1287201]
- Davidian M, Gallant AR. The nonlinear mixed effects model with a smooth random effects density. *Biometrika*. 1993; 80:475–488.
- Eddington AS. The correction of statistics for accidental error. *Monthly Notices of the Royal Astronomical Society*. 1940; 100:354–361.
- Freedman LS, Fainberg V, Kipnis V, Midthune D, Carroll RJ. A new method for dealing with measurement error in explanatory variables of regression models. *Biometrics*. 2004; 60:172–181. [PubMed: 15032787]

- Freedman LS, Midthune D, Carroll RJ, Kipnis V. A comparison of regression calibration, moment reconstruction and imputation for adjusting for covariate measurement error in regression. *Statistics in Medicine*. 2008; 27:5195–5216. [PubMed: 18680172]
- Freudenheim JL, Marshall JR. The problem of profound mismeasurement and the power of epidemiologic studies of diet and cancer. *Nutrition and Cancer*. 1988; 11:243–250. [PubMed: 3217262]
- Gallant AR, Nychka DW. Semiparametric maximum likelihood estimation. *Econometrica*. 1987; 55:363–390.
- Garcia-Vera M, Sanz J. How many self-measured blood pressure readings are needed to estimate hypertensive patients' true blood pressure. *Journal of Behavioral Medicine*. 1999; 22:93–113. [PubMed: 10196731]
- Gheorghiadu M, Abraham WT, Albert NM, Greenberg BH, O'Connor CM, She L, Stough WG, Yancy CW, Young JB, Fonarow GC. Systolic blood pressure at admission, clinical characteristics, and outcomes in patients hospitalized with acute heart failure. *Journal of the American Medical Association*. 2006; 296:2217–2226. [PubMed: 17090768]
- Gleser, LJ. *Statistical Analysis of Measurement Error Models and Application*. Brown, PJ.; Fuller, WA., editors. Providence: American Mathematical Society; 1990.
- Huang X, Stefanski LA, Davidian M. Latent-model robustness in structural measurement error models. *Biometrika*. 2006; 93:53–64.
- Li E, Zhang D, Davidian M. Conditional estimation for generalized linear models when covariate are subject-specific parameters in a mixed model for longitudinal parameters. *Biometrics*. 2004; 60:1–7. [PubMed: 15032767]
- Liu K, Stammler J, Dyer A, Mc Keever J, McKeever P. Statistical methods to assess and minimize the role of intra-individual variability in obscuring the relationship between dietary lipids and serum cholesterol. *Journal of Chronic Diseases*. 1978; 31:399–418. [PubMed: 711832]
- Louis TA. Estimating a population of parameter values using Bayes and empirical Bayes methods. *Journal of the American Statistical Association*. 1984; 79:393–398.
- Marshall TP. Blood pressure variability: the challenge of variability. *American Journal of Hypertension*. 2008; 21:3–4. [PubMed: 18091733]
- Pickering TG, Hall JE, Appel LJ, Falkner BE, Graves J, Hill MN, Jones DW, Kurtz T, Sheps SG, Roccella EJ. Recommendations for blood pressure measurement in humans and experimental animals: Part 1: Blood pressure measurement in humans: A statement for professionals for the subcommittee of professional and public education of the American Heart Association Council on High Blood Pressure Research. *Hypertension*. 2005; 45:142–161. [PubMed: 15611362]
- Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Statistics in Medicine*. 1989; 8:1051–1070. [PubMed: 2799131]
- Shen W, Louis TA. Triple-goal estimates in two stage hierarchical models. *Journal of the Royal Statistical Society, Series B*. 1998; 60:455–471.
- Shohat, JA.; Tamarkin, JD. *The Problem of Moments*. Providence, R.I.: American Mathematical Society; 1943.
- Stefanski LA. Unbiased estimation of a nonlinear function of a normal mean with application to measurement error models. *Communications in Statistics, Series A*. 1989; 18:4335–4358.
- Stefanski LA, Boos DD. *The Calculus of M-Estimation*. American Statistical Association. 2002; 56:29–38.
- Stefanski LA, Carroll RJ. Covariate measurement error in logistic regression. *The Annals of Statistics*. 1985; 13:1335–1351.
- Stefanski LA, Carroll RJ. Conditional scores and optimal scores for generalized linear measurement error models. *Biometrika*. 1987; 74:703–716.
- Stefanski LA, Novick SJ, Devanarayan V. Estimating a nonlinear function of a normal mean. *Biometrika*. 2005; 92:732–736.
- Stulajter F. Nonlinear estimators of polynomials in mean values of a Gaussian stochastic process. *Kybernetika*. 1978; 14:206–220.

- Tsiatis AA, Davidian M. A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error. *Biometrika*. 2001; 88:447–458.
- Tukey JW. Named and faceless values: an initial exploration in memory of Prasanta C. Mahalanobis. *Sankhya A*. 1974; 36:125–176.
- Wang CY. Corrected score estimator for joint modeling of longitudinal and failure time data. *Statistica Sinica*. 2006; 16:235–253.
- Wang CY, Wang N, Wang S. Regression analysis when covariates are regression parameters of a random effects model for observed longitudinal measurements. *Biometrics*. 2000; 56:487–495. [PubMed: 10877308]
- Zhang D, Davidian M. Linear mixed models with flexible distributions of random effects for longitudinal data. *Biometrics*. 2001; 57:795–802. [PubMed: 11550930]

Figure 1.

Kernel Density Estimation; solid line: KDE of X , dark-dashed line: KDE of $\hat{X}_4$, light-dotted lines: KDE of $\hat{X}_2$ and W ; $n = 2000$ and $RR = 0.50$.

Figure 2.

Boxplots of $\hat{\beta}_X$ from $B = 500$ simulated data sets, for three distributions of X , where $P(Y = 1|X,Z) = F(\beta_0 + \beta_X X + \beta_Z Z)$ with true values $(\beta_0, \beta_X, \beta_Z) = (-1.5, 1, 1)$, $\sigma_u^2 = 1.0$, and $n = 2000$. Method: W, naive; RC, regression calibration; MR, moment reconstruction; MAI, moment adjusted imputation with $M = (4, 2, 2)$; CS, conditional score.

Figure 3.

Boxplots of $\hat{\beta}_X$ from $B = 500$ simulated data sets and three distributions of X , where $\lambda(t|X,Z) = \lambda_0(t) \exp(\beta_X X + \beta_Z Z)$ with true values $\{\lambda_0(t), \beta_X, \beta_Z\} = (0.2, 0.7, 0.7)$ or $\{\lambda_0(t), \beta_X, \beta_Z\} = (1, 0.3, 0.3)$, $\sigma_u^2 = 1.0$, and $n = 2000$. Method: W, naive; RC, regression calibration; MR, moment reconstruction; MAI, moment adjusted imputation with $M = (4, 2, 2, 2)$; CS, conditional score.

Figure 4.
Kernel Density Estimate of systolic blood pressure for the OPTIMIZE-HF study; light-dotted lines: KDE of W, dark-dashed line: KDE of $\hat{X}_4$

Table 1

Simulation results for three latent variable distributions, $f_X(x)$; two reliability ratios (RR), $B = 500$ simulated data sets, and $n = 1000$. Statistics reported: (a) $MSE(W)/MSE(\hat{X})$, where $MSE(\hat{X}) = B^{-1} \sum_{b=1}^B \sum_{i=1}^n n^{-1} (\hat{X}_{i,b} - X_{i,b})^2$ (coefficient of variation ≈ 0.001), and (b) $ISE(G_W)/ISE(G_{\hat{X}})$, where, $ISE(G_{\hat{X}}) = B^{-1} \sum_{b=1}^B \int \{G_{\hat{X},b}(t) - G_{X,b}(t)\}^2 dt$ for $G_X(t) = n^{-1} \sum_{i=1}^n I_{(X_i \leq t)}$, $-\infty < t < \infty$ (coefficient of variation ≈ 0.02). Adjusted data $\hat{X}$: RC, regression calibration; MAI matching 2, 4 or 6 respectively; and SNP, semi-nonparametric. cdf_{SNP} is the estimated cumulative density calculated by integrating the estimated SNP density.

Distribution	RR	$\hat{X}_{RC}$	$\hat{X}_2$	$\hat{X}_4$	$\hat{X}_6$	$\hat{X}_{SNP}$	cdf_{SNP}
(a) $\frac{MSE(W)}{MSE(\hat{X})}$							
Normal	0.75	1.33	1.24	1.24	1.23	1.32	-
	0.50	1.99	1.71	1.70	1.55	1.99	-
Chi Sq df=4	0.75	1.33	1.24	1.38	1.37	1.43	-
	0.50	1.99	1.71	1.88	1.79	2.15	-
Bimodal	0.75	1.33	1.24	1.50	1.43	1.64	-
	0.50	2.00	1.71	1.79	1.64	2.15	-
(b) $\frac{ISE(G_W)}{ISE(G_{\hat{X}})}$							
Normal	0.75	1.09	7.90	7.03	4.20	1.08	5.99
	0.50	1.34	23.72	11.65	0.82	1.33	9.64
Chi Sq df=4	0.75	1.39	2.41	6.96	5.15	0.61	1.08
	0.50	1.74	4.39	10.99	3.15	1.19	1.91
Bimodal	0.75	0.81	1.32	5.05	1.94	2.35	2.71
	0.50	0.86	1.90	4.13	1.29	1.48	4.99

Table 2

Parameter estimates for the OPTIMIZE-HF data analysis. Confidence intervals, based on bootstrap standard errors, are given parenthesis. Method: Unadjusted; RC, regression calibration; MAI, moment adjusted imputation.

	Unadjusted	RC	MAI
(a) Odds ratios for logistic regression (5)			
β_1	1.21 (1.17, 1.25)	1.28 (1.22, 1.33)	1.36 (1.44, 1.52)
(b) Hazard ratios for Cox model (5)			
β_1	1.18 (1.10, 1.26)	1.24 (1.13, 1.36)	1.38 (1.22, 1.54)
β_2	1.08 (1.01, 1.15)	1.10 (1.00, 1.21)	1.06 (0.94, 1.17)