ABSTRACT

It has been proven that pooling of samples is a more efficient and cost-effective way to collect data in human biomonitoring. However, this data collection method has led to difficulties in estimation because each pooled sample is assumed to be a weighted sum of log-normal distributions, which has an intractable density function. As such, parameters based on pooled samples cannot be estimated using maximum likelihood estimation. To solve this problem, an iterative bias correction method is proposed. The general idea of the method proposed is to first obtain naïve estimates from a set of p estimating equations, then adjust these estimates iteratively so that they eventually become asymptotically unbiased and consistent. When applied to two demographic groups in the NHANES 2003-2004 dataset, the method proposed produces estimates that are nearly unbiased. The estimated standard errors of the regression parameter estimates are only slightly larger than the estimated standard errors obtained based on individual samples. By assuming a chi-square distribution, confidence bands are plotted for all demographic groups in the NHANES 2003-2004 dataset, which are observed to be very close to those obtained from individual samples. Subsequent application of the method proposed in a simulation study based on two demographic groups also produces empirical averages of estimates that are nearly unbiased. Empirical noncoverages obtained from the simulation study are also close to 5% when a chi-square distribution is assumed. Therefore, we can conclude that the method proposed is performing well, given that only a negligible loss of accuracy is sacrificed for a significant amount of cost saved.