2019

Parameter Estimation for Data with Lower Limit of Detection Values under the Truncated Model – EM Solutions

Lutfiyya NaQiyba Muhammad
Medical University of South Carolina

Parameter Estimation for Data with Lower Limit of Detection Values Under the Truncated Model – EM Solutions

Lutfiyya NaQiyba Muhammad

A dissertation proposal submitted to the faculty of the Medical University of South Carolina in partial fulfillment of the requirement for the degree of Doctor of Philosophy in the College of Graduate Studies.

Department of Public Health Sciences

2019

Approved by:

_____
Paul J. Nietert, Ph.D.
Co-Chair, Advisory Committee

_____
Viswanathan Ramakrishnan, Ph.D.
Co-Chair, Advisory Committee

_____
Bethany J. Wolf, Ph.D.

_____
Diane L. Kamen, M.D.

_____
Jim C. Oates, M.D.

ACKNOWLEDGMENTS

I am grateful that I had the opportunity to interact with my dissertation committee members through the Core Center for Clinical Research in addition to my dissertation research. Thank you to my co-mentors, Drs. Paul J. Nietert and Viswanathan Ramakrishnan, for guiding me as I developed the concepts in this dissertation. The many discussions I had with my remaining committee members, Drs. Bethany J. Wolf, Diane L. Kamen, and Jim C. Oates, shaped the methods that I constructed for my dissertation. I am forever appreciative for the skills that I learned from each of you.

Thank you to my parents, NaQiyba and Shayaa Muhammad, for believing in me. NaQiyba Muhammad, you are my biggest supporter, and I could never ask for a better role model. Shariyf, Khaliyq, Haniyfa, and Rafiyq Muhammad I am blessed to have such loving and dependable siblings. Thank you to Levonia Coard, my grandmother, for teaching me to be resilient and hardworking.

I have so many amazing friends that have motivated me throughout this process. I would like to thank Emily Durham, Devin Durham, Joelle Zambrano, Johana Lambert, Kelby Killoy, and Wendy Pusser. I am fortunate to have met you all during my first semester. I am thankful for the encouragement I received from Brittany Brinson, Jessica Lieu, Kerrie Swanson, Samantha Burke, Sheriff Shittu, Joshua Burney, Raisa Habersham, Truth Price, Afua Akhi-Gbade, Kiana Anthony, and Angela Murro.

I am appreciative for the support that I received from you all throughout my doctoral studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

LLOD – lower limit of detection

EM – Expectation Maximization

ABSTRACT

Computing unbiased parameter estimates from a distribution using a sample with observations appearing below a lower limit of detection (LLOD) can be challenging. Frequently, LLOD observations are excluded from calculations for parameter estimates, or the LLOD observations are replaced with arbitrary values (LLOD, LLOD/2, LLOD/$\sqrt{2}$) prior to the calculations. Despite the frequent use of these simple approaches, the approaches are known to provide biased parameter estimates. Alternative approaches include implementing a left truncation or left censoring approach. In the first dissertation aim, we will explore and establish a general theoretical relationship between accurately estimating parameters under left truncated and left censored models. Estimation methods under both models require iterative algorithms. The left truncation approach is applied through an Expectation-Maximization (EM) algorithm. While the left censoring approach is implemented by the Newton-Raphson method. We conclude in the first aim that the left truncation and left censoring approaches yielded equivalent parameter estimates. Computationally, we favored the left truncation approach that is implemented through an EM algorithm. The left truncation approach for estimating parameters is utilized in the remaining aims. In the second aim of this dissertation, we propose an EM algorithm for estimating parameters from a normal distribution when there are multiple LLOD values present. The third aim includes solutions to an EM algorithm for estimating bivariate normal distribution parameters. In the third aim, the data under the left truncation approach can be categorized into 24 scenarios. The construction of the EM algorithm includes the scenarios. All dissertation aims are motivated by toxicology and serology data collected in the Systemic Lupus Erythematosus in Gullah Health study.

1. INTRODUCTION

1.1. Lower Limit of Detection Definition and Related Statistical Classifications

Data below a lower limit of detection (LLOD) and lower limit of quantitation (LLOQ) frequently occur in laboratory and environmental pollutant settings when the concentration of an analyte is below the level at which an analytical tool can reliably measure.[1] In order to explicitly define LLOD and LLOQ, the limit of blank (LOB) must first be defined. The LOB is the maximum detectable signal of an analyte produced by a laboratory tool in repeated samples of a serum that does not contain any of the analyte of interest.[1] In other words, there may be a positive quantitative measurement from a sample despite the analyte of interest being absent from the sample. The LLOD is the lowest level at which a measured analyte can be considered to be recognizably different from the LOB level.[1,2] The LLOQ is defined as the lowest "level at which measurements have sufficient precision for quantitative determination."[3] Several methods have been proposed to estimate the LLOD and LLOQ values in situations where the limits themselves are unknown.[2,4] It is plausible for the LLOD and LLOQ to occur at the same value, but typically the LLOD is smaller than the LLOQ.[1] Since LLOD and LLOQ values can be equivalent, similar statistical approaches are used to analyze data with observations appearing below a LLOD and a LLOQ. Therefore, in this dissertation, the phrase LLOD will describe both LLOD and LLOQ methods because both limits can be the same.

When a LLOD exists, the observations above the LLOD are considered to be fully observed. The fully observed data are classified as being from a left truncated distribution or observations below the LLOD value are classified as left censored observations. A left truncation value is required for both data classifications. The left truncation value is the LLOD value when it is

known, or it may be estimated when the proportion below the LLOD is known. A left truncated distribution is a distribution that has a support restricted by a lower bound at the LLOD value. In the left censoring framework, observations below the LLOD are considered censored. Also, it is assumed that none of the observations above the LLOD are censored.

This dissertation will propose novel methods that are applicable to data with observations that are below LLOD. The remaining of the first chapter will unify the current left truncation and left censoring methods and address gaps in the current literature related to dissertation objectives. Lastly, the data motivation and specific aims of this dissertation are stated.

1.2. Current Left Truncation and Left Censoring Literature Related to Dissertation

Two main approaches for analyzing data with below the LLOD observations are based on the two data classifications mentioned above: left truncation and left censoring. Statisticians typically make a preference to use one of the two approaches, and then use the selected approach through the use of an iterative algorithm to estimate parameters from the underlying (non-truncated) distribution. However, it does not appear that anyone has provided a formal proof that demonstrates the similarities of both approaches for estimating parameters from the underlying distribution. The first objective of this dissertation is to provide a parallel between the two approaches when there is one left truncation value.

The majority of the existing statistical literature on left truncation and left censoring approaches focus on one left truncation value in a single random variable. A commonly used continuous left truncated distribution for data with observations below the LLOD is the left truncated normal

distribution. The MLEs of the parameters from the left truncated normal distribution have been established under various constraints such as whether or not the left truncation value is known.[5-7] A normal distribution with left censored observations is often considered for continuous left censored methods. MLEs of parameters from a normal distribution with right censored observations were first constructed by Gupta in 1952.[8] Shortly following Gupta's approach, MLEs of parameters from a normal distribution with left censored observations were identified by Cohen.[9] These initial researchers, however, did not examine multiple truncation values appearing on one side of a distribution.

The second dissertation objective is devoted to providing an estimation method for parameters from a normal distribution, which is the underlying distribution, when the sample data is from a left truncated normal distribution with multiple distinct left truncation values. This estimation method will benefit studies involving chemical concentrations and biomarker assays in which datasets may consist of observations that have been generated over time with multiple LLOD values.[10,11] LLOD values can vary throughout time for some equipment, even within a 24 hour period, due to the calibration and detection sensitivity of the laboratory analytical tools.[12,13] Therefore, combining datasets from data collected at different times can result in a dataset with multiple LLOD values for a single variable.[11,14] The terms *doubly* and *multiply* truncated describes situations with two or more truncation values, respectively, occurring simultaneously in a variable.[5,15,16] The normal distribution with *left censored* observations has a theoretical MLE framework that includes doubly and multiply left truncation values in a single random variable.[15-17] However, doubly and multiply truncation values are not included in the *left truncated* framework for estimating parameters from a normal distribution.

The third dissertation objective addresses an estimation method for parameters from an underlying multivariate normal (MVN) distribution in the presence of one distinct left truncation value for each of the univariate components of the MVN random variable. Current left *censored* statistical approaches for estimating parameters from a multivariate normal distribution have been extended to include variables with at least one left truncation value, but continuous *truncated* multivariate statistical approaches have not. For example, Haiying and colleagues (2011) developed a multiple imputation method based on a left censored likelihood to estimate the parameters of a multivariate normal distribution when an observed sample includes a LLOD value for each variable.[18] Similarly, Hoffman and Johnson (2011, 2015) proposed a censored likelihood with unstructured covariance parameters and a censored pseudo-likelihood to account for left truncation values while estimating MLEs of parameters of multivariate normal and log-normal distributions.[19,20]

1.3. Motivating Example

This dissertation will explore statistical methods to account for continuous data with observations that are below the LLOD. The motivation of this dissertation is the Systemic Lupus Erythematosus in Gullah Health (SLEIGH) study. The objective of the SLEIGH study is to assess environmental and genetic factors in the progression of autoimmunity.[21] SLEIGH included a toxicology and serology component in which contaminant concentrations for 13 perfluorinated chemicals (PFCs) and 8 polybrominated diphenyl ethers (PBDEs) were measured in the serum of 86 participants with systemic lupus erythematosus (SLE) and 139 control participants. Exposure to PFCs and PBDEs have been shown to have adverse health effects such

as disruption of hormone systems and delayed cognitive development.[22-25] The amounts of PFCs and PBDEs present in the study participants' serum samples were measured in the chemistry laboratory at the Wadsworth Center of the New York State Department of Health. The wet weight of the PFCs and PBDEs were measured on a continuous scale in nanogram/gram (ng/g). Additional information regarding the quality control and assurance of the serum samples have been previously published.[26] The serum samples were collected once from study participants, but not all of the serum samples were measured in the laboratory at the same time. The serum samples were measured in multiple batches which resulted in more than one LLOD value for most of the PFC and PBDE variables.

Table 1 below describes how two contaminants, perfluorodecanoic acid (PFDA) and perfluorohexane sulfonate (PFHxS), were recorded in the SLEIGH study for the first batch of data. The LLOD value for PFDA was 0.03 ng/g, and for PFHxS, the LLOD value was 0.14 ng/g. Observations for PFDA and PFHxS could be both be below the LLOD value (e.g. Participant ID 1) or both fully observed (e.g. Participant ID 4). Either contaminant could be below the LLOD value while the other contaminant was fully observed (e.g. Participant IDs 2 and 3).

Table 1 Contaminants PFDA and PFHxS Data for One Batch

Pseudo Study Participant ID	Batch	PFDA	PFHxS
1	1	<0.03	<0.14
2	1	<0.03	0.24
3	1	1.14	<0.14
4	1	0.97	0.94

Table 2 illustrates PFDA and PFHxS data from two separate batches. By examining solely one contaminant variable such as PFDA, one will notice that PFDA had two LLOD values. The first batch LLOD value was 0.03 ng/g, while the second batch LLOD value was 0.10 ng/g. There were eight situations in which the below the LLOD observations could occur in the SLEIGH study when considering two contaminant variables together such as PFDA and PFHxS. For any given study participant, the eight situations were 1) the PFDA and PFHxS observations for batch 1 were both below their respective LLODs, 2) the PFDA observation was below its batch 1 LLOD, and the PFHxS observation was fully measured in the first batch, 3) the PFDA observation was fully measured in batch 1, and the PFHxS observation was below its batch 1 LLOD, 4) the PFDA and PFHxS observations in the first batch were both fully measured, 5) the PFDA and PFHxS observations for batch 2 were both below their respective LLODs, 6) the PFDA observation was below its batch 2 LLOD, and the PFHxS observation was fully measured in the second batch, 7) the PFDA observation was fully measured in batch 2, and the PFHxS observation was below its batch 2 LLOD, and 8) the PFDA and PFHxS observations in the second batch were both fully measured.

Table 2 Contaminants PFDA and PFHxS Data with Two Batches

Pseudo Study Participant ID	Batch	PFDA	PFHxS
1	1	<0.03	<0.14
2	1	<0.03	0.24
3	1	1.14	<0.14
4	1	0.97	0.94
5	2	<0.10	<0.16
6	2	<0.10	1.01
7	2	0.78	<0.16
8	2	0.34	0.22

1.4. Specific Aims

The presence of LLOD values in the contaminant variables from the SLEIGH study allows the data to be an ideal dataset for developing estimation methods for underlying (non-truncated) distribution parameters. Aims of the dissertation are as follows.

1. Establish the equivalence of left truncation and left censoring framework for applications involving limits of detection.

2. Develop a method for estimating the mean and variance of a single normal random variable with multiple lower limits of detection arising from different batches (i.e. one variable with multiple batches).

3. Develop an approach for estimating the mean and covariance matrices for multivariate normal random variables, with each marginal distribution having one lower limit of detection arising from a single batch (i.e. multiple variables with one batch).

The chapters related to the dissertation aims are written as individual manuscripts, and sections of the chapters reiterate the introduction and descriptions of literature on existing statistical approaches. The second chapter of this dissertation reviews literature related to truncated distributions and distributions with censored observations. The third chapter describes the first aim and establishes a parallel between left truncation and left censoring methods. The fourth chapter focuses on the second aim that includes estimation methods of parameters from a

univariate normal distribution when multiple left truncation values are present. The fifth chapter examines the third aim which is an estimation methods of multiple continuous variables with a single left truncation value for each variable. Results from simulations and an application to data from the SLEIGH study are presented. The sixth chapter summarizes the procedures related to the aims and outlines future research topics that would extend the current work.

2. REVIEW OF LITERATURE

2.1. Underlying Distributional Parameter Estimates from a Sample with LLOD Observations

Formulas for estimating underlying distributional parameters (such as a sample mean and variance) exclude LLOD observations from the calculation because the LLOD observations are considered to be missing data. This approach is known as a complete case analysis. Statisticians have shown that complete case analysis lead to biased parameter estimates.[27,28] In particular, the formulas will overestimate the true mean and underestimate the true variance when there are below LLOD observations.

A common method to computationally account for LLOD observations while estimating parameters from an underlying distribution is to apply a replacement method. For the replacement method, LLOD observations are replaced with the LLOD value itself, LLOD/$\sqrt{2}$, or LLOD/2.[29-37] Several organizations and government agencies have provided guidance for computing the LLOD value if the value is not specified by a laboratory instrument.[2,38] The United States Environmental Protection Agency recommends these replacement methods when 15% or less of the observations are below LLOD.[34] It has been shown that replacing the below LLOD value with the LLOD value itself, LLOD/$\sqrt{2}$, or LLOD/2 can lead to inaccurate inferences about the parameter estimate.[29,35]

Approaches listed in statistical literature as alternatives to the complete case analysis and replacement method are either to classify the data above the LLOD value as data following a left truncated distribution or to classify the observations below the LLOD value as left censored

observations. A left truncated distribution is a distribution that has a support restricted to a lower bound at the LLOD value. In the left censoring framework, any observation below the fixed LLOD value is censored. There are not any closed-form solutions for obtaining maximum likelihood estimators (MLEs) of parameters from an underlying (non-truncated) distribution when truncation or censoring is present.[6] Therefore, imputation methods or iterative algorithms must be joined with left truncation or left censoring methods in order to estimate the underlying distributional parameters.

Imputation methods to handle LLOD observations often include regression models such as the Tobit model or a quantile regression model.[39-47] For data with LLOD observations, the Tobit model is a type of censored regression model that accounts for the LLOD observations to be between zero and the LLOD value while adjusting for the variance.[40,47-49] Quantile regression expresses the outcome variable as a quantile such as the median (i.e. 50^{th} percentile). Quantiles depend on the rank of the outcome variable rather than the exact value of the observations in a sample.[40] Since the exact value of an LLOD observation is unknown, expressing the outcome as a quantile can be ideal for estimating a parameter using a sample with LLOD observations.

Frequent iterative algorithms combined with left truncation and left censoring methods are the Newton-Raphson method and expectation-maximization (EM) algorithm. In general, the Newton-Raphson method and EM algorithm can efficiently compute MLEs of parameters from an underlying distribution when a sample contains missing observations.[50,51] LLOD observations are included in the context of missing data since the value of the LLOD observations are not exactly known due to the observation being less than the LLOD value.[50]

2.2. Applications of Newton-Raphson Method

The Newton-Raphson method has become a recommended iterative procedure for estimating parameters from an underlying distribution under censored models.[52-55] For example, the scientific agency United States Geological Survey (2012) recommended the implementation of the Newton-Raphson method for obtaining MLEs of parameters within the context of a left censored model.[52] Swan (1969) provided solutions and suggested the use of the Newton-Raphson method under censored and grouped data for estimating the parameters of a normal distribution.[54] Also, Singh and Nocerion in 2002 applied the Newton-Raphson method with a left censored model to estimate parameters of a normal distribution from a sample with LLOD observations.[55]

Infrequently, Newton-Raphson is applied to estimate parameters from truncated distributions. Cohen first suggested the use of the Newton-Raphson method in 1950 for estimating parameters from a truncated normal distribution, but Halperin (1952) did not recommend the application of New-Raphson to estimate parameters from a truncated normal distribution.[5,56] Halperin stated that a convergence issue related to the initial value being specified incorrectly can occur when estimating parameters from a truncated normal distribution.[56] An inaccurate initial value (i.e. value outside the support of the truncated distribution) can produce irrational parameter estimates.[51,56] There have been extensive statistical methodology constructed for initial value selection in the context of the Newton-Raphson method.[57-61] The Newton-Raphson method is applied occasionally for estimating parameters from a truncated normal distribution despite Halperin's advice. For instance, Hattaway's thesis in 2010 explored the mean and variance

estimates of a truncated obtained from the Newton-Raphson method in the context of grades from a statistics course.[62] Also Alaesa's thesis (2017) compared MLEs obtained from Newton-Raphson method and other methods such as a method based on spatial linear combinations.[63] Findings from this thesis illustrated a preference to MLEs obtained from Newton-Raphson method relative to the other methods.

2.3. Applications of EM Algorithm

Applications of the EM algorithm have been established separately under censored models and truncated models. Wolynetz (1979) provided EM algorithm solutions for estimating normal distribution parameters when a sample contains censored observations.[64] Park (2003) developed EM algorithms to estimate parameters from normal, Rayleigh, and Laplace distributions for censored samples.[65] Particularly for left truncated models, the EM algorithm has been used to calculate MLEs of parameters from univariate and multivariate normal distributions using observed data from truncated distributions.[66,67] Expectations required for EM algorithms under the left truncated model were initially constructed by Cohen.[5,6,9]

There have been several EM algorithms that combined truncated and censored models.[68-70] Lee and Scott (2010) developed EM algorithm solutions to estimate parameters from a Guassian mixture model for data with truncation and censoring present. In 2018, Lodhi et al. constructed an EM algorithm under a censored model to estimate parameters from a truncated normal distribution.[69] Zunxiong et al. used clustering methods to develop an EM algorithm for

estimating parameters from a multivariate Gaussian mixture model when truncation and censoring occurs concurrently.[71]

2.4. Truncated Normal Distribution

A distribution that is often used for continuous data with below LLOD observations is a left truncated normal distribution. The moments of a truncated normal distribution were originally presented by Pearson and Lee.[72,73] Cohen established the MLEs for the mean and variance of a truncated normal distribution with left and right truncation occurring simultaneously (doubly truncated) or separately (singly left or right truncated).[5,6] Hald formed the MLEs of parameters from a truncated normal distribution when the truncation value is known.[7] Charts were constructed by Halperin for computing the MLEs of parameters from a right truncated normal distribution when the truncation value is unknown and known.[56] The probability density function (pdf) of a *truncated* normal distribution is constructed from the pdf and cumulative distribution function (cdf) of a normal distribution.

Normal Distribution

A random variable X has a normal distribution with mean μ and variance σ^2. The pdf is

$f\left(x\right)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left\{-\frac{1}{2\sigma^2}\left(x-\mu\right)^2\right\}$ where $x\in\mathbb{R}$, $\mu\in\mathbb{R}$, and $\sigma^2\in\mathbb{R}^+$ The simplest form of the normal distribution is the standard normal distribution. In a standard normal distribution, the mean $\mu=0$ and variance $\sigma^2=1$. The standard normal pdf of random variable W is

$\phi(w)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}w^2}$, where $w\in\mathbb{R}$. The cumulative distribution function

of the standard normal distribution is defined as, $\Phi(w) = P(W \le w) = \int_{-\infty}^{w} \phi(z)\, dz$. All normal distributions can be expressed as a standard normal distribution. The pdf of X can be expressed in terms of a standard normal distribution as the following, $f(x) = \frac{1}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)$.

Left Truncated Normal Distribution

Suppose a random variable Y_{LT} has a support that is restricted by a left truncation value, t_l. The support of Y_{LT} is $Y_{LT} \in (t_l, \infty)$. Y_{LT} follows a left truncated normal (LTN) distribution with mean μ, variance σ^2, and a left truncation value t_l. The pdf of Y_{LT} is expressed as

$f_{LTN}(y_{LT}) = \frac{1}{1-\Phi\left(\frac{t_l-\mu}{\sigma}\right)} \frac{1}{\sigma}\phi\left(\frac{y_{LT}-\mu}{\sigma}\right)$, where $y_{LT} \in (t_l, \infty)$, $\mu \in \mathbb{R}$, and $\sigma^2 \in \mathbb{R}^+$. The mean of Y_{LT} is $\mu_{LT} = \mu + \sigma \frac{\phi\left(\frac{t_l-\mu}{\sigma}\right)}{1-\Phi\left(\frac{t_l-\mu}{\sigma}\right)}$ and variance of Y_{LT} is $\sigma_{LT}^2 = \sigma^2\left[1+\left(\frac{t_l-\mu}{\sigma}\right)\frac{\phi\left(\frac{t_l-\mu}{\sigma}\right)}{1-\Phi\left(\frac{t_l-\mu}{\sigma}\right)} - \left(\frac{\phi\left(\frac{t_l-\mu}{\sigma}\right)}{1-\Phi\left(\frac{t_l-\mu}{\sigma}\right)}\right)^2\right]$.

Right Truncated Normal Distribution

Similarly, by letting the support of $Y_{RT} \in (-\infty, t_r)$, Y_{RT} follows a right truncated normal (RTN) distribution. The pdf of Y_{RT} is $f_{RTN}(y_{RT}) = \frac{1}{\Phi\left(\frac{t_r - \mu}{\sigma}\right)} \frac{1}{\sigma} \phi\left(\frac{y_{RT} - \mu}{\sigma}\right)$,

where $y_{RT} \in (-\infty, t_r)$, $\mu \in \mathbb{R}$, and $\sigma^2 \in \mathbb{R}^+$. The mean of Y_{RT} is $\mu_{RT} = \mu - \sigma \frac{\phi\left(\frac{t_r - \mu}{\sigma}\right)}{\Phi\left(\frac{t_r - \mu}{\sigma}\right)}$,

and variance of Y_{RT} is $\sigma_{RT}^2 = \sigma^2 \left[1 - \left(\frac{t_r - \mu}{\sigma}\right) \frac{\phi\left(\frac{t_r - \mu}{\sigma}\right)}{\Phi\left(\frac{t_r - \mu}{\sigma}\right)} - \left(\frac{\phi\left(\frac{t_r - \mu}{\sigma}\right)}{\Phi\left(\frac{t_r - \mu}{\sigma}\right)} \right)^2 \right]$.

2.5. Truncated Multivariate Normal Distribution

The truncated multivariate normal distribution was first derived by Birnbaum and Meyer in 1953.[74] With Birnbaum and Meyer's derivation as a guidance, the pdf and moments of the truncated bivariate normal distribution, a 2-dimensional truncated multivariate normal distribution, were formed by Weiler.[75] In 1961 Rosenbaum derived the moments of the left truncated bivariate normal distribution.[76] Also in 1961, Tallis constructed the moment generating function of the left truncated multivariate normal distribution with left truncation occurring in each of the variables.[77] The moments of the doubly multivariate normal distributions were formed with truncation appearing on the left and right.[78,79]

The pdf of the truncated multivariate normal distribution is defined by employing the multivariate normal distribution pdf. The simulation studies and SLEIGH study data applications in the subsequent dissertation chapters involve an underlying bivariate normal distribution which is a special case of the multivariate normal distribution. Therefore the following pdfs and cdfs are expressed using two variables.

Bivariate Normal Distribution

Let there be a random vector $\mathbf{Y} = \begin{bmatrix} Y_1 & Y_2 \end{bmatrix}'$, a mean vector $\boldsymbol{\mu} = \begin{bmatrix} \mu_{y_1} & \mu_{y_2} \end{bmatrix}'$, and a covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2_{y_1} & \rho\sigma_{y_1}\sigma_{y_2} \\ \rho\sigma_{y_1}\sigma_{y_2} & \sigma^2_{y_2} \end{bmatrix}$. The elements of $\boldsymbol{\Sigma}$ include the variance of Y_1 denoted as $\sigma^2_{y_1}$, $\sigma^2_{y_2}$ is the variance of Y_2, the covariance of Y_1 and Y_2 is represented as $\rho\sigma_{y_1}\sigma_{y_2}$, ρ is the correlation between Y_1 and Y_2, and the standard deviations of are Y_1 and Y_2 represented by σ_{y_1} and σ_{y_2}. The covariance matrix $\boldsymbol{\Sigma}$ is positive semi-definite. The joint distribution of Y_1 and Y_2 is a bivariate normal (BVN) distribution. The pdf of the bivariate normal distribution is,

$$f_{BVN}\left(y_1, y_2\right) = \left(2\pi\sqrt{|\boldsymbol{\Sigma}|}\right)^{-1} e^{-\frac{1}{2}\left[(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right]},$$

where $y_1 \in \mathbb{R}$, $y_2 \in \mathbb{R}$, $|\boldsymbol{\Sigma}|$ is the determinant of the covariance matrix, and $\boldsymbol{\Sigma}^{-1}$ is the inverse of the covariance matrix.

The conditional distribution of Y_1 given Y_2 is a normal distribution with mean $\mu_{y_1|y_2} = \mu_{y_1} + \frac{\sigma_{y_1}}{\sigma_{y_2}}\rho\left(y_2 - \mu_{y_2}\right)$ with variance $\sigma^2_{y_1|y_2} = \left(1-\rho^2\right)\sigma^2_{y_1}$. Likewise, the conditional

distribution of Y_2 given Y_1 is a normal distribution with mean $\mu_{y_2|y_1} = \mu_{y_2} + \frac{\sigma_{y_2}}{\sigma_{y_1}} \rho \left(y_1 - \mu_{y_1} \right)$ with variance $\sigma^2_{y_2|y_1} = \left(1-\rho^2\right) \sigma^2_{y_2}$.

Left Truncated Bivariate Normal Distribution

Suppose there are two random variables Y_{1LT} and Y_{2LT}. The support of Y_{1LT} is $Y_{1LT} \in (t_{y_1}, \infty)$, and the support of Y_{2LT} is $Y_{2LT} \in (t_{y_2}, \infty)$. Jointly Y_{1LT} and Y_{2LT} follow a left truncated bivariate normal (LTBVN) distribution with a mean vector $\boldsymbol{\mu} = \begin{bmatrix} \mu_{y_1} & \mu_{y_2} \end{bmatrix}'$, a covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2_{y_1} & \rho\sigma_{y_1}\sigma_{y_2} \\ \rho\sigma_{y_1}\sigma_{y_2} & \sigma^2_{y_2} \end{bmatrix}$, and left truncation vector $\mathbf{t} = \begin{bmatrix} t_{y_1} & t_{y_2} \end{bmatrix}'$. With the use the BVN pdf , the LTBVN pdf is expressed as,

$$f_{LTBVN}\left(y_{1LT}, y_{2LT}\right) = \frac{1}{\int_{\frac{t_{y_1}-\mu_{y_1}}{\sigma}}^{\infty} \int_{\frac{t_{y_2}-\mu_{y_2}}{\sigma}}^{\infty} f_{BVN}\left(y_{1LT}, y_{2LT}\right) dy_{1LT} dy_{2LT}} f_{BVN}\left(y_{1LT}, y_{2LT}\right),$$

where $t_{y_1} \le y_{1LT} \le \infty$ and $t_{y_2} \le y_{2LT} \le \infty$. The marginal mean of Y_{1LT} is

$$\mu_{y_{1LT}} = \mu_{y_1} + \sigma_{y_1} \frac{\phi\left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right)}{1-\Phi\left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right)}$$ and the marginal variance of Y_{1LT} is

$$\sigma^2_{y_{1LT}} = \sigma^2_{y_1}\left[1+\left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right)\frac{\phi\left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right)}{1-\Phi\left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right)} - \left(\frac{\phi\left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right)}{1-\Phi\left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right)}\right)^2\right].$$ Similarly the marginal

mean of Y_{2LT} is $\mu_{y_{2LT}} = \mu_{y_2} + \sigma_{y_2} \dfrac{\phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}{1 - \Phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}$ and

$\sigma^2_{y_{2LT}} = \sigma^2_{y_2}\left[1 + \left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)\dfrac{\phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}{1 - \Phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)} - \left(\dfrac{\phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}{1 - \Phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}\right)^2\right]$ is the marginal variance

of Y_{2LT}.

Let ρ_{LT} represent the correlation between Y_{1LT} and Y_{2LT}. The conditional distribution of Y_{1LT} given Y_{2LT} is a LTN distribution with mean $\mu_{y_{1LT}|y_{2LT}} = \mu_{y_{1LT}} + \dfrac{\sigma_{y_{1LT}}}{\sigma_{y_{2LT}}} \rho_{LT}\left(y_{2LT} - \mu_{y_{2LT}}\right)$, and variance $\sigma^2_{y_{1LT}|y_{2LT}} = \left(1 - \rho^2_{LT}\right)\sigma^2_{y_{1LT}}$. Also the conditional distribution of Y_{2LT} given Y_{1LT} is a LTN distribution with mean $\mu_{y_{2LT}|y_{1LT}} = \mu_{y_{2LT}} + \dfrac{\sigma_{y_{2LT}}}{\sigma_{y_{1LT}}} \rho_{LT}\left(y_{1LT} - \mu_{y_{1LT}}\right)$, and variance $\sigma^2_{y_{2LT}|y_{1LT}} = \left(1 - \rho^2_{LT}\right)\sigma^2_{y_{2LT}}$.

Right Truncated Bivariate Normal Distribution

The joint distribution of random variables Y_{1RT} and Y_{2RT} is a right truncated bivariate normal (RTBVN) distribution with mean vector $\boldsymbol{\mu} = \begin{bmatrix} \mu_{y_1} & \mu_{y_2} \end{bmatrix}'$, a covariance matrix

$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{y_1}^2 & \rho\sigma_{y_1}\sigma_{y_2} \\ \rho\sigma_{y_1}\sigma_{y_2} & \sigma_{y_2}^2 \end{bmatrix}$, and right truncation vector $\mathbf{t} = \begin{bmatrix} t_{y_1} & t_{y_2} \end{bmatrix}'$. Each random variable has a support of $Y_{1RT} \in (-\infty, t_{y_1})$ and $Y_{2RT} \in (-\infty, t_{y_2})$. The RTBVN pdf is denoted as,

$$f_{RTBVN}\left(y_{1RT}, y_{2RT}\right) = \frac{1}{\int_{-\infty}^{\frac{t_{y_1}-\mu_{y_1}}{\sigma}} \int_{-\infty}^{\frac{t_{y_2}-\mu_{y_2}}{\sigma}} f_{BVN}\left(y_{1RT}, y_{2RT}\right) dy_{1RT} dy_{2RT}} f_{BVN}\left(y_{1RT}, y_{2RT}\right)$$

where $t_{y_1} \le y_{1LT} \le \infty$ and $-\infty \le y_{2RT} \le t_{y_2}$.

The marginal distributions of Y_{1RT} and Y_{2RT} are RTN distributions. The marginal mean of Y_{1RT} is

$$\mu_{y_{1RT}} = \mu_{y_1} - \sigma_{y_1} \frac{\phi\left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right)}{\Phi\left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right)} \quad \text{and}$$

$$\sigma_{y_{1RT}}^2 = \sigma_{y_1}^2 \left[1 - \left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right) \frac{\phi\left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right)}{\Phi\left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right)} - \left(\frac{\phi\left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right)}{\Phi\left(\frac{t_{y_1}-\mu_{y_1}}{\sigma_{y_1}}\right)} \right)^2 \right] \quad \text{is the marginal variance}$$

of Y_{1RT}. Marginally, the mean of Y_{2RT} is $\mu_{y_{2RT}} = \mu_{y_2} - \sigma_{y_2} \frac{\phi\left(\frac{t_{y_2}-\mu_{y_2}}{\sigma_{y_2}}\right)}{\Phi\left(\frac{t_{y_2}-\mu_{y_2}}{\sigma_{y_2}}\right)}$ and the variance of is expressed as $\sigma_{y_{2RT}}^2 = \sigma_{y_2}^2 \left[1 - \left(\frac{t_{y_2}-\mu_{y_2}}{\sigma_{y_2}}\right) \frac{\phi\left(\frac{t_{y_2}-\mu_{y_2}}{\sigma_{y_2}}\right)}{\Phi\left(\frac{t_{y_2}-\mu_{y_2}}{\sigma_{y_2}}\right)} - \left(\frac{\phi\left(\frac{t_{y_2}-\mu_{y_2}}{\sigma_{y_2}}\right)}{\Phi\left(\frac{t_{y_2}-\mu_{y_2}}{\sigma_{y_2}}\right)} \right)^2 \right]$.

The conditional distribution of Y_{1RT} given Y_{2RT} is a RTN distribution with mean

$$\mu_{y_{1RT}|y_{2RT}} = \mu_{y_{1RT}} + \frac{\sigma_{y_{1RT}}}{\sigma_{y_{2RT}}} \rho_{RT} \left(y_{2RT} - \mu_{y_{2RT}} \right) \text{ and variance } \sigma^2_{y_{1RT}|y_{2RT}} = \left(1 - \rho^2_{RT}\right) \sigma^2_{y_{1RT}}$$

where ρ_{RT} is the correlation between Y_{1RT} and Y_{2RT}. Correspondingly the conditional distribution of Y_{2RT} given Y_{1RT} follows a RTN distribution with mean

$$\mu_{y_{2RT}|y_{1RT}} = \mu_{y_{2RT}} + \frac{\sigma_{y_{2RT}}}{\sigma_{y_{1RT}}} \rho_{RT} \left(y_{1RT} - \mu_{y_{1RT}} \right) \text{ and variance } \sigma^2_{y_{2RT}|y_{1RT}} = \left(1 - \rho^2_{RT}\right) \sigma^2_{y_{2RT}} .$$

3. ESTABLISH THE EQUIVALENCE OF LEFT TRUNCATION AND LEFT CENSORING FRAMEWORK FOR APPLICATIONS INVOLVING LIMITS OF DETECTION

INTRODUCTION

Unobserved data below a lower limit of detection (LLOD) occur when a measured analyte is below the level of accuracy that an analytical tool can recognize.[1] Standard formulas for estimating parameters (e.g. sample mean and variance) ignore the observations that are below the LLOD and are thus biased.[80] For example, ignoring observations below the LLOD causes the sample mean to be overestimated and the sample variance to be underestimated. A frequent remedy used in such situations is to replace unobserved data below the LLOD with the LLOD itself, LLOD/$\sqrt{2}$, or LLOD/2. The United States Environmental Protection Agency recommends these replacement methods when 15% or less of the observations are below the LLOD.[34] Despite the popularity of these replacement methods, they can lead to inaccurate statistical inferences about the parameters.[29]

An alternative to replacement methods described above is to assume a model that accounts for the unobserved data below the LLOD. In the literature, such data are treated either as coming from a left truncated distribution, or the unobserved observations are treated as left censored. Both approaches require the LLOD value to be known and specified. A left truncated model assumes a distribution that has a support restricted by the LLOD value. In the left censored model, observations below the LLOD are considered censored. However, the literature lacks guidance with respect to which method might be more appropriate in general or if they are equivalent.

In this brief report, we will demonstrate the equivalence of the two approaches in terms of estimating the parameters of the underlying distribution (continuous or discrete) using theory based on maximum likelihood estimators (MLEs). We have organized this report as follows. First, a review of left truncation and left censoring approaches for data with LLOD are described. Second, a theoretical rationale paralleling the MLEs of parameters from distributions either as left truncated or left censored, is provided. This is followed by a numerical illustration of their equivalence through Monte-Carlo simulations. In the subsequent section an application using BDE-153, a polybrominated diphenyl ether contaminant, from an observational cohort study, is presented. Lastly, we discuss and summarize our findings.

BELOW THE LLOD DATA: TWO LIKELIHOODS

Suppose we have a random sample of n_1 observations that are fully observed, and we know n_2 are not observed because they are below the LLOD. The total sample size is $n = n_1 + n_2$. Let, t_l denote the left truncation value which is equal to the LLOD, and y_i represent the i^{th} observation. Also let $f(y_i;\boldsymbol{\theta})$ be a probability density function, $F(t_l;\boldsymbol{\theta})$ is the corresponding cumulative distribution function, and $\boldsymbol{\theta}$ is a $p \times 1$ vector of p parameters of the distribution. The likelihood under the left truncated model can be written,

$$L_t(\boldsymbol{\theta};\boldsymbol{y}) = \prod_{i=1}^{n_1} \frac{f(y_i;\boldsymbol{\theta})}{1-F(t_l;\boldsymbol{\theta})}. \quad (1)$$

The corresponding log-likelihood is,

$$l_t(\boldsymbol{\theta}) = \sum_{i=1}^{n_1} \ln f(y_i;\boldsymbol{\theta}) - n_1 \ln\left(1-F(t_l;\boldsymbol{\theta})\right). \quad (2)$$

For the left censored model, let δ_i be an indicator for which,

$$\delta_i = \begin{cases} 1 \text{ if } y_i > t_l \\ 0 \text{ if } y_i \le t_l . \end{cases}$$

That is, $\delta_i = 1$ for all observations that are above LLOD (not 'censored'), and 0 when they are not observed. The likelihood under the left censored model is,

$$L_c(\boldsymbol{\theta}; \boldsymbol{y}) = \prod_{i=1}^{n} f(y_i; \boldsymbol{\theta})^{\delta_i} F(t_l; \boldsymbol{\theta})^{1-\delta_i}$$

$$= \left[\prod_{i}^{n_1} f(y_i; \boldsymbol{\theta}) \right] \left(F(t_l; \boldsymbol{\theta}) \right)^{n_2} . \quad (3)$$

The corresponding log-likelihood is

$$l_c(\boldsymbol{\theta}) = \sum_{i=1}^{n_1} \ln f(y_i; \boldsymbol{\theta}) + n_2 \ln F(t_l; \boldsymbol{\theta}) . \quad (4)$$

In general, the choice between truncation and censoring analysis methods may be based on whether the information about the left truncation value is known and/or the number of observations below the LLOD is known. That is, the four scenarios regarding the left truncation value and the number of observations below the LLOD could be tabulated as in Table 1.

Table 1 Method of choice under various scenarios

	Left Truncation Value, t_l, Unknown	Left Truncation Value, t_l, Known
Number of Below the LLOD Observations Unknown, n_2	1. Neither Method	2. Left Truncation Methods
Number of Below the LLOD Observations Known, n_2	3. Left Censoring and Left Truncation Methods	4. Left Censoring and Left Truncation Methods

Scenario 1 in Table 1 occurs when the left truncation value and the number of observations below the LLOD are both unknown. In this scenario, there is inadequate information to fully

define the likelihoods in equations (1) and (3) because they require specification of n_1, n_2, and/or t_l. In scenario 2, the observed data likelihood from the left truncated model can still be defined. However, the censored data likelihood, which depends on n_2, cannot be defined. In scenario 3, both models can be applied by treating t_l as an unknown additional parameter that could be estimated from the same likelihoods. For instance, since the number of observations below LLOD is known, under the truncated model the proportion of observations below the LLOD could provide the information necessary to estimate the LLOD.

In scenario 4 the left truncation value and number of below the LLOD observations are both known. This is the most common scenario in the applications discussed in the introduction section. We will establish the equivalence of the two likelihoods below and argue, for scenarios 3 and 4, that the two methods will yield identical MLEs.

THEORETICAL EQUIVALENCE OF MLES

The MLEs of $\boldsymbol{\theta}$ are acquired from estimating equations under standard regularity conditions by equating the first derivative of the log likelihoods to 0. Thus the estimating equation for the left truncated model is,

$$\sum_{i=1}^{n_1} \frac{df(y_i;\boldsymbol{\theta})}{d\boldsymbol{\theta}} \frac{1}{f(y_i;\boldsymbol{\theta})} + n_1 \frac{dF(t_l;\boldsymbol{\theta})}{d\boldsymbol{\theta}} \frac{1}{1-F(t_l;\boldsymbol{\theta})} = 0, \tag{5}$$

and the left censored model estimating equation is,

$$\sum_{i=1}^{n_1} \frac{df(y_i;\boldsymbol{\theta})}{d\boldsymbol{\theta}} \frac{1}{f(y_i;\boldsymbol{\theta})} + n_2 \frac{dF(t_l;\boldsymbol{\theta})}{d\boldsymbol{\theta}} \frac{1}{F(t_l;\boldsymbol{\theta})} = 0. \tag{6}$$

The equations (5) and (6) will produce similar MLEs when the left hand side of the equations (5) and (6) are equal. Suppose $\boldsymbol{\theta}$ for the left truncated model is $\boldsymbol{\theta}_l$ while $\boldsymbol{\theta}_c$ is for the left censored model. Equating (5) and (6) and simplifying leads to the condition for equality to be

$$n_2 \frac{1}{F(t_l; \boldsymbol{\theta}_c)} = n_1 \frac{1}{1-F(t_l; \boldsymbol{\theta}_l)} \quad . \tag{7}$$

This equality holds under expectations. That is, since $E\left(n_1\right) = n\left(1-F(t_l; \boldsymbol{\theta}_c)\right)$ and $E\left(n_2\right) = nF(t_l; \boldsymbol{\theta}_c)$ both sides will be n. Also for any given $\boldsymbol{\theta}$, $\frac{n_2}{n}$ is the MLE of the proportion of the distribution below t_l, $\hat{F}(t_l; \boldsymbol{\theta})$, and $\frac{n_1}{n}$ is the MLE of the proportion of the distribution above t_l, $1-\hat{F}(t_l; \boldsymbol{\theta})$, where $\hat{F}$ is the MLE of the cumulative distribution function evaluated at t_l. Therefore,

$$\frac{n_2}{n_1} = \frac{\hat{F}(t_l; \boldsymbol{\theta})}{1-\hat{F}(t_l; \boldsymbol{\theta})}$$

is a consistent estimator of $\frac{F(t_l; \boldsymbol{\theta}_c)}{1-F(t_l; \boldsymbol{\theta}_c)}$. In other words, for a large sample (n) the probability that the equation (7) will hold is 1.

Likewise we can show that the MLE of $\boldsymbol{\theta}$ from likelihoods under right truncated and right censored models are equivalent for scenarios 3 and 4. The MLE of $\boldsymbol{\theta}$ is also comparable in the presence of doubly truncated and doubly censoring likelihoods. Doubly truncated or doubly censored data occurs when there are observations that appear below an LLOD and above an upper limit of detection (ULOD) simultaneously.[5] Suppose there is a random sample of n_1

observations that are fully observed, n_2 are not observed because they are below the LLOD, and n_3 are not observed due to being above the ULOD. The total sample size is $n = n_1 + n_2 + n_3$. Let t_l represent the left truncation value, t_r is the right truncation value, and y_i denote the i^{th} observation. Also let $f(y_i;\boldsymbol{\theta})$ denote a probability density function where $t_l < y_i < t_r$, $F(t_r;\boldsymbol{\theta})$ and $F(t_l;\boldsymbol{\theta})$ denote cumulative distribution functions, and $\boldsymbol{\theta}$ is a vector with *p* parameters of the distribution. The likelihood under the doubly truncated model is,

$$L_{dt}(\boldsymbol{\theta};\boldsymbol{y}) = \prod_{i=1}^{n_1} \frac{f(y_i;\boldsymbol{\theta})}{F(t_r;\boldsymbol{\theta}) - F(t_l;\boldsymbol{\theta})}. \tag{8}$$

The corresponding log-likelihood is denoted as,

$$l_{dt}(\boldsymbol{\theta}) = \sum_{i=1}^{n_1} \ln f(y_i;\boldsymbol{\theta}) - n_1 \ln\left(F(t_r;\boldsymbol{\theta}) - F(t_l;\boldsymbol{\theta})\right). \tag{9}$$

The estimating equation of the log-likelihood under the doubly truncated model is,

$$\sum_{i=1}^{n_1} \frac{df(y_i;\boldsymbol{\theta})}{d\boldsymbol{\theta}} \frac{1}{f(y_i;\boldsymbol{\theta})} - n_1 \frac{1}{F(t_r;\boldsymbol{\theta}) - F(t_l;\boldsymbol{\theta})}\left(\frac{dF(t_r;\boldsymbol{\theta})}{d\boldsymbol{\theta}} - \frac{dF(t_l;\boldsymbol{\theta})}{d\boldsymbol{\theta}}\right) = 0. \tag{10}$$

Before displaying the likelihood under the doubly censored model, two indicators ω_i and γ_i must be defined. The two indicators represent the following,

$$\omega_i = \begin{cases} 1 \text{ if } t_l < y_i < t_r \\ 0 \text{ if otherwise} \end{cases} \text{ and } \gamma_i = \begin{cases} 1 \text{ if } y_i \le t_l \\ 0 \text{ if } y_i > t_l \end{cases}.$$

The likelihood under the doubly censored model is,

$$L_{dc}(\boldsymbol{\theta};\boldsymbol{y}) = \prod_{i=1}^{n} f(y_i;\boldsymbol{\theta})^{\omega_i} F(t_l;\boldsymbol{\theta})^{\gamma_i} \left(1 - F(t_r;\boldsymbol{\theta})\right)^{1-\omega_i-\gamma_i}$$

$$= \left(\prod_{i}^{n_1} f(y_i;\boldsymbol{\theta}) \right) \left(F(t_l;\boldsymbol{\theta}) \right)^{n_2} \left(1 - F(t_r;\boldsymbol{\theta}) \right)^{n_3} \tag{11}$$

The log-likelihood under the doubly censored model is denoted as,

$$l_{dc}(\boldsymbol{\theta}) = \sum_{i=1}^{n_1} \ln f(y_i;\boldsymbol{\theta}) + n_2 \ln F(t_l;\boldsymbol{\theta}) + n_3 \ln \left(1 - F(t_r;\boldsymbol{\theta}) \right). \tag{12}$$

The estimating equation under the doubly censored model is,

$$\sum_{i=1}^{n_1} \frac{df(y_i;\boldsymbol{\theta})}{d\boldsymbol{\theta}} \frac{1}{f(y_i;\boldsymbol{\theta})} + n_2 \frac{1}{F(t_l;\boldsymbol{\theta})} \frac{dF(t_l;\boldsymbol{\theta})}{d\boldsymbol{\theta}} - n_3 \frac{1}{1 - F(t_r;\boldsymbol{\theta})} \frac{dF(t_r;\boldsymbol{\theta})}{d\boldsymbol{\theta}} = 0 . \tag{13}$$

Although the equivalence of the two approaches can be shown in general, for simplicity we demonstrate it for the case where the assumption $F(t_l;\boldsymbol{\theta}) = 1 - F(t_r;\boldsymbol{\theta})$ is made. Under this assumption, let $\boldsymbol{\theta}$ for the doubly truncated model and doubly censored model be denoted as $\boldsymbol{\theta}_{dt}$ and $\boldsymbol{\theta}_{dc}$ respectively. By equating (10) and (13) we have the following equality,

$$\left(n_2 - n_3 \right) \frac{1}{F(t_l;\boldsymbol{\theta}_{dc})} = n_1 \frac{1}{F(t_r;\boldsymbol{\theta}_{dt}) - F(t_l;\boldsymbol{\theta}_{dt})} .$$

This equality holds due to $E(n_1) = n \left[F(t_r;\boldsymbol{\theta}) - F(t_l;\boldsymbol{\theta}) \right]$, $E(n_2) = nF(t_l;\boldsymbol{\theta})$, and $E(n_3) = n \left[1 - F(t_r;\boldsymbol{\theta}) \right]$. Also for any given $\boldsymbol{\theta}$, $\frac{n_1}{n}$ is the MLE of the proportion of the distribution between t_l and t_r (i.e. $\hat{F}(t_r;\boldsymbol{\theta}) - \hat{F}(t_l;\boldsymbol{\theta})$), $\frac{n_2}{n}$ is the MLE of the proportion of the distribution below t_l, and $\frac{n_3}{n}$ is the MLE of the proportion of the distribution above t_r.

COMPUTATIONAL ADVANTAGE OF LEFT TRUNCATION APPROACH

For obtaining the MLE of the parameters denoted by $\widehat{\boldsymbol{\theta}}$, the Newton-Raphson method is typically utilized for a likelihood under the left censored model, while for the left truncated model the Expectation-Maximization (EM) algorithm is often used. Recall that for estimating $\widehat{\boldsymbol{\theta}}$ using the Newton-Raphson method, an equation is iteratively computed until there is a small difference between $\boldsymbol{\theta}_{new}$, the current value for $\boldsymbol{\theta}$, and the previous value for $\boldsymbol{\theta}$ which is denoted as $\boldsymbol{\theta}_{old}$. The equation for the Newton-Raphson method is

$$\boldsymbol{\theta}_{new} = \boldsymbol{\theta}_{old} + \left(\frac{d^2 l(\boldsymbol{\theta})}{d\boldsymbol{\theta}} |_{\boldsymbol{\theta}=\boldsymbol{\theta}_{old}} \right)^{-1} \frac{dl(\boldsymbol{\theta})}{d\boldsymbol{\theta}} |_{\boldsymbol{\theta}=\boldsymbol{\theta}_{old}},$$

where $\frac{dl(\boldsymbol{\theta})}{d\boldsymbol{\theta}} |_{\boldsymbol{\theta}=\boldsymbol{\theta}_{old}}$ is a p dimensional vector of the first derivatives of the log-likelihood of $\boldsymbol{\theta}$ with respect to $\boldsymbol{\theta}$ and $\frac{d^2 l(\boldsymbol{\theta})}{\boldsymbol{\theta}} |_{\boldsymbol{\theta}=\boldsymbol{\theta}_{old}}$ is a $p \times p$ matrix of the second and partial derivatives of the log-likelihood of $\boldsymbol{\theta}$.

The EM algorithm involves iterating between two steps for estimating $\widehat{\boldsymbol{\theta}}$.[50] In the first step, known as the E-step, the first moment of the log-likelihood of $\boldsymbol{\theta}$ given the observed data denoted as $E\left(l(\boldsymbol{\theta}) | \boldsymbol{y}_{observed}\right)$ is computed.[66] The M-step, maximizes $E\left(l(\boldsymbol{\theta}) | \boldsymbol{y}_{observed}\right)$ with respect to $\boldsymbol{\theta}$. Commonly, $E\left(l(\boldsymbol{\theta}) | \boldsymbol{y}_{observed}\right)$ is maximized by finding the first derivative of $E\left(l(\boldsymbol{\theta}) | \boldsymbol{y}_{observed}\right)$ with respect to $\boldsymbol{\theta}$ and forming an estimating equation by setting the first derivative to zero. The estimating equation is used to solve for $\widehat{\boldsymbol{\theta}}$.

When below LLOD data arise in complex study designs, such as longitudinal studies, clustered data or multivariate data, the Newton-Raphson approach could become computationally intense given the need for repeated matrix inversion along with the need for first, second, and partial derivatives of the log-likelihood with respect to $\boldsymbol{\theta}$.[51] The advantages of the EM algorithm are well known. Basically, it does not require partial derivatives for estimating multiple parameters, and the convergence is faster at the beginning stages of the algorithms. The actual computation times for the Newton-Raphson method and EM algorithm are included within the discussion section.

SIMULATION STUDY

To numerically illustrate that left truncation and left censoring approaches produce equivalent parameter estimates, a left truncated normal distribution and a normal distribution with left censored observations are utilized in this simulation study. The data is generated from a normal distribution with $\mu = 5$ and $\sigma^2 = 4$. The total sample sizes are $n = 50$, 100, and 500. The left truncation value is computed by $t_l = z_k \times \sigma + \mu$ where k is the proportion of observations that are below the LLOD and z_k is the quantile associated with $F(t_l; \mu, \sigma^2) = k$. The values for t_l are 2.437, 2.927, 3.317, and 3.951, which allow for $k \times 100\% = 10\%, 15\%, 20\%,$ and 30% the observations to appear below the LLOD, respectively. Any generated observations less than t_l are considered to be below the LLOD observations.

The MLEs of μ and σ^2 are found by applying the EM algorithm and the Newton-Raphson method for the left truncation and left censoring approaches, respectively. Observed data is distributed as a left truncated normal distribution in the EM algorithm. It has been shown that the

first and second moments are necessary in the E-Step of the EM algorithm for estimating the parameters of a normal distribution.[66] The inclusion of the second moment improves the estimate of the variance. The Newton-Raphson method incorporates a normal distribution that treats the below the LLOD observations as left censored.

The results of the simulation study represent an average of 1,000 simulations. The MLE of μ and σ^2 are denoted as $\widehat{\mu}$ and $\widehat{\sigma^2}$. The mean square error (MSE) of $\widehat{\mu}$ was computed as $\text{MSE}(\widehat{\mu}) = \text{V}(\widehat{\mu}) + \text{Bias}(\widehat{\mu})$where $\text{V}(\widehat{\mu})$ is the variance of $\widehat{\mu}$, and the bias of $\widehat{\mu}$ is $\text{Bias}(\widehat{\mu}) = \widehat{\mu} - 5$. Similarly, the mean square error of $\widehat{\sigma^2}$ (MSE($\widehat{\sigma^2}$) was also computed. The simulations were conducted in R version 3.3.2.[81] The Newton-Raphson method was implemented using the R function *maxLik* function within the *maxLik* package.[82] The EM algorithm and Newton-Raphson method code are included in the appendices. The formulas of the EM algorithm steps that are included in the code were previously published.[65,66]

Table 2 Results of the simulation study examining the MLE at different percentages of observations being below the LLOD. Prior to any observations appearing below the LLOD when n = 50, $\widehat{\mu}$ = 4.995 and $\widehat{\sigma^2}$ = 3.911; n = 100, $\widehat{\mu}$= 5.000 and $\widehat{\sigma^2}$ = 3.954; n = 500, $\widehat{\mu}$ = 4.998 and $\widehat{\sigma^2}$ = 3.984.

Total Sample Size n	Percent of Below the LLOD Observations	Left Truncation Approach Estimate (MSE)	Left Censoring Approach Estimate (MSE)	Left Truncation Approach 95% Confidence Interval Coverage, $\left\|\frac{\hat{\mu}-\mu}{\sqrt{Var(\hat{\mu})}}\right\| \geq 1.96$	Left Censoring Approach 95% Confidence Interval Coverage, $\left\|\frac{\hat{\mu}-\mu}{\sqrt{Var(\hat{\mu})}}\right\| \geq 1.96$
50	10%	$\widehat{\mu}$ = 5.003 (0.086) $\widehat{\sigma^2}$ = 3.868 (0.864)	$\widehat{\mu}$ = 4.993 (0.085) $\widehat{\sigma^2}$ = 3.924 (0.822)	93.8%	94.4%
100	10%	$\widehat{\mu}$ = 5.004 (0.043) $\widehat{\sigma^2}$ = 3.932 (0.424)	$\widehat{\mu}$ = 4.999 (0.042) $\widehat{\sigma^2}$ = 3.960 (0.411)	93.9%	94.6%
500	10%	$\widehat{\mu}$ = 4.998 (0.008) $\widehat{\sigma^2}$ = 3.987 (0.086)	$\widehat{\mu}$ = 4.998 (0.008) $\widehat{\sigma^2}$ = 3.989 (0.081)	94.6%	95.3%
50	15%	$\widehat{\mu}$ = 5.004 (0.090) $\widehat{\sigma^2}$ = 3.868(0.947)	$\widehat{\mu}$ = 4.991 (0.088) $\widehat{\sigma^2}$ = 3.937 (0.899)	93.6%	93.4%
100	15%	$\widehat{\mu}$ = 5.005 (0.044) $\widehat{\sigma^2}$ = 3.931 (0.456)	$\widehat{\mu}$ = 4.999 (0.043) $\widehat{\sigma^2}$ = 3.963 (0.432)	92.9%	94.1%
500	15%	$\widehat{\mu}$ = 4.998 (0.008) $\widehat{\sigma^2}$ = 3.990 (0.095)	$\widehat{\mu}$ = 4.998 (0.008) $\widehat{\sigma^2}$ = 3.991(0.086)	94.4%	94.7%
50	20%	$\widehat{\mu}$ = 5.007 (0.093) $\widehat{\sigma^2}$ = 3.862 (1.033)	$\widehat{\mu}$ = 4.989 (0.091) $\widehat{\sigma^2}$ = 3.950 (0.987)	92.7%	93.3%
100	20%	$\widehat{\mu}$ = 5.006 (0.045) $\widehat{\sigma^2}$ =3.930 (0.495)	$\widehat{\mu}$ = 4.998 (0.044) $\widehat{\sigma^2}$ = 3.969 (0.475)	92.5%	93.0%
500	20%	$\widehat{\mu}$ = 4.998 (0.009) $\widehat{\sigma^2}$ = 3.991 (0.105)	$\widehat{\mu}$ = 4.996 (0.008) $\widehat{\sigma^2}$ = 3.998 (0.092)	93.7 %	94.1%
50	30%	$\widehat{\mu}$ = 5.017 (0.100) $\widehat{\sigma^2}$ = 3.831 (1.188)	$\widehat{\mu}$ = 4.989 (0.097) $\widehat{\sigma^2}$ = 3.951 (1.146)	90.4%	92.4%
100	30%	$\widehat{\mu}$ = 5.010 (0.050) $\widehat{\sigma^2}$ = 3.917 (0.592)	$\widehat{\mu}$ = 4.997 (0.047) $\widehat{\sigma^2}$ = 3.975 (0.529)	91.6%	92.7%
500	30%	$\widehat{\mu}$ = 4.998 (0.010) $\widehat{\sigma^2}$ = 3.990 (0.130)	$\widehat{\mu}$ = 4.997 (0.009) $\widehat{\sigma^2}$ = 3.996 (0.111)	92.7%	93.3%

Table 2 displays the results from the simulation study. The results of $\widehat{\mu}$, $\widehat{\sigma^2}$, and the MSE$(\widehat{\mu})$ were nearly identical as the total sample size increases from 50 to 500 regardless of the percentage of below the LLOD observations. With the total sample size of 100, there were minimal differences between the estimates for $\widehat{\mu}$ found using the left truncation and left censoring methods. Also we listed the 95% confidence interval coverage for the approaches in Table 2, and the left truncation approach has a slightly larger 95% coverage interval than the left censoring approach. The 95% confidence interval coverage improved as the sample size increased.

APPLICATION IN SLEIGH STUDY

Data with below LLOD observations appeared in the Systemic Lupus Erythematosus in Gullah Health (SLEIGH) study. SLEIGH is an observational cohort study of African American Gullah patients with systemic lupus erythematosus (SLE) and control participants. The Gullah population are a unique group of African Americans, whose ancestors resided on the South Carolina and Georgia coastal barrier islands. SLEIGH included a toxicology and serology component in which serum samples were collected from the participants.[83] The concentration of a polybrominated diphenyl ether contaminant, BDE-153, present in the serum samples of study participants was measured on a continuous scale of nanogram per gram (ng/g) in a laboratory.

The SLEIGH study consisted of 65 study participants with SLE and 123 controls that had serum samples evaluated for contaminant BDE-153. The log base 10 scale of the BDE-153 contaminant data follows a normal distribution. Often ecological pollutant data are not normally distributed, but the data is normally distributed once a log transformation is applied.[84-87] The observed data

likelihood is assumed to be a left truncated normal distribution in the EM algorithm, while a likelihood of the normal distribution with left censored observations was used for the Newton-Raphson method. In this study the LLOD for the BDE-153 contaminant is -5.991 log ng/g (0.0025 ng/g). Approximately, 21.5% of the participants with SLE and 11.4% of the controls have BDE-153 observations that are below -5.991 log ng/g. The EM algorithm and Newton-Raphson method were used to obtain the MLEs of μ and σ^2 for BDE-153, separately for SLE and control participants.

Table 3 MLE results of contaminant BDE-153 by systemic lupus erythematosus (SLE) disease status with a truncation value of -5.991 log ng/g (0.0025 ng/g).

Percent of Observations Below the LLOD	Left Truncation Approach	Left Censoring Approach
Participants with SLE, $n = 65$		
21.5%	$\widehat{\mu}$ = -3.972 log ng/g $\widehat{\sigma^2}$ = 2.766 log ng/g	$\widehat{\mu}$ = -3.972 log ng/g $\widehat{\sigma^2}$ = 2.766 log ng/g
Control Participants, $n = 123$		
11.4%	$\widehat{\mu}$ = -3.695 log ng/g $\widehat{\sigma^2}$ = 1.822 log ng/g	$\widehat{\mu}$ = -3.695 log ng/g $\widehat{\sigma^2}$ = 1.822 log ng/g

The MLEs for the mean and the variance of BDE-153 by SLE and control participants are shown in Table 3. The MLEs computed by the two procedures are essentially identical. The left truncation and left censoring approaches produced $\widehat{\mu}$ = - 3.972 log ng/g and $\widehat{\sigma^2}$ = 2.766 log ng/g for participants with SLE. For control participants, the MLEs from the left truncation and left censoring approaches were $\widehat{\mu}$ = -3.695 log ng/g and $\widehat{\sigma^2}$ = 1.822 log ng/g.

DISCUSSION

In this paper, we explained how estimation methods relying on left truncation and left censoring approaches coincide in the estimation of parameters from any underlying distribution when there exists unobserved data because they are below LLOD and the number of such occurrences is known. When the left truncation value is unknown, it could be estimated using the proportion of observations below the LLOD. A simulation study and an application to real data were presented to further illustrate the theoretical relationship of the MLEs. The two approaches are implemented by the use of numerical algorithms to estimate the parameters. Although the advantages of one algorithm versus another is hard to establish in a straightforward design presented here, as the parameter space increases, the truncation approach paired with the EM algorithm is expected to have a significant computational advantage in comparison to the censoring approach using the Newton-Raphson method. Specifically, even in this simple situation, the left truncation approach using the EM algorithm was uniformly faster than the left censoring method in all situations in R version 3.3.2.[81] For example, in the simulation scenario of 30% of the observations appearing below the LLOD and $n = 500$, the computation time for the left truncation approach that relied on the EM algorithm was 4.658 seconds on a laptop computer with a processor of i7-7500U CPU @ 2.70 GHz and 8 GB of RAM while the time was 7.181 seconds (54.165% greater) for the left censoring approach under the Newton-Raphson method. In our simulations and real data application, an EM algorithm under the left censored model was not considered, because an EM solution under that likelihood is not available. Although EM algorithms for time-to-event parametric survival models have previously been described, no EM algorithms applicable to the situation described here currently exist in the literature.[64,88-90] Also, the EM algorithm provided under the left truncated model requires both first and second

moments in the E-step.[66] While the EM algorithm achieved faster convergence, when the underlying models get more complex, properly specifying the second moments could be considerably more complicated. For example, the E-step may involve cross product moments which involves all patterns of observed and incomplete data. We have established the equivalence between both approaches in general, however, we plan to extend the left truncation approach to include more complex models.

The censoring that appears in this paper is different than censoring occurring in a typical survival analysis framework. Unlike the typical survival analysis framework, observations that are considered to be censored in this paper have a common censoring point. Therefore, we do not include censoring within the survival analysis context in the theoretical relationship established in this paper.

In our study, we did not consider multiple LLODs occurring simultaneously for a given variable while showing the relationship between the MLEs of parameters using the truncation and censoring approaches. There are several left censoring methods established to account for multiple LLODs arising in a single variable, but these methods have yet to be compared with left truncation approaches.[15,16,19,20] Therefore future studies may consider examining the theoretical relationship of estimating parameters using left truncation and left censoring approaches with multiple left truncation values occurring in a single variable. For a single left truncation value, we were able to conclude that left truncation and left censoring approaches yield equivalent parameter estimates of an underlying distribution.

4. DEVELOP A METHOD FOR ESTIMATING THE MEAN AND VARIANCE OF A SINGLE NORMAL RANDOM VARIABLE WITH MULTIPLE LOWER LIMITS OF DETECTION ARISING FROM DIFFERENT BATCHES (I.E. ONE VARIABLE WITH MULTIPLE BATCHES)

INTRODUCTION

The lowest quantity of a substance that can be measured by an analytical tool is often called the lower limit of detection (LLOD).[1] Frequently statistical analysis involving data with observations below an LLOD has an assumption that the observations above the LLOD are from a left truncated distribution or that the LLOD observations are left censored. The support of the left truncated distribution is constrained to being greater than the LLOD, which also may be referred to as the left truncation value. In the left censoring framework, any observation below the fixed LLOD value is considered left censored. Statistical analysis methods to account for one left truncation value have been developed relying either on left truncated and left censored models. A commonly used continuous left truncated distribution is a left truncated normal distribution. The maximum likelihood estimators (MLEs) of the parameters from the left truncated normal distribution have been established under various assumptions such as whether or not the left truncation value is known.[5-7] For continuous left censored methods, a normal distribution with left censored observations is often considered, and the processes for estimating MLEs in this context have previously been outlined.[8,9]

When LLOD observations occur in a dataset, MLEs of parameters from an underlying distribution can be achieved by incorporating left truncated or left censored modeling strategies. Most estimation methods previously described for the left truncated framework assume only one LLOD within a univariate distribution, however the fact that it is possible for multiple LLOD

values to occur within a given experiment.[10,11] On the other hand, prior work using the left censored framework for normal distributions has included the situation involving multiple LLOD values occurring for a single random variable.[15-17,91,92] For example, Peng (2010) assembled confidence intervals for MLEs of parameters from a normal distribution with multiple left censored values, [17] and Aboueissa and Stoline (2009) derived methods for finding MLEs of parameters from a normal distribution with multiple left censored values.[15,16,92]

In an earlier manuscript, we established in general truncated and censored model approaches yield equivalent MLEs of the parameters from an underlying distribution when there is one LLOD value. We also argued the advantage of left truncated model in complex designs because it lends itself to the implementation of an EM algorithm more easily. In terms of convergence in complex designs, EM has a proven computational advantage in comparison to other iterative algorithms such as the Newton-Raphson method. The application of the left censored model through the Newton-Raphson method requires first, second, and partial derivatives of the log-likelihood of the parameters with respect to the parameters, but the left truncated model approach does not require derivatives. The EM algorithm applied using the log-likelihood of the left truncated model requires first and second moments of the distribution of the data appearing below the LLOD value. This advantage motivated us to extend the left truncated model and EM algorithm framework to handle multiple LLOD values.

To our knowledge, the statistics literature has not addressed the situation of using truncated models to obtain MLEs of parameters from an underlying normal distribution in which multiple LLOD values exist. Developing such a methodological framework could be beneficial for many

chemical substance and biomarker assay studies where the outcome may consist of data that has been generated over time with multiple LLOD values.[10,11] LLOD values vary throughout time, even within a 24 hour period, due to the calibration and detection sensitivity of the laboratory analytical tool.[12,13] Therefore, pooling data sets produced at different time periods for a variable can result in multiple LLOD values.[11,14]

The objective of this paper is to propose an estimation approach for finding MLEs for parameters (i.e. mean and variance) of a normal distribution for an outcome variable that has multiple LLOD values in the observed sample. The estimation technique relies on an expectation-maximization (EM) algorithm that incorporates a linear mixed model (LMM) in the M-step. Such an approach has been suggested in the past; Dempster, Laird, and Rubin proposed that observations that are not completely recorded due to a truncation value should be included in the context of missing data, and that the EM algorithm is an applicable statistical approach for estimating the MLEs of parameters when there are observations excluded from sample due to truncation value(s).[50] However, the implementation of their recommendation has not appeared in the literature so far with multiple LLOD values. After describing our version of the EM algorithm, a simulation study is included to compare the performance of the proposed EM algorithm with existing methods that account for multiple LLOD values. Finally, a real data application is presented.

METHODS

Consider data of a substance collected in *m* batches of data for an experiment. The collected substance is an outcome of interest for the experimental study. Suppose some data within each batch are unobserved but known to be less than a batch specific LLOD value. The number of

observations within the total sample size for the experiment is denoted as *n*, and *n* can be separated into components based on the *m* batches and whether an observation is below or above the LLOD. Let $j = 1,\ldots, m$ batches, $k = 1$ if the observation is above the batch specific LLOD, and $k = 2$ if the observation is below the LLOD. Each batch has a separate LLOD value which we will also refer to as a truncation value, and denote as t_j . The total sample size is $n = \sum_{j=1}^{m}\sum_{k=1}^{2} n_{jk}$ where n_{jk} is the sample size for the j^{th} batch that is either above or below the LLOD value.

Our proposed method for estimating parameters incorporates an LMM. Recall that an LMM is written as $\mathbf{Y} = \mathbf{X\beta} + \mathbf{Z\gamma} + \boldsymbol{\varepsilon}$. Let $\mathbf{Y}$ be an $n \times 1$ outcome vector, and each observation in $\mathbf{Y}$ is denoted as Y_{ij} where $i = 1, 2, ..., n$. $\mathbf{X}$ is a $n \times p$ matrix of p predictors, $\boldsymbol{\beta}$ is a vector of p fixed effects, $\mathbf{Z}$ is a $n \times q$ design matrix for the q random effects, $\boldsymbol{\gamma}$ is a vector of q random effects, and $\boldsymbol{\varepsilon}$ is a vector of the residuals. We assume that $\boldsymbol{\gamma} \sim \mathbf{N}(\mathbf{0}, \mathbf{G})$ where $\mathbf{G}$ is the covariance matrix of the random effects. Also we assume that $\boldsymbol{\varepsilon} \sim \mathbf{N}(\mathbf{0}, \mathbf{R})$ $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{R})$, and σ^2 is the residual variance such that $\mathbf{R} = \sigma^2\mathbf{I}_n$ .The variance for each observation is denoted as $\mathbf{V} = \mathbf{ZGZ}' + \mathbf{R}$. The distribution of $\mathbf{Y}$ is a multivariate normal distribution with mean $\mathbf{X\beta}$ and covariance matrix $\mathbf{V}$. The MLEs of the mean and covariance matrix are found by maximizing the log-likelihood.

The likelihood of the parameters $\mathbf{X\beta}$ and $\mathbf{V}$ given $\mathbf{Y}$ is expressed using the trace function as,

$$\mathbf{L}\left(\mathbf{X\beta, V}\,|\,\mathbf{Y}\right)=\frac{1}{\left(2\pi\right)^{n/2}\left|\mathbf{V}\right|^{1/2}}\exp\left\{-\frac{1}{2}\left(\mathbf{Y-X\beta}\right)'\,\mathbf{V}^{-1}\left(\mathbf{Y-X\beta}\right)\right\}$$

$$=\frac{1}{\left(2\pi\right)^{n/2}\left|\mathbf{V}\right|^{1/2}}\exp\left\{-\frac{1}{2}\,\text{trace}\left\{\mathbf{V}^{-1}\left(\mathbf{Y-X\beta}\right)\left(\mathbf{Y-Xb}\right)'\right\}\right\}$$

$$=\frac{1}{\left(2\pi\right)^{n/2}\left|\mathbf{V}\right|^{1/2}}\exp\left\{-\frac{1}{2}\,\text{trace}\left\{\mathbf{V}^{-1}\left(\mathbf{YY}'-2\mathbf{Y}\left(\mathbf{X\beta}\right)'+\left(\mathbf{X\beta}\right)'\,\mathbf{X\beta}\right)\right\}\right\}\quad.\qquad(1)$$

EM Algorithm

We are interested in estimating the mean and the variance of the total sample size for an outcome rather than a mean and variance for each batch separately. Due to our interest, we assume a constant mean and variance across *m* batches. We also assume observations within a batch are correlated, but the batches are independent of each other. We make an assumption that the underlying distribution of the batches of data follows a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ of $m\times 1$ dimension and a covariance matrix $\boldsymbol{\Sigma}$ of dimension $m\times m$. The density function of the multivariate normal distributed is

$f_{MVN}(\mathbf{y})=\frac{1}{\left(2\pi\right)^{n/2}\left|\boldsymbol{\Sigma}\right|^{1/2}}\exp\left\{-\frac{1}{2}\left(\mathbf{y-\mu}\right)'\,\boldsymbol{\Sigma}^{-1}\left(\mathbf{y-\mu}\right)\right\}$ where $\mathbf{y}\in\mathbb{R}^n$, and the corresponding cumulative distribution function is denoted as $F_{MVN}\left(\mathbf{y}\right)$. Since $\mathbf{Y}$ includes observations appearing below the LLOD, we cannot accurately estimate parameters of the multivariate normal distribution using $\mathbf{Y}$ as it is originally observed. Our proposed EM algorithm treats components of $\mathbf{Y}$ as data from either a multivariate left or right truncated normal distribution based on if the observation is above or below an LLOD. Observations above the LLOD for each batch follows a multivariate left truncated normal distribution with parameters $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ and $\mathbf{t}$, where

$\mathbf{t} = \left[t_1, t_2, ..., t_m \right]$ is a vector of the truncation values. The density function of the multivariate left truncated normal distribution is, $f_{MLTN}\left(\mathbf{y}\right) = \frac{f_{MVN}\left(\mathbf{y}\right)}{F_{MVN}\left(\mathbf{t}\right)}$, where $y_{ij} \geq t_j$ for all j. Furthermore in our method, the observations below the LLOD for each batch are assumed to follow a multivariate right truncated normal distribution $\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \mathbf{t}\right)$. The density function of the multivariate right truncated normal distribution is,

$f_{MRTN}\left(\mathbf{y}\right) = \frac{f_{MVN}\left(\mathbf{y}\right)}{F_{MVN}\left(\mathbf{t}\right)}$, where $y_{ij} \leq t_j$ for all j.

E-Step

A previous study demonstrated that first and second moments of a truncated normal distribution are necessary in the E-Step of an EM algorithm for accurately estimating the variance parameter from a univariate normal distribution.[66] The equations for the moments from the multivariate truncated normal distribution with left and right truncation occurring simultaneously in all variables have been derived by Manjunath and Wilhelm, but an example is provided for one direct of truncation occurring solely.[79] We adapt the formulas by Manjunath and Wilhelm for moments of a multivariate right truncated normal distribution for the E-Step of our method. The moments can be computed using the *tmvtnorm* R package.[76,93-96] The moments by Manjunath and Wilhelm are extensions of previous moments formulas with various truncation scenarios.[77,78] The formulas for the first and second moment of the response for the j^{th} batch are,

$E\left(Y_j\right) = \mu_j - \sum_{d=1}^{m} \sigma_{j,d}\left(F\left(t_j\right)\right)$ and

$$E\left(Y_j^2\right)=\sigma_{j,j}-\sum_{d=1}^{m}\sigma_{j,d}\frac{\sigma_{j,d}\left[\left(\frac{t_d-\mu_d}{\sigma_{d,d}}\right)F\left(t_j\right)\right]}{\sigma_{d,d}}$$
$$+\sum_{d=1}^{m}\sigma_{j,d}\sum_{w\neq d}\left(\sigma_{j,w}-\frac{\sigma_{d,w}\sigma_{j,d}}{\sigma_{d,d}}\right)\left[F\left(t_d,t_w\right)-F\left(t_d\right)-F\left(t_w\right)\right].$$

Here, $F\left(t_j\right)=\int_{-\infty}^{t_1}\ldots\int_{-\infty}^{t_{j-1}}\int_{-\infty}^{t_{j+1}}\ldots\int_{-\infty}^{t_m}f_{MLTN}\left(\mathbf{y}_{-ij}\right)d\mathbf{y}_{-ij}$ where $\mathbf{y}_{-ij}$ is a subset of $\mathbf{y}$ with the ij^{th} observation removed. Also let $F\left(t_j,t_{j'}\right)==\int_{-\infty}^{t_1}\ldots\int_{-\infty}^{t_{j-1}}\int_{-\infty}^{t_{j+1}}\ldots\int_{-\infty}^{t_{j'-1}}\int_{-\infty}^{t_{j'+1}}\ldots\int_{-\infty}^{t_m}f_{MLTN}\left(\mathbf{y}_{-ij,-ij'}\right)d\mathbf{y}_{-ij,-ij'}$ where $\mathbf{y}_{-ij,-ij'}$ is a subset of $\mathbf{y}$, the observations ij and ij' are excluded, and $j\neq j'$.

For the $\mathbf{Y}$ vector, the appropriate first moment replaces observations below the LLOD depending on which batch the observations belongs to. The simplified formula in equation (1) has $\mathbf{YY}'$ which is an element-wise multiplication matrix that will include the second moment. The elements of $\mathbf{YY}'$ are replaced with the batch specific second moment only if the element in the $\mathbf{Y}$ vector is an observation below the LLOD and the observation appears on the diagonal of $\mathbf{YY}'$. An example is below to illustrate how the moments replace elements in $\mathbf{Y}$ and $\mathbf{YY}'$. Let $\mathbf{Y}_{\text{observed}}$ denote the response vector as it is originally observed. Also let an asterisk denote the observations that are below the LLOD in $\mathbf{Y}_{\text{observed}}$. In this example $\mathbf{Y}_{\text{observed}}$, $\mathbf{Y}$, and $\mathbf{YY}'$ are as follows.

$$\mathbf{Y}_{\text{observed}} = \begin{bmatrix} Y_{11}^{*} \\ Y_{21} \\ \cdot \\ \cdot \\ \cdot \\ Y_{nm}^{*} \end{bmatrix}$$

$$\mathbf{Y} = \begin{bmatrix} E\left(Y_1\right) \\ Y_{21} \\ \cdot \\ \cdot \\ \cdot \\ E\left(Y_m\right) \end{bmatrix}$$

$$\mathbf{YY}' = \begin{bmatrix} E\left(Y_1^2\right) & E\left(Y_1\right)\times Y_{21} & \cdot\ \cdot\ \cdot & E\left(Y_1\right)\times E\left(Y_m\right) \\ Y_{21}\times E\left(Y_1\right) & Y_{21}\times Y_{21} & \cdot\ \cdot\ \cdot & Y_{21}\times E\left(Y_m\right) \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ E\left(Y_m\right)\times E\left(Y_1\right) & E\left(Y_m\right)\times Y_{21} & \cdot\ \cdot\ \cdot & E\left(Y_m^2\right) \end{bmatrix}$$

M-Step

Based on the LMM, the log-likelihood of the parameters $\mathbf{X\beta}$ and $\mathbf{V}$ given $\mathbf{Y}$ is,

$\ln\left(\mathbf{L}\left(\mathbf{X\beta},\, \mathbf{V} \mid \mathbf{Y}\right)\right) = -\frac{n}{2}\ln\left(2\pi\right) - \frac{1}{2}\ln\left|\mathbf{V}\right| - \frac{1}{2}tr\left\{\mathbf{V}^{-1}\left(\mathbf{YY}' - 2\mathbf{Y}\left(\mathbf{X\beta}\right)' + \left(\mathbf{X\beta}\right)'\mathbf{X\beta}\right)\right\}$. In our proposed EM algorithm $\mathbf{X}$ is a $n \times 1$ vector for the intercept, $\boldsymbol{\beta}$ is a scalar representing the intercept, $\mathbf{Z}$ is a $n \times 2$ matrix where the columns indicates which batch the observation belongs to, $\boldsymbol{\gamma}$ is the random batch effect, $\mathbf{G} = \mathrm{s}_{\text{batch}}^2$, and $\mathbf{R} = \sigma^2\mathbf{I}_n$. By the inclusion of $\mathbf{G}$ and $\mathbf{R}$, our method incorporates the assumption that the batches are independent of each other, but

observations within a batch are correlated. The proposed EM algorithm finds $\widehat{\mathbf{X}\boldsymbol{\beta}}$ by maximizing the log-likelihood. For estimating $\widehat{\mathbf{V}}$ from an LMM, our EM algorithm maximizes the log-likelihood of based on the conditional derivation of the residual maximum likelihood approach.[97] The log-likelihood of $\mathbf{V}$ that is maximized in the proposed EM algorithm is,

$$\ln\left(\mathbf{L}\left(\mathbf{V}\,|\,\mathbf{Y}\right)\right)=-\frac{1}{2}\left[\ln\left(\det\left(\mathbf{V}\right)\right)+\text{trace}\left\{\ln\left(\det\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)+\left(\mathbf{V}^{-1}-\mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{V}^{-1}\right)\mathbf{Y}\mathbf{Y}'\right)\right\}\right].$$

SIMULATION STUDY

In our simulation study, *m* = 2 batches. We generated data under two scenarios. In the first scenario, a random batch effect was produced by generating two numbers from a normal distribution with $\mu=5$ and $s^2_{\text{batch}}=1$. The two numbers are then used as the means of two normal distributions with $\sigma^2=4$ to generate two batches with a sample size 100 each so that the total sample size is $n=200$. The parameters for the first scenario are $\mu=5$, $\sigma^2=4$, and $s^2_{\text{batch}}=1$. For the second scenario, the parameters are $\mu=10$, $\sigma^2=4$, and $s^2_{\text{batch}}=1$. In both simulation scenarios, various percentages of the observations in each batch are removed to resemble an observation appearing below the LLOD. For batch 1 the percentage of observations below the LLOD is denoted as $t_1\%$, and similarly $t_2\%$ represents the percentage of observations below the LLOD in batch 2. Percentages of observations below the LLOD per batch considered in the simulation study are the following: $t_1\%=5\%$, $t_2\%=10\%$; $t_1\%=10\%$, $t_2\%=15\%$; $t_1\%=20\%$, $t_2\%=25\%$; and $t_1\%=30\%$, $t_2\%=35\%$. The results are an average of 1000 simulations.

The performance of the proposed EM algorithm is compared with simple substitution methods and an EM algorithm with the first moment of a multivariate right truncated normal distribution only. The simple substitution methods included are substituting the LLOD observations with LLOD itself, LLOD/2, and LLOD/√2.[29,34] The MLE of σ^2_{batch} cannot be estimated from the substitution methods. We compare our proposed EM algorithm to an EM algorithm with the first moment of a multivariate right truncated normal distribution only because we wanted to explore advantages of including the second moment in an EM algorithm when there are multiple LLOD values since the second moment is recommended in the case of one LLOD value.[66]

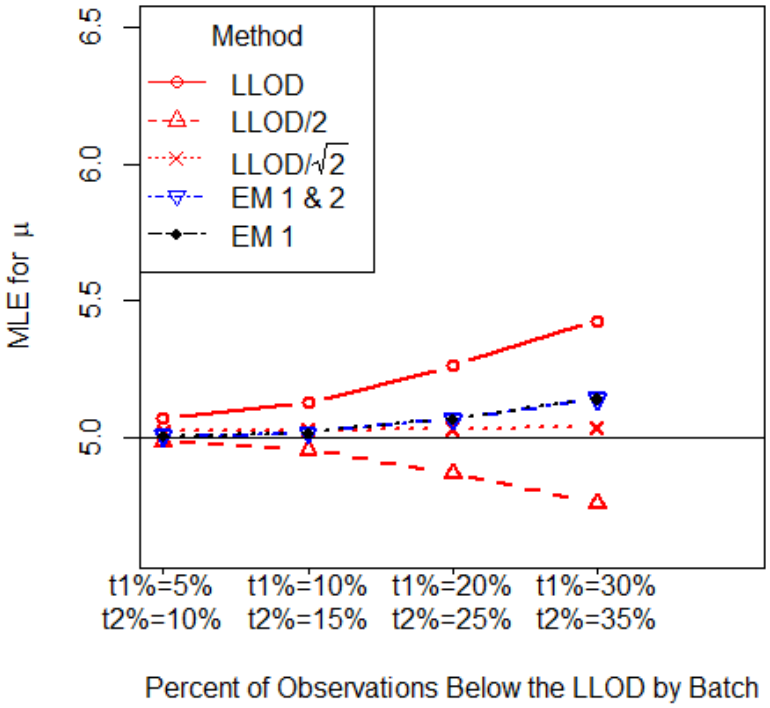

Figure 1 Plot of $\widehat{\mu}$ from the first simulation scenario with $\mu = 5$, $\sigma^2 = 4$, and $s^2_{batch} = 1$. A horizontal line is at $\mu = 5$. In the legend, "EM 1" represents an EM algorithm with the first moment of the multivariate right truncated normal distribution. Our proposed EM algorithm that includes first and second moments of the multivariate right truncated normal distribution is labeled as "EM 1 & 2".

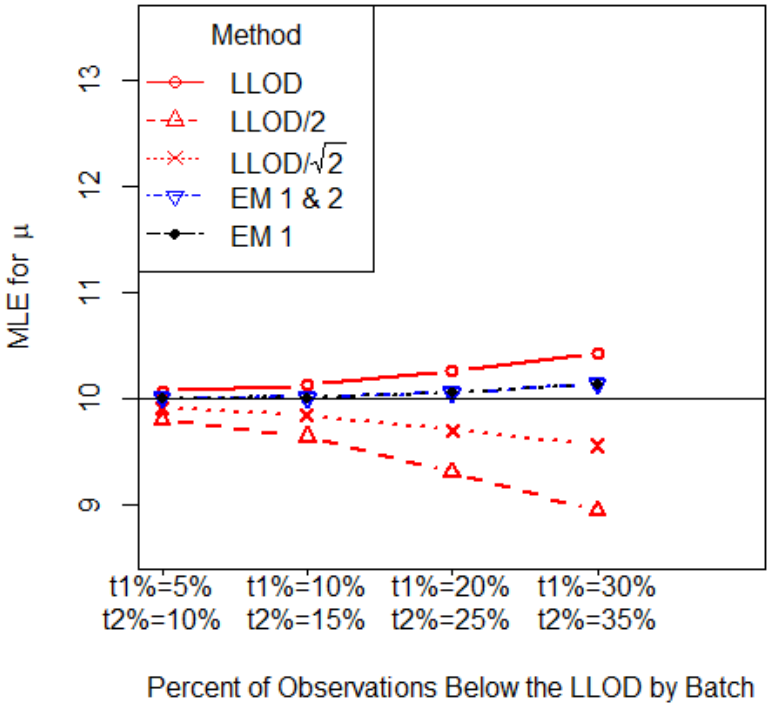

Figure 2 Plot of $\widehat{\mu}$ from the second simulation scenario with $\mu = 10$, $\sigma^2 = 4$, and $s^2_{batch} = 1$. A horizontal line is at $\mu = 10$. In the legend, "EM 1" represents an EM algorithm with the first moment of the multivariate right truncated normal distribution. Our proposed EM algorithm that includes first and second moments of the multivariate right truncated normal distribution is labeled as "EM 1 & 2".

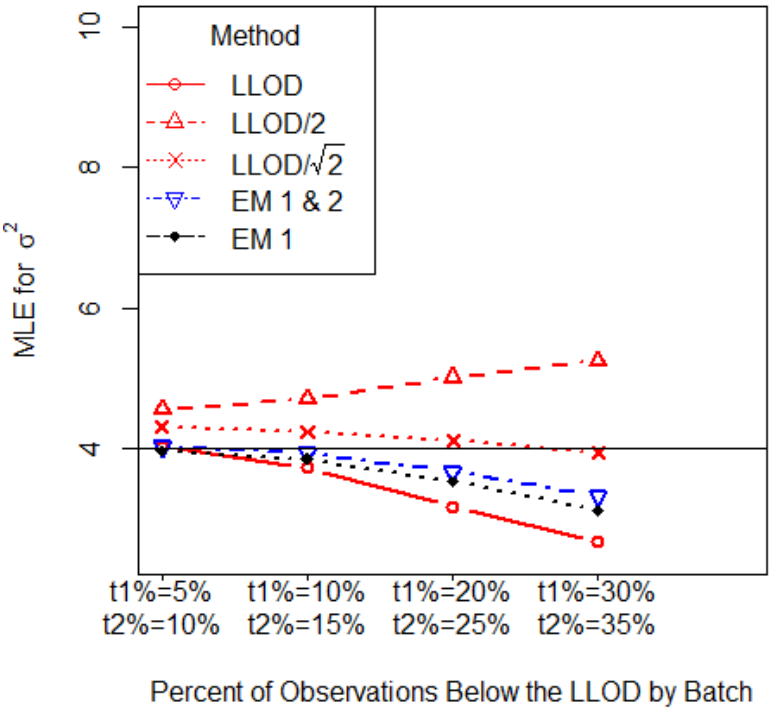

Figure 3 Plot of $\widehat{\sigma^2}$ from the first simulation scenario with $\mu = 5$, $\sigma^2 = 4$, and $s^2_{batch} = 1$. A horizontal line is at $\sigma^2 = 4$. In the legend, "EM 1" represents an EM algorithm with the first moment of the multivariate right truncated normal distribution. Our proposed EM algorithm that includes first and second moments of the multivariate right truncated normal distribution is labeled as "EM 1 & 2".

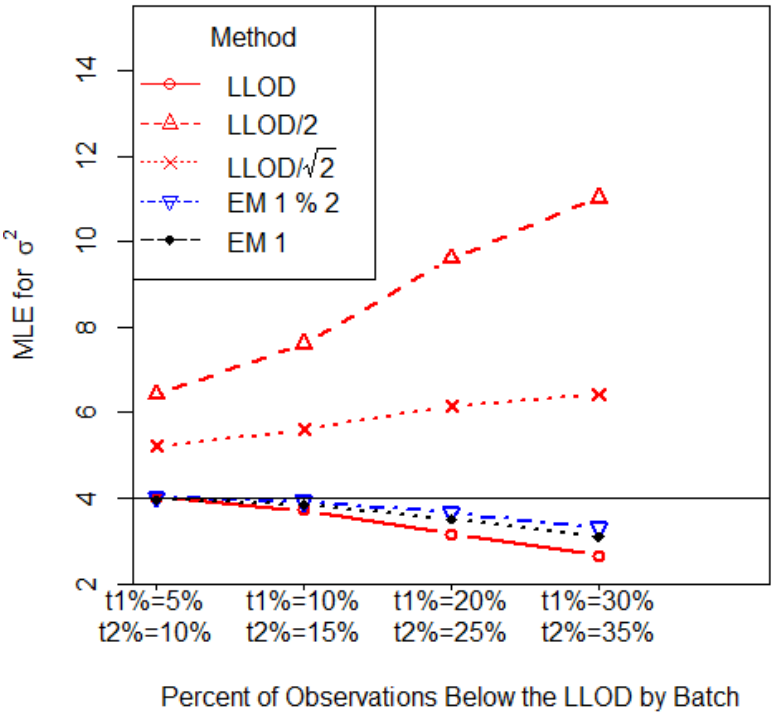

Figure 4 Plot of $\widehat{\sigma^2}$ from the second simulation scenario with $\mu = 10$, $\sigma^2 = 4$, and $s^2_{batch} = 1$. A horizontal line is at $\sigma^2 = 4$. In the legend, "EM 1" represents an EM algorithm with the first moment of the multivariate right truncated normal distribution. Our proposed EM algorithm that includes first and second moments of the multivariate right truncated normal distribution is labeled as "EM 1 & 2".

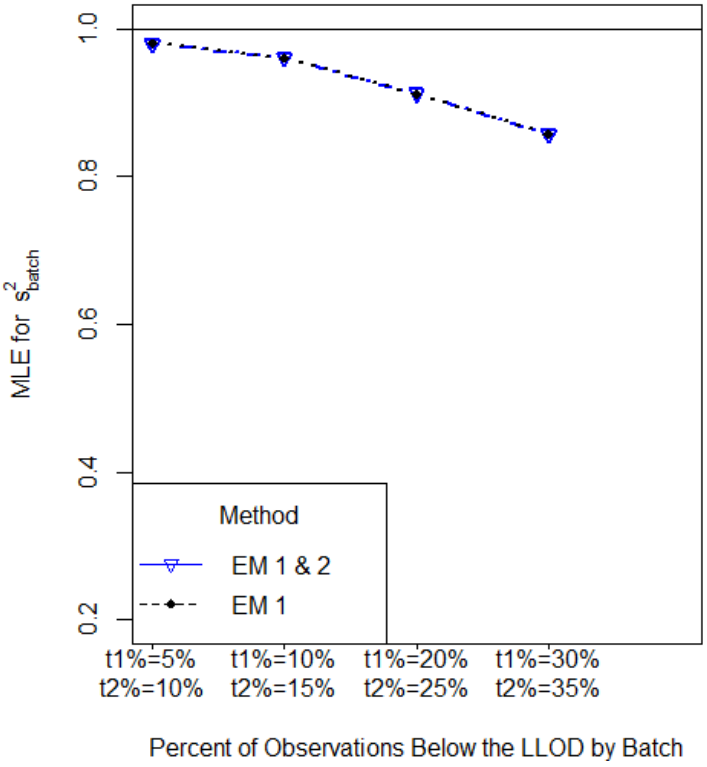

Figure 5 Plot of $\widehat{s^2_{batch}}$ from the first simulation scenario with $\mu = 5$, $\sigma^2 = 4$, and $s^2_{batch} = 1$. A horizontal line is at $s^2_{batch} = 1$. In the legend, "EM 1" represents an EM algorithm with the first moment of the multivariate right truncated normal distribution. Our proposed EM algorithm that includes first and second moments of the multivariate right truncated normal distribution is labeled as "EM 1 & 2".

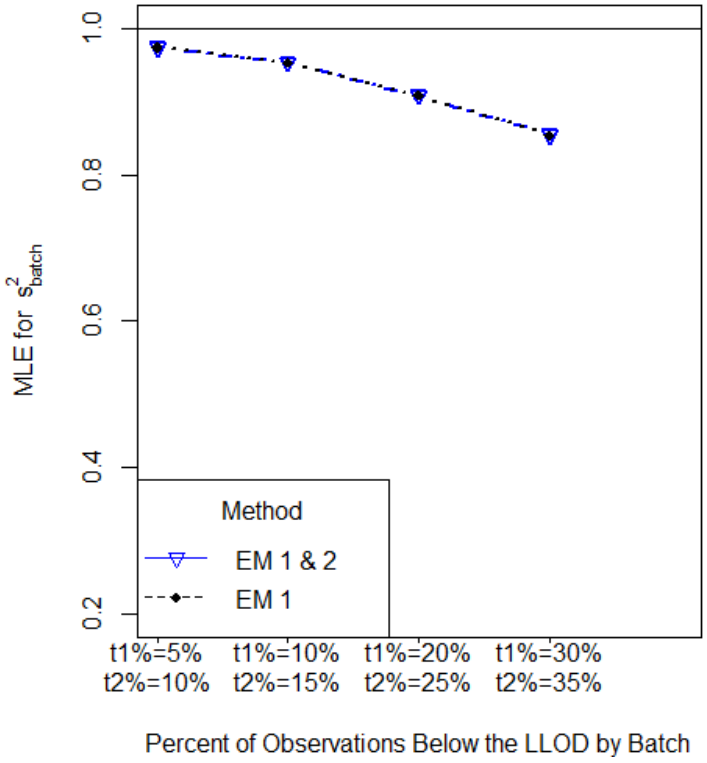

Figure 6 Plot of $\widehat{s^2_{batch}}$ from the second simulation scenario with $\mu = 10$, $\sigma^2 = 4$, and $s^2_{batch} = 1$. A horizontal line is at $s^2_{batch} = 1$. In the legend, "EM 1" represents an EM algorithm with the first moment of the multivariate right truncated normal distribution. Our proposed EM algorithm that includes first and second moments of the multivariate right truncated normal distribution is labeled as "EM 1 & 2".

- $\widehat{\mu}$ Results from Figures 1 and 2
    - For both simulation scenarios, $\widehat{\mu}$ from both EM algorithms perform equally well, even as the percentages of observations below the LLOD in each batch increases.
    - The difference in the performance of the replacement methods, especially LLOD/√2, for $\widehat{\mu}$ across the simulation scenarios illustrated how sensitive these methods are to changes in the mean relative to the variance and the LLOD value.
    - In Figure 2, $\widehat{\mu}$ from both EM algorithms perform better than all of the substitution methods.
- $\widehat{\sigma^2}$ Results from Figures 3 and 4
    - The unreliability of the replacement methods for estimating σ^2 is illustrated from the differences in the performance of $\widehat{\sigma^2}$ across the simulation scenarios.
    - Our proposed EM algorithm is the best method in the second simulation scenario (Figure 4).
- $\widehat{s^2_{batch}}$ Results from Figures 5 and 6
    - The substitution methods are excluded from the figures because these methods do not compute $\widehat{s^2_{batch}}$.
    - The results of $\widehat{s^2_{batch}}$ from each EM algorithm are comparable.

APPLICATION TO SLEIGH DATA

Systemic Lupus Erythematosus in Gullah Health (SLEIGH) is an observational cohort study of African American Gullah participants with systemic lupus erythematosus (SLE) and control

participants. Further details about the SLEIGH study has been published.[21,26,83] In the toxicology and serology component of the SLEIGH study, perfluorinated chemicals (PFCs) and polybrominated diphenyl ethers (PBDEs) contaminant levels were measured from serum samples of 86 participants with systemic lupus erythematosus and 139 control participants. The wet weight of the PFCs and PBDEs were measured on a continuous scale in nanogram/gram (ng/g) using an analytical laboratory tool. The serum samples were measured in multiple batches which resulted in the majority of the contaminants having more than one lower limit of detection (LLOD) value. Further details about the SLEIGH study, the quality assurance, and quality control of the samples collected are published.[21,26,83] Our proposed EM algorithm will be used to estimate the mean and variance of contaminants perfluorohexane sulfonate (PFHxS) and perfluorodecanoic acid (PFDA).

Contaminants PFHxS and PFDA are presented on a log (base 10) ng/g scale because the observed data for each contaminant was not normally distributed. Frequently ecological data including ecological pollutant data are not normally distributed, and the data undergoes a log transformation in order to be normally distributed.[84-87] For SLE participants, PFHxS has two batches with LLOD values at $t_1 =$ - 2.303 log ng/g (0.10 ng/g in the original scale) and $t_2 =$ - 1.966 log ng/g (0.14 ng/g). The sample size for batch 1 is 10 and 76 for batch 2. There are 10% and 28.9% of the observations appearing below the LLOD in batch 1 and 2, respectively, for PFHxS. The two LLOD values for control participants with PFDA measurements are $t_1 =$ – 1.833 log ng/g (0.16 ng/g) and $t_2 =$ - 3.507 log ng/g (0.14 ng/g). For batch 1 of PFDA the sample size is 70, and 2.9% of the observations appear below the LLOD. There are 7.4% of the

68 observations in batch 2 appearing below the LLOD for PFDA. Histograms of the contaminants by batch in the log ng/g scale are displayed in Figures 7 - 10.

Based on our proposed EM algorithm, the sample mean of PFHxS for those with SLE in the SLEIGH study is – 0.502 log ng/g with a sample variance of 1.714 log ng/g and sample batch variance of 0.205 log ng/g. For control participants in the SLEIGH study, the sample mean of PFDA is – 0.414 log ng/ g, sample variance is 1.248 log ng/g, and sample batch variance is 0.288 log ng/g.

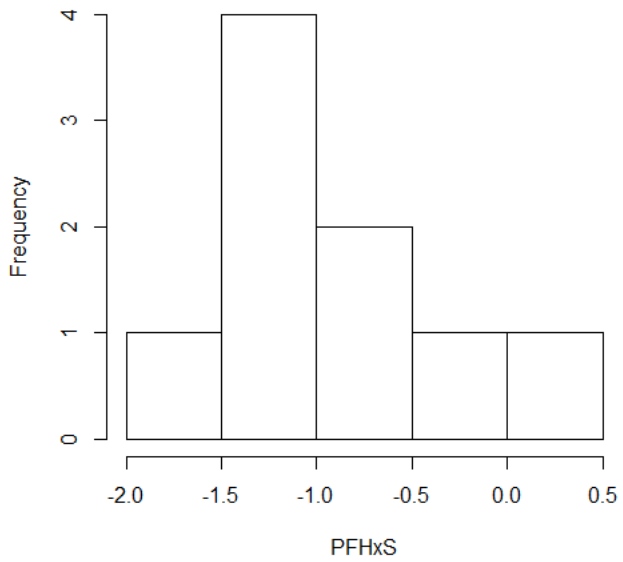

Figure 7 Histogram of log transformed PFHxS for batch 1.

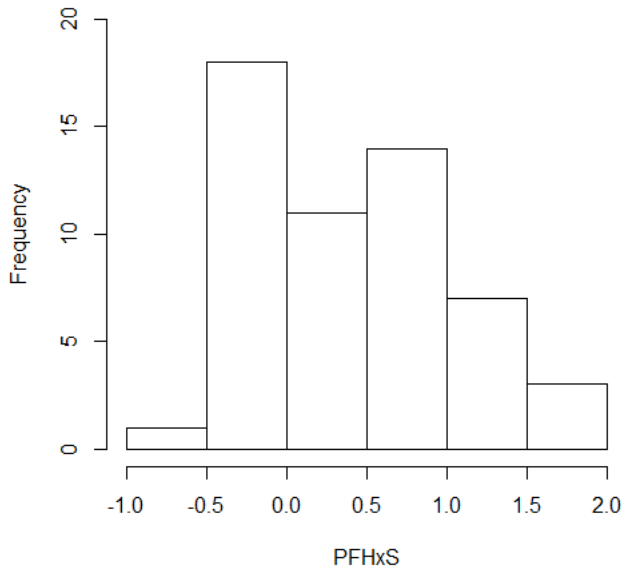

Figure 8 Histogram of log transformed PFHxS for batch 2.

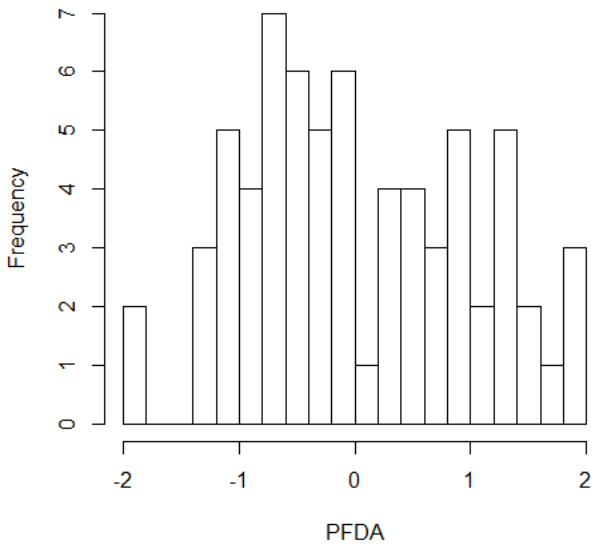

Figure 9 Histogram of log transformed PFDA for batch 1.

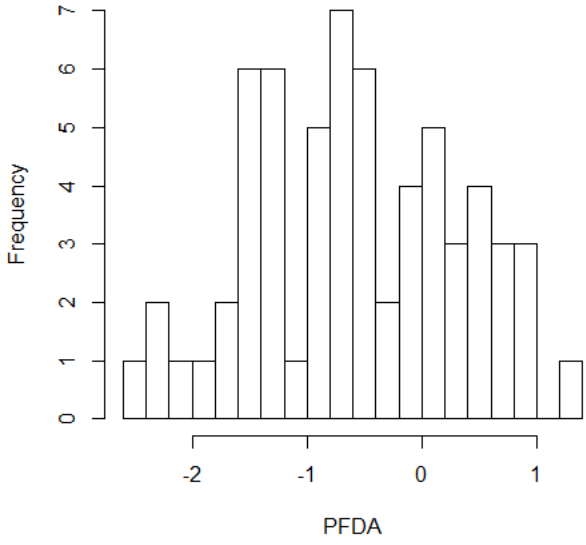

Figure 10 Histogram of log transformed PFDA for batch 2.

DISCUSSION

In this paper, we proposed a method for estimating MLEs of a normal distribution for an outcome variable that has multiple LLOD values in the observed sample. The random effect for the batches of data in our EM algorithm allowed us to incorporate the correlated observations within each batch. In the simulation study, we compared our method with substitution methods and an EM algorithm with only the first moment of a multivariate right truncated normal distribution. Substitution methods are recommended by the United States Environmental Protection Agency when 15% or less of the observations appear below the LLOD.[34] The substitution methods have been shown to produce biased estimates even when the percent occurring below the LLOD is small.[29,98] The results from our simulation study confirmed that MLEs from the substitution methods are biased and sensitive to the underlying mean and variance. Our method is more consistent and less sensitive to changes occurring in the underlying mean and variance relative to the substitution methods. We are able to conclude from the simulation study results that the variance estimate from our EM algorithm is less biased in comparison to the EM algorithm with the first moment of a right truncated multivariate normal distribution. In general the underlying mean and variance are unknown for real data applications, and we therefore recommend that our proposed method should be applied rather than the substitution methods and an EM algorithm with the first moment only.

5. DEVELOP AN APPROACH FOR ESTIMATING THE MEAN AND COVARIANCE MATRICES FOR MULTIVARIATE NORMAL RANDOM VARIABLES, WITH EACH MARGINAL DISTRIBUTION HAVING ONE LOWER LIMIT OF DETECTION ARISING FROM A SINGLE BATCH (I.E. MULTIPLE VARIABLES WITH ONE BATCH)

INTRODUCTION

Parameter estimation difficulties arise when sample data contains observations that appear below a lower limit of detection (LLOD) value. Directly estimating parameters such as the sample mean and variance from data with LLOD observations is analogous to conducting a complete case analysis.[27,28] The LLOD observations are not included in the calculation, and the estimated parameters are biased estimates of the underlying distribution.[27,28,99]

For simple cases with one variable, several methods have been proposed in the literature. These methods to include LLOD observations is to substitute the LLOD observations with the LLOD value, LLOD/2, or LLOD/$\sqrt{2}$ prior to computing the parameter estimates.[29-37] Substitution methods are recommended by the United States Environmental Protection Agency when 15% or less of the data appears below the LLOD.[34] The substitution estimation methods have been applied to multiple variables, but, similarly to parameter estimates from one variable, these estimation methods result in inaccurate inferences about the underlying distributional parameters.[29,35,100] In a previous article a better alternative for parameter estimation under a truncated model using an EM-algorithm was introduced for one variable with LLOD observations.[66] However, multiple correlated variables, each with an LLOD, has not been adequately addressed.

Data collected for two continuous variables with an LLOD in each variable could be treated as a truncated bivariate normal distribution. In this article the objective is to provide an EM-solution for estimation of parameters from a bivariate normal distribution. This would require expressions for moments of the truncated bivariate normal distribution. We will first discuss the various components required for the E-step and M-step of the algorithm. Then the proposed EM algorithm is compared with other existing methods in a simulation study and real data application.

METHODS

Let $\mathbf{Y_1}$ and $\mathbf{Y_2}$ represent two random vectors of size n. Y_{ij} denotes the i^{th} observation from the j^{th} random variable where i=1,…,n and j=1 or 2. The observations within each random vector can be categorized as either being above or below a variable specific LLOD value. The LLOD value is also known as a truncation value. The truncation value for the j^{th} random variable is denoted as t_j. Observations above t_j are fully measured to be an exact value while observations below t_j are simply recorded as "$< t_j$". The underlying distribution of $\mathbf{Y_1}$ and $\mathbf{Y_2}$ is a bivariate normal distribution with a two dimensional mean vector $\boldsymbol{\mu} = \begin{bmatrix} \mu_{y_1} & \mu_{y_2} \end{bmatrix}'$ and a 2×2 covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2_{y_1} & \rho\sigma_{y_1}\sigma_{y_2} \\ \rho\sigma_{y_1}\sigma_{y_2} & \sigma^2_{y_2} \end{bmatrix}$.

Data above t_j in $\mathbf{Y_1}$ and $\mathbf{Y_2}$ jointly follow a left truncated bivariate normal distribution with mean vector $\boldsymbol{\mu}$, covariance matrix $\boldsymbol{\Sigma}$, and a 2 dimensional vector of truncation values $\mathbf{t} = \begin{bmatrix} t_1 & t_2 \end{bmatrix}'$, where support (t_1, ∞) x (t_2, ∞). Also, the data below t_j in $\mathbf{Y_1}$ and $\mathbf{Y_2}$ jointly

follow a right truncated bivariate normal distribution ($\boldsymbol{\mu}, \boldsymbol{\Sigma}; \mathbf{t}$) with the support (-∞, t_1) x (-∞, t_2). (Details of the probability distribution function (pdf) and moments of the distributions are listed in the Appendix.)

EM Algorithm

The proposed EM algorithm for estimating parameters of the bivariate normal distribution will be developed under a linear model (LM) framework, in anticipation of extensions to mixed models. For this, first $\mathbf{Y_1}$ implement to the LM, $\mathbf{Y_1}$ and $\mathbf{Y_2}$ are stacked into a single vector $\mathbf{Y}$, which is $2n \times 1$ dimensional. The elements in $\mathbf{Y}$ are sorted by the observation and variable number. That is $\mathbf{Y} = \{ Y_{11}, Y_{12}, Y_{21}, Y_{22}, ..., Y_{n1}, Y_{n2} \}$. The matrix form of the LM for our proposed EM algorithm is then denoted as $\mathbf{Y} = \mathbf{X\beta} + \boldsymbol{\varepsilon}$, where $\mathbf{X}$ is a $2n \times 2$ design matrix, in this case with each column is a vector of 1's and 0's representing the corresponding variable, $\boldsymbol{\beta}$ is a 2×1 vector, $\boldsymbol{\varepsilon}$~N $\mathbf{0}, \mathbf{V}$ and $\mathbf{V}$ is a block diagonal matrix such that $\mathbf{V} = \mathbf{I}_{2n} \otimes \boldsymbol{\Sigma}$ where $\otimes$ is a Kronecker product. The likelihood of $\mathbf{X\beta}$ and $\mathbf{V}$ given $\mathbf{Y}$ is simplified as,

$$\begin{aligned}
\mathbf{L}\left(\mathbf{X\beta}, \mathbf{V} \mid \mathbf{Y}\right) &= \frac{1}{\left(2\pi\right)^{n/2} \left|\mathbf{V}\right|^{1/2}} \exp\left\{ -\frac{1}{2} \left(\mathbf{Y} - \mathbf{X\beta}\right)' \mathbf{V}^{-1} \left(\mathbf{Y} - \mathbf{X\beta}\right) \right\} \\
&= \frac{1}{\left(2\pi\right)^{n/2} \left|\mathbf{V}\right|^{1/2}} \exp\left\{ -\frac{1}{2} tr\left\{ \mathbf{V}^{-1} \left(\mathbf{Y} - \mathbf{X\beta}\right) \left(\mathbf{Y} - \mathbf{X\beta}\right)' \right\} \right\} \\
&= \frac{1}{\left(2\pi\right)^{n/2} \left|\mathbf{V}\right|^{1/2}} \exp\left\{ -\frac{1}{2} tr\left\{ \mathbf{V}^{-1} \left(\mathbf{YY}' - 2\mathbf{Y} \left(\mathbf{X\beta}\right)' + \left(\mathbf{X\beta}\right)' \mathbf{X\beta} \right) \right\} \right\} \quad .
\end{aligned}$$

$\mathbf{YY}'$ is an element-wise multiplication matrix with dimension $2n \times 2n$. MLEs of the mean $\widehat{\mathbf{X\beta}}$ and covariance matrix $\mathbf{V}$ are found by maximizing the log-likelihood.

E-Step

Any observation below t_j is treated as a missing observation. The E-Step of our proposed EM algorithm replaces observations below t_j in $\mathbf{Y}$ and $\mathbf{YY}'$ with either the marginal, conditional, or joint moments from a truncated bivariate normal distribution. The moments that are replacing the observations below t_j in $\mathbf{Y}$ and $\mathbf{YY}'$ depend on the paired observations within $\mathbf{Y_1}$ and $\mathbf{Y_2}$. Paired observations in $\mathbf{Y_1}$ and $\mathbf{Y_2}$ can occur in four scenarios as shown in table 1. In this table, observations below t_j for j = 1, 2, are considered unobserved. The corresponding moments needed in the E-Step are also listed in Table 1.

Table 1 Scenarios of paired observations in $\mathbf{Y}_1$ and $\mathbf{Y_2}$ where $i = i'$ and $j \neq j'$. The moments required for each scenario are based on if the corresponding paired observation is below or above the truncation value.

Scenarios ($i = i'$ and $j \neq j'$)	Moments Required for Observations Appearing Below t_j or $t_{j'}$ in $\mathbf{Y}$
1) $Y_{ij} > t_j$ and $Y_{i'j'} > t_{j'}$	N/A
2) $Y_{ij} \leq t_j$ and $Y_{i'j'} \leq t_{j'}$	$E\left(Y_{ij}\right) = \mu_{y_{jRT}}$ and $E\left(Y_{i'j'}\right) = \mu_{y_{j'RT}}$
3) $Y_{ij} \leq t_j$ and $Y_{i'j'} > t_{j'}$	$E\left(Y_{ij}\right) = \mu_{y_{jRT}\mid y_{j'LT}}$
4) $Y_{ij} > t_j$ and $Y_{i'j'} \leq t_{j'}$	$E\left(Y_{i'j'}\right) = \mu_{y_{j'RT}\mid y_{jLT}}$

In scenario 2 from Table 1, Y_{ij} and $Y_{i'j'}$ are both below the truncation value. The support of Y_{ij} and $Y_{i'j'}$ are $Y_{ij} \in (-\infty, t_j)$ and $Y_{i'j'} \in (-\infty, t_{j'})$, respectively. The joint distribution of Y_{ij} and $Y_{i'j'}$ is a right truncated bivariate normal distribution with mean vector $\boldsymbol{\mu} = \begin{bmatrix} \mu_{y_j} & \mu_{y_{j'}} \end{bmatrix}'$, a covariance

matrix $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2_{y_j} & \rho\sigma_{y_j}\sigma_{y_{j'}} \\ \rho\sigma_{y_j}\sigma_{y_{j'}} & \sigma^2_{y_{j'}} \end{bmatrix}$, and truncation vector $\mathbf{t} = \begin{bmatrix} t_j & t_{j'} \end{bmatrix}'$ but with different supports as mentioned above. The pdf, moments, and R package to compute the moments of a multivariate truncated normal distribution with left and right truncation occurring simultaneously were provided by Manjunath and Wilhelm.[79,93,94] By allowing the left truncation value to be $-\infty$ and the dimension be equal to 2 in Manjunath and Wilhelm's formulas, the pdf and moments of a right truncated bivariate normal distribution are constructed. Marginally, the mean of Y_{ij} in scenario 2 is

$$\mu_{y_{jRT}} = \mu_{y_j} + \sigma_{y_j} \frac{\phi\left(\frac{t_j - \mu_{y_j}}{\sigma_{y_j}}\right)}{1 - \Phi\left(\frac{t_{y_j} - \mu_{y_j}}{\sigma_{y_j}}\right)}, \tag{1}$$

and the variance is

$$\sigma^2_{y_{jRT}} = \sigma^2_{y_j}\left[1 + \left(\frac{t_j - \mu_{y_j}}{\sigma_{y_j}}\right)\frac{\phi\left(\frac{t_j - \mu_{y_j}}{\sigma_{y_j}}\right)}{1 - \Phi\left(\frac{t_j - \mu_{y_j}}{\sigma_{y_j}}\right)} - \left(\frac{\phi\left(\frac{t_j - \mu_{y_j}}{\sigma_{y_j}}\right)}{1 - \Phi\left(\frac{t_j - \mu_{y_j}}{\sigma_{y_j}}\right)}\right)^2\right]. \tag{2}$$

In formulas (1) and (2), ϕ and Φ denote the pdf and cumulative distribution function of a normal distribution, respectively. Also the marginal mean and variance of $Y_{i'j'}$ are

$$\mu_{y_{j'RT}} = \mu_{y_{j'}} + \sigma_{y_{j'}} \frac{\phi\left(\frac{t_{j'} - \mu_{y_{j'}}}{\sigma_{y_{j'}}}\right)}{1 - \Phi\left(\frac{t_{j'} - \mu_{y_{j'}}}{\sigma_{y_{j'}}}\right)} \tag{3}$$

and $\sigma^2_{y_{j',RT}} = \sigma^2_{y_{j'}} \left[1 + \left(\frac{t_{j'} - \mu_{y_{j'}}}{\sigma_{y_{j'}}} \right) \frac{\phi\left(\frac{t_{j'} - \mu_{y_{j'}}}{\sigma_{y_{j'}}} \right)}{1 - \Phi\left(\frac{t_{j'} - \mu_{y_{j'}}}{\sigma_{y_{j'}}} \right)} - \left(\frac{\phi\left(\frac{t_{j'} - \mu_{y_{j'}}}{\sigma_{y_{j'}}} \right)}{1 - \Phi\left(\frac{t_{j'} - \mu_{y_{j'}}}{\sigma_{y_{j'}}} \right)} \right)^2 \right]$. (4)

Observations belonging to scenarios 3 and 4 are replaced with conditional moments of a truncated bivariate normal distribution. In scenario 3, the support of Y_{ij} and $Y_{i'j'}$ are $Y_{ij} \in (-\infty, t_j)$ and $Y_{i'j'} \in (t_{j'}, \infty)$. Jointly Y_{ij} and $Y_{i'j'}$ follow a truncated bivariate normal distribution $(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \mathbf{t})$ with truncation occurring on the right side for Y_{ij} and the left side for $Y_{i'j'}$. The pdf and moments of the truncated bivariate normal distribution with Y_{ij} right truncated and $Y_{i'j'}$ left truncated are in the Appendix. The conditional distribution of this distribution is a truncated univariate normal distribution.[101] Specifically, the conditional mean and variance of Y_{ij} given $Y_{i'j'}$ in scenario 3 are

$$\mu_{y_{jRT}|y_{j',LT}} = \mu_{y_{jRT}} + \frac{\sigma_{y_{jRT}}}{\sigma_{y_{j',LT}}} \rho_T \left(y_{i'j'} - \mu_{y_{j',LT}} \right) \quad (5)$$

and $\sigma^2_{y_{jRT}|y_{j',LT}} = \left(1 - \rho_T^2\right) \sigma^2_{y_{jRT}}$ (6)

where ρ_T is the correlation between $\mathbf{Y_1}$ and $\mathbf{Y_2}$. The formulas for $\mu_{y_{jRT}}$ and $\sigma^2_{y_{jRT}}$ are in equations (1)and (2). Likewise the distribution of Y_{ij} and $Y_{i'j'}$ in scenario 4 is a truncated bivariate normal distribution, but Y_{ij} is truncated on the left side while $Y_{i'j'}$ is right truncated. The conditional mean and variance in scenario 4 includes equations (3) and (4). In scenario 4, the conditional mean and variance are

$$\mu_{y_{j',RT}|y_{jLT}} = \mu_{y_{j',RT}} + \frac{\sigma_{y_{j',RT}}}{\sigma_{y_{jLT}}} \rho_T \left(y_{ij} - \mu_{y_{jLT}} \right) \quad (7)$$

and $\sigma^2_{y_{j'RT}|y_{jLT}} = \left(1-\rho_T^2\right)\sigma^2_{y_{j'RT}}$. (8)

The formulas for $\mu_{y_{j'LT}}$, $\sigma^2_{y_{j'LT}}$, $\mu_{y_{jLT}}$, and $\sigma^2_{y_{jLT}}$ are included in the Appendix.

The corresponding moments for scenarios in $\mathbf{YY}'$ are listed in Table 2. The four scenarios in $\mathbf{Y}$ from Table 1 result in 24 scenarios in $\mathbf{YY}'$. In the Appendix the formula by Manjunath and Wilhelm for $\mu_{y_{jRT}y_{j'RT}}$, which is the cross product moment of Y_{ij} and $Y_{i'j'}$ where $i = i'$ and $j \neq j'$, is listed. [79] This moment is from a right truncated bivariate normal distribution, and it appears in Table 2. In our proposed EM algorithm, equations 1 – 4 and $\mu_{y_{jRT}y_{j'RT}}$ are computed by the R package *tmvtnorm*.[93,94]

M-Step

Based on the LM, the log-likelihood of the parameters $\mathbf{X\beta}$ and $\mathbf{V}$ given $\mathbf{Y}$ is,

$$\ln\left(\mathbf{L}\left(\mathbf{X\beta}, \mathbf{V} \mid \mathbf{Y}\right)\right) = -\frac{n}{2}\ln\left(2\pi\right) - \frac{1}{2}\ln\left|\mathbf{V}\right| - \frac{1}{2}\,\text{trace}\left\{\mathbf{V}^{-1}\left(\mathbf{YY}' - 2\mathbf{Y}\left(\mathbf{X\beta}\right)' + \left(\mathbf{X\beta}\right)'\,\mathbf{X\beta}\right)\right\}.$$

The proposed EM algorithm finds $\widehat{\mathbf{X\beta}}$ by maximizing the log-likelihood. For estimating $\mathbf{V}$ from an LM, our EM algorithm maximizes the log-likelihood of $\mathbf{V}$ based on the conditional derivation of the restricted maximum likelihood approach.[97] The M-step of our propose method is implemented through the *mle2* function within the R package *bbmle*.[102] The log-likelihood of $\mathbf{V}$ that is maximized to compute the MLEs in the proposed EM algorithm is,

$$\ln\left(\mathbf{L}\left(\mathbf{V} \mid \mathbf{Y}\right)\right) = -\frac{1}{2}\left[\ln\left(\det\left(\mathbf{V}\right)\right) + \text{trace}\left\{\ln\left(\det\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right) + \left(\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\,\mathbf{X}'\mathbf{V}^{-1}\right)\mathbf{YY}'\right)\right\}\right].$$

Table 2 This table includes moments for all possible scenarios in the $\mathbf{YY}'$ matrix. Observation number is denoted as *i* and *j* is the variable number. Moments that replace elements in the $\mathbf{YY}'$ matrix depend on the corresponding paired observation that could be either below or above the truncation value. If the corresponding observation is above the truncation value, a conditional moments replaces the element in $\mathbf{YY}'$. Marginal moments replace $\mathbf{YY}'$ elements if the corresponding observation is below the truncation value.

	$Y_{ij} > t_j \,\&\, Y_{i'j'} > t_{j'}$	$Y_{ij} \le t_j \;\&\; Y_{i'j'} \le t_{j'}$	$Y_{ij} > t_j \;\&\; Y_{i'j'} \le t_{j'}$	$Y_{ij} \le t_j \;\&\; Y_{i'j'} > t_{j'}$
$i = i' \,\&\, j = j'$	Y_{ij}^2	If $Y_{ij} \le t_{j'}$, $E\left(Y_{ij}^2 \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \sigma^2_{y_{jRT}} + \left[\mu_{y_{jRT}}\right]^2$ If $Y_{ij} > t_{j'}$, $E\left(Y_{ij}^2 \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \sigma^2_{y_{jRT}\mid y_{jLT}} + \left[\mu_{y_{jRT}\mid y_{jLT}}\right]^2$	N/A	
$i = i' \,\&\, j \neq j'$	$Y_{ij}Y_{ij'}$	$E\left(Y_{ij}Y_{ij'} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \mu_{y_{jRT}y_{j'RT}}$	$E\left(Y_{ij}Y_{ij'} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = Y_{ij} \times \mu_{y_{j'RT}\mid y_{j'LT}}$	$E\left(Y_{ij}Y_{ij'} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \mu_{y_{jRT}\mid y_{jLT}} \times Y_{ij'}$
$i \neq i' \,\&\, j = j'$	$Y_{ij}Y_{i'j}$	If $Y_{i'j'} > t_{j'} \;\&\; Y_{ij} > t_{j'}$, $E\left(Y_{ij}Y_{i'j} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \mu_{y_{jRT}\mid y_{jLT}} \times \mu_{y_{jRT}\mid y_{jLT}}$	If $Y_{i'j'} > t_{j'}$, $E\left(Y_{ij}Y_{i'j} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = Y_{ij} \times \mu_{y_{jRT}\mid y_{jLT}}$	If $Y_{ij} > t_{j'}$, $E\left(Y_{ij}Y_{i'j} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \mu_{y_{jRT}\mid y_{jLT}} \times Y_{i'j}$
		If $Y_{i'j'} \le t_{j'} \;\&\; Y_{ij} \le t_{j'}$, $E\left(Y_{ij}Y_{i'j} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \sigma^2_{y_{jRT}} + \left[\mu_{y_{jRT}}\right]^2$	If $Y_{i'j'} \le t_{j'}$, $E\left(Y_{ij}Y_{i'j} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = Y_{ij} \times \mu_{y_{jRT}}$	If $Y_{ij} \le t_{j'}$, $E\left(Y_{ij}Y_{i'j} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \mu_{y_{jRT}} \times Y_{i'j}$
$i \neq i' \,\&\, j \neq j'$	$Y_{ij}Y_{i'j'}$	If $Y_{ij} > t_{j'} \;\&\; Y_{i'j} > t_j$, $E\left(Y_{ij}Y_{i'j'} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \mu_{y_{jRT}\mid y_{jLT}} \times \mu_{y_{j'RT}\mid y_{j'LT}}$ If $Y_{ij} \le t_{j'} \;\&\; Y_{i'j} > t_j$, $E\left(Y_{ij}Y_{i'j'} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \mu_{y_{jRT}} \times \mu_{y_{j'RT}\mid y_{j'LT}}$ If $Y_{ij} > t_{j'} \;\&\; Y_{i'j} \le t_j$, $E\left(Y_{ij}Y_{i'j'} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \mu_{y_{jRT}\mid y_{jLT}} \times \mu_{y_{j'RT}}$ If $Y_{ij} \le t_{j'} \;\&\; Y_{i'j} \le t_j$, $E\left(Y_{ij}Y_{i'j'} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \mu_{y_{jRT}} \times \mu_{y_{j'RT}}$	If $Y_{i'j} > t_j$, $E\left(Y_{ij}Y_{i'j'} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = Y_{ij} \times \mu_{y_{j'RT}\mid y_{j'LT}}$ If $Y_{i'j} \le t_j$, $E\left(Y_{ij}Y_{i'j'} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = Y_{ij} \times \mu_{y_{j'RT}}$	If $Y_{ij} > t_{j'}$, $E\left(Y_{ij}Y_{i'j'} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \mu_{y_{jRT}\mid y_{jLT}} \times Y_{i'j'}$ If $Y_{ij} \le t_{j'}$, $E\left(Y_{ij}Y_{i'j'} \mid Y_{ij}; i = 1, \ldots, n \;\&\; j = 1, 2\right) = \mu_{y_{jRT}} \times Y_{i'j'}$

SIMULATION STUDY

The following simulation study is conducted to determine the performance of our proposed method in comparison to simpler methods. The simpler methods are substitution methods (LLOD value, LLOD/2, and LLOD/$\sqrt{2}$) and an EM algorithm that only includes the first marginal moments (equations (1) and (3)) of a bivariate right truncated distribution in the E-Step. The M-step of the EM algorithm with the first marginal moments is implemented by the *lm* function in R.[81] The *lm* function requires only the $\mathbf{Y}$ and $\mathbf{X}$ matrices prior to computing the MLEs from a linear model. In our method, we must specify the negative log-likelihood function of $\mathbf{X\beta}$ and $\mathbf{V}$ given $\mathbf{Y}$ and the log-likelihood through the *mle2* R function.[102] By including an EM algorithm with only the first marginal moment in our simulation study, we will be able to determine whether the second and conditional moments included in our method are necessary for computing MLEs of parameters from a bivariate normal distribution.

In our simulation study we generated data under different correlations $(\rho = 0.20; 0.50; 0.80)$. For each correlation level, two data vectors of size 100 are generated from a bivariate normal distribution with $\boldsymbol{\mu} = \begin{bmatrix} 2 & 3 \end{bmatrix}'$ and $\boldsymbol{\Sigma} = \begin{bmatrix} 8 & \rho\sqrt{8}\sqrt{14} \\ \rho\sqrt{8}\sqrt{14} & 14 \end{bmatrix}$. Percentages of the observations in each data vector are removed to replicate observation appearing below an LLOD value. For each variable, the percentage of observations below the LLOD is denoted as t_j %. The percentages considered in our simulation study are t_1 % = 5%, t_2 % = 5%; t_1 % = 5%, t_2 % = 10%; t_1 % = 10%, t_2 % = 15%; and t_1 % = 20%, t_2 % = 25%. The simulation size is 1000, and the results reported are an average of the 1000 simulations. R version 3.3.2 was used to conduct the simulation study.[81]

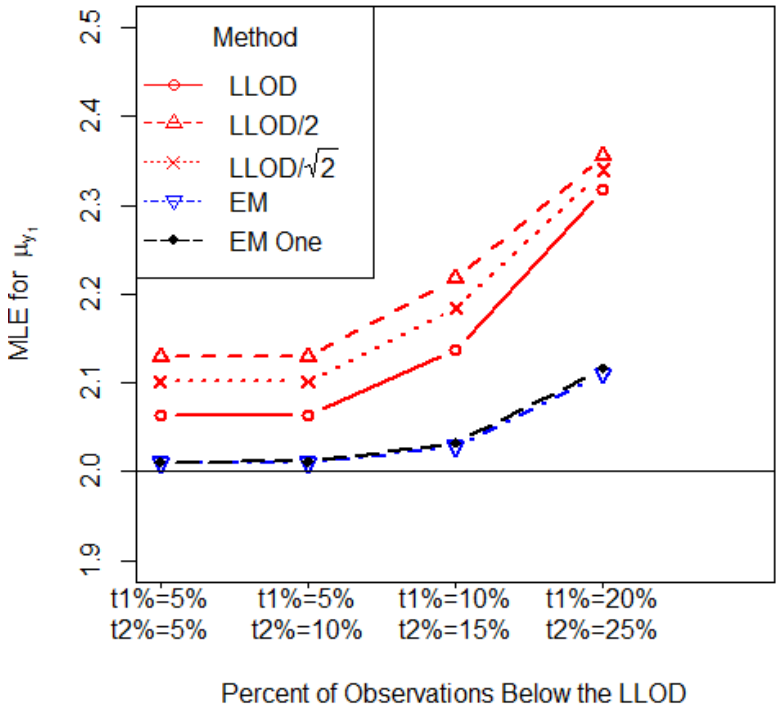

Figure 1 Simulation results for MLE of μ_{y_1} under data generated with $\rho = 0.20$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

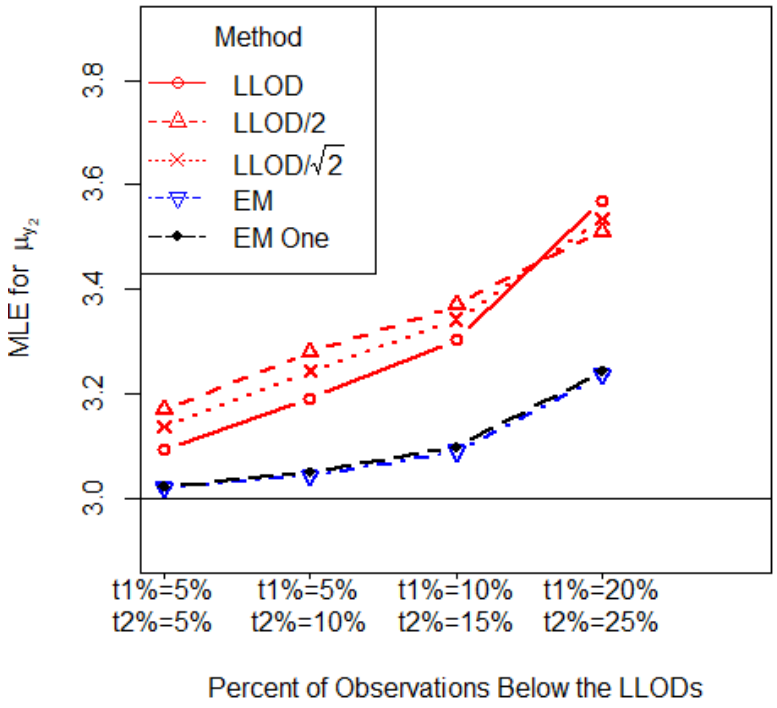

Figure 2 Simulation results for MLE of μ_{y_2} under data generated with $\rho = 0.20$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

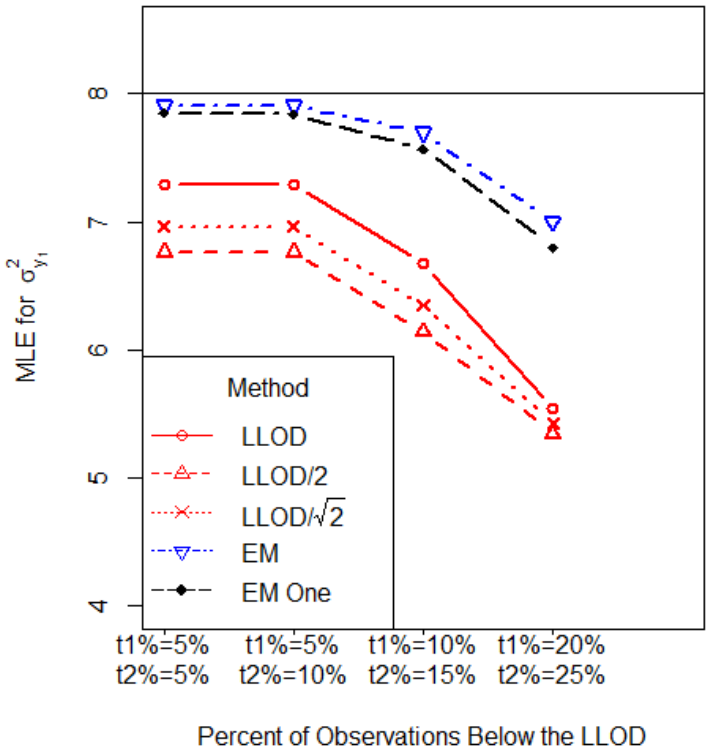

Figure 3 Simulation results for MLE of $\sigma^2_{y_1}$ under data generated with $\rho = 0.20$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

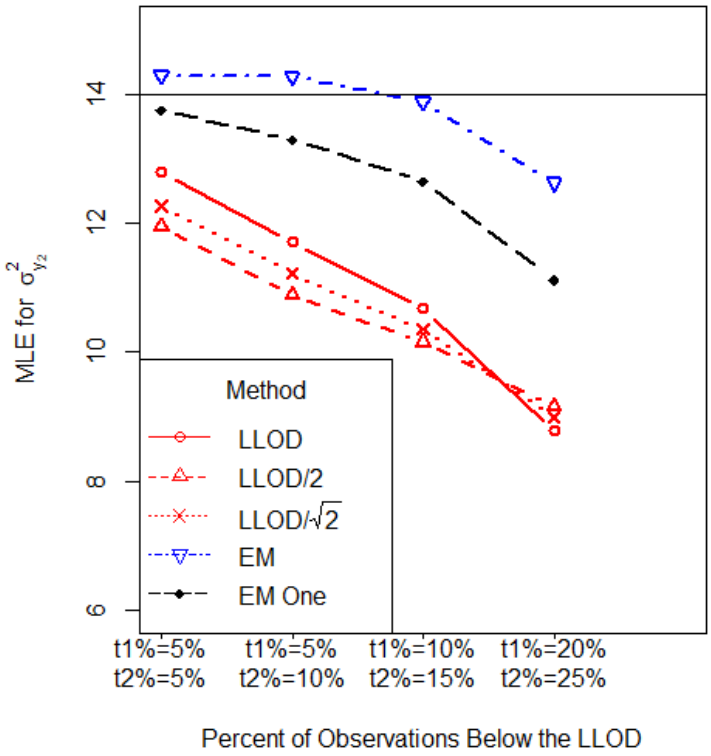

Figure 4 Simulation results for MLE of $\sigma^2_{y_2}$ under data generated with $\rho = 0.20$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

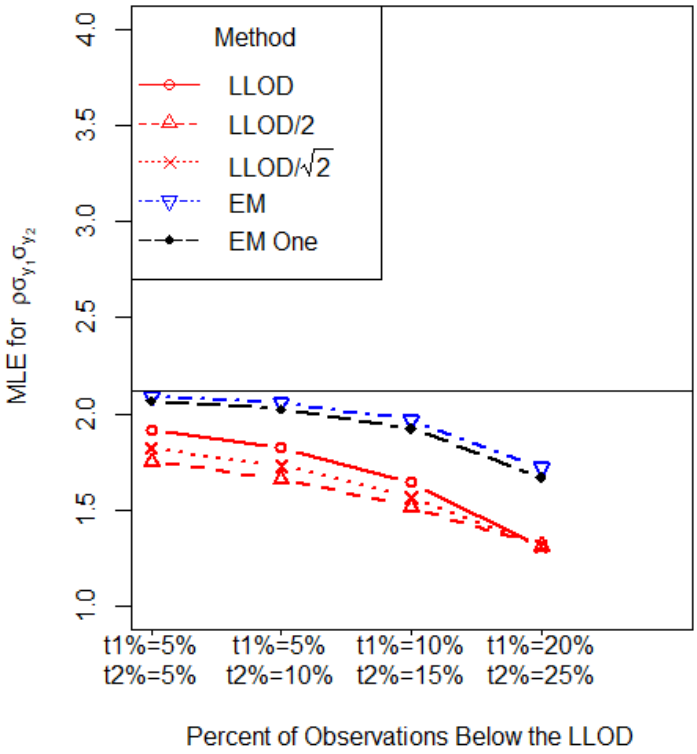

Figure 5 Simulation results for MLE of $\rho\sigma_{y_1}^2\sigma_{y_2}^2$ under data generated with $\rho = 0.20$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

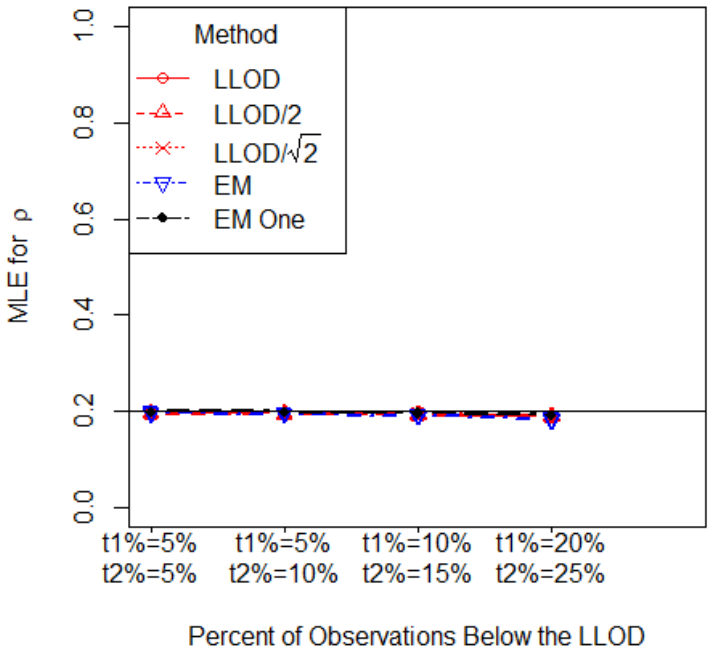

Figure 6 Simulation results for MLE of ρ under data generated with $\rho = 0.20$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

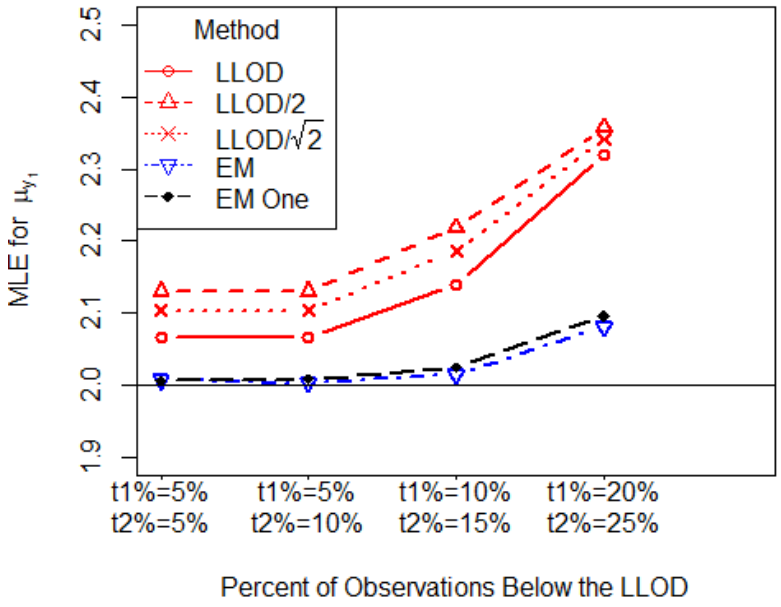

Figure 7 Simulation results for MLE of μ_{y_1} under data generated with $\rho = 0.50$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

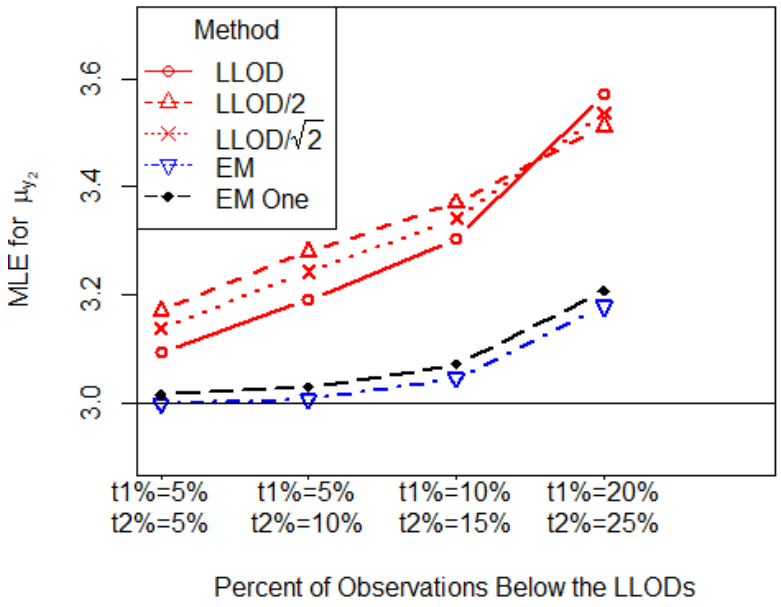

Figure 8 Simulation results for MLE of μ_{y_2} under data generated with $\rho = 0.50$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

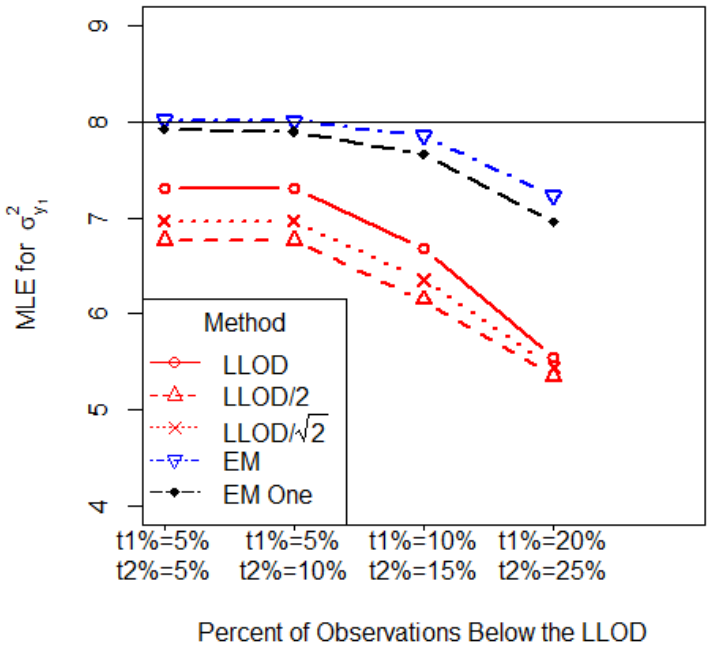

Figure 9 Simulation results for MLE of $\sigma^2_{y_1}$ under data generated with $\rho = 0.50$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

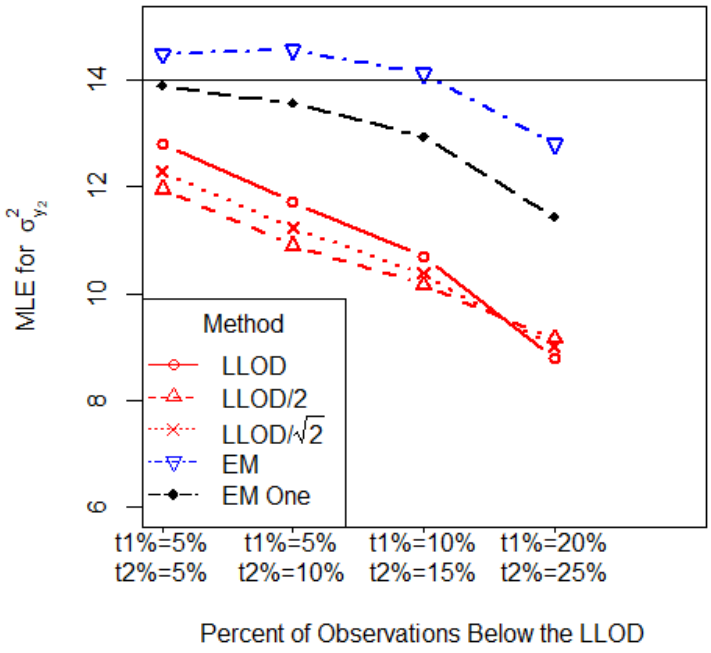

Figure 10 Simulation results for MLE of $\sigma^2_{y_2}$ under data generated with $\rho = 0.50$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

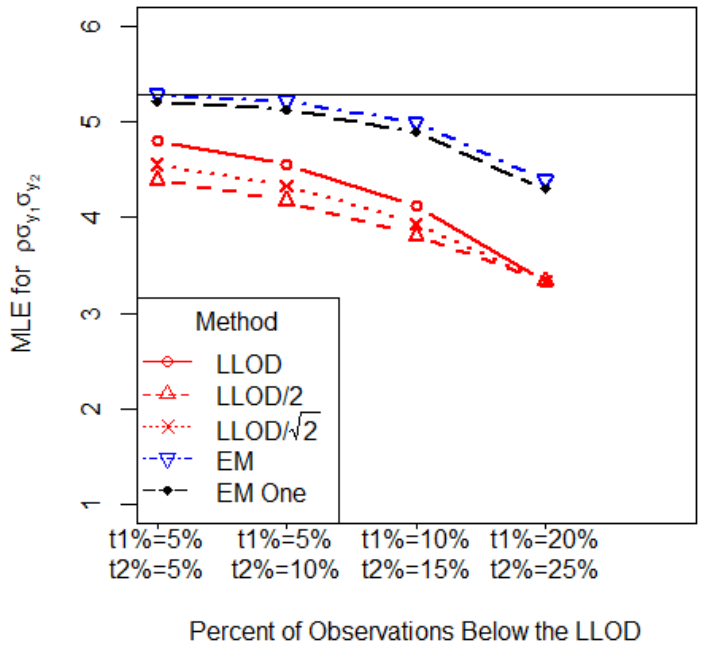

Figure 11 Simulation results for MLE of $\rho\sigma_{y_1}^2\sigma_{y_2}^2$ under data generated with $\rho = 0.50$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

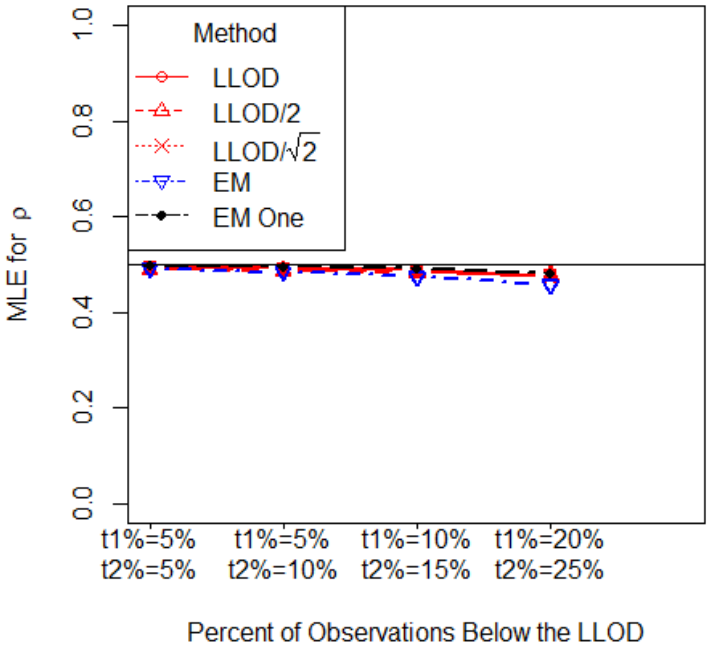

Figure 12 Simulation results for MLE of ρ under data generated with $\rho = 0.50$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

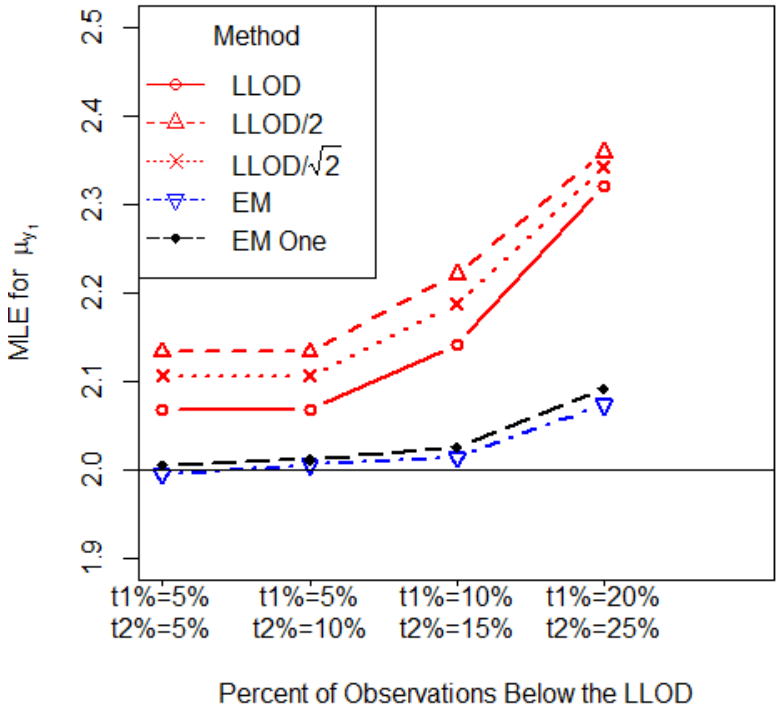

Figure 13 Simulation results for MLE of μ_{y_1} under data generated with $\rho = 0.80$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

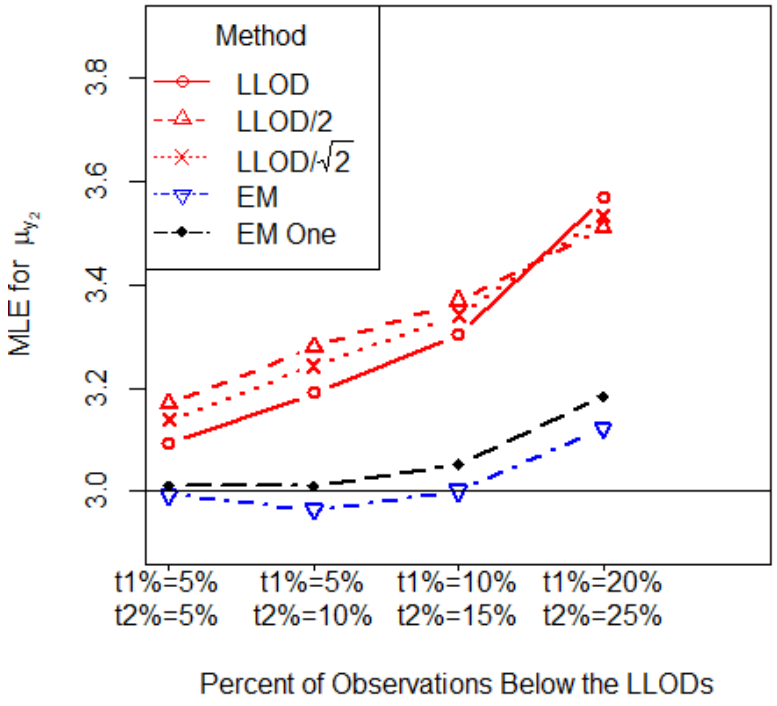

Figure 14 Simulation results for MLE of μ_{y_2} under data generated with $\rho = 0.80$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

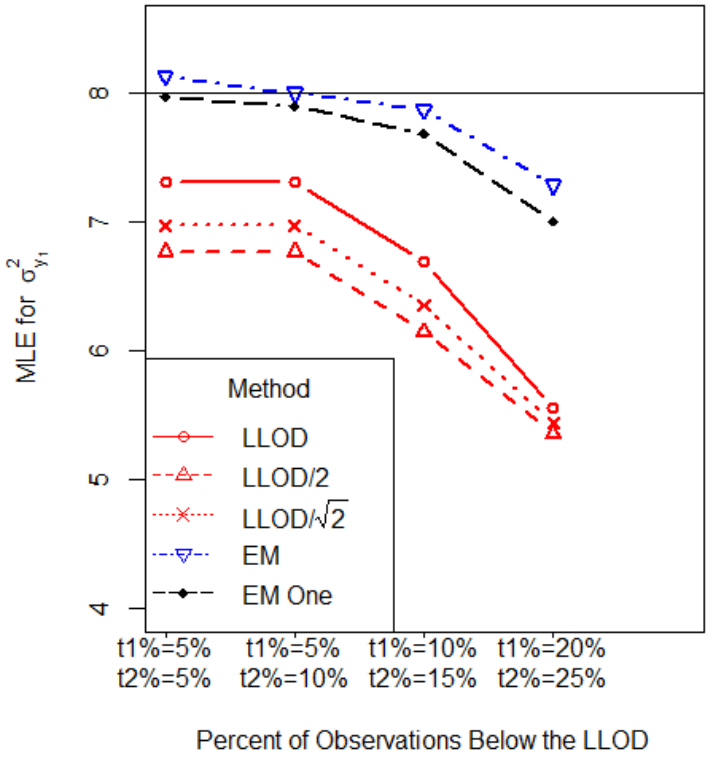

Figure 15 Simulation results for MLE of $\sigma^2_{y_1}$ under data generated with $\rho = 0.80$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

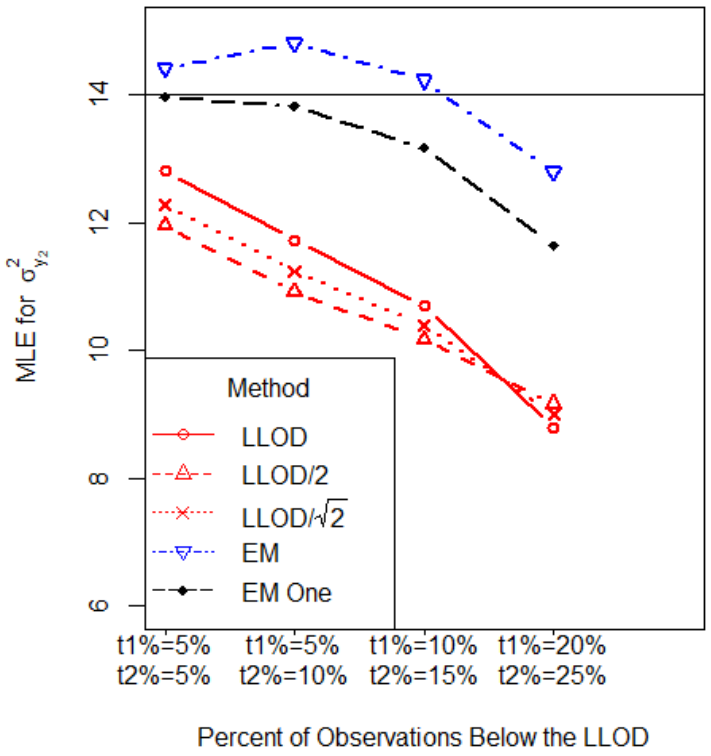

Figure 16 Simulation results for MLE of $\sigma^2_{y_2}$ under data generated with $\rho = 0.80$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

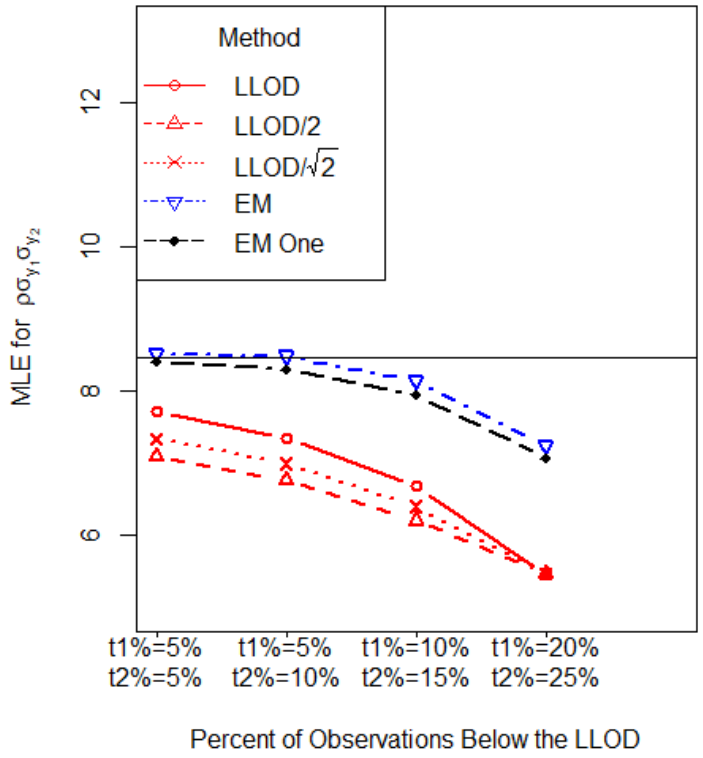

Figure 17 Simulation results for MLE of $\rho\sigma_{y_1}^2\sigma_{y_2}^2$ under data generated with $\rho = 0.80$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

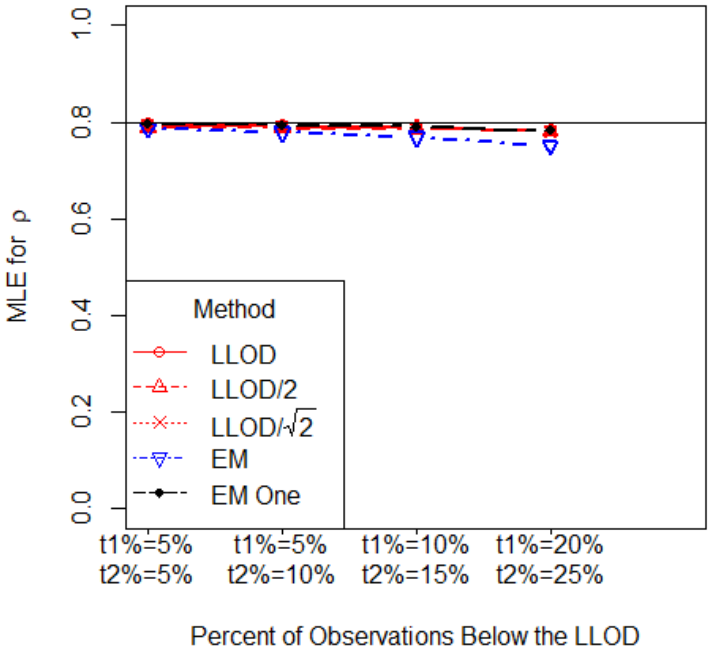

Figure 18 Simulation results for MLE of ρ under data generated with $\rho = 0.80$. The horizontal line represents the true value for the parameter. Our EM algorithm is labeled as 'EM' while the EM algorithm with only the first marginal moments is labeled as 'EM One'.

The results from the simulation study are displayed in Figures 1 – 18. In each figure, the horizontal line represents the values of the parameters that the bivariate normal distribution data was generated from. There are six figures for each parameter from a bivariate normal distribution under every correlation level.

$\widehat{\boldsymbol{\mu}}$ Results:

- The results for $\widehat{\boldsymbol{\mu}}$ are in Figures 1, 2, 7, 8, 13, and 14.
- Regardless of the correlation value, the substitution methods consistently overestimated the $\widehat{\boldsymbol{\mu}}$ for every t_j %.
    - The LLOD/2 substitution method performed the poorest for $\widehat{\boldsymbol{\mu}}$.
- The estimates for $\boldsymbol{\mu}$ are similar for both EM algorithms when the correlation is small (Figures 1 and 2), but our method performed better than the EM algorithm with the first marginal moment for the simulation scenarios with higher correlation values (Figures 7, 8, and 13).
- In Figure 14, the EM algorithm only with the first marginal moment only performed better than our method for one truncation scenario (t_1 % = 5% and t_2 % = 10%).

$\widehat{\boldsymbol{\Sigma}}$ Results:

- The results for $\widehat{\boldsymbol{\Sigma}}$ are in Figures 3 – 6, 9 – 12, and 15 – 18.
- The substitution methods underestimated $\widehat{\boldsymbol{\Sigma}}$ except when estimating ρ in Figures 6, 12, and 18.
- All of the estimation methods produced similar estimates for ρ in Figures 6, 12, and 18.

- For $\widehat{\sigma^2_{y_1}}$ and $\widehat{\sigma^2_{y_2}}$, our proposed method is the best method with the exception of a few cases.
    - In Figures 10 and 14, the EM algorithm with the first marginal moment slightly performed better than our method for $\widehat{\sigma^2_{y_2}}$ at the first two truncation scenarios (t_1% = 5% and t_2% = 5%; t_1% = 5% and t_2% = 10%) only.
    - In Figure 10 at t_1% = 5% and t_2% = 5%, the EM algorithm with the first marginal moment barely is closer to $\sigma^2_{y_1}$ = 8 than our method.
- Our method estimated the covariance the best even as the percentage of LLOD observations increased for each variable (Figures 6, 12, and 18).

DATA APPLICATION

The Systemic Lupus Erythematosus in Gullah Health (SLEIGH) study included a toxicology and serology component in which perfluorinated contaminant levels were recorded from serum samples. SLEIGH is an observational cohort study of African American Gullah participants. In the toxicology and serology component of SLEIGH, there are 86 systemic lupus erythematosus (SLE) and 139 control participants. Additional information regarding the SLEIGH study and quality assurance of the serum samples have been published.[21,26,83] In this data application we will analyze the contaminant data for PFDA and PFOA from SLE participants solely. The wet weight of PFDA and PFOA were measured on a continuous scale in nanogram/gram (ng/g) using a laboratory instrument. Commonly, ecological data are log transformed in order that they can be assumed to be normally distributed.[84-87] The MLEs of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ for PFDA and PFOA are presented on a log base 10 scale. There are 15 out 86 PFDA observations (17.442%) appearing

below the LLOD value which is -3.507 log ng/g (0.03 ng/g) for SLE participants. The LLOD value for PFOA is -3.219 log ng/g (0.04 ng/g), and 2 out of 86 observations (2.326%) appear below the LLOD value for SLE participants.

Table 3 SLEIGH study results from all methods are listed this table. Y_1 and Y_2 represents PFDA and PFOA, respectively. Estimates are in log ng/g scale.

LLOD	LLOD/2	LLOD/$\sqrt{2}$	EM	EM One
$\widehat{\mu}_{y_1} = -1.111$	$\widehat{\mu}_{y_1} = -0.805$	$\widehat{\mu}_{y_1} = -0.932$	$\widehat{\mu}_{y_1} = -1.218$	$\widehat{\mu}_{y_1} = -1.182$
$\widehat{\mu}_{y_2} = 0.547$	$\widehat{\mu}_{y_2} = 0.584$	$\widehat{\mu}_{y_2} = 0.568$	$\widehat{\mu}_{y_2} = 0.542$	$\widehat{\mu}_{y_2} = 0.542$
$\widehat{\sigma}^2_{y_1} = 1.852$	$\widehat{\sigma}^2_{y_1} = 0.817$	$\widehat{\sigma}^2_{y_1} = 1.137$	$\widehat{\sigma}^2_{y_1} = 2.453$	$\widehat{\sigma}^2_{y_1} = 2.223$
$\widehat{\sigma}^2_{y_2} = 1.134$	$\widehat{\sigma}^2_{y_2} = 0.909$	$\widehat{\sigma}^2_{y_2} = 0.988$	$\widehat{\sigma}^2_{y_2} = 1.169$	$\widehat{\sigma}^2_{y_2} = 1.168$
$\widehat{\rho\sigma_{y_1}\sigma_{y_2}} = 0.906$	$\widehat{\rho\sigma_{y_1}\sigma_{y_2}} = 0.505$	$\widehat{\rho\sigma_{y_1}\sigma_{y_2}} = 0.657$	$\widehat{\rho\sigma_{y_1}\sigma_{y_2}} = 1.062$	$\widehat{\rho\sigma_{y_1}\sigma_{y_2}} = 1.003$
$\widehat{\rho} = 0.625$	$\widehat{\rho} = 0.586$	$\widehat{\rho} = 0.620$	$\widehat{\rho} = 0.627$	$\widehat{\rho} = 0.622$

The SLEIGH study results for the substitution methods, the proposed EM algorithm, and EM algorithm with the first marginal moments are in Table 3. PFDA and PFOA are labeled as Y_1 and Y_2 respectively in the table. The MLEs for $\boldsymbol{\mu}$ is larger and smaller for $\boldsymbol{\Sigma}$ when comparing the substitution methods to the two EM algorithms. Since PFOA had very few observations below the LLOD value, our EM algorithm and the EM algorithm with the first marginal moments produced similar results for μ_{y_2} and $\sigma^2_{y_2}$ (EM: $\mu_{y_2} = 0.542$ and $\sigma^2_{y_1} = 2.453$; EM One: $\mu_{y_2} = 0.542$ and $\sigma^2_{y_2} = 1.168$). With the exception of the LLOD/2 method, $\widehat{\rho}$ is similar for each of the methods.

DISCUSSION

Our EM-solution under the truncated model for obtaining MLEs of bivariate normal parameters involved marginal, conditional, and joint moments from a truncated bivariate normal distributions. In the simulation study we compared our method with substitution methods and an EM algorithm with only the first marginal moments. By including the EM algorithm with the first marginal moment only, we were able to determine that the joint, conditional, and second marginal moments in our method are beneficial for estimating bivariate normal distribution parameters. The simulation study illustrated that our method produced the least amount of bias for MLEs in most of the simulation scenarios in comparison to the other methods.

Existing methods developed for estimating parameters in the context of left truncation occurring in multiple variables are quite different than our EM algorithm version. Our proposed EM algorithm includes conditional moments of a truncated bivariate normal distribution, while Jin et al. constructed an EM algorithm that did not include conditional moments but did include the generalized method of moments of a truncated multivariate normal distribution.[103] There is another EM solution that includes conditional moments, but the conditional moments are in the context of multivariate normal mixture model.[68] The four scenarios that occur in $\mathbf{Y}$ that yielded to the use of conditional moments in our method have been considered in an imputation method, but the imputation method was constructed for discrete data below an LLOD.[104]

We applied our method to toxicology and serology data, but we believe our proposed estimation method is also beneficial for computing MLEs of parameters from the bivariate normal distribution for studies related to assays and environmental pollutants. Existing literature shows

research interests in estimating parameters in the context of multiple assays and environmental pollutants. [105-111] Our estimation method only focused on estimating parameters from a bivariate normal distribution. A possible future direction of our method is to extend the proposed EM algorithm to incorporate more than two variables.

6. CONCLUSIONS AND FUTURE DIRECTIONS

This dissertation explored statistical estimation methods to account for continuous data with observations appearing below a lower limit of detection (LLOD) value. The Systemic Lupus Erythematosus in Gullah Health (SLEIGH) study motivated the aims of this dissertation. Many contaminant variables collected in the SLEIGH study had one or more LLOD values. Each dissertation aim is summarized below, and future directions are briefly discussed.

AIM 1 CONCLUSIONS

In aim 1 we theoretically showed that for any distribution the maximum likelihood estimates (MLEs) for the parameters of the underlying distribution are equivalent regardless if a left truncation or left censoring approach is applied to data with lower limit of detection (LLOD) observations. Simulation studies and a real data application numerically illustrated the theoretical relationship between left truncation and left censoring approaches. In order to estimate the parameters from an underlying distribution, the two approaches are implemented using iterative algorithms. The left truncation approach was implemented through an expectation-maximization (EM) algorithm while the left censoring approach was applied with the Newton-Raphson method.

Despite both approaches producing identical results, we discussed the advantage of left truncation relative to left censoring. When the parameter space increases, implementation of the left truncation approach is less intensive computationally than the left censoring approach. This is due to the need of first, second, and partial derivatives of the log-likelihood of the parameters with respect to the parameters in order to implement the left censoring approach with the

Newton-Raphson method. The left truncation approach applied with an EM algorithm does not need derivatives, but it does involve the moments of a log-likelihood of the parameters given the observed data. This advantage motived the use of the left truncation approach and EM algorithm in aims 2 and 3.

AIM 2 CONCLUSIONS

The objective of aim 2 was to estimate parameters from a normal distribution when a sample consists of multiple LLOD values. The multiple LLOD values were a result of the contaminant data from the SLEIGH study being measured in separate batches by a laboratory device. The proposed estimation method was an EM algorithm that included moments from a truncated distribution. Specifically, the first and second moments of a multivariate right truncated normal distribution were computed in the E-step of the algorithm. The M-step maximized the log-likelihood of parameters from a linear mixed model using a residual maximum likelihood (REML) approach. Without REML, the parameters in the covariance matrix was severely underestimated. The log-likelihood of the parameters from a linear mixed model had to be simplified using the trace of a matrix so that the second moment could be included in $\mathbf{YY}'$ matrix.

Results from the simulation studies in aim 2 and a previous study illustrated that the variance parameter estimate improves when the second moment is included in an EM algorithm.[66] We are also able to conclude that the second moment does not have an impact on the estimate of the mean parameter from a normal distribution. Lastly, in the simulation study we considered different values for the mean parameter of the normal distribution. Due to the consideration of

different parameter values in the simulation, we concluded that our EM algorithm is more consistent and less sensitive in comparison to substitution methods and an EM algorithm with only the first moment included.

AIM 3 CONCLUSIONS

Aim 3 involved constructing an EM algorithm that would estimate parameters from a bivariate normal distribution when data for each variable includes LLOD observations. Additionally, each variable had a separate LLOD value. The E-step included the marginal (first and second), conditional, and joint moments of a truncated bivariate normal distribution. We discussed which moment was necessary based on if either or both of the paired observations appeared above or below their respective LLOD values. In the M-step we maximized the log-likelihood of parameters from a linear model using REML.

In the simulation studies and data application we compared our EM algorithm with the substitution methods and an EM algorithm that only included the first marginal moments of a right truncated bivariate normal distribution. Various correlation levels were considered in the simulation study. Regardless of the correlation level, our proposed method performed better than the compared methods more often than not. We concluded that the inclusion of the second marginal, conditional, and joint moments are beneficial for estimating parameters from a bivariate normal distribution.

FUTURE DIRECTION RELATED TO AIMS 2 AND 3

One of the primary focus in the future is to extend our approach, which only focused on estimation, to hypothesis testing. At convergence for the proposed EM algorithms, the necessary statistics required for hypothesis testing from a linear mixed model (aim 2) and linear model (aim 3) are obtained. However, we believe the degrees of freedom would need an adjustment to account for the observations appearing below the LLOD value. One possible adjustment may only be to reduce the degrees of freedom by the number of first moments computed in the response vector, $\mathbf{Y}$, that are unique because the LLOD observations that fall below the same LLOD value are replaced with the same first moment. A second possible adjustment is to reduce the degrees of freedom by the total number of observations appearing below the LLOD value. A simulation study is necessary for determining the appropriate degrees of freedom adjustment.

Another direction for future work is related to generalizing the findings from aims 2 and 3 such that an EM algorithm could be used to estimate parameters from a multivariate normal distribution with data for each variable including multiple LLOD values. There are interests in computing MLEs of parameters from multivariate distributions for studies related to assays and environmental contaminants.[105-111] Frequently, multiple LLOD values occur in data collected for assays and environmental pollutants.[10,11]

The development of an EM algorithm that includes methods from aims 2 and 3 will require information about the corresponding observations from different variables. The use of marginal and conditional moments in the E-step will be based on the information about the corresponding observations from different variables. Formulas for marginal and conditional moments from a

truncated multivariate normal distribution have been previously derived.[79,93,94,112,113] The M-step of the EM algorithm will maximize the log-likelihood of a linear mixed model. Similarly to aims 2 and 3, the log-likelihood of the linear mixed model includes $\mathbf{Y}$ and $\mathbf{YY}'$. Below in Table 1, we provide an example of three variables with just one LLOD for each variable and the moments that should be included in for $\mathbf{Y}$. The scenarios are constructed according to which observation appears below or above the variable specific LLOD value. $\mathbf{Y}$ and $\mathbf{YY}'$ become more complex if we were to extend Table 1 to include more variables with 2 or more LLOD values.

Table 1 Example of scenarios and moments associated with developing an EM algorithm with three random variables. Each of the random variables have a separate LLOD value. Observations below the variable specific LLOD value are replaced with either a marginal or conditional moment in the **Y** vector.

Scenarios For Three Variables with One LLOD Value Each	Moment Required for Observations Appearing below LLOD value in **Y**
1) $Y_{i1} > t_1$ and $Y_{i2} > t_2$ and $Y_{i3} > t_3$	N/A
2) $Y_{i1} \leq t_1$ and $Y_{i2} \leq t_2$ and $Y_{i3} \leq t_3$	Marginal Moment
3) $Y_{i1} \leq t_1$ and $Y_{i2} > t_2$ and $Y_{i3} > t_3$	Conditional Moment
4) $Y_{i1} \leq t_1$ and $Y_{i2} \leq t_2$ and $Y_{i3} > t_3$	Conditional Moment
5) $Y_{i1} \leq t_1$ and $Y_{i2} > t_2$ and $Y_{i3} \leq t_3$	Conditional Moment
6) $Y_{i1} > t_1$ and $Y_{i2} \leq t_2$ and $Y_{i3} \leq t_3$	Conditional Moment
7) $Y_{i1} > t_1$ and $Y_{i2} > t_2$ and $Y_{i3} \leq t_3$	Conditional Moment
8) $Y_{i1} > t_1$ and $Y_{i2} \leq t_2$ and $Y_{i3} > t_3$	Conditional Moment

FUTURE DIRECTIONS RELATED TO DISCRETE VARIABLES IN SLEIGH

The data for the contaminants was not the only data appearing below an LLOD in the SLEIGH study. The SLEIGH study also included discrete variables for cell counts in urine samples that were left truncated. The cell count variables represent the number of red blood cells per high power field, white blood cells per high power field, and epithelial cells per high power field. High power field is the area related to greatest magnification level obtained from a microscope.[114,115] The majority of the observations for the cell counts are completely observed, but there are observations that are recorded as below the LLOD value at various left truncation values. For example, the normal range of red blood cells per high power field is 0-4, but in the SLEIGH study, normal ranges of red blood cells per high power field are occasionally recorded as <1 or <4. Due to LLOD observations occurring in cell count variables from the SLEIGH study, we will describe two future directions briefly.

First Future Direction Related to Discrete Variables in SLEIGH

One future direction of this dissertation could be to estimate the mean, λ, from a Poisson distribution when a sample consists of several observations appearing below two different LLOD values. The data with observations appearing from two different LLOD values can be thought of as data arising from a bivariate Poisson (BP) distribution. The BP distribution was introduced by Campbell in 1934 and by Aitken in 1936 as a Poisson correlation function.[116,117] The BP distribution is constructed using three independent random variables. For example, let X_1, X_2, and X_3 be independent random variables where $X_1 \sim \text{Poisson}(\lambda_1)$, $X_2 \sim \text{Poisson}(\lambda_2)$, and $X_3 \sim \text{Poisson}(\lambda_3)$. Also let $Y_1 = X_1 + X_3$ and $Y_2 = X_2 + X_3$. The distributions of Y_1 and Y_2 are as follow, $Y_1 \sim \text{Poisson}(\lambda_1 + \lambda_3)$ and $Y_2 \sim \text{Poisson}(\lambda_2 + \lambda_3)$. The joint distribution of Y_1 and Y_2 is

$f_{BP}(y_1, y_2) \sim BP(\lambda_1, \lambda_2, \lambda_3)$ where $Cov(Y_1, Y_2) = \lambda_3$ and $\lambda_3 \geq 0$. The probability mass function of the BP distribution is simplified as,

$$
\begin{aligned}
f_{BP}(y_1, y_2) &= \sum_{r=0}^{\min(y_1, y_2)} \left(\frac{e^{-\lambda_1 - \lambda_3} (\lambda_1 - \lambda_3)^{y_1 - r}}{(y_1 - r)!} \right) \left(\frac{e^{-\lambda_2 - \lambda_3} (\lambda_2 - \lambda_3)^{y_2 - r}}{(y_2 - r)!} \right) \left(\frac{e^{-\lambda_3} \lambda_3^{\,r}}{r!} \right) \\
&= e^{-(\lambda_1 + \lambda_2 + \lambda_3)} \frac{\lambda_1^{\,y_1}}{y_1!} \frac{\lambda_2^{\,y_2}}{y_2!} \sum_{r=0}^{\min(y_1, y_2)} \frac{\lambda_3^{\,r}}{r!} \left(\frac{y_1 (\lambda_1)^{-r}}{(y_1 - r)!} \right) \left(\frac{y_2 (\lambda_2)^{-r}}{(y_2 - r)!} \right) \\
&= e^{-(\lambda_1 + \lambda_2 + \lambda_3)} \frac{\lambda_1^{\,y_1}}{y_1!} \frac{\lambda_2^{\,y_2}}{y_2!} \sum_{r=0}^{\min(y_1, y_2)} \binom{y_1}{r} \binom{y_2}{r} r! \left(\frac{\lambda_3}{\lambda_1 \lambda_2} \right)^r .
\end{aligned}
$$

There are several things to note about the BP distribution. First, the random variables X_1, X_2, X_3, Y_1, and Y_2 must all be positive. Second, the support of x_3 is $0 \leq x_3 \leq \min(y_1, y_2)$. Therefore the summation in the BP distribution goes from 0 to $\min(y_1, y_2)$. Lastly, recall that

$$
\frac{y_1!}{r!(y_1 - r)!} = \binom{y_1}{r} .
$$

For the future EM algorithm, computation of the first and second moments of Y_1 and Y_2 from a right truncated bivariate Poisson (RTBP) distribution are required in the E-step. The moments of Y_1 and Y_2 were established by Ahmad in 1968.[118] In the M-Step, maximizing the pseudo log-likelihood of a Poisson generalized linear mixed model is required for estimating the parameter of a Poisson distribution.

Second Future Direction Related to Discrete Variables in SLEIGH

A second future direction is to estimate the parameters (i.e. λ_1, λ_2, λ_3) from a BP distribution when observations occasionally appear below the variables' specific LLOD values. Similarly to aim 3, we believe that the E-step of an EM algorithm would require the marginal, conditional, and joint moments of the BP distribution within the context of truncation. Marginal and joint moments of a truncated BP distribution exists, but the conditional moment of a truncated BP distribution for Y_1 appearing below the LLOD value given Y_2 appearing above the LLOD value does not exist.[118-122]

Without any truncation occurring, the conditional distribution of Y_1 given Y_2 from the BP distribution is called the Charlier series (CS) distribution.[123-126] The CS distribution and moments were initially constructed by Ong in 1988.[123] Ding et al. (2015) constructed the probability mass function and moments of a truncated multivariate CS distribution, but the truncation considered only occurs at zero for each variable.[124] Currently, there are not any other statistical methodology articles about a truncated CS distribution. We believe that the theoretical work by Ding et al. (2015) for constructing a zero truncated multivariate CS can be extended to solve the conditional moment of a truncated BP distribution for Y_1 appearing below the LLOD value given Y_2 appearing above the LLOD value.

FINAL REMARKS

This dissertation addressed methods for estimating of parameters within the context of LLOD values. Existing statistical methods related to handling LLOD observations prior to estimating parameters from an underlying distribution were discussed in chapters 1 and 2. Chapter 3 addressed the first aim of this dissertation. In the first aim we established the equivalence of left truncation and left censoring approaches for computing MLEs of parameters from an underlying distribution. Despite the equivalence of the two approaches, we discussed the advantage of the left truncation approach relative to the left censoring approach. In chapters 4 and 5, we constructed estimation methods utilizing the left truncation approach to address the second and third dissertation aims, respectively. Chapter 4 incorporated a method for estimating parameters from a normal distribution when observations appeared below multiple LLOD values. The method proposed in Chapter 5 was an EM algorithm that computed MLEs of parameters from a bivariate normal distribution. Data for each variable included observations appearing below a variable specific LLOD value. In general we believe the estimation methods proposed in this dissertation are beneficial to statistical methods and applications related to data with detection limits.

7. APPENDIX

7.1. APPENDIX FOR R CODE FROM CHAPTER 3

EM Algorithm R code

```
EM.TN.steps<-function(inc.data, mu, sigma.sq, n.missing, pdf.cdf.ratio,t.z.score){
  #E-step
  sigma<-sqrt(sigma.sq)
  #First and second moment of a right truncated normal distribution
  exp.value<-mu-sigma*pdf.cdf.ratio
  exp.value2<-sigma.sq*(1-t.z.score*pdf.cdf.ratio-pdf.cdf.ratio^2)+exp.value^2

  #combining dataset with expected value
  complete.data<-c(inc.data, rep(exp.value,n.missing))

  n.complete.data<-length(complete.data)


  #M-step

  mu.new<-(1/n.complete.data)*(sum(inc.data)+n.missing*exp.value)
  sigma.sq.new<-(1/n.complete.data)*(sum(inc.data^2)+n.missing*exp.value2)-mu.new^2

  list(mu.new=mu.new, sigma.sq.new=sigma.sq.new,exp.value=exp.value,
       complete.data=complete.data,inc.data=inc.data)

}

EM.TN.algorithm<-function(inc.data, mu, sigma.sq, n.missing, pdf.cdf.ratio, tolerance,
iteration.stop, t.z.score){
  mu.current<-NULL
  mu.current[1]<-mu
  sigma.sq.current<-NULL
  sigma.sq.current[1]<-sigma.sq

  iteration<-1
  stop<-1

  while(stop){
    EM.TN.run<-EM.TN.steps(inc.data=inc.data, mu=mu.current[iteration],
sigma.sq=sigma.sq.current[iteration],
                           n.missing=n.missing, pdf.cdf.ratio=pdf.cdf.ratio, t.z.score=t.z.score)
    #Difference between current and new values
    mu.diff<-mu.current[iteration]-EM.TN.run$mu.new
    sigma.sq.diff<-sigma.sq.current[iteration]-EM.TN.run$sigma.sq.new
    stop<-ifelse(mu.diff<tolerance & sigma.sq.diff<tolerance | iteration>iteration.stop, 0, 1)
```

```
   #Updating values
   iteration<-iteration+1
   mu.current[iteration]<-EM.TN.run$mu.new
   sigma.sq.current[iteration]<-EM.TN.run$sigma.sq.new

  }

  list(mu.mle=mu.current[iteration], sigma.sq.mle=sigma.sq.current[iteration],
mu.all=mu.current, sigma.sq.mle=sigma.sq.current,
      exp.value.mle=EM.TN.run$exp.value)

}

##########################################################################
##########################################################################
##########################################################################

#Example for 10% appearing below the limit of detection

set.seed(2031578)
#Simulation size
n.sim<-1000
#Sample size
n.total<-100
#Percentage below the lower limit of detection
trunc.prob<-.10
#Parameter values
mean.true<-5
var.true<-4
#Setting up vectors
mu.mle.loop<-matrix(data=NA, nrow = n.sim, ncol=1)
sigma.sq.mle.loop<-matrix(data=NA, nrow = n.sim, ncol=1)
#mean of the expectation value from the loop
expect.loop<-matrix(data=NA, nrow=n.sim,ncol=
prob<-matrix(data=NA,nrow = n.sim, ncol=1)
missing.value<-matrix(data=NA,nrow = n.sim, ncol=1)
loop.data<-matrix(data=NA, nrow=n.total, ncol=n.sim)
mean.loop<-matrix(data=NA, nrow=n.sim, ncol=1)
var.loop<-matrix(data=NA, nrow=n.sim, ncol=1)
mean.trunc<-matrix(data=NA, nrow=n.sim, ncol=1)
var.trunc<-matrix(data=NA, nrow=n.sim, ncol=1)
pdf.cdf.ratio.vec<-matrix(data = NA, nrow=n.sim, ncol=1)
count<-1

while(count <= nrow(mu.mle.loop)){
  #Generating data from normal distribution
```

```
  loop.data[,count]<- matrix(rnorm(n.total, mean = mean.true, sd = sqrt(var.true)))
  mean.loop[count]<-mean(loop.data[,count])
   var.loop[count]<-var(loop.data[,count])*(nrow(loop.data)-1)/nrow(loop.data)
  #This matrix is for truncating the data
  truncating<-matrix(loop.data[,count])
  #Truncating value
  trunc.value<-qnorm(trunc.prob)*sqrt(var.true)+mean.true
  #All truncated data in this dataset below
  trunc.loop.data<-subset(truncating,truncating[,1] > trunc.value)
  mean.trunc[count]<-mean(trunc.loop.data)
  var.trunc[count]<-var(trunc.loop.data)*((length(trunc.loop.data)-1)/length(trunc.loop.data))
  prob[count]<-1-length(trunc.loop.data)/nrow(loop.data)
  missing.value[count]<-nrow(loop.data)-length(trunc.loop.data)
  #Z-score
  z.score<-qnorm(prob[count])
 pdf.cdf.ratio.vec[count]<-dnorm(z.score)/pnorm(z.score)

  em.algorithm.run<-EM.TN.algorithm(inc.data=trunc.loop.data, mu=mean.trunc[count],
sigma.sq=var.trunc[count],
                            n.missing=missing.value[count], pdf.cdf.ratio=pdf.cdf.ratio.vec[count],
                            tolerance=0.00000001, iteration.stop=30, t.z.score=z.score)

  mu.mle.loop[count,1]<-em.algorithm.run$mu.mle
  sigma.sq.mle.loop[count,1]<-em.algorithm.run$sigma.sq.mle
  expect.loop[count]<-em.algorithm.run$exp.value.mle

  #Updating the count for the while loop
  count<-count+1

  list(mu.mle.loop=mu.mle.loop, sigma.sq.mle.loop=sigma.sq.mle.loop,
expect.loop=expect.loop)

}

# Estimates for mu and sigma square
mean(mu.mle.loop)
mean(sigma.sq.mle.loop)

#MSE for mu hat
mean.sim.mu<-mean(mu.mle.loop)
var.sim.mu<-var(mu.mle.loop)
bias.sim.mu<-mean.sim.mu-mean.true
mse.sim.mu<-var.sim.mu+(bias.sim.mu^2)


#MSE for sigma square hat
```

```
mean.sim.sigma.sq<-mean(sigma.sq.mle.loop)
var.sim.sigma.sq<-var(sigma.sq.mle.loop)
bias.sim.sigma.sq<-mean.sim.sigma.sq-var.true
mse.sim.sigma.sq<-var.sim.sigma.sq+(bias.sim.sigma.sq^2)
```

Newton Raphson Method R Code

```
#Function for log-likelhood with censored observations

library(maxLik)
cens.loglike<-function(param){
  mu<-param[1]
  sigma.sq<-param[2]
#Components of the log-likelihood
  a<--0.5*n.observed*log(2*pi)
  b<-n.observed*log(sqrt(sigma.sq))
  c<-sum(0.5*(y - mu)^2/sigma.sq)
  d<-n.censored*log(pnorm(t,mean = mu,sd=sqrt(sigma.sq)))

# log-likelihood
  ll<-ifelse(n.censored==0, a-b-c, a-b-c+d)

#Returning the log-likelihood value
  ll
}

##########################################################################
##########################################################################
##########################################################################

#Example for 10% censored
set.seed(2031578)
#Simulation size
sim.num<-1000
#Parameter values
mean.true<-5
var.true<-4
#Sample size
n.total<-100
mu.hat.sim<-matrix(data=NA, nrow = sim.num, ncol=1)
sigma.hat.sim<-matrix(data=NA, nrow = sim.num, ncol=1)
n1.obs<-matrix(data=NA, nrow = sim.num, ncol=1)
n2.cens<-matrix(data=NA, nrow = sim.num, ncol=1)
#Proportion below the limit of detection
censor.prob<-.10
count<-1

while( count <= nrow(mu.hat.sim)){
#Vector of all y observations
  y.all <- rnorm(n.total, mean=mean.true, sd=sqrt(var.true))
```

```
#Left truncation value
  t<-qnorm(censor.prob)*sqrt(var.true)+mean.true
#Vector of observation above the left truncation value
y<-subset(y.all,y.all > t)
  n.observed<-length(y)
  n.censored<-length(y.all)-n.observed
  mu.start<-mean(y)
  sigma.sq.start<-var(y)

# Newton-Raphson method
  NR.sim<-maxLik(cens.loglike, start=c(mu=mu.start,sigma.sq=sigma.sq.start),method = 'NR')
  mu.hat.sim[count]<-NR.sim$estimate[1]
  sigma.hat.sim[count]<-NR.sim$estimate[2]
  n1.obs[count]<-n.observed
  n2.cens[count]<-n.censored
  count<-count+1

  list(mu.hat.sim=mu.hat.sim, sigma.hat.sim=sigma.hat.sim, n1.obs=n1.obs,n2.cens=n2.cens)

}

#Estimates for mu and sigma square
mean(mu.hat.sim)
mean(sigma.hat.sim)
#MSE for mu hat
mean.sim.mu<-mean(mu.hat.sim)
var.sim.mu<-var(mu.hat.sim)
bias.sim.mu<-mean.sim.mu-mean.true
mse.sim.mu<-var.sim.mu+(bias.sim.mu^2)


#MSE for sigma square hat
mean.sim.sigma.sq<-mean(sigma.hat.sim)
var.sim.sigma.sq<-var(sigma.hat.sim)
bias.sim.sigma.sq<-mean.sim.sigma.sq-var.true
mse.sim.sigma.sq<-var.sim.sigma.sq+(bias.sim.sigma.sq^2)
```

7.2 Appendix for Formulas from Chapter 5

In the main text of this chapter, we let paired observations be denoted as Y_{ij} and $Y_{i'j'}$ where $i = i'$, $i = 1,\ldots, n$ observations, $j \neq j'$, $j = 1^{st}$ variable, $j' = 2^{nd}$ variable. For simplicity, formulas in this Appendix refers to the first and second paired random variables as Y_1 and Y_2.

Bivariate Normal Distribution

Consider two random variables, Y_1 and Y_2. The joint distribution of Y_1 and Y_2 is a bivariate normal (BVN) distribution with mean vector $\boldsymbol{\mu} = \begin{bmatrix} \mu_{y_1} & \mu_{y_2} \end{bmatrix}'$, and a covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2_{y_1} & \rho\sigma_{y_1}\sigma_{y_2} \\ \rho\sigma_{y_1}\sigma_{y_2} & \sigma^2_{y_2} \end{bmatrix}$. The elements of $\boldsymbol{\Sigma}$ include the variance of Y_1 denoted as $\sigma^2_{y_1}$, $\sigma^2_{y_2}$ is the variance of Y_2, the covariance of Y_1 and Y_2 is represented as $\rho\sigma_{y_1}\sigma_{y_2}$, ρ is the correlation between Y_1 and Y_2, and the standard deviations of are Y_1 and Y_2 represented by σ_{y_1} and σ_{y_2}. The covariance matrix $\boldsymbol{\Sigma}$ is positive semi-definite. The probability density function (pdf) of the bivariate normal distribution is,

$$f_{BVN}\left(y_1, y_2\right) = \left(2\pi\sqrt{|\boldsymbol{\Sigma}|}\right)^{-1} e^{-\frac{1}{2}\left[(\boldsymbol{y}-\boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y}-\boldsymbol{\mu})\right]},$$

where $y_1 \in \mathbb{R}$, $y_2 \in \mathbb{R}$, $|\boldsymbol{\Sigma}|$ is the determinant of the covariance matrix, and $\boldsymbol{\Sigma}^{-1}$ is the inverse of the covariance matrix.

The conditional distribution of Y_1 given Y_2 is a normal distribution with mean $\mu_{y_1|y_2} = \mu_{y_1} + \frac{\sigma_{y_1}}{\sigma_{y_2}} \rho \left(y_2 - \mu_{y_2} \right)$ with variance $\sigma^2_{y_1|y_2} = \left(1-\rho^2\right) \sigma^2_{y_1}$. Likewise, the conditional distribution of Y_2 given Y_1 is a normal distribution with mean $\mu_{y_2|y_1} = \mu_{y_2} + \frac{\sigma_{y_2}}{\sigma_{y_1}} \rho \left(y_1 - \mu_{y_1} \right)$ with variance $\sigma^2_{y_2|y_1} = \left(1-\rho^2\right) \sigma^2_{y_2}$.

Left Truncated Bivariate Normal Distribution

Suppose there are two random variables Y_{1LT} and Y_{2LT}. The support of Y_{1LT} is $Y_{1LT} \in (t_{y_1}, \infty)$, and the support of Y_{2LT} is $Y_{2LT} \in (t_{y_2}, \infty)$. Jointly Y_{1LT} and Y_{2LT} follow a left truncated bivariate normal (LTBVN) distribution with a mean vector $\boldsymbol{\mu} = \begin{bmatrix} \mu_{y_1} & \mu_{y_2} \end{bmatrix}'$, a covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2_{y_1} & \rho\sigma_{y_1}\sigma_{y_2} \\ \rho\sigma_{y_1}\sigma_{y_2} & \sigma^2_{y_2} \end{bmatrix}$, and left truncation vector $\boldsymbol{t} = \begin{bmatrix} t_{y_1} & t_{y_2} \end{bmatrix}'$. With the use the BVN pdf , the LTBVN pdf is expressed as,

$$f_{LTBVN}\left(y_{1LT}, y_{2LT}\right) = \frac{1}{\int_{\frac{t_{y_1}-\mu_{y_1}}{\sigma}}^{\infty} \int_{\frac{t_{y_2}-\mu_{y_2}}{\sigma}}^{\infty} f_{BVN}\left(y_{1LT}, y_{2LT}\right) dy_{1LT} dy_{2LT}} f_{BVN}\left(y_{1LT}, y_{2LT}\right),$$

where $t_{y_1} \le y_{1LT} \le \infty$ and $t_{y_2} \le y_{2LT} \le \infty$.

The following marginal formulas appear in Table 2 of the main text. Let the standard normal pdf of random variable W is $\phi(w) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}w^2}$, where $w \in \mathbb{R}$. The cumulative distribution function (cdf) of the standard normal distribution is defined as, $\Phi\left(w\right) = P\left(W \le w\right) = \int_{-\infty}^{w} \phi\left(z\right) dz$.

Using the pdf and cdf of the standard normal distribution, we can express the marginal mean

variance of each variable. The marginal mean of Y_{1LT} is $\mu_{y_{1LT}} = \mu_{y_1} + \sigma_{y_1} \dfrac{\phi\left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)}{1 - \Phi\left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)}$ and

the marginal variance of Y_{1LT} is

$\sigma^2_{y_{1LT}} = \sigma^2_{y_1}\left[1 + \left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)\dfrac{\phi\left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)}{1 - \Phi\left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)} - \left(\dfrac{\phi\left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)}{1 - \Phi\left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)}\right)^2\right]$. Similarly the marginal

mean of Y_{2LT} is $\mu_{y_{2LT}} = \mu_{y_2} + \sigma_{y_2} \dfrac{\phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}{1 - \Phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}$ and

$\sigma^2_{y_{2LT}} = \sigma^2_{y_2}\left[1 + \left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)\dfrac{\phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}{1 - \Phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)} - \left(\dfrac{\phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}{1 - \Phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}\right)^2\right]$ is the marginal variance

of Y_{2LT}.

Let ρ_{LT} represent the correlation between Y_{1LT} and Y_{2LT}. The conditional distribution of Y_{1LT} given

Y_{2LT} is a LTN distribution with mean $\mu_{y_{1LT}|y_{2LT}} = \mu_{y_{1LT}} + \dfrac{\sigma_{y_{1LT}}}{\sigma_{y_{2LT}}}\rho_{LT}\left(y_{2LT} - \mu_{y_{2LT}}\right)$, and variance

$\sigma^2_{y_{1LT}|y_{2LT}} = \left(1 - \rho^2_{LT}\right)\sigma^2_{y_{1LT}}$. Also the conditional distribution of Y_{2LT} given Y_{1LT} is a LTN

distribution with mean $\mu_{y_{2LT}|y_{1LT}} = \mu_{y_{2LT}} + \frac{\sigma_{y_{2LT}}}{\sigma_{y_{1LT}}} \rho_{LT} \left(y_{1LT} - \mu_{y_{1LT}} \right)$, and variance

$\sigma^2_{y_{2LT}|y_{1LT}} = \left(1 - \rho^2_{LT}\right) \sigma^2_{y_{2LT}}$.

Right Truncated Bivariate Normal Distribution

The joint distribution of random variables Y_{1RT} and Y_{2RT} is a right truncated bivariate normal

(RTBVN) distribution with mean vector $\boldsymbol{\mu} = \begin{bmatrix} \mu_{y_1} & \mu_{y_2} \end{bmatrix}'$, a covariance matrix

$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma^2_{y_1} & \rho\sigma_{y_1}\sigma_{y_2} \\ \rho\sigma_{y_1}\sigma_{y_2} & \sigma^2_{y_2} \end{bmatrix}$, and right truncation vector $\boldsymbol{t} = \begin{bmatrix} t_{y_1} & t_{y_2} \end{bmatrix}'$. Each random variable has

a support of $Y_{1RT} \in (-\infty, t_{y_1})$ and $Y_{2RT} \in (-\infty, t_{y_2})$. The RTBVN pdf is denoted as,

$$f_{RTBVN}\left(y_{1RT}, y_{2RT}\right) = \frac{1}{\int_{-\infty}^{\frac{t_{y_1} - \mu_{y_1}}{\sigma}} \int_{-\infty}^{\frac{t_{y_2} - \mu_{y_2}}{\sigma}} f_{BVN}\left(y_{1RT}, y_{2RT}\right) dy_{1RT} dy_{2RT}} f_{BVN}\left(y_{1RT}, y_{2RT}\right)$$

where $t_{y_1} \le y_{1LT} \le \infty$ and $-\infty \le y_{2RT} \le t_{y_2}$.

The marginal distributions of Y_{1RT} and Y_{2RT} are RTN distributions. The marginal mean of Y_{1RT} is

$$\mu_{y_{1RT}} = \mu_{y_1} - \sigma_{y_1} \frac{\phi\left(\frac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)}{\Phi\left(\frac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)} \quad \text{and}$$

$$\sigma^2_{y_{1RT}} = \sigma^2_{y_1} \left[1 - \left(\frac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right) \frac{\phi\left(\frac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)}{\Phi\left(\frac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)} - \left(\frac{\phi\left(\frac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)}{\Phi\left(\frac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)} \right)^2 \right] \quad \text{is the marginal variance}$$

of Y_{1RT}. Marginally, the mean of Y_{2RT} is $\mu_{y_{2RT}} = \mu_{y_2} - \sigma_{y_2} \dfrac{\phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}{\Phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}$ and the variance of

is expressed as $\sigma^2_{y_{2RT}} = \sigma^2_{y_2}\left[1 - \left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)\dfrac{\phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}{\Phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)} - \left(\dfrac{\phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}{\Phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}\right)^2\right]$.

The conditional distribution of Y_{1RT} given Y_{2RT} is a RTN distribution with mean

$\mu_{y_{1RT}|y_{2RT}} = \mu_{y_{1RT}} + \dfrac{\sigma_{y_{1RT}}}{\sigma_{y_{2RT}}}\rho_{RT}\left(y_{2RT} - \mu_{y_{2RT}}\right)$ and variance $\sigma^2_{y_{1RT}|y_{2RT}} = \left(1 - \rho^2_{RT}\right)\sigma^2_{y_{1RT}}$

where ρ_{RT} is the correlation between Y_{1RT} and Y_{2RT}. Correspondingly the conditional distribution of Y_{2RT} given Y_{1RT} follows a RTN distribution with mean

$\mu_{y_{2RT}|y_{1RT}} = \mu_{y_{2RT}} + \dfrac{\sigma_{y_{2RT}}}{\sigma_{y_{1RT}}}\rho_{RT}\left(y_{1RT} - \mu_{y_{1RT}}\right)$ and variance $\sigma^2_{y_{2RT}|y_{1RT}} = \left(1 - \rho^2_{RT}\right)\sigma^2_{y_{2RT}}$.

In Table 2 from the main text, the joint mean of Y_{1RT} and Y_{2RT} is denoted as $\mu_{y_{1RT}y_{2RT}}$. In our proposed method, $\mu_{y_{1RT}y_{2RT}}$ is computed by the *mtmvnorm* function in the *tmvtnorm* package by Manjunath and Wilhelm.[93,94] Jointly, the mean of Y_{1RT} and Y_{2RT} is found using Manjunath and Wilhelm's formula number 16 in their paper by allowing the lower truncation value be $-\infty$ and

the dimension be 2.[79] The mean of Y_{1RT} and Y_{2RT} is $\mu_{y_{1RT}y_{2RT}} = E\left(X_r, X_s\right)$ where

$$E\left(X_r, X_s\right) = \sigma_{r,s} + \sum_{k=1}^{2} \sigma_{r,k} \frac{\sigma_{s,k}\left(-\infty\left(F_k\left(-\infty\right)\right) - t_{x_k}\left(F_k\left(t_{x_k}\right)\right)\right)}{\sigma_{k,k}}$$
$$+\sum_{k=1}^{2} \sigma_{r,k} \sum_{q \neq k} \left(\sigma_{s,q} - \frac{\sigma_{k,q}\sigma_{s,k}}{\sigma_{k,k}}\right) \times$$
$$\left[\left(F_{k,q}\left(-\infty, -\infty\right) - F_{k,q}\left(-\infty, t_{x_q}\right)\right) - \left(F_{k,q}\left(t_{x_k}, -\infty\right) - F_{k,q}\left(t_{x_k}, t_{x_q}\right)\right)\right],$$

X_r denotes Y_{1RT}, X_s symbolizes Y_{2RT}, $\sigma_{r,s}$ is the r^{th} row and s^{th} column of the covariance matrix $\boldsymbol{\Sigma}$, $F_k\left(x\right)$ is the k^{th} marginal density of the bivariate normal distribution, $F_{k,q}\left(x_r, x_s\right)$ is the bivariate marginal density. Further details about $E\left(X_r, X_s\right), F_k\left(x\right)$, and $F_{k,q}\left(x_r, x_s\right)$ are explained in Manjunath and Wilhelm's paper.[79]

Truncated Bivariate Normal Distribution with Truncation Occurring in Opposite Directions

Suppose there are two random variables Y_{1RT} and Y_{2LT}. The support of Y_{1RT} is $Y_{1RT} \in (-\infty, t_{y_1})$. Y_{2LT} support is $Y_{2LT} \in (t_{y_2}, \infty)$. Jointly Y_{1RT} and Y_{2LT} follow a truncated bivariate normal distribution (TBVN) with a mean vector $\boldsymbol{\mu} = \begin{bmatrix} \mu_{y_1} & \mu_{y_2} \end{bmatrix}'$, a covariance matrix

$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{y_1}^2 & \rho\sigma_{y_1}\sigma_{y_2} \\ \rho\sigma_{y_1}\sigma_{y_2} & \sigma_{y_2}^2 \end{bmatrix}$, and truncation vector $\boldsymbol{t} = \begin{bmatrix} t_{y_1} & t_{y_2} \end{bmatrix}'$. The TBVN pdf is denoted as,

$$f_{TBVN}\left(y_{1RT}, y_{2LT}\right) = \frac{1}{\int_{-\infty}^{\frac{t_{y_1} - \mu_{y_1}}{\sigma}} \int_{\frac{t_{y_2} - \mu_{y_2}}{\sigma}}^{\infty} f_{BVN}\left(y_{1RT}, y_{2LT}\right) dy_{1RT} dy_{2LT}} f_{BVN}\left(y_{1RT}, y_{2LT}\right)$$

where $-\infty \leq y_{1RT} \leq t_{y_1}$ and $t_{y_2} \leq y_{2LT} \leq \infty$.

The marginal distribution of Y_{1RT} is a RTN with mean $\mu_{y_{1RT}} = \mu_{y_1} - \sigma_{y_1} \dfrac{\phi\left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)}{\Phi\left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)}$

and variance $\sigma^2_{y_{1RT}} = \sigma^2_{y_1}\left[1 - \left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)\dfrac{\phi\left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)}{\Phi\left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)} - \left(\dfrac{\phi\left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)}{\Phi\left(\dfrac{t_{y_1} - \mu_{y_1}}{\sigma_{y_1}}\right)}\right)^2\right]$.

Y_{2LT} follows a LTN with mean $\mu_{y_{2LT}} = \mu_{y_2} + \sigma_{y_2} \dfrac{\phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}{1 - \Phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}$

and variance $\sigma^2_{y_{2LT}} = \sigma^2_{y_2}\left[1 + \left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)\dfrac{\phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}{1 - \Phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)} - \left(\dfrac{\phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}{1 - \Phi\left(\dfrac{t_{y_2} - \mu_{y_2}}{\sigma_{y_2}}\right)}\right)^2\right]$.

The conditional mean and variance of Y_{1RT} given Y_{2LT} are required in Table 2 of the main text.

The conditional mean and variance are $\mu_{y_{1RT}|y_{2LT}} = \mu_{y_{1RT}} + \dfrac{\sigma_{y_{1RT}}}{\sigma_{y_{2LT}}} \rho_T \left(y_{2LT} - \mu_{y_{2LT}}\right)$ and

$\sigma^2_{y_{1RT}|y_{2LT}} = \left(1 - \rho_T^2\right)\sigma^2_{y_{1RT}}$ where ρ_T is the correlation between Y_{1RT} and Y_{2LT}.

In a similar fashion we can determine the conditional mean and variance of Y_{2RT} given Y_{1LT}. This conditional mean and variance appears in Table 2 of the main text. The support of Y_{1LT} is

$Y_{1LT} \in (t_{y_1}, \infty)$. Y_{2RT} support is $Y_{2RT} \in (-\infty, t_{y_2})$. The joint distribution of Y_{1LT} and Y_{2RT} follow a truncated bivariate normal distribution (TBVN) with a mean vector $\boldsymbol{\mu} = \begin{bmatrix} \mu_{y_1} & \mu_{y_2} \end{bmatrix}'$, a covariance matrix $\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{y_1}^2 & \rho\sigma_{y_1}\sigma_{y_2} \\ \rho\sigma_{y_1}\sigma_{y_2} & \sigma_{y_2}^2 \end{bmatrix}$, and truncation vector $\boldsymbol{t} = \begin{bmatrix} t_{y_1} & t_{y_2} \end{bmatrix}'$. This TBVN pdf is denoted as,

$$f_{TBVN}\left(y_{1LT}, y_{2RT}\right) = \frac{1}{\int_{\frac{t_{y_1}-\mu_{y_1}}{\sigma}}^{\infty} \int_{-\infty}^{\frac{t_{y_2}-\mu_{y_2}}{\sigma}} f_{BVN}\left(y_{1LT}, y_{2RT}\right) dy_{1LT} dy_{2RT}} f_{BVN}\left(y_{1LT}, y_{2RT}\right)$$

where $t_{y_1} \le y_{1LT} \le \infty$ and $-\infty \le y_{2RT} \le t_{y_2}$. The conditional mean and variance of Y_{2RT} given Y_{1LT} are $\mu_{y_{2RT}|y_{1LT}} = \mu_{y_{2RT}} + \frac{\sigma_{y_{2RT}}}{\sigma_{y_{1LT}}} \rho_T \left(y_{1LT} - \mu_{y_{1LT}}\right)$ and $\sigma^2_{y_{2RT}|y_{1LT}} = \left(1 - \rho_T^2\right) \sigma^2_{y_{2RT}}$ where ρ_T is the correlation between Y_{1LT} and Y_{2RT}.

8. REFERENCES

1. Armbruster DA, Pry T. Limit of blank, limit of detection and limit of quantitation. *The Clinical Biochemist Reviews.* 2008;29(Suppl 1):S49-S52.
2. ICH. Validation of analytical procedures: Text and methodology q2(r1). In:2005.
3. Woolley CF, Hayes MA, Mahanti P, Douglass Gilman S, Taylor T. Theoretical limitations of quantification for noncompetitive sandwich immunoassays. *Anal Bioanal Chem.* 2015;407(28):8605-8615.
4. Alankar S, Vipin BG. Methods for the determination of limit of detection and limit of quantitation of the analytical methods. *chronicles of young scientists.* 2011;2(1):21-25.
5. Cohen AC. Estimating the mean and variance of normal populations from singly truncated and doubly truncated samples. *The Annals of Mathematical Statistics.* 1950;21(4):557-569.
6. Cohen AC. On estimating the mean and standard deviation of truncated normal distributions. *Journal of the American Statistical Association.* 1949;44(248):518-525.
7. Hald A. Maximum likelihood estimation of the parameters of a normal distribution which is truncated at a known point. *Scandinavian Actuarial Journal.* 1949;1949(1):119-134.
8. Gupta AK. Estimation of the mean and standard deviation of a normal population from a censored sample. *Biometrika.* 1952;39(3/4):260.
9. Cohen AC. Simplified estimators for the normal distribution when samples are singly censored or truncated. *Technometrics.* 1959;1(3):217-237.
10. Whitcomb BW, Schisterman EF. Assays with lower detection limits: implications for epidemiological investigations. *Paediatr Perinat Epidemiol.* 2008;22(6):597-602.
11. Lee L, Helsel D. Statistical analysis of water-quality data containing multiple detection limits: S-language software for regression on order statistics. *Computers & Geosciences.* 2005;31(10):1241-1248.
12. Lawson GM. Defining limit of detection and limit of quantitation as applied to drug of abuse testing: striving for a consensus. *Clinical chemistry.* 1994;40(7 Pt 1):1218-1219.
13. Saadati N, Abdullah MP, Zakaria Z, Sany SBT, Rezayi M, Hassonizadeh H. Limit of detection and limit of quantification development procedures for organochlorine pesticides analysis in water and sediment matrices. *Chem Cent J.* 2013;7(1):63-63.
14. Flikkema RM. *Statistical Methodology for Data with Multiple Limits of Detection* [Dissertations]2016.
15. Aboueissa AE-MA. Maximum likelihood estimators of population parameters from multiply censored samples. *Environmetrics.* 2009;20(3):312-330.
16. Aboueissa AE-MA, Stoline MR. Maximum likelihood estimators of population parameters from doubly left-censored samples. *Environmetrics.* 2006;17(8):811-826.
17. Peng C. Interval estimation of population parameters based on environmental data with detection limits. *Environmetrics.* 2010;21(6):645-658.
18. Haiying C, Quandt SA, Grzywacz JG, Arcury TA. A Distribution-Based Multiple Imputation Method for Handling Bivariate Pesticide Data with Values below the Limit of Detection. *Environmental Health Perspectives.* 2011;119(3):351-356.
19. Hoffman HJ, Johnson RE. Pseudo-likelihood estimation of multivariate normal parameters in the presence of left-censored data. *Journal of Agricultural, Biological, and Environmental Statistics.* 2015;20(1):156-171.

20. Hoffman HJ, Johnson RE. Estimation of multiple trace metal water contaminants in the presence of left-censored and missing data. *Journal of Environmental Statistics.* 2011;2(2).
21. Gilkeson G, James J, Kamen D, et al. The United States to Africa lupus prevalence gradient revisited. *Lupus.* 2011;20(10):1095-1103.
22. Suja F, Pramanik BK, Zain SM. Contamination, bioaccumulation and toxic effects of perfluorinated chemicals (PFCs) in the water environment: a review paper. *Water Science and Technology.* 2009;60(6):1533-1544.
23. Costa LG, Giordano G, Tagliaferri S, Caglieri A, Mutti A. Polybrominated diphenyl ether (PBDE) flame retardants: environmental contamination, human body burden and potential adverse health effects. *Acta bio-medica : Atenei Parmensis.* 2008;79(3):172-183.
24. McDonald TA. A perspective on the potential health risks of PBDEs. *Chemosphere.* 2002;46(5):745-755.
25. DeWitt JC. Toxicological effects of perfluoroalkyl and polyfluoroalkyl substances. In: Cham: Humana Press; 2015.
26. Gribble MO, Bartell SM, Kannan K, Wu Q, Fair PA, Kamen DL. Longitudinal measures of perfluoroalkyl substances (PFAS) in serum of Gullah African Americans in South Carolina: 2003-2013. *Environ Res.* 2015;143(Pt B):82-88.
27. Harel O, Perkins N, Schisterman EF. The use of multiple imputation for data subject to limits of detection. *Sri Lankan J Appl Stat.* 2014;5(4):227-246.
28. Little RJ, Rubin DB. *Statistical analysis with missing data.* New York: Wiley; 2002.
29. Helsel DR. Fabricating data: How substituting values for nondetects can ruin results, and what can be done about it. *Chemosphere.* 2006;65(11):2434-2439.
30. 13 methods for data below the reporting limit. In: Helsel DR, Hirsch RM, eds. *Studies in Environmental Science.* Vol 49. Elsevier; 1992:357-376.
31. Hopke PK, Liu C, Rubin DB. Multiple imputation for multivariate data with missing and below-threshold measurements: Time-series concentrations of pollutants in the arctic. *Biometrics.* 2001;57(1):22-33.
32. Horton NJ, Lipsitz SR. Multiple imputation in practice: comparison of software packages for regression models with missing variables. *The American Statistician.* 2001;55(3):244-254.
33. Nie L, Chu H, Liu C, Cole SR, Vexler A, Schisterman EF. Linear regression with an independent variable subject to a detection limit. *Epidemiology (Cambridge, Mass).* 2010;21 Suppl 4(Suppl 4):S17-S24.
34. USEPA. Guidance for data quality assessment: Practical methods for data analysis. In. Washington, DC Office of Environmental Information; 2000:154-155.
35. Cohen MA, Ryan PB. Observations less than the analytical limit of detection: A new approach. *JAPCA.* 1989;39(3):328-329.
36. Helsel DR. *Nondetects and data analysis: Statistics for censored environmental data.* John Wiley & Sons; 2005.
37. Barr Dana B, Landsittel D, Nishioka M, et al. A survey of laboratory and statistical issues related to farmworker exposure studies. *Environmental Health Perspectives.* 2006;114(6):961-968.
38. Guidance for industry: Q2b validation of analytical procedures: Methodology. In: US Food and Drug Administration; 1996.

39. Tobin J. Estimation of relationships for limited dependent variables. *Econometrica: journal of the Econometric Society.* 1958:24-36.
40. Smith D, Silver E, Harnly M. *Environmental samples below the limits of detection - comparing regression methods to predict environmental concentrations.* 2006.
41. Lee M, Rahbar MH, Brown M, et al. A multiple imputation method based on weighted quantile regression models for longitudinal censored biomarker data with missing values at early visits. *BMC Medical Research Methodology.* 2018;18(1):8.
42. Eilers PHC, Röder E, Savelkoul HFJ, van Wijk RG. Quantile regression for the statistical analysis of immunological data with many non-detects. *BMC Immunol.* 2012;13:37-37.
43. Lee M, Kong L. Quantile regression for longitudinal biomarker data subject to left censoring and dropouts. *Communications in Statistics - Theory and Methods.* 2014;43(21):4628-4641.
44. Lee M, Kong L, Weissfeld L. Multiple imputation for left-censored biomarker data based on Gibbs sampling method. *Statistics in medicine.* 2012;31(17):1838-1848.
45. Tellez-Plaza M, Navas-Acien A, Crainiceanu C, Guallar E. A tobit model to address the instrumental limit of detection in the study of blood cadmium and peripheral arterial disease in us adults. *Epidemiology.* 2009;20(6):S187.
46. N. WM, Andreas Z. Multiple censored data in dentistry: A new statistical model for analyzing lesion size in randomized controlled trials. *Biometrical Journal.* 2015;57(3):384-394.
47. Uh H-W, Hartgers FC, Yazdanbakhsh M, Houwing-Duistermaat JJ. Evaluation of regression methods when immunological measurements are constrained by detection limits. *BMC Immunol.* 2008;9:59-59.
48. Lubin JH, Colt JS, Camann D, et al. Epidemiologic evaluation of measurement data in the presence of detection limits. *Environmental health perspectives.* 2004;112(17):1691-1696.
49. Lorimer MF, Kiermeier A. Analysing microbiological data: Tobit or not Tobit? *International Journal of Food Microbiology.* 2007;116(3):313-318.
50. Dempster A, M. Laird N, B. Rubin D. Maximum likelihood from incomplete data via the em algorithm1977:1-38.
51. Wilks DS. Chapter 4 - parametric probability distributions. In: Wilks DS, ed. *International Geophysics.* Vol 100. Academic Press; 2011:71-131.
52. Donato DI. *Computing maximum-likelihood estimates for parameters of the National Descriptive Model of Mercury in Fish.* Reston, VA2012. 2012-1181.
53. Helsel D. *Statistics for censored environmental data using Minitab and R.* 2012.
54. Swan AV. Algorithm as 16: Maximum likelihood estimation from grouped and censored normal data. *Journal of the Royal Statistical Society Series C (Applied Statistics).* 1969;18(1):110-114.
55. Singh A, Nocerino J. Robust estimation of mean and variance using environmental data sets with below detection limit observations. *Chemometrics and Intelligent Laboratory Systems.* 2002;60(1):69-86.
56. Halperin M. Estimation in the truncated normal distribution. *Journal of the American Statistical Association.* 1952;47(259):457-465.
57. Kornerup P, Muller J-M. Choosing starting values for certain Newton–Raphson iterations. *Theoretical Computer Science.* 2006;351(1):101-110.

58. T. Fike C. *Starting approximations for square root calculation on IBM System 360.* Vol 91966.
59. Eve J. Starting approximations for the iterative calculation of square roots. *The Computer Journal.* 1963;6(3):274-276.
60. Sterbenz PH, Fike CT. Optimal starting approximations for newton's method. *Mathematics of Computation.* 1969;23(106):313-318.
61. Wilson MW. Optimal starting approximations for generating square root for slow or no divide. *Commun ACM.* 1970;13(9):559-560.
62. Hattaway JT. Parameter Estimation and Hypothesis Testing for the Truncated Normal Distribution with Applications to Introductory Statistics Grades. In. Bringham Young University: All Theses and Dissertations. 2053; 2010.
63. Alaesa MSI. *Comparison among some methods of estimating the parameters of truncated normal distribution*. Jordan, Zarqa University; 2017.
64. Wolynetz MS. Algorithm as 139: Maximum likelihood estimation in a linear model from confined and censored normal data. *Journal of the Royal Statistical Society Series C (Applied Statistics).* 1979;28(2):195-206.
65. Park C, Lee SB. Parameter estimation from censored samples using the expectation-maximization algorithm2012.
66. Bee M. On maximum likelihood estimation of operational loss distributions. In:2005.
67. McLachlan G, Krishnan T. *The EM algorithm and extensions.* Vol 382: John Wiley & Sons; 2007.
68. Lee G, Scott C. EM algorithms for multivariate Gaussian mixture models with truncated and censored data. *Computational Statistics & Data Analysis.* 2012;56(9):2816-2829.
69. Lodhi C, Mani Tripathi Y, Kumar Rastogi M. Estimating the parameters of a truncated normal distribution under progressive type II censoring. *Communications in Statistics - Simulation and Computation.* 2019:1-25.
70. Emura T, Shiu S-K. Estimation and model selection for left-truncated and right-censored lifetime data with application to electric power transformers analysis. *Communications in Statistics - Simulation and Computation.* 2016;45(9):3171-3189.
71. Zunxiong Liu YC, Shanshan Tian, Zheng Xu. Multivariate Gaussian Mixture Model Based Clustering with Truncated and Censored Data. *Journal of Information and Computational Science.* 2015;12(2):775-785.
72. Karl P, Alice L. On the generalised probable error in multiple normal correlation. *Biometrika.* 1908;6(1):59-68.
73. Lee A. Table of the gaussian "tail" functions; when the "tail" is larger than the body. *Biometrika.* 1914;10(2/3):208-214.
74. Birnbaum ZW, Meyer PL. On the effect of truncation in some or all co-ordinates of a multi-normal population. *Journal of the Indian Society of Agricultural Statistics Indian Society of Agricultural Statistics.* 1953;5:17-28.
75. Weiller H. Means and standard deviations of a truncated normal bivariate distribution. *Australian Journal of Statistics.* 1959;1(3):73-81.
76. Rosenbaum S. Moments of a truncated bivariate normal distribution. *Journal of the Royal Statistical Society Series B (Methodological).* 1961;23(2):405-408.
77. Tallis GM. The moment generating function of the truncated multi-normal distribution. *Journal of the Royal Statistical Society Series B (Methodological).* 1961;23(1):223-229.

78. Muthén B. Moments of the censored and truncated bivariate normal distribution. *British Journal of Mathematical and Statistical Psychology.* 1990;43(1):131-143.
79. Manjunath B, Wilhelm S. *Moments calculation for the doubly truncated multivariate normal density.* 2012.
80. Jain RB, Wang RY. Limitations of maximum likelihood estimation procedures when a majority of the observations are below the limit of detection. *Analytical Chemistry.* 2008;80(12):4767-4772.
81. *R: A language and environment for statistical computing* [computer program]. Vienna, Austria: R Foundation for Statistical Computing; 2016.
82. Henningsen A, Toomet O. maxLik: A package for maximum likelihood estimation in R. *Computational Statistics.* 2011;26(3):443-458.
83. Kamen DL, Barron M, Parker TM, et al. Autoantibody prevalence and lupus characteristics in a unique African American population. *Arthritis and rheumatism.* 2008;58(5):1237-1247.
84. Eberhardt LL, Gilbert RO, Hollister HL, Thomas JM. Sampling for contaminants in ecological systems. *Environmental Science & Technology.* 1976;10(9):917-925.
85. McDonald JH. *Handbook of biological statistics.* 3rd ed. Baltimore, Maryland: Sparky House Publishing; 2014.
86. Sokal RR, Rohlf FJ. *Biometry: The principles and practice of statistics in biological research.* W. H. Freeman; 1981.
87. Zar JH. *Biostatistical analysis.* Prentice-Hall; 1974.
88. Palarea-Albaladejo J, Martín-Fernández JA. Values below detection limit in compositional chemical data. *Analytica Chimica Acta.* 2013;764:32-43.
89. Lynn HS. Maximum likelihood inference for left-censored HIV RNA data. *Statistics in medicine.* 2001;20(1):33-45.
90. Liu Y, Brown SD. Imputation of left-censored data for cluster analysis. *Journal of Chemometrics.* 2014;28(3):148-160.
91. Helsel DR, Cohn TA. Estimation of descriptive statistics for multiply censored water quality data. *Water Resources Research.* 1988;24(12):1997-2004.
92. Aboueissa AE-MA, Stoline MR. Estimation of the mean and standard deviation from normally distributed singly-censored samples. *Environmetrics.* 2004;15(7):659-673.
93. *tmvtnorm: Truncated multivariate normal and student t distribution* [computer program]. Version R package version 1.4-102015.
94. Wilhelm S, Manjunath B. *tmvtnorm: A package for the truncated multivariate normal distribution.* Vol 22010.
95. Leppard P, Tallis GM. Algorithm as 249: Evaluation of the mean and covariance of the truncated multinormal distribution. *Applied Statistics.* 1989;38(3):543.
96. Lee L-F. The determination of moments of the doubly truncated multivariate normal tobit model. *Economics Letters.* 1983;11(3):245-250.
97. Verbyla A. A conditional derivation of residual maximum likelihood. *Australian Journal of Statistics.* 1990;32:227-230.
98. Newman MC, Dixon PM, Looney BB, Pinder JE. Estimating mean and variance for environmental samples with below detection limit observations. *Journal of the American Water Resources Association.* 1989;25(4):905.
99. Barr DR, Sherrill ET. Mean and variance of truncated normal distributions. *The American Statistician.* 1999;53(4):357-361.

100. Chen H, Quandt SA, Grzywacz JG, Arcury TA. A distribution-based multiple imputation method for handling bivariate pesticide data with values below the limit of detection. *Environmental health perspectives.* 2011;119(3):351-356.
101. Yu J-w, Tian G-L. *Efficient algorithms for generating truncated multivariate normal distributions.* Vol 272011.
102. Bolker B. Package 'bbmle'. 2017.
103. Jin B-s, Han J-j, Ding S, Miao B-q. Em algorithm of the truncated multinormal distribution with linear restriction on the variables. *Acta Mathematicae Applicatae Sinica, English Series.* 2018;34(1):155-162.
104. Lyles RH, Fan D, Chuachoowong R. Correlation coefficient estimation involving a left censored laboratory assay variable. *Statistics in medicine.* 2001;20(19):2921-2933.
105. LaFleur B, Lee W, Billhiemer D, Lockhart C, Liu J, Merchant N. Statistical methods for assays with limits of detection: Serum bile acid as a differentiator between patients with normal colons, adenomas, and colorectal cancer. *Journal of Carcinogenesis.* 2011;10(1):12-12.
106. Iturria SJ. Statistical inference for relative potency in bivariate dose-response assays with correlated responses. *Journal of Biopharmaceutical Statistics.* 2005;15(2):343-351.
107. Vink MA, Berkhof J, van de Kassteele J, van Boven M, Bogaards JA. A bivariate mixture model for natural antibody levels to human papillomavirus types 16 and 18: Baseline estimates for monitoring the herd effects of immunization. *PLOS ONE.* 2016;11(8):e0161109.
108. Vølund A. Multivariate bioassay. *Biometrics.* 1980;36(2):225-236.
109. Krajden M, Minor J, Cork L, Comanor L. Multi-measurement method comparison of three commercial hepatitis B virus DNA quantification assays. *Journal of Viral Hepatitis.* 1998;5(6):415-422.
110. Wagner BJ. Simultaneous parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modelling. *Journal of Hydrology.* 1992;135(1-4):275-303.
111. Di Leonardo R, Adelfio G, Bellanca A, Chiodi M, Mazzola S. Analysis and assessment of trace element contamination in offshore sediments of the Augusta Bay (SE Sicily): A multivariate statistical approach based on canonical correlation analysis and mixture density estimation approach. *Journal of Sea Research.* 2014;85:428-442.
112. Arismendi JC. Multivariate truncated moments. *Journal of Multivariate Analysis.* 2013;117:41-75.
113. Horrace WC. Some results on the multivariate truncated normal distribution. *Journal of Multivariate Analysis.* 2005;94(1):209-221.
114. Sovoda R, Donev K, Qu Z. A low count load method for accurate quantitative assessments of high power fields. *American Journal of Clinical Pathology.* 2016;146(suppl_1).
115. Meuten DJ, Moore FM, George JW. Mitotic count and the field of view area: Time to standardize. *Veterinary Pathology.* 2015;53(1):7-9.
116. Campbell JT. The Poisson Correlation Function. *Proceedings of the Edinburgh Mathematical Society.* 1934;4(01):18.
117. Aitken AC. A further note on multivariate selection. *Proceedings of the Edinburgh Mathematical Society.* 1936;5(1):37-40.

118. Ahmad M. Truncated multivariate Poisson distribution. *Retrospective Theses and Dissertations.* 1968;3272. https://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=4271&context=rtd.
119. Patil SA, Patel DI, Kovner JL. On bivariate truncated poisson distribution. *Journal of Statistical Computation and Simulation.* 1977;6(1):49-66.
120. Hamdan MA. Estimation in the Truncated Bivariate Poisson Distribution. *Technometrics.* 1972;14(1):37-45.
121. Cohen AC. Estimation of the Poisson Parameter from Truncated Samples and from Censored Samples. *Journal of the American Statistical Association.* 1954;49(265):158-168.
122. Cohen AC. Estimating the Poisson Parameter from Samples That Are Truncated on the Right. *Technometrics.* 1961;3(3):433-438.
123. Ong SH. A Discrete Charlier Series Distribution. *Biometrical Journal.* 1988;30(8):1003-1009.
124. Ding X, Ju D, Tian G-L. Multivariate zero-truncated/adjusted Charlier series distributions with applications. *Journal of Statistical Distributions and Applications.* 2015;2(1):5.
125. Papageorgiou H, Loukas S, Loukas S. A bivariate discrete charlier series distribution. *Biometrical Journal.* 1995;37(1):105-117.
126. Kitano M, Shimizu K, Ong S-H. *The generalized Charlier series distribution as a distribution with two-step recursion.* Vol 752005.