A unified framework for estimation in lognormal models (DRAFT)

Fengqing Zhang and Jiangtao Gou

February 8, 2019

[ABSTRACT] Lognormal models have broad applications in various research areas such as economics, actuarial science, biology, environmental science and psychology. The estimation problem in lognormal models have been extensively studied. In this paper, we summarize all the existing estimators for lognormal models, which belong to twelve estimator families. As some estimators were only proposed for the independent and identical distribution setting, we further generalize these estimators to accommodate the general loglinear regression setting. Additionally, we propose nineteen new estimators based on different optimization criteria. Mostly importantly, we present a unified framework for all the existing and newly proposed estimators. Simulation studies are conducted to compare and evaluate the different estimators in terms of bias, mean absolute error, mean squared error, mean cubic error, and mean fourth power of error. Data from the Economic Research Service in the United States Department of Agriculture are used to demonstrate the application of the various estimators using a lognormal linear regression model. Findings from our study provide comparative benchmarks on all the available estimators for lognormal models, which is helpful for researchers to select the estimator most appropriate for the study at hand.

1 Introduction

A lognormal distribution occurs if the variability is based on a variety of independent forces acting multiplicatively (Limpert et al., 2001). The lognormal model has been applied in economics since Bachelier (1900)'s work. In econometric research, a production process is expressed in terms of a lognormal regression model, which is known as a Cobb-Douglas function model:

$$X_i = \prod_{j=1}^{p} U_{ij}^{\beta_j} e^{\varepsilon_i},\ i = 1, \ldots, n, \tag{1}$$

where X_i is economic output in the ith observation, U_{ij} represent the jth explanatory variables in the ith observation, β_j's are unknown parameters, and $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ are independent normal errors. The model parameters are estimated from the logarithmic form of the model equations $\log X_i = \sum_{j=1}^{p} \beta_j \log U_{ij} + \varepsilon_i$. It is often of interest to estimate $\mathbb{E}[\widehat{X}|U = u]$, which is $\mathbb{E}[\exp(\sum_{j=1}^{p} \widehat{\beta}_j \log u_j)]$, as a descriptive measure of central tendency for a lognormal population in econometrics (Goldberger, 1968; Rukhin, 1988).

Lognormal models are also widely applied in various branches of natural, social and applied sciences (Limpert et al., 2001), including actuarial science (Serfling, 2002), astronomy (Fontenla et al., 2007) biology (Koch, 1966; Loper et al., 1984; Baur, 1997), chemistry (Siano, 1972), ecology (Preston, 1981), environmental science (Biondini, 1976; de Valk and Cai, 2018), geology (Krige, 1966), linguistics (Herdan, 1958), medicine (Kondo, 1977), physics (Pirjol, 2011), psychology (Ulrich and Miller, 1993; Zhang et al., 2018), and sociology (Marwell et al., 1988).

The estimation problem in lognormal models has a long history. In this article, we start with a summary and further extension of nineteen existing estimators which belong to twelve estimator families. Some estimators were proposed only for the independent and identically distributed setting, so we generalized these estimators and extended them to accommodate the general loglinear regression model. All existing estimators and their extensions are included in a unified framework where all estimators share the same expression $\exp(a\widehat{\mu} + b\psi S^2/2)$ with different parameter ψ's. We then compare the expansions of these ψ's in different estimators, and categorize these estimators into several groups based on the leading terms of ψ's expansions. Based on a variety of optimization criteria with some approximation methods, we further obtain five new families of estimators, including nineteen new estimators. The total thirty-eight estimators are compared and evaluated via simulation studies, and several ranking lists are provided based on estimators' behaviors on error rate minimization. We also demonstrate the application of the various estimators using a lognormal linear regression model with a data set from the Economic Research Service (ERS) in the United States Department of Agriculture (USDA). Mathematical derivations of parameter ψ's expansions and the newly proposed estimators are given in the appendix.

2 The unified framework

A random variable X is lognormally distributed with parameter μ and σ^2 if $\log(X)$ has the normal distribution with mean μ and variance σ^2. Lognormally distributed data exhibit noticeable skewness. In this case, log-transforming the data will make the normal distribution-based statistical methods fit their corresponding assumptions. In practice, by applying the log transformation on each observation, we make the data normal. Next, we calculate the statistics on the transformed data. Finally, we back-transform these numbers, in order to informatively report the results in original units.

Suppose the original data $\{X_i\}_{i=1}^n$ follow the lognormal distribution. For the log-transformed data $\{\log(X_i)\}_{i=1}^n$, assume we have available an estimator $\widehat{\mu}$ of μ and an estimator S^2 of σ^2, so the log-transformed data are summarized in the form $\widehat{\mu}\pm S$. We further suppose the estimator $\widehat{\mu}$ follows a normal distribution with mean μ and variance $d\sigma^2$, where the positive constant d is known. We also suppose mS^2/σ^2 has a chi-squared distribution with m degrees of freedom, where S^2 and $\widehat{\mu}$ are independent based on Basu (1955)'s theorem. In the independent and identically distributed (i.i.d.) case, we have $\widehat{\mu} = \overline{\log(X)}$, $d = 1/n$ and $m = n-1$. In a lognormal linear model $\log(X) = T\beta + \varepsilon$ where $T = \{t_{ij}\}_{n\times p}$, $\beta = (\beta_1, \ldots, \beta_p)^\intercal$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^\intercal$ with $\varepsilon_i \sim N(0, \sigma^2)$ i.i.d., suppose we aim to estimate $\mu = t_0^\intercal\beta$ for some specified predictor vector $t_0 = (t_{01}, \ldots, t_{0p})^\intercal$, we have $\widehat{\mu} = t_0^\intercal\widehat{\beta}$, $d = t_0^\intercal(T^\intercal T)^{-1}t_0$ and $m = n - p$.

Consider a parametric function in the original scale we are interested in estimating

$$\theta(a, b) = \exp(a\mu + b\sigma^2/2), \tag{2}$$

where constants a and b are known. Specifically, $\theta(1, 1)$ is the mean of the lognormal distribution, $\theta(2, 4)$ is the second moment, $\theta(2, 4) - \theta(2, 2)$ is the variance, and $(\theta(0, 2) - 1)^{1/2}$ is the coefficient of variation. For simplicity, when parameters a and b do not need to be specified, we will simply write θ rather than $\theta(a, b)$.

A naive estimator of θ is

$$\widehat{\theta}_{\text{naive}} = \exp(a\widehat{\mu} + bS^2/2), \tag{3}$$

which is calculated by directly back-transforming $a\widehat{\mu} + bS^2/2$ to the original scale. Since statistics $\widehat{\mu}$ and S^2 are from the log-transformed data, the naive estimator comes with an apparent shortcoming—the bias from the back-transformation (Smith, 1993; Packard, 2013).

We propose a family of estimators of θ with different ψ's:

$$\widehat{\theta}_\psi = \exp(a\widehat{\mu} + b\psi S^2/2), \tag{4}$$

where ψ represents various corrections on S^2. The naive estimator in (3) is a member of this family where $\psi = 1$. Some well-known estimators are special cases of this general estimator in (4), and we will construct new estimators under this framework.

2.1 Review and generalization of existing estimators

1. Maximum likelihood estimator (ML) and quasi maximum likelihood estimator (QML). These two simple estimators are widely used in practice, and a general discussion can be found in Lawless (1982) and Cohn et al. (1989). The QML estimator is the same as the native estimator in (3), where $\psi = 1$. The ML estimator uses the maximum likelihood estimator of σ^2 instead of S^2, where $\widehat{\sigma^2_{ML}} = mS^2/n$. So the ML estimator is a member of $\widehat{\theta}_\psi$ family where $\psi = m/n$.

2. Simple adjustment estimator (SA). Note that $\mathbf{E}[\exp(a\widehat{\mu})] = \exp(a\mu + a^2 d\sigma^2/2)$, where $\widehat{\mu} \sim (0, d\sigma^2)$. A simple adjustment is followed by using

$$\text{SA: } \psi = 1 - a^2 d/b, \tag{5}$$

which makes $\exp(b\psi\sigma^2/2)\cdot\mathbf{E}[\exp(a\widehat{\mu})] = \exp(a\mu + b\sigma^2/2)$. Meulenberg (1965) and Kennedy (1981) have applied this simple adjustment in econometrics.

3. Finney's estimator (F). Finney (1941) derived an estimator of mean and that of variance of a lognormal distribution for the independent and identically distributed case. Laurent (1963) showed that Finney's estimator is a minimum variance unbiased estimator (MVUE) following from the Rao-Blackwell theorem. Dhrymes (1962), Heien (1968) and Bradu and Mundlak (1970) adopted Finney's approach to accommodate the lognormal linear models. Aitchison and Brown (1957) and Crow and Shimizu (1988) provided comprehensive summaries of Finney's estimator. The general form of Finney's estimator is

$$\begin{aligned} \widehat{\theta}_{\mathrm{F}} &= \exp(a\widehat{\mu})\cdot \sum_{k=0}^{+\infty} (\Gamma(m/2)/\Gamma(m/2+k))\cdot \left(m(b-a^2d)S^2\right)^k/(k!2^{2k}) \\ &= \exp(a\widehat{\mu}) \cdot \Gamma\left(m/2\right) \cdot \left(-m(b-a^2d)S^2/4\right)^{\frac{1-m/2}{2}} \cdot J_{m/2-1}\left(\sqrt{-m(b-a^2d)S^2}\right), \end{aligned} \tag{6}$$

where $J_\alpha(x) = \sum_{i=0}^{+\infty} \frac{(-1)^i}{i!\Gamma(i+\alpha+1)} \left(\frac{x}{2}\right)^{2i+\alpha}$ denotes the Bessel function of the first kind of order α. In addition, we can simplify the expression of Finney's estimator by defining a function

$$\Psi^F_\omega(t) = \Gamma\left(\omega\right)\left(-\omega t\right)^{\frac{1-\omega}{2}} J_{\omega-1}\left(2\sqrt{-\omega t}\right),$$

and it follows that $\widehat{\theta}_F = \exp(a\widehat{\mu}) \cdot \Psi^F_{m/2}((b-a^2d)S^2/2)$. To include Finney's estimator in the $\widehat{\theta}_\psi$ family, the corresponding ψ can be written as follows.

$$\text{F: } \psi = (2/(bS^2))\cdot \log \Psi^F_{m/2}((b-a^2d)S^2/2). \tag{7}$$

4. Zellner's estimator (Z). Zellner (1971) proposed a conditionally minimal mean squared error estimator of the log-normal mean of an independent sample. This estimator is also a Bayes estimator with respect to a diffuse prior, minimizing a relative quadratic loss function $((\widehat{\theta}_\psi - \theta)/\theta)^2$. We generalize Zellner's estimator to include arbitrary parameter a and b. This estimator belongs to the $\widehat{\theta}_\psi$ family, and ψ is shown below.

$$\text{Z: } \psi = 1 - 3a^2d/b. \tag{8}$$

5. Evans and Shaban's estimator (ES). Evans and Shaban (1974, 1976) considered the estimator $\exp(a\widehat{\mu})g(S^2)$, and calculated the mean squared error $\mathbb{E}[(\exp(a\widehat{\mu})g(S^2)-\theta)^2] = \theta^2\left[\exp(-(b-2a^2d)\sigma^2)\cdot\mathbb{E}(g(S^2) - \exp((b-3a^2d)\sigma^2/2))^2 + 1 - \exp(-a^2d\sigma^2)\right]$. Next, they chose $g(S^2)$ as an unbiased estimator of $\exp((b-3a^2d)\sigma^2/2)$ in order to minimize the MSE, and obtained Evans and Shaban's estimator:

$$\widehat{\theta}_{\text{ES}} = \exp(a\widehat{\mu}) \cdot \Psi^F_{m/2}((b-3a^2d)S^2/2). \tag{9}$$

The corresponding ψ parameter in the $\widehat{\theta}_\psi$ family is

$$\text{ES: } \psi = (2/(bS^2))\cdot \log \Psi^F_{m/2}((b-3a^2d)S^2/2). \tag{10}$$

6. Rukhin's estimator (R). Rukhin (1986) improved Evans and Shaban's estimator by approximately minimizing $\mathbb{E}(g(S^2)-\exp((b-3a^2d)\sigma^2/2))^2$ instead of making $g(S^2)$ an unbiased estimator of $\exp((b-3a^2d)\sigma^2/2)$. Rukhin's estimator using Finney's function Ψ^F (R-F) is shown in (11).

$$\widehat{\theta}_{\text{R-F}} = \exp(a\widehat{\mu}) \cdot \Psi^F_{m/2}\left((m/(m+2))\cdot(b-3a^2d)S^2/2\right), \tag{11}$$

Rukhin (1986) also suggested a simple version (R-S) using an exponential function instead of Finney's function. Both R-F and R-S estimated can be included in the $\widehat{\theta}_\psi$ family, as shown below.

$$\begin{aligned} &\text{R-S: } \psi = (m/(m+2))\cdot(1-3a^2d/b), \\ &\text{R-F: } \psi = (2/(bS^2))\cdot \log \Psi^F_{m/2}\left((m/(m+2))\cdot(b-3a^2d)S^2/2\right). \end{aligned} \tag{12}$$

Meanwhile, Rukhin noticed that $(m/2)^k\cdot(\Gamma(m/2+k)/\Gamma(m/2+2k))\cdot S^{2k}$ is the best estimator under quadratic loss of σ^{2k} when considering all estimators proportional to S^{2k}. Based on this estimator of σ^{2k}, Rukhin provided a locally optimal estimator (R-LO), and we generalize the R-LO estimator for any a and b:

$$\begin{aligned} \widehat{\theta}_{\text{R-LO}} &= \sum_{k=0}^{+\infty}(\Gamma(m/2+k)/\Gamma(m/2+2k))\cdot\left(m(b-3a^2d)S^2\right)^k/(k!2^{2k}) \\ &= \exp(a\widehat{\mu}) \cdot \Psi^R_{m/2}((b-3a^2d)S^2/2), \end{aligned} \tag{13}$$

where Rukhin's function $\Psi_\omega^R(t)$ is defined as

$$\Psi_\omega^R(t) = \sum_{k=0}^{+\infty} (\Gamma(\omega+k)/\Gamma(\omega+2k)) \cdot (\omega t)^k/k!.$$

The ψ parameter of R-LO estimator in the $\widehat{\theta}_\psi$ family is shown below.

$$\text{R-LO: } \psi = (2/(bS^2)) \cdot \log \Psi_{m/2}^R((b-3a^2d)S^2/2). \tag{14}$$

Rukhin (1986) also provided a Bayes estimator for the loss function $(\widehat{\theta}_\psi/\theta - 1)^2$ that was consider by Zellner (1971). Rukhin considered prior densities which are uniform in μ and generalized inversed gamma distributed with parameter ν and γ in σ^2, and got a Bayes estimator (R-B):

$$\widehat{\theta}_{\text{R-B}} = \exp(a\widehat{\mu}) \cdot \left(\frac{\sqrt{\gamma^2 - (b-3a^2d)}}{\gamma} \right)^\nu \cdot \frac{K_\nu\left(\sqrt{m\left(\gamma^2 - (b-3a^2d)\right)S^2}\right)}{K_\nu\left(\sqrt{m\gamma^2 S^2}\right)}, \tag{15}$$

where K_ν is a modified Bessel function of the second kind. Rukhin found $\nu = m/2+2$ by assuming small mS^2, and proposed $\gamma = \sqrt{9(b-3a^2d)/8}$ by assuming large mS^2. This combination provided an estimator $\widehat{\theta}_\psi$ with parameter ψ:

$$\begin{aligned} \text{R-B: } \psi = (2/(bS^2)) \cdot \Big[&\log K_{m/2+2}\left(\sqrt{m\,(b-3a^2d)\,S^2/8}\right) \\ &- \log K_{m/2+2}\left(\sqrt{9m\,(b-3a^2d)\,S^2/8}\right) - (m/2+2)\log 3 \Big]. \end{aligned} \tag{16}$$

7. El-Shaarawi and Viveros' estimator (EV). El-Shaarawi and Viveros (1997) proposed an estimator of the rth moment with respect to X, where $a = r$ and $b = r^2$. Their estimator can be generalized to include any parameter a and b, belonging to the family of $\widehat{\theta}_\psi$ estimator, where

$$\text{EV: } \psi = 1 - a^2d/b - bS^2/(2m) - b^2S^4/(3m^2). \tag{17}$$

8. Zhou's estimator (Zh). Based on Zellner (1971) and Evans and Shaban (1974), Zhou (1998) refined Zellner's conditionally minimal mean squared error estimator of the log-normal mean of an independent sample. Zhou's estimator can be viewed as a slightly modification of Evans and Shaban's estimator, and can be generalized to include arbitrary parameter a and b.

$$\widehat{\theta}_{\text{Zh}} = \exp(a\widehat{\mu}) \cdot \Psi_{m/2}^F((b-4a^2d)S^2/2), \tag{18}$$

This estimator also belongs to the $\widehat{\theta}_\psi$ family, where ψ is

$$\text{Zh: } \psi = (2/(bS^2)) \cdot \log \Psi_{m/2}^F((b-4a^2d)S^2/2). \tag{19}$$

9. Shen and Zhu's estimators (SZ). Shen et al. (2006) and Shen and Zhu (2008) proposed two estimators with an assumption that $d = O(1/n)$ for mean estimation. One estimator approximately minimizes the mean squared error, and the other minimize the bias approximately. Shen and Zhu called the first estimator MM estimator and the second MB estimator. Both MM and MB estimators can be generalized for arbitrary a and b.

$$\begin{aligned} \text{SZ-MM: } \psi &= \frac{m}{m + 2 + 3a^2dm/b + 3bS^2/2}, \\ \text{SZ-MB: } \psi &= \frac{m}{m + a^2dm/b + bS^2/2}. \end{aligned} \tag{20}$$

10. Longford's estimators (L). Longford (2009) derived three estimators of the expectation, median and mode of the lognormal distribution. These estimators aim to be approximately unbiased (UB) or minimize the mean squared error (MS). We generalize Longford's three estimators for the situation with arbitrary a and b.

$$\begin{aligned} \text{L-UB: } \psi &= \left[m/(bS^2)\right] \cdot \left[1 - \exp(-(b - a^2d)S^2/m)\right], \\ \text{L-MS: } \psi &= \frac{m}{bS^2} \cdot \frac{1 - \exp(-(b - 3a^2d)S^2/(m+2))}{2 - \exp(-(b - 3a^2d)S^2/(m+2))}. \end{aligned} \tag{21}$$

11. Fabrizi and Trivisano's estimators (FT). Fabrizi and Trivisano (2012) followed Rukhin (1986)'s method to construct Bayes estimators. Their original estimators included the modified Bessel function of the second kind K_ν. In order to choose hyperparameters, Fabrizi and Trivisano (2012) simplified the expression by assuming S^2 is small. The parameter ψ of this approximated estimator is

$$\text{FT: } \psi = (1 - 3a^2d/b)/(1 + bS^2/m), \tag{22}$$

where the hyperparameters are chosen to minimize the MSE.

12. Gou and Tamhane's estimators (GT). Gou and Tamhane (2017) provided three estimators which can be included in the family of $\widehat{\theta}_\psi$ estimators: (1) Finney's type (GT-F), (2) Evans and Shaban's type (GT-ES), and (3) Rukhin's type (GT-R). Estimator GT-F is the same as L-UB, which is a generalization of Longford (2009)'s approximately unbiased estimator, as shown in (21). Other two types of estimators GT-ES and GT-R are listed below.

$$\begin{aligned} \text{GT-ES: } \psi &= \left[m/bS^2\right] \cdot \left[1 - \exp(-(b - 3a^2d)S^2/m)\right], \\ \text{GT-R: } \psi &= \left[m/bS^2\right] \cdot \left[1 - \exp(-(b - 3a^2d)S^2/(m+2))\right]. \end{aligned} \tag{23}$$

2.2 Estimator expansions and comparisons

We compare nineteen aforementioned estimators under the unified framework with different ψ's, based on their Taylor series of polynomials of degree two for argument $1/m$. In addition, we suppose $d = O(1/m)$, since it is often true that $d = O(1/m)$ in statistical estimation problems. These ψ's expansions are included in Table 1, where the linear term and quadratic

Table 1: Summary of existing estimators and their generalizations: expansions of ψ up to $O(1/m^2)$ terms

	ψ: $1+O(1/m)$	ψ: $O(1/m^2)$
QML	1	–
ML	$1-p/m$	$+p^2/m^2$
SA	$1-a^2d/b$	–
F	$1-a^2d/b-bS^2/(2m)$	$+a^2dS^2/m+bS^2/m^2+2b^2S^4/(3m^2)$
Z	$1-3a^2d/b$	–
ES	$1-3a^2d/b-bS^2/(2m)$	$+3a^2dS^2/m+bS^2/m^2+2b^2S^4/(3m^2)$
R-S	$1-3a^2d/b-2/m$	$+6a^2d/(bm)+4/m^2$
R-F	$1-3a^2d/b-2/m-bS^2/(2m)$	$+4/m^2+6a^2d/(bm)+3a^2dS^2/m$ $+3bS^2/m^2+2b^2S^4/(3m^2)$
R-LO	$1-3a^2d/b-2/m-3bS^2/(2m)$	$+6a^2d/(bm)+9a^2dS^2/m+14/(3m^2)$ $+16bS^2/m^2+17b^2S^4/(3m^2)$
EV	$1-a^2d/b-bS^2/(2m)$	$-b^2S^4/(3m^2)$
Zh	$1-4a^2d/b-bS^2/(2m)$	$+4a^2dS^2/m+bS^2/m^2+2b^2S^4/(3m^2)$
SZ-MM	$1-3a^2d/b-2/m-3bS^2/(2m)$	$+[3a^2d/b+2/m+3bS^2/(2m)]^2$
SZ-MB	$1-a^2d/b-bS^2/(2m)$	$+[a^2d/b+bS^2/(2m)]^2$
L-UB/GT-F	$1-a^2d/b-bS^2/(2m)$	$+a^2dS^2/m+b^2S^4/(6m^2)$
L-MS	$1-3a^2d/b-2/m-3bS^2/(2m)$	$+6a^2d/(bm)+9a^2dS^2/m+4/m^2$ $+6bS^2/m^2+13b^2S^4/(6m^2)$
FT	$1-3a^2d/b-bS^2/m$	$+3a^2dS^2/m+b^2S^4/m^2$
GT-ES	$1-3a^2d/b-bS^2/(2m)$	$+3a^2dS^2/m+b^2S^4/(6m^2)$
GT-R	$1-3a^2d/b-2/m-bS^2/(2m)$	$+6a^2d/(bm)+3a^2dS^2/m+4/(3m^2)$ $+2bS^2/m^2+b^2S^4/(6m^2)$

terms of $1/m$ are listed in two separate columns. The derivations of some ψ's expansions are provided in the appendix.

The expansion of Rukhin's Bayes estimator (R-B) is not included in Table 1, because calculating the expansion of the modified Bessel function of the second kind K_ν requires an additional assumption: mS^2 is small, which is equivalent to $S^2 = o(1/m)$. This assumption is not necessary for all other estimators, so we exclude estimator R-B from the list of expansion. The expansion of parameter ψ in estimator R-B in (16) is $\psi_{\text{R-B}} = 1 - 3a^2d/b - 2/m + o(1/m)$ that shares the same leading terms in all types of Rukhin's estimators, as shown in Table 1.

All estimators in Table 1 can be summarized and categorized by retaining only the linear terms. A general form of ψ has the following expression in (24).

$$\psi = 1 - \frac{\tau_1 \cdot a^2 d}{b} - \frac{\tau_2}{m} - \frac{\tau_3 \cdot bS^2}{2m} + o(1/m), \tag{24}$$

where τ_1, τ_2 and τ_3 are coefficients. For example, Finney's estimator (F) has a set of coefficients $(\tau_1, \tau_2, \tau_3) = (1, 0, 1)$, and Longford's MS estimator (L-MS) has $(\tau_1, \tau_2, \tau_3) = (3, 2, 3)$. All nineteen estimators can be grouped using eleven different (τ_1, τ_2, τ_3) triplets, as shown in Table 2.

Table 2: Model-based summary of existing estimators and their generalizations: $\psi = 1 - \tau_1 \cdot a^2d/b - \tau_2/m - \tau_3 \cdot bS^2/(2m) + o(1/m)$

(τ_1, τ_2, τ_3)	Estimators
$(0, 0, 0)$	QML
$(0, p, 0)$	ML
$(1, 0, 0)$	SA
$(1, 0, 1)$	F, EV, SZ-MB, L-UB/GT-F
$(3, 0, 0)$	Z
$(3, 0, 1)$	ES, GT-ES
$(3, 0, 2)$	FT
$(3, 2, 0)$	R-S, R-B
$(3, 2, 1)$	R-F, GT-R
$(3, 2, 3)$	R-LO, SZ-MM
$(4, 0, 1)$	Zh

2.3 Newly proposed estimators

The unified framework of estimator $\widehat{\theta}_\psi = \exp(a\widehat{\mu} + b\psi S^2/2)$ in (4) includes both simple and complicated forms of ψ for existing lognormal estimators and their generalizations. In this section, we propose nineteen new estimators based on various optimization criteria, where the parameter ψ is assumed to follow a simple linear function of $1/m$. The optimization criteria include: (1) unbiasedness, which requires that $\mathbb{E}[\widehat{\theta}_\psi - \theta] = 0$ holds approximately, (2) minimization of the mean squared error (MSE) which is $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^2]$, (3) generalized unbiasedness with cubic function, which requires that $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^3] = 0$ is approximately true, (4) minimization of the mean fourth power of error (M4E) which is $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^4]$. The details of generalized unbiased estimator and minimized mean fourth power of error estimator are presented in the following sections when introducing estimator sets in (28) and (29).

1. Approximately unbiased estimators (ZG-1/2). Suppose $\psi = \lambda_1 + \lambda_2/m + \lambda_3 S^2/m$, where λ_1, λ_2 and λ_3 are undetermined coefficients. The estimator $\widehat{\theta}_\psi = \exp(a\widehat{\mu} + b\psi S^2/2)$ is expanded and its expected value is calculated. It is unbiased up to $O(1/m)$ terms if

$$\mathbb{E}[\widehat{\theta}_\psi] = \theta + o(1/m)$$

   is satisfied. Based on two different assumptions that $d = O(1/m)$ and $d = O(1)$, we obtain two $\widehat{\theta}_\psi$ estimators, as shown in (25), and the deviation details are included in the appendix.

$$\begin{aligned} &\text{ZG-1: } \psi = 1 - a^2 d/b - bS^2/(2m), \quad \text{when } d = O(1/n), \\ &\text{ZG-2: } \psi = 1 - a^2 d/b - (b - a^2 d)S^2/(2m), \quad \text{when } d = O(1). \end{aligned} \tag{25}$$

2. Approximately minimized mean squared error estimators (ZG-3/4/5/6). The mean squared error (MSE) is a popular measure of the quality of an estimator in the statistical literature. We calculate the expansion of the MSE up to $O(1/m^2)$ terms, and find the estimator $\widehat{\theta}_\psi$ that approximately minimize the leading terms of the MSE. Assume that $\psi = \lambda_1 + \lambda_2/m$. Two estimators are followed under the assumptions that $d = O(1/n)$ and $d = O(1)$. Details of deviations are presented in the appendix.

$$\begin{aligned} &\text{ZG-3: } \psi = 1 - 3a^2 d/b - 2/m - 3bS^2/(2m), \text{ when } d = O(1/n), \\ &\text{ZG-4: } \psi = 1 - 3a^2 d/b - 2(1 - 3a^2 d/b)/m \\ &\qquad\qquad - 3b(1 - 3a^2 d/b)^2 S^2/(2m), \text{ when } d = O(1). \end{aligned} \tag{26}$$

   Instead of assuming $\psi = \lambda_1 + \lambda_2/m$, if we assume that $\psi = \lambda_1 + \lambda_2/m + \lambda_3 S^2/m$, we derive two additional estimators, as shown in (27).

$$\begin{aligned} &\text{ZG-5: } \psi = 1 - 3a^2 d/b - 4/m - 3bS^2/(2m), \text{ when } d = O(1/n), \\ &\text{ZG-6: } \psi = 1 - 3a^2 d/b - 4(1 - 3a^2 d/b)/m \\ &\qquad\qquad - 3b(1 - 3a^2 d/b)^2 S^2/(2m), \text{ when } d = O(1). \end{aligned} \tag{27}$$

3. Approximately cubic-function-induced generalized unbiased estimator (ZG-7/8/9). The regular unbiased estimator satisfies $\mathbb{E}[\widehat{\theta}_\psi] = \theta$ or equivalently $\mathbb{E}[\widehat{\theta}_\psi - \theta] = 0$. The generalized unbiased estimator can be defined in two ways, either

$$\mathbb{E}[\zeta(\widehat{\theta}_\psi)] = \zeta(\theta)$$

   or

$$\mathbb{E}[\zeta(\widehat{\theta}_\psi - \theta)] = 0,$$

where $\zeta(\cdot)$ is a transformation function (Hong, 2017). When $\zeta(\cdot)$ is an identity function, we return to the definition of the regular unbiased estimator. In this article, we consider a cubic transformation, and compute $\mathbb{E}[\widehat{\theta}_{\psi}^{k}]$ $(k = 1, 2, 3)$ up to $O(1/m^2)$ terms, and find estimators approximately satisfy

$$\mathbb{E}[(\widehat{\theta}_{\psi} - \theta)^3] = o(1/m^2).$$

We only consider the estimators under the assumption that $d = O(1/m)$, since there is closed-form solution when assuming $d = O(1)$. Three $\widehat{\theta}_{\psi}$ estimators in (28) are found under different approximations, and the appendix includes more deviation details.

$$\begin{aligned}
&\text{ZG-7: } \psi = 1 - \frac{3a^4d^2m^2 + 2b^3S^2/3 + 3a^2dmb^2S^2 + 3b^4S^4/4}{m\cdot(a^2dmb + b^3S^2/2)}, \text{ when } d = O(1/n),\\
&\text{ZG-8: } \psi = 1 - 6a^2d/b - 4/(3m) - 3bS^2/(2m), \text{ when } d = O(1/n),\\
&\text{ZG-9: } \psi = 1 - 3a^2d/b - bS^2/m\cdot\left(3/2 + 2b/(3a^2dm)\right), \text{ when } d = O(1/n).
\end{aligned} \tag{28}$$

4. Approximately minimized mean fourth power of error estimators (ZG-10/11/12/13/14). Besides the mean squared error (MSE), the mean fourth power of error (M4E) is another option of quality measure of an estimator (Wilson, 1923). We expand $\mathbb{E}[(\widehat{\theta}_{\psi} - \theta)^4]$ up to $O(1/m^3)$ terms and calculate the solutions that reach the minimum M4E approximately. Similarly, we assume $d = O(1/m)$ additional in order to have closed-form solutions. The deviation details for estimators in (29) are provided in the appendix.

$$\begin{aligned}
&\text{ZG-10: } \psi = 1 - 5a^2d/b - 2/m, \text{ when } d = O(1/n),\\
&\text{ZG-11: } \psi = 1 - 10a^2d/b - 10/(3m) - 5bS^2/(2m), \text{ when } d = O(1/n),\\
&\text{ZG-12: } \psi = 1 - \frac{5a^4d^2m^2 + 5b^3S^2/3 + 5a^2dmb^2S^2 + 5b^4S^4/4 + 2a^2dmb}{m\cdot(a^2dmb + b^3S^2/2)}, \text{ when } d = O(1/n),\\
&\text{ZG-13: } \psi = 1 - 5a^2d/b - 2/m - 5bS^2/(2m) - 2b^2S^2/(3a^2dm^2), \text{ when } d = O(1/n),\\
&\text{ZG-14: } \psi = 1 - 5a^2d/b - 10/(3m) - 5bS^2/(2m), \text{ when } d = O(1/n).
\end{aligned} \tag{29}$$

More generally, a quality measure of an estimator can be expressed as

$$\mathbb{E}[\eta(\widehat{\theta}_{\psi} - \theta)], \tag{30}$$

where $\eta(\cdot)$ is a convex function.

5. Estimators with the simple form based on existing estimators (ZG-15/16/17/18/19). Table 2 includes the linear terms of $1/m$ of parameter ψ for all existing estimators and their generalization, which is $\psi = 1 - \tau_1\cdot a^2d/b - \tau_2/m - \tau_3\cdot bS^2/(2m)$. Consider estimators with ψ in this simple form, as represented by (τ_1, τ_2, τ_3) triplets. For example, the SA estimator has this simple form $(\tau_1, \tau_2, \tau_3) = (1, 0, 0)$ and the Z estimator has the form $(\tau_1, \tau_2, \tau_3) = (3, 0, 0)$. We purpose five new estimators based on the leading

terms of ψ's expansions in Table 2, as shown in (31).

$$\begin{aligned}
&\text{ZG-15: } \psi = 1 - 3a^2 d/b - bS^2/(2m), \text{ when } d = O(1/n),\\
&\text{ZG-16: } \psi = 1 - 3a^2 d/b - 2/m, \text{ when } d = O(1/n),\\
&\text{ZG-17: } \psi = 1 - 3a^2 d/b - 2/m - bS^2/(2m), \text{ when } d = O(1/n),\\
&\text{ZG-18: } \psi = 1 - 4a^2 d/b - bS^2/(2m), \text{ when } d = O(1/n),\\
&\text{ZG-19: } \psi = 1 - 3a^2 d/b - bS^2/m, \text{ when } d = O(1/n),
\end{aligned} \tag{31}$$

The new estimator ZG-15 is based on the ES and GT-ES estimators, ZG-16 is based on R-S and R-B, ZG-17 is based on R-F and GT-R, ZG-18 is based on the Zh estimator, and ZG-19 is based on the FT estimator.

3 Numerical comparisons of estimators

In this section we compare all thirty-eight estimators introduced in Section 2. We include five measures of estimator quality into consideration: bias, mean absolute error (MAE), mean squared error (MSE), mean cubic error (MCE), and mean fourth power of error (M4E). The MAE and MCE belong to the general quality measure in (30), where MAE $= \mathbb{E}[|\widehat{\theta}_\psi - \theta|]$ and MCE $= \mathbb{E}[|\widehat{\theta}_\psi - \theta|^3]$. Meanwhile, in order to keep all five measure under the same scale, we report the root mean squared error $\sqrt{\text{MSE}}$, the cube root mean cubic error $\sqrt[3]{\text{MCE}}$ and the fourth root mean fourth power of error $\sqrt[4]{\text{M4E}}$, instead of MSE, MCE and M4E themselves.

Simulated data are independently log-normal distributed with the number of replica 10^9 for each scenario. We apply all estimators one by one to estimate the means of the log-normally distributed random samples with $\mu = 1$, $\sigma^2 \in \{0.25, 1\}$, and sample size $n \in \{5, 20\}$. The simulation setting is chosen since it is unlikely that σ^2 is greater than one in real applications (Shen and Zhu, 2008). Results under different combinations of σ^2 and n are reported in Table 3-6, where the top five $\widehat{\theta}_\psi$ estimators with smallest values of each measure of the estimation quality are shown in bold.

Based on the magnitudes of all five estimator quality measures under four simulation settings, we obtain twenty rank lists summarized in Figure 1. Thirty-eight estimators are sorted based on their average ranks and presented in Figure 1, where the high ranked estimator are in the top and the estimators with low ranks are put in the bottom. The L-MS estimator has the highest average rank, which is 10.65 out of 38. The newly proposed ZG-4 estimator has the second highest average rank, which is 11.50 out of 38. There is a clear separation between the top fourteen estimators and the other twenty-four estimators. The fourteen estimators with high average ranks include eight existing estimators and their generalizations (Rukhin (1986)'s R-S, R-F and R-B, Zhou (1998)'s Zh, Shen and Zhu (2008)'s SZ-MM, Longford (2009)'s L-MS, Fabrizi and Trivisano (2012)'s FT, and Gou and Tamhane (2017)'s GT-R estimators) and six newly proposed estimators (ZG-4/7/9/15/18/19).

4 Empirical illustration

We demonstrate the application of the various estimators under the unified framework, using a logistic regression model introduced in (1) in Section 1. This regression model was obtained

Table 3: Simulated bias, MAE, $\sqrt{\mathrm{MSE}}$, $\sqrt[3]{\mathrm{MCE}}$ and $\sqrt[4]{\mathrm{M4E}}$ for estimating the mean, where the top five $\widehat{\theta}_\psi$ estimators with smallest values of each measure of the estimation quality are shown in bold. The true value $\theta = 3.08022$, the number of replica 10^9, parameter $\mu = 1$ and $\sigma^2 = 0.25$, sample size $n = 5$.

Estimator	bias	MAE	$\sqrt{\mathrm{MSE}}$	$\sqrt[3]{\mathrm{MCE}}$	$\sqrt[4]{\mathrm{M4E}}$
QML	0.09089	0.60309	0.78638	0.97085	1.16614
ML	**0.00798**	0.57656	0.73808	0.89457	1.05756
SA	**0.00798**	0.57656	0.73808	0.89457	1.05756
F	**0.00001**	0.57341	0.73236	0.88517	1.04329
Z	-0.14836	0.55480	0.68890	0.80789	0.92523
ES	-0.15022	0.55464	0.68846	0.80703	0.92383
R-S	-0.19784	0.55677	0.68563	0.79675	0.90365
R-F	-0.19865	0.55679	0.68557	0.79656	0.90328
R-LO	-0.10307	0.55632	0.69633	0.82375	0.95172
R-B	-0.19927	0.55682	**0.68553**	0.79642	0.90300
EV	-0.01355	0.56809	0.72281	0.86963	1.01998
Zh	-0.22256	0.55944	0.68620	**0.79402**	0.89636
SZ-MM	-0.13094	**0.55419**	0.69013	0.81178	0.93245
SZ-MB	0.00834	0.57481	0.73514	0.88954	1.04914
L-UB/GT-F	**-0.00433**	0.57172	0.72932	0.88019	1.03578
L-MS	-0.20138	0.55690	**0.68538**	0.79594	0.90206
FT	-0.16113	**0.55401**	0.68633	0.80273	0.91664
GT-ES	-0.15119	**0.55455**	0.68823	0.80659	0.92311
GT-R	-0.19906	0.55681	0.68554	0.79646	0.90308
ZG-1	-0.01185	0.56889	0.72423	0.87195	1.02352
ZG-2	**-0.00792**	0.57033	0.72680	0.87609	1.02963
ZG-3	-0.37409	0.60311	0.72381	0.81952	0.90335
ZG-4	-0.23031	0.56025	0.68635	**0.79313**	**0.89402**
ZG-5	-0.53213	0.68589	0.80752	0.90010	0.97705
ZG-6	-0.30092	0.57558	0.69765	0.79740	**0.88807**
ZG-7	-0.22266	0.55867	**0.68523**	**0.79279**	0.89477
ZG-8	-0.51186	0.67359	0.79484	0.88746	0.96476
ZG-9	-0.22890	0.56014	0.68642	**0.79340**	0.89450
ZG-10	-0.45947	0.64042	0.75929	0.85085	0.92821
ZG-11	-0.88764	0.95553	1.09015	1.19179	1.27546
ZG-12	-0.54568	0.69835	0.82330	0.91898	0.99890
ZG-13	-0.55030	0.70204	0.82791	0.92450	1.00535
ZG-14	-0.62611	0.75057	0.87682	0.97253	1.05162
ZG-15	-0.16624	**0.55375**	**0.68539**	0.80080	0.91339
ZG-16	-0.32763	0.58354	0.70456	0.80225	**0.88977**
ZG-17	-0.34341	0.58943	0.71009	0.80686	**0.89287**
ZG-18	-0.23911	0.56148	0.68691	**0.79261**	**0.89199**
ZG-19	-0.18378	**0.55398**	**0.68367**	0.79621	0.90499

Table 4: Simulated bias, MAE, $\sqrt{\text{MSE}}$, $\sqrt[3]{\text{MCE}}$ and $\sqrt[4]{\text{M4E}}$ for estimating the mean, where the top five $\widehat{\theta}_\psi$ estimators with smallest values of each measure of the estimation quality are shown in bold. The true value $\theta = 3.08022$, the number of replica 10^9, parameter $\mu = 1$ and $\sigma^2 = 0.25$, sample size $n = 20$.

Estimator	bias	MAE	$\sqrt{\text{MSE}}$	$\sqrt[3]{\text{MCE}}$	$\sqrt[4]{\text{M4E}}$
QML	0.02190	0.29374	0.37148	0.43974	0.50339
ML	0.00232	0.29069	0.36631	0.43211	0.49306
SA	0.00232	0.29069	0.36631	0.43211	0.49306
F	**0.00001**	0.29020	0.36553	0.43097	0.49152
Z	-0.03643	0.28732	0.35954	0.42110	0.47713
ES	-0.03825	0.28712	0.35917	0.42049	0.47626
R-S	-0.06740	0.28727	0.35756	0.41644	0.46912
R-F	-0.06888	0.28723	0.35741	0.41615	0.46866
R-LO	-0.03725	0.28696	0.35902	0.42039	0.47619
R-B	-0.06942	0.28721	0.35736	0.41605	0.46849
EV	**-0.00058**	0.29008	0.36532	0.43067	0.49111
Zh	-0.05720	0.28691	0.35773	0.41737	0.47105
SZ-MM	-0.06212	**0.28675**	**0.35722**	0.41640	0.46948
SZ-MB	0.00066	0.29026	0.36565	0.43115	0.49178
L-UB/GT-F	**-0.00026**	0.29015	0.36544	0.43083	0.49134
L-MS	-0.07227	0.28715	**0.35711**	**0.41554**	0.46765
FT	-0.04116	**0.28681**	0.35859	0.41958	0.47492
GT-ES	-0.03846	0.28709	0.35912	0.42042	0.47615
GT-R	-0.06905	0.28722	0.35739	0.41612	0.46860
ZG-1	**-0.00055**	0.29008	0.36534	0.43069	0.49114
ZG-2	**-0.00041**	0.29011	0.36539	0.43076	0.49124
ZG-3	-0.08497	0.28777	**0.35713**	**0.41465**	0.46556
ZG-4	-0.07668	0.28732	**0.35705**	**0.41515**	0.46683
ZG-5	-0.12446	0.29309	0.36142	0.41700	**0.46516**
ZG-6	-0.11039	0.29087	0.35949	**0.41568**	**0.46473**
ZG-7	-0.04831	**0.28620**	0.35735	0.41754	0.47190
ZG-8	-0.12805	0.29377	0.36205	0.41750	0.46547
ZG-9	-0.04885	**0.28612**	**0.35722**	0.41734	0.47162
ZG-10	-0.11429	0.29181	0.36043	0.41652	**0.46541**
ZG-11	-0.24439	0.33344	0.40217	0.45574	0.50034
ZG-12	-0.13125	0.29399	0.36213	0.41739	**0.46509**
ZG-13	-0.13177	0.29405	0.36218	0.41741	**0.46507**
ZG-14	-0.15392	0.29944	0.36747	0.42213	0.46885
ZG-15	-0.03926	0.28701	0.35896	0.42017	0.47578
ZG-16	-0.07663	0.28772	0.35757	0.41578	0.46759
ZG-17	-0.07942	0.28771	0.35739	**0.41536**	0.46686
ZG-18	-0.05841	0.28684	0.35756	0.41708	0.47060
ZG-19	-0.04208	**0.28672**	0.35841	0.41928	0.47449

Table 5: Simulated bias, MAE, $\sqrt{\text{MSE}}$, $\sqrt[3]{\text{MCE}}$ and $\sqrt[4]{\text{M4E}}$ for estimating the mean, where the top five $\widehat{\theta}_\psi$ estimators with smallest values of each measure of the estimation quality are shown in bold. The true value $\theta = 4.48169$, the number of replica 10^9, parameter $\mu = 1$ and $\sigma^2 = 1$, sample size $n = 5$.

Estimator	bias	MAE	$\sqrt{\text{MSE}}$	$\sqrt[3]{\text{MCE}}$	$\sqrt[4]{\text{M4E}}$
QML	0.85907	2.35398	4.03844	7.58932	18.0216
ML	**0.21233**	1.96057	2.93783	4.60895	7.78279
SA	**0.21233**	1.96057	2.93783	4.60895	7.78279
F	**0.00000**	1.82303	2.54207	3.56506	4.95488
Z	-0.77284	1.64735	2.01623	2.44618	3.05737
ES	-0.81239	**1.64044**	1.99292	2.38691	2.93074
R-S	-1.03303	1.66029	1.96202	2.26032	2.65045
R-F	-1.04904	1.66111	1.95893	2.24933	2.62438
R-LO	-0.60052	1.64777	2.06334	2.57057	3.27073
R-B	-1.05852	1.66205	**1.95780**	2.24419	2.61190
EV	-0.49148	**1.60645**	2.02180	2.50531	3.13074
Zh	-1.16157	1.68542	1.96394	**2.21887**	2.53147
SZ-MM	-0.82491	**1.62394**	1.96393	2.32908	2.81606
SZ-MB	**-0.04607**	1.78116	2.43272	3.30625	4.42965
L-UB/GT-F	**-0.12595**	1.74838	2.35248	3.14927	4.17352
L-MS	-1.09736	1.66655	**1.95426**	**2.22524**	2.56473
FT	-0.98412	**1.63873**	**1.94429**	2.24695	2.63726
GT-ES	-0.83511	**1.63703**	1.98097	2.35691	2.86835
GT-R	-1.05793	1.66166	**1.95739**	**2.24362**	2.61080
ZG-1	-0.34805	1.64255	2.11161	2.68072	3.40878
ZG-2	-0.25237	1.68013	2.19519	2.83672	3.65118
ZG-3	-2.16940	2.31156	2.52857	2.68633	2.81799
ZG-4	-1.30656	1.72892	1.98674	**2.20566**	**2.45257**
ZG-5	-2.57546	2.65137	2.83951	2.97421	3.08209
ZG-6	-1.60351	1.87840	2.10769	2.28402	**2.45528**
ZG-7	-1.78984	2.04327	2.28981	2.47748	2.65008
ZG-8	-2.52754	2.60947	2.80123	2.93865	3.04902
ZG-9	-1.91161	2.14106	2.39305	2.58307	2.75184
ZG-10	-2.10802	2.23197	2.42164	2.55683	2.66850
ZG-11	-3.30263	3.32617	3.45888	3.55334	3.62711
ZG-12	-2.75168	2.81858	3.00975	3.14625	3.25442
ZG-13	-2.79762	2.86269	3.05554	3.19310	3.30178
ZG-14	-2.87290	2.92509	3.10118	3.22682	3.32601
ZG-15	-1.13614	1.67002	**1.94990**	**2.20593**	2.51742
ZG-16	-1.62228	1.88462	2.11041	2.28278	**2.44792**
ZG-17	-1.84578	2.04434	2.26034	2.41947	2.56059
ZG-18	-1.45009	1.80029	2.04576	2.24394	**2.45237**
ZG-19	-1.40942	1.79620	2.05375	2.26633	**2.49534**

Table 6: Simulated bias, MAE, $\sqrt{\text{MSE}}$, $\sqrt[3]{\text{MCE}}$ and $\sqrt[4]{\text{M4E}}$ for estimating the mean, where the top five $\widehat{\theta}_\psi$ estimators with smallest values of each measure of the estimation quality are shown in bold. The true value $\theta = 4.48169$, the number of replica 10^9, parameter $\mu = 1$ and $\sigma^2 = 1$, sample size $n = 20$.

Estimator	bias	MAE	$\sqrt{\text{MSE}}$	$\sqrt[3]{\text{MCE}}$	$\sqrt[4]{\text{M4E}}$
QML	0.17660	1.01768	1.35042	1.70545	2.09520
ML	0.05544	0.98019	1.27923	1.59368	1.94215
SA	0.05544	0.98019	1.27923	1.59368	1.94215
F	**0.00004**	0.95932	1.23967	1.52837	1.84588
Z	-0.17655	0.93541	1.17989	1.42108	1.68712
ES	-0.21854	0.92633	1.16018	1.38507	1.62900
R-S	-0.35477	0.92835	1.14220	1.33677	1.54140
R-F	-0.38765	0.92603	1.13395	1.31887	1.50888
R-LO	-0.25346	**0.91813**	1.14320	1.35452	1.57946
R-B	-0.39896	0.92561	1.13168	1.31362	1.49910
EV	-0.02001	0.95140	1.22460	1.50295	1.80728
Zh	-0.32379	0.92339	1.14003	1.33879	1.54829
SZ-MM	-0.40974	0.91882	**1.12104**	1.29677	1.47209
SZ-MB	**-0.00259**	0.95746	1.23634	1.52275	1.83710
L-UB/GT-F	**-0.00851**	0.95599	1.23333	1.51771	1.82979
L-MS	-0.46058	0.92522	**1.12222**	**1.28952**	1.45246
FT	-0.28400	**0.91549**	1.13491	1.33754	1.55054
GT-ES	-0.22481	0.92489	1.15711	1.37945	1.61975
GT-R	-0.39243	0.92563	1.13267	1.31616	1.50392
ZG-1	**-0.01683**	0.95288	1.22741	1.50780	1.81487
ZG-2	**-0.01325**	0.95414	1.22982	1.51180	1.82083
ZG-3	-0.59291	0.94236	1.12641	**1.27309**	**1.40422**
ZG-4	-0.51020	0.92883	**1.12007**	**1.27811**	**1.42664**
ZG-5	-0.79768	1.01640	1.18994	1.32105	**1.42987**
ZG-6	-0.68936	0.97649	1.15592	1.29511	**1.41516**
ZG-7	-0.42998	**0.90764**	**1.10369**	**1.26992**	1.43015
ZG-8	-0.81584	1.02483	1.19755	1.32762	1.43510
ZG-9	-0.46407	**0.90688**	**1.09836**	**1.25752**	**1.40706**
ZG-10	-0.61277	0.96580	1.15265	1.30302	1.44078
ZG-11	-1.43321	1.46181	1.59118	1.68669	1.76388
ZG-12	-0.94433	1.08710	1.25239	1.37469	1.47354
ZG-13	-0.97277	1.10325	1.26672	1.38736	1.48464
ZG-14	-1.02371	1.13827	1.29953	1.41811	1.51344
ZG-15	-0.24438	0.92109	1.14865	1.36378	1.59404
ZG-16	-0.40669	0.93120	1.13808	1.32143	1.50986
ZG-17	-0.47016	0.93043	1.12761	1.29508	1.45860
ZG-18	-0.35333	0.92019	1.13108	1.32060	1.51634
ZG-19	-0.31069	**0.91134**	1.12515	1.31895	1.51914

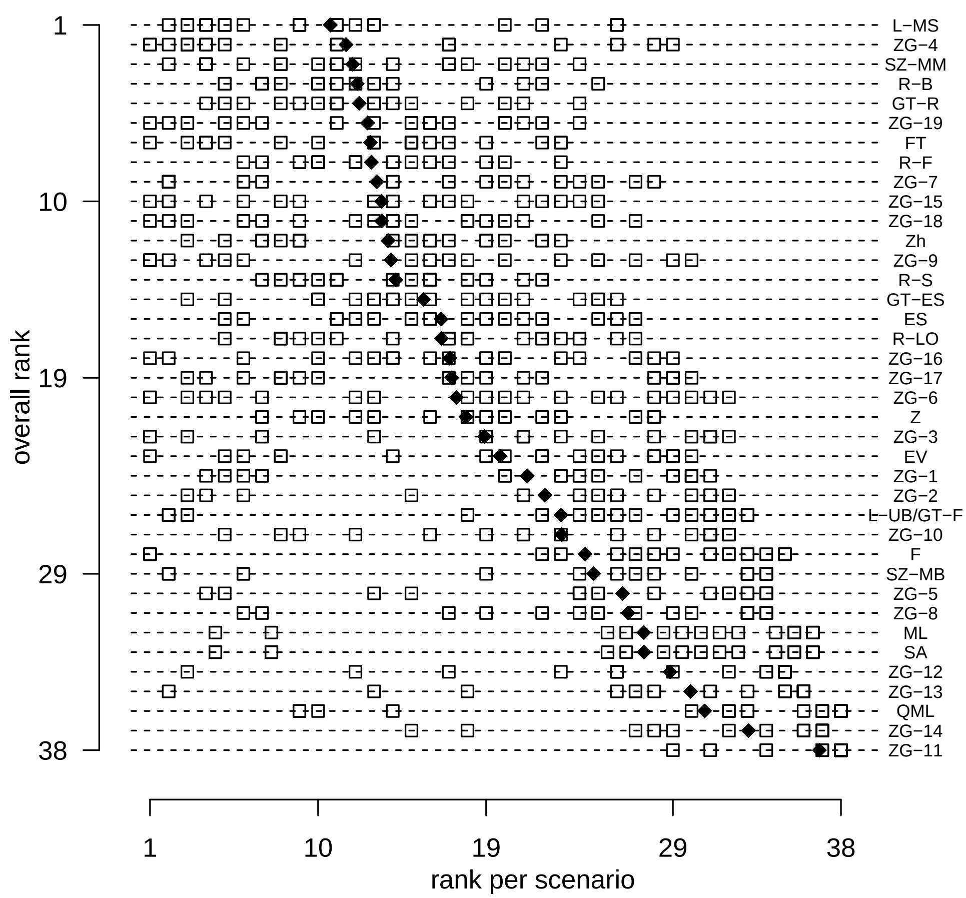

Figure 1: Comparisons of thirty-eight estimators based on summarizing twenty ranking lists. Hollow squares □ represent the ranks of each estimator in various combination of simulation settings and estimator quality measures, and solid diamonds ♦ indicate the average ranks for each estimator.

and employed in Bradu and Mundlak (1970), which is

$$X = e^{\beta_1} U^{\beta_2} e^{\varepsilon},$$

where

$$X = \frac{\text{production of cotton}}{\text{production of wheat}} \text{ in year } i,$$
$$U = \frac{\text{price of cotton}}{\text{price of wheat}} \text{ in year } (i-1),$$

and $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ is the error term. The data are for agriculture of the United States for the period 2008-2015 and are presented in Table 7. Both cotton data and wheat data were published by the Economic Research Service (ERS) in the United States Department of Agriculture (USDA). The cotton data set is available at `www.ers.usda.gov/data-products/cotton-wool-and-textile-data/` with quantities in ten thousand 480-pound bales and prices in dollars per ten pounds. The wheat data set can be downloaded from `catalog.data.gov/dataset/wheat-data` with quantities in million bushels and prices in dollars per bushel.

Table 7: Data for agriculture of the United States (2008-2015)

Products	2008	2009	2010	2011	2012	2013	2014	2015
	Production (cotton: ten thousand 480-pound bales, wheat: million bushels)							
cotton	1282.5	1218.3	1810.2	1557.3	1731.4	1290.9	1631.9	1288.8
wheat	2511.9	2208.9	2163.0	1993.1	2252.3	2135.0	2026.3	2061.9
	Price (cotton: dollars per ten pounds, wheat: dollars per bushel)							
cotton	4.91	6.48	8.46	9.35	7.57	8.38	6.57	6.45
wheat	6.78	4.87	5.70	7.24	7.77	6.87	5.99	4.89

Let $x = \log(X)$ and $u = \log(U)$. The model equation is described in logarithmic form by $x = \beta_1 + \beta_2 u + \varepsilon$, and is estimated for the period 2008-2015:

$$\widehat{x} = -0.4308(0.0345) + 0.6178(0.1351)u; \quad R^2 = 0.807, s^2 = 0.00631; \quad n = 7, m = 5,$$

where the numbers in parentheses are the standard errors of the corresponding regression coefficients.

Next we consider a prediction of X using the estimated model. The observations of productions and prices of cotton and wheat for the period 2016-2017 are shown in Table 8 and were not included in calculating the lognormal regression model.

Predictions provided by thirty-eight $\widehat{\theta}_\psi$ estimators are listed in Table 9. The corresponding biases $\widehat{X} - X_{\text{obs}}$ are also reported. In year 2016, the observed ratio of cotton production to wheat production in the United States is 0.7437, and the predictions range from 0.7670 to 0.7737. In 2017, the observed ratio is 1.2027, and the predictions range from 0.9112 to 0.9349. Note that Rukhin (1986)'s R-B estimator only has a prediction for year 2016. The

Table 8: Data for agriculture of the United States (2016-2017)

Year	cotton production	cotton price	wheat production	wheat price
2016	1717.0	6.97	2308.7	3.89
2017	2092.3	7.10	1739.6	4.72

reason is that the computing of R-B estimation involves evaluating two modified Bessel functions of the second kind, and the calculation requires that $d < b/(3a^2)$. For mean estimation in this illustration, we have $a = 1$ and $b = 1$, with $d = 0.2092 < 1/3$ for year 2016 and $d = 0.7492 > 1/3$ for year 2017.

5 Concluding remarks

A unified framework for analyeses of the lognormal distribution and of lognoral regressions is presented to include existing estimators and their generalizations. Nineteen new estimators are proposed under this unified framework. The efficiencies of all existing and newly derived estimators are evaluated. Meanwhile, we provide estimator rankings as shown in Figure 1. Our list of rankings is consistent with some existing estimator performance reviews. For example, Shen and Zhu (2008) summarized that the performance of the EV estimator is between SZ-MM and SZ-MB. According to this list of rankings, Longford (2009)'s L-MS estimator and our newly purposed ZG-4 estimator are the top two estimators, and the expression of ZG-4 is simpler than that of the L-MS estimator.

With different optimization criteria, various estimators can be obtained under the unified framework. In this article we focus on efficient fitting of lognormal distributions. In addition, estimators' robustness is also important, which can be measured using their breakdown points (BP) or gross error sensitivity (GES) (Serfling, 2002). We will include robustness criteria into consideration, and derive new estimators optimizing some combinations of efficiency and robustness in a separate paper.

Appendix

Derivation of Expansions in Table 1. We include the calculations of expanding ψ in estimator F, ES, R-F, R-LO, Zh, L-UB/GT-F, GT-ES, and GT-R in this proof.

Estimator F: Gou and Tamhane (2017) showed that

$$\Psi^F_{m/2}(t) = \exp\left(t - t^2/m + (8t^3/3 + 2t^2)/m^2 + o(1/m^2)\right).$$

Let $t = bS^2/2 - a^2dS^2/2$, where $bS^2/2$ is an $O(1)$ term and $a^2dS^2/2$ is an $O(1/m)$ term. It follows that

$$\begin{aligned}&\Psi^F_{m/2}(b(1 - a^2d/b)S^2/2)\\ &= \exp\left(bS^2/2\cdot \left[1 - a^2d/b - bS^2/(2m) + a^2dS^2/m + bS^2/m^2 + 2b^2S^4/(3m^2)\right] + o(1/m^2)\right).\end{aligned}$$

Table 9: Predicted ratios of cotton production to wheat production (2016-2017)

	Year 2016		Year 2017	
	Observed: 0.7437		Observed: 1.2027	
Estimator	Prediction	Bias	Prediction	Bias
QML	0.7737	0.0300	0.9349	-0.2678
ML	0.7730	0.0293	0.9341	-0.2686
SA	0.7732	0.0295	0.9327	-0.2700
F	0.7732	0.0295	0.9327	-0.2700
Z	0.7722	0.0285	0.9283	-0.2744
ES	0.7719	0.0282	0.9294	-0.2733
R-S	0.7719	0.0282	0.9294	-0.2733
R-F	0.7719	0.0282	0.9294	-0.2733
R-LO	0.7727	0.0290	0.9325	-0.2702
R-B	0.7719	0.0282	N/A	N/A
EV	0.7732	0.0295	0.9327	-0.2700
Zh	0.7717	0.0280	0.9261	-0.2766
SZ-MM	0.7725	0.0288	0.9328	-0.2699
SZ-MB	0.7733	0.0296	0.9337	-0.2690
L-UB/GT-F	0.7732	0.0295	0.9327	-0.2700
L-MS	0.7719	0.0282	0.9294	-0.2733
FT	0.7722	0.0285	0.9283	-0.2744
GT-ES	0.7722	0.0285	0.9283	-0.2744
GT-R	0.7719	0.0282	0.9294	-0.2733
ZG-1	0.7732	0.0295	0.9327	-0.2700
ZG_2	0.7732	0.0295	0.9327	-0.2700
ZG-3	0.7712	0.0275	0.9272	-0.2756
ZG-4	0.7718	0.0281	0.9298	-0.2729
ZG-5	0.7703	0.0266	0.9260	-0.2767
ZG-6	0.7715	0.0278	0.9313	-0.2715
ZG-7	0.7722	0.0285	0.9283	-0.2744
ZG-8	0.7700	0.0263	0.9210	-0.2817
ZG-9	0.7722	0.0285	0.9283	-0.2744
ZG-10	0.7702	0.0265	0.9228	-0.2799
ZG-11	0.7670	0.0233	0.9112	-0.2915
ZG-12	0.7702	0.0265	0.9228	-0.2799
ZG-13	0.7702	0.0265	0.9228	-0.2799
ZG-14	0.7696	0.0259	0.9220	-0.2807
ZG-15	0.7722	0.0285	0.9283	-0.2744
ZG-16	0.7712	0.0275	0.9272	-0.2755
ZG-17	0.7712	0.0275	0.9272	-0.2755
ZG-18	0.7717	0.0280	0.9261	-0.2766
ZG-19	0.7722	0.0285	0.9283	-0.2744

Estimator ES: Let $t = bS^2/2 - 3a^2dS^2/2$. Similarly, it follows that

$$\Psi^F_{m/2}(b(1-3a^2d/b)S^2/2)$$
$$= \exp\left(bS^2/2\cdot\left[1-3a^2d/b-bS^2/(2m)+3a^2dS^2/m+bS^2/m^2+2b^2S^4/(3m^2)\right]+o(1/m^2)\right).$$

Estimator R-F: Note that in estimator R-F, we have

$$\begin{aligned} t &= (m/(m+2))\cdot\left(b-3a^2d\right)S^2/2 \\ &= bS^2/2 - bS^2/m - 3a^2dS^2/2 + 2bS^2/m^2 + 3a^2dS^2/m + o(1/m^2), \end{aligned}$$

which follows that

$$\begin{aligned} \Psi^F_{m/2}((m/(m+2))\cdot b(1-3a^2d/b)S^2/2) = \exp\big(bS^2/2\cdot\big[1-2/m-3a^2d/b-bS^2/(2m) \\ +4/m^2+6a^2d/(bm)+3a^2dS^2/m+3bS^2/m^2+2b^2S^4/(3m^2)\big]+o(1/m^2)\big). \end{aligned}$$

Estimator R-LO: We defined Rukhin's function by

$$\Psi^R_\omega(t) = \sum\nolimits_{k=0}^{+\infty}(\Gamma(\omega+k)/\Gamma(\omega+2k))\cdot(\omega t)^k/k! = 1+\sum\nolimits_{j=1}^{+\infty}\left[\prod\nolimits_{i=j}^{2j-1}(1+i/\omega)\right]^{-1}\cdot t^j/j!,$$

and consider the expansion of $\Psi^R_{1/x}(t)$ around $x=0$. Direct calculations show that

$$\left.\frac{d}{dx}\Psi^R_{1/x}(t)\right|_{x=0} = -\sum_{j=1}^{+\infty}\frac{3j-1}{2}\cdot\frac{t^j}{(j-1)!} = -\left(\frac{3t^2}{2}+t\right)e^t,$$

and

$$\left.\frac{d^2}{dx^2}\Psi^R_{1/x}(t)\right|_{x=0} = -\sum_{j=1}^{+\infty}\frac{27j^3+10j^2-15j+4}{12}\cdot\frac{t^j}{(j-1)!} = \left(\frac{9t^4}{4}+\frac{43t^3}{3}+17t^2+\frac{7t}{3}\right)e^t.$$

It follows that

$$\Psi^R_{m/2}(t) = \exp\left(t-(3t^2+2t)/m+(68t^3/3+32t^2+14t/3)/m^2+o(1/m^2)\right).$$

Let $t = bS^2/2 - 3a^2dS^2/2$, we have

$$\begin{aligned} \Psi^R_{m/2}(b(1-3a^2d/b)S^2/2) = \exp\big(bS^2/2\cdot\big[1-3a^2d/b-2/m-3bS^2/(2m) \\ +6a^2d/(bm)+9a^2dS^2/m+14/(3m^2)+16bS^2/m^2+17b^2S^4/(3m^2)\big]+o(1/m^2)\big). \end{aligned}$$

Estimator Zh: Here $t = bS^2/2 - 2a^2dS^2$. Direct calculation gives

$$\Psi^F_{m/2}(b(1-4a^2d/b)S^2/2)$$
$$= \exp\left(bS^2/2\cdot\left[1-4a^2d/b-bS^2/(2m)+4a^2dS^2/m+bS^2/m^2+2b^2S^4/(3m^2)\right]+o(1/m^2)\right).$$

Estimator L-UB/GT-F: Gou and Tamhane (2017) gave that

$$\exp\left((m/2)\cdot(1-\exp(-2t/m))\right) = \exp\left(t-t^2/m+2t^3/(3m^2)+o(1/m^2)\right).$$

It follows that

$$\widehat{\theta}_{\text{L-UB}} = \exp\left(bS^2/2\cdot \left[1 - a^2d/b - bS^2/(2m) + a^2dS^2/m + b^2S^4/(6m^2)\right] + o(1/m^2)\right)$$

by plugging $t = bS^2/2 - a^2dS^2/2$.

Estimator L-MS: A direct calculation shows that

$$\exp\left(\frac{m}{2}\cdot\frac{1-\exp(-2t/m)}{2-\exp(-2t/m)}\right) = \exp\left(t - 3t^2/m + 26t^3/(3m^2) + o(1/m^2)\right).$$

Estimator $\widehat{\theta}_{\text{L-MS}}$ can be achieved by plugging in $t = (m/(m+2))\cdot(b-3a^2d)\,S^2/2$, and the corresponding expansion is

$$\begin{aligned}\widehat{\theta}_{\text{L-MS}} = \exp\big(bS^2/2\cdot&\left[1 - 3a^2d/b - 2/m - 3bS^2/(2m)\right.\\ &\left.+6a^2d/(bm) + 9a^2dS^2/m + 4/m^2 + 6bS^2/m^2 + 13b^2S^4/(6m^2)\right] + o(1/m^2)\big).\end{aligned}$$

Estimator GT-ES: With the expansion of $\exp\left((m/2)\cdot(1-\exp(-2t/m))\right)$, let $t = (b-3a^2d)S^2/2$, we have

$$\widehat{\theta}_{\text{GT-ES}} = \exp\left(bS^2/2\cdot \left[1 - 3a^2d/b - bS^2/(2m) + 3a^2dS^2/m + b^2S^4/(6m^2)\right] + o(1/m^2)\right).$$

Estimator GT-R: Similarly, we obtain

$$\begin{aligned}\widehat{\theta}_{\text{GT-R}} = \exp\big(bS^2/2\cdot&\left[1 - 3a^2d/b - 2/m - bS^2/(2m)\right.\\ &\left.+6a^2d/(bm) + 3a^2dS^2/m + 4/(3m^2) + 2bS^2/m^2 + b^2S^4/(6m^2)\right] + o(1/m^2)\big).\end{aligned}$$

This computation is made using the expansion of $\exp\left((m/2)\cdot(1-\exp(-2t/m))\right)$ and letting $t = (m/(m+2))\cdot(b-3a^2d)\,S^2/2$. $\square$

Derivation of Equation (25). Let $\psi = \lambda_0 + \lambda_1/m + \lambda_2 S^2/m$, where λ_i $(i = 0,1,2)$ does not depend on m and σ^2. Note that

$$\begin{aligned}&\mathbb{E}\left[\exp\left(a\widehat{\mu} + b\psi S^2/2\right)\right] = \exp\left(a\mu + a^2d\sigma^2/2\right)\cdot\mathbb{E}\left[\exp\left(bS^2\left(\lambda_0+\lambda_1/m\right)/2 + bS^4\lambda_2/(2m)\right)\right]\\ &= \exp\left(a\mu + a^2d\sigma^2/2\right)\cdot\mathbb{E}\left[\exp\left(bS^2\left(\lambda_0+\lambda_1/m\right)/2\right)\cdot\left(1 + bS^4\lambda_2/(2m) + o(1/m)\right)\right]\\ &= \exp\left(a\mu + a^2d\sigma^2/2\right)\cdot\left(1 - b\sigma^2(\lambda_0+\lambda_1/m)/m\right)^{-m/2}\\ &\quad\cdot\left(1 + b\sigma^4\lambda_2/(2m)\cdot\left(1 - b\sigma^2(\lambda_0+\lambda_1/m)/m\right)^{-2} + o(1/m)\right)\\ &= \exp\left(a\mu + b\sigma^2/2\right)\cdot\left(1 + (a^2d - b(1-\lambda_0))\sigma^2/2 + b\lambda_1\sigma^2/(2m) + b^2\lambda_0\sigma^4/(4m) + o(1/m)\right)\\ &\quad\cdot\left(1 + b\lambda_2\sigma^4/(2m) + o(1/m)\right)\\ &= \theta\cdot\left(1 + a^2d\sigma^2/2 - b(1-\lambda_0)\sigma^2/2 + b\lambda_1\sigma^2/(2m) + b(b\lambda_0 + 2\lambda_2)\sigma^4/(4m) + o(1/m)\right).\end{aligned}$$

Suppose $d = O(1/n)$, and consider $\mathbb{E}[\widehat{\theta}_\psi] = \theta$. It follows that

$$1 - \lambda_0 = 0, \quad a^2d + b\lambda_1/m = 0, \quad b\lambda_0 + 2\lambda_2 = 0,$$

and the solution is $\lambda_0 = 1$, $\lambda_1 = -a^2dm/b$, $\lambda_2 = -b/2$. We then have

$$\psi = 1 - a^2d/b - bS^2/(2m).$$

Another situation is when $d = O(1)$, and we need

$$a^2 d - b(1-\lambda_0) = 0, \quad \lambda_1 = 0, \quad b\lambda_0 + 2\lambda_2 = 0,$$

and the solution is $\lambda_0 = 1 - a^2 d/b$, $\lambda_1 = 0$, $\lambda_2 = -b(1 - a^2 d/b)/2$. It follows that

$$\psi = 1 - a^2 d/b - (b - a^2 d) S^2/(2m).$$

These two $\widehat{\theta}_\psi$'s are approximately unbiased. □

Derivation of Equation (26). Let $\psi = \lambda_0 + \lambda_1/m$, where λ_0 and λ_1 do not depend on m. We first calculate $\mathbb{E}\left[\exp\left(2a\widehat{\mu} + b\psi S^2\right)\right]$.

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(2a\widehat{\mu} + b\psi S^2\right)\right] = \exp\left(2a\mu + 2a^2 d\sigma^2\right) \cdot \mathbb{E}\left[\exp\left(bS^2\left(\lambda_0 + \lambda_1/m\right)\right)\right] \\
&= \exp\left(2a\mu + 2a^2 d\sigma^2\right) \cdot \left(1 - 2b\lambda_0\sigma^2/m - 2b\lambda_1\sigma^2/m^2\right)^{-m/2} \\
&= \theta^2 \cdot \exp\left(b\sigma^2(\lambda_0 - 1) + 2a^2 d\sigma^2 + b\lambda_1\sigma^2/m + b^2\lambda_0^2\sigma^4/m \right. \\
&\qquad \left. + 2b^2\lambda_0\lambda_1\sigma^4/m^2 + 4b^3\lambda_0^3\sigma^6/(3m^2) + o(1/m^2)\right)
\end{aligned}$$

Similarly, we calculate $\mathbb{E}\left[\exp\left(a\widehat{\mu} + b\psi S^2/2\right)\right]$, and the result is

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(a\widehat{\mu} + b\psi S^2/2\right)\right] = \theta \cdot \exp\left(b\sigma^2(\lambda_0 - 1)/2 + a^2 d\sigma^2/2 + b\lambda_1\sigma^2/(2m) + b^2\lambda_0^2\sigma^4/(4m) \right. \\
&\qquad \left. + b^2\lambda_0\lambda_1\sigma^4/(2m^2) + b^3\lambda_0^3\sigma^6/(6m^2) + o(1/m^2)\right)
\end{aligned}$$

The mean squared error (MSE) of $\widehat{\theta}_\psi$ is

$$\begin{aligned}
\mathbb{E}[(\widehat{\theta}_\psi - \theta)^2] = \theta^2 \cdot &\left[1 + \exp\left(b\sigma^2(\lambda_0 - 1) + 2a^2 d\sigma^2 + b\lambda_1\sigma^2/m + b^2\lambda_0^2\sigma^4/m \right.\right. \\
&\quad \left. + 2b^2\lambda_0\lambda_1\sigma^4/m^2 + 4b^3\lambda_0^3\sigma^6/(3m^2)\right) \\
&- 2\exp\left(b\sigma^2(\lambda_0 - 1)/2 + a^2 d\sigma^2/2 + b\lambda_1\sigma^2/(2m) + b^2\lambda_0^2\sigma^4/(4m) \right. \\
&\quad \left.\left. + b^2\lambda_0\lambda_1\sigma^4/(2m^2) + b^3\lambda_0^3\sigma^6/(6m^2)\right) + o(1/m^2)\right]
\end{aligned}$$

Suppose $d = O(1/n)$, and consider the $O(1)$ terms in the expression of $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^2]$, we have

$$\lambda_0 = \arg\min_{\lambda_0} \left[\exp\left(b\sigma^2(\lambda_0 - 1)\right) - 2\exp\left(b\sigma^2(\lambda_0 - 1)/2\right) + 1\right] = 1.$$

With $d = O(1/n)$ and $\lambda_0 = 1$, the MSE is

$$\begin{aligned}
&\mathbb{E}[(\widehat{\theta}_\psi - \theta)^2] = \theta^2 \cdot \left[1 + \exp\left(2a^2 d\sigma^2 + b\lambda_1\sigma^2/m + b^2\lambda_0^2\sigma^4/m + 2b^2\lambda_0\lambda_1\sigma^4/m^2 + 4b^3\lambda_0^3\sigma^6/(3m^2)\right)\right. \\
&\quad \left. - 2\exp\left(a^2 d\sigma^2/2 + b\lambda_1\sigma^2/(2m) + b^2\lambda_0^2\sigma^4/(4m) + b^2\lambda_0\lambda_1\sigma^4/(2m^2) + b^3\lambda_0^3\sigma^6/(6m^2)\right) + o(1/m^2)\right] \\
&= \theta^2 \cdot \left[a^2 d\sigma^2 + b^2\lambda_0^2\sigma^4/(2m) + 7a^4 d^2\sigma^4/4 + 7b^4\lambda_0^4\sigma^8/(16m^2) + 7a^2 d b^2\lambda_0^2\sigma^6/(4m) \right. \\
&\qquad \left. + b^2\sigma^4/m^2 \cdot (\lambda_1^2/4 + \lambda_1(\lambda_0 + 3b\lambda_0^2\sigma^2/4 + 3ma^2 d/(2b))) + o(1/m^2)\right]
\end{aligned}$$

The undetermined coefficient λ_1 appears in $O(1/m^2)$ terms. Omitting these $o(1/m^2)$ terms, choosing

$$\lambda_1 = -2 - 3b\sigma^2/2 - 3ma^2 d/b$$

minimize the MSE. We then have an estimator $\widehat{\theta}_\psi$ with

$$\psi = 1 - 3a^2d/b - 2/m - 3bS^2/(2m)$$

by replacing σ^2 with its estimator S^2.

If $d = O(1)$, we find the solution of λ_0 based on the $O(1)$ terms in the expression of $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^2]$, which is

$$\begin{aligned}\lambda_0 &= \arg\min_{\lambda_0} \left[\exp\left(b\sigma^2(\lambda_0 - 1) + 2a^2d\sigma^2\right) - 2\exp\left(b\sigma^2(\lambda_0-1)/2 + a^2d\sigma^2/2\right) + 1\right] \\ &= 1 - 3a^2d/b.\end{aligned}$$

With $d = O(1)$ and $\lambda_0 = 1 - 3a^2d/b$, the MSE is

$$\begin{aligned}&\mathbb{E}[(\widehat{\theta}_\psi - \theta)^2] = \theta^2 \cdot \left[1 + \exp\left(-a^2d\sigma^2 + b\lambda_1\sigma^2/m + b^2\lambda_0^2\sigma^4/m + 2b^2\lambda_0\lambda_1\sigma^4/m^2 + 4b^3\lambda_0^3\sigma^6/(3m^2)\right)\right. \\ &\quad \left. -2\exp\left(-a^2d\sigma^2 + b\lambda_1\sigma^2/(2m) + b^2\lambda_0^2\sigma^4/(4m) + b^2\lambda_0\lambda_1\sigma^4/(2m^2) + b^3\lambda_0^3\sigma^6/(6m^2)\right) + o(1/m^2)\right] \\ &= \theta^2 \cdot \left[1 + \exp(-a^2d\sigma^2) \cdot \left(\exp(b\sigma^2(\lambda_1 + b\lambda_0^2\sigma^2)/m + b^2\sigma^4\lambda_0(2\lambda_1 + 4b\lambda_0^2\sigma^2/3)/m^2)\right.\right. \\ &\quad \left.\left. -2\exp(b\sigma^2(\lambda_1/2 + b\lambda_0^2/4\sigma^2)/m + b^2\sigma^4\lambda_0(\lambda_1/2 + b\lambda_0^2\sigma^2/6)/m^2)\right) + o(1/m^2)\right] \\ &= \theta^2 \cdot \left[1 + \exp(-a^2d\sigma^2) \cdot \left(-1 + b^2\lambda_0^2\sigma^4/(2m)\right.\right. \\ &\quad \left.\left. + b^2\sigma^4/m^2 \cdot \left(\lambda_1^2/4 + \lambda_0\lambda_1 + 3b\lambda_0^2\sigma^2\lambda_1/4 + b\lambda_0^3\sigma^3 + 7b^2\lambda_0^4\sigma^4/16\right)\right) + o(1/m^2)\right].\end{aligned}$$

Minimizing the MSE up to $O(1/m^2)$ terms, the solution of λ_1 is

$$\lambda_1 = -2\lambda_0 - 3b\lambda_0^2\sigma^2/2.$$

By replacing σ^2 with its estimator S^2, the corresponding ψ is

$$\psi = 1 - 3a^2d/b - 2(1 - 3a^2d/b)/m - 3b(1 - 3a^2d/b)^2S^2/(2m).$$

These two $\widehat{\theta}_\psi$'s approximately minimize the MSE. $\square$

Derivation of Equation (27). Let $\psi = \lambda_0 + \lambda_1/m + \lambda_2 S^2/m$, where λ_i $(i = 0, 1, 2)$ does not depend on m and σ^2. A calculation shows that

$$\begin{aligned}&\mathbb{E}\left[\exp\left(2a\widehat{\mu} + b\psi S^2\right)\right] = \exp\left(2a\mu + b\sigma^2\right) \cdot \exp\left(b\sigma^2(\lambda_0 - 1) + 2a^2d\sigma^2\right. \\ &\quad \left. + b\sigma^2(\lambda_1 + \lambda_2\sigma^2 + b\lambda_0^2\sigma^2)/m + b\sigma^4(2b\lambda_0\lambda_1 + 4b^2\lambda_0^3\sigma^2/3 + 2\lambda_2 + 4b\lambda_0\lambda_2\sigma^2)/m^2 + o(1/m^2)\right).\end{aligned}$$

Similarly, we have

$$\begin{aligned}&\mathbb{E}\left[\exp\left(a\widehat{\mu} + b\psi S^2/2\right)\right] = \exp\left(a\mu + b\sigma^2/2\right) \cdot \exp\left(b\sigma^2(\lambda_0 - 1)/2 + a^2d\sigma^2/2\right. \\ &\quad \left. + b\sigma^2(\lambda_1 + \lambda_2\sigma^2 + b\lambda_0^2\sigma^2/2)/(2m) + b\sigma^4(b\lambda_0\lambda_1/2 + b^2\lambda_0^3\sigma^2/6 + \lambda_2 + b\lambda_0\lambda_2\sigma^2)/m^2 + o(1/m^2)\right).\end{aligned}$$

The MSE of $\widehat{\theta}_\psi$ is followed that

$$\begin{aligned}\mathbb{E}[(\widehat{\theta}_\psi - \theta)^2] = \theta^2 \cdot &\left[1 + \exp\left(b\sigma^2(\lambda_0 - 1) + 2a^2d\sigma^2 + b\sigma^2/m \cdot (\lambda_1 + \lambda_2\sigma^2 + b\lambda_0^2\sigma^2)\right.\right. \\ &\quad \left. + b\sigma^4/m^2 \cdot (2b\lambda_0\lambda_1 + 4b^2\lambda_0^3\sigma^2/3 + 2\lambda_2 + 4b\lambda_0\lambda_2\sigma^2)\right)\end{aligned}$$

$$-2\exp\big(b\sigma^2/2\cdot(\lambda_0-1)+a^2d\sigma^2/2+b\sigma^2/m\cdot(\lambda_1/2+\lambda_2\sigma^2/2+b\lambda_0^2\sigma^2/4)$$
$$+b\sigma^4/m^2\cdot(b\lambda_0\lambda_1/2+b^2\lambda_0^3\sigma^2/6+\lambda_2+b\lambda_0\lambda_2\sigma^2)\big)+o(1/m^2)\big].$$

The coefficient λ_0 can be determined by considering minimizing the $O(1)$ terms of the MSE. We obtain the solution $\lambda_0=1$ when $d=O(1/n)$, and $\lambda_0=1-3a^2d/b$ when $d=O(1)$.

When $d=O(1/n)$ and $\lambda_0=1$, the MSE is

$$\begin{aligned}
\mathbb{E}[(\widehat{\theta}_\psi-\theta)^2] &= \theta^2\cdot\big[1+\exp\big(2a^2d\sigma^2+b\sigma^2/m\cdot(\lambda_1+\lambda_2\sigma^2+b\lambda_0^2\sigma^2)\\
&\quad +b\sigma^4/m^2\cdot(2b\lambda_0\lambda_1+4b^2\lambda_0^3\sigma^2/3+2\lambda_2+4b\lambda_0\lambda_2\sigma^2)\big)\\
&\quad -2\exp\big(a^2d\sigma^2/2+b\sigma^2/m\cdot(\lambda_1/2+\lambda_2\sigma^2/2+b\lambda_0^2\sigma^2/4)\\
&\quad +b\sigma^4/m^2\cdot(b\lambda_0\lambda_1/2+b^2\lambda_0^3\sigma^2/6+\lambda_2+b\lambda_0\lambda_2\sigma^2)\big)+o(1/m^2)\big]\\
&= \theta^2\cdot\big[a^2d\sigma^2+b^2\lambda_0^2\sigma^4/(2m)\\
&\quad +7a^4d^2\sigma^4/4+7b^4\lambda_0^4\sigma^8/(16m^2)+7a^2db^2\lambda_0^2\sigma^6/(4m)+b^3\lambda_0^3\sigma^6/m^2\\
&\quad +b^2\sigma^4/m^2\cdot\big(\lambda_1^2/4+\sigma^4\lambda_2^2/4+\sigma^2\lambda_1\lambda_2/2\\
&\quad +\lambda_1(\lambda_0+3b\lambda_0^2\sigma^2/4+3ma^2d/(2b))+\lambda_2(2\lambda_0\sigma^2+3b\lambda_0^2\sigma^4/4+3ma^2d\sigma^2/(2b))\big)+o(1/m^2)\big]\\
&= \theta^2\cdot\big[a^2d\sigma^2+b^2\lambda_0^2\sigma^4/(2m)\\
&\quad +7a^4d^2\sigma^4/4+7b^4\lambda_0^4\sigma^8/(16m^2)+7a^2db^2\lambda_0^2\sigma^6/(4m)+b^3\lambda_0^3\sigma^6/m^2\\
&\quad +b^2\sigma^4/m^2\cdot\big(\lambda_1^2/4+\lambda_0\lambda_1+3ma^2d\lambda_1/(2b)\\
&\quad +\sigma^2\cdot(\lambda_1\lambda_2/2+3b\lambda_0^2\lambda_1/4+2\lambda_0\lambda_2+3ma^2d\lambda_2/(2b))+\sigma^4\cdot(\lambda_2^2/4+3b\lambda_0^2\lambda_2/4)\big)+o(1/m^2)\big],
\end{aligned}$$

where the $O(1/m^2)$ terms involving coefficient λ_1 and λ_2 are

$$\theta^2\cdot b^2\sigma^4/m^2\cdot\big(\lambda_1^2/4+\lambda_0\lambda_1+3ma^2d\lambda_1/(2b)$$
$$+\sigma^2\cdot(\lambda_1\lambda_2/2+3b\lambda_0^2\lambda_1/4+2\lambda_0\lambda_2+3ma^2d\lambda_2/(2b))+\sigma^4\cdot(\lambda_2^2/4+3b\lambda_0^2\lambda_2/4)\big).$$

Assuming λ_2 does not depend on σ^2, the coefficient λ_2 can be found via

$$\lambda_2=\arg\min_{\lambda_2}\big[\lambda_2^2/4+3b\lambda_0^2\lambda_2/4\big]=-3b\lambda_0^2/2.$$

If we want a coefficient λ_1 that does not depend on σ^2, and approximately minimize the MSE at the same time, we can take

$$\lambda_1=\arg\min_{\lambda_1}\big[\lambda_1^2/4+\lambda_0\lambda_1+3ma^2d\lambda_1/(2b)\big]=-2\lambda_0-3ma^2d/b,$$

and this solution $(\lambda_0,\lambda_1,\lambda_2)=(1,-2-3ma^2d/b,-3b/2)$ will result the estimator $\widehat{\theta}_\psi$ (ZG-3) in (26).

Another solution can be found by solving

$$\begin{cases}(\partial/\partial\lambda_1)(\lambda_1\lambda_2/2+3b\lambda_0^2\lambda_1/4+2\lambda_0\lambda_2+3ma^2d\lambda_2/(2b))=\lambda_2/2+3b\lambda_0^2/4=0,\\ (\partial/\partial\lambda_2)(\lambda_1\lambda_2/2+3b\lambda_0^2\lambda_1/4+2\lambda_0\lambda_2+3ma^2d\lambda_2/(2b))=\lambda_1/2+2\lambda_0+3ma^2d/(2b)=0,\end{cases}$$

and the solution is $(\lambda_1,\lambda_2)=(-4\lambda_0-3ma^2d/b,-3b\lambda_0^2/2)$, and the corresponding parameter ψ is

$$\psi=1-3a^2d/b-4/m-3bS^2/(2m).$$

If we relax the assumption on λ_1, and allow that λ_1 depends on σ^2, then the undetermined coefficient λ_1 can be found via

$$\begin{aligned}\lambda_1 &= \arg\min_{\lambda_1}\left[\lambda_1^2/4+\lambda_0\lambda_1+3ma^2d\lambda_1/(2b)+\sigma^2\cdot(\lambda_1\lambda_2/2+3b\lambda_0^2\lambda_1/4+2\lambda_0\lambda_2+3ma^2d\lambda_2/(2b))\right]\\ &= -2\lambda_0-3ma^2d/b-\lambda_2\sigma^2-3b\lambda_0^2\sigma^2/2.\end{aligned}$$

Since $\lambda_2=-3b\lambda_0^2/2$, it is followed that $\lambda_1=-2\lambda_0-3ma^2d/b$. This gives estimator ZG-3 as well.

When $d=O(1)$ and $\lambda_0=1-3a^2d/b$, the MSE can be calculated as

$$\begin{aligned}\mathbb{E}[(\widehat{\theta}_\psi-\theta)^2] &= \theta^2\cdot\Big[1+\exp\big(-a^2d\sigma^2+b\sigma^2/m\cdot(\lambda_1+\lambda_2\sigma^2+b\lambda_0^2\sigma^2)\\ &\qquad +b^2\sigma^4/m^2\cdot(2\lambda_0\lambda_1+4b\lambda_0^3\sigma^2/3+2\lambda_2/b+4\lambda_0\lambda_2\sigma^2)\big)\\ &\quad -2\exp\big(-a^2d\sigma^2+b\sigma^2/m\cdot(\lambda_1/2+\lambda_2\sigma^2/2+b\lambda_0^2\sigma^2/4)\\ &\qquad +b^2\sigma^4/m^2\cdot(\lambda_0\lambda_1/2+b\lambda_0^3\sigma^2/6+\lambda_2/b+\lambda_0\lambda_2\sigma^2)\big)+o(1/m^2)\Big]\\ &= \theta^2\cdot\Big[1+\exp\big(-a^2d\sigma^2\big)\cdot\big(-1+b^2\lambda_0^2\sigma^4/(2m)+b^3\lambda_0^3\sigma^6/m^2+7b^4\lambda_0^4\sigma^8/(16m^2)\\ &\quad +b^2\sigma^4/m^4\cdot\big[\lambda_0\lambda_1+\lambda_1^2/4+\sigma^2(2\lambda_0\lambda_2+\lambda_1\lambda_2/2+3b\lambda_0^2\lambda_1/4)\\ &\qquad +\sigma^4(\lambda_2^2/4+3b\lambda_0\lambda_2/4)\big]\big)+o(1/m^2)\Big].\end{aligned}$$

Consider these $O(1/m^2)$ terms involving in λ_1 and λ_2 in the equation above, which is

$$\lambda_0\lambda_1+\lambda_1^2/4+\sigma^2(2\lambda_0\lambda_2+\lambda_1\lambda_2/2+3b\lambda_0^2\lambda_1/4)+\sigma^4(\lambda_2^2/4+3b\lambda_0\lambda_2/4).$$

Since λ_1 and λ_2 do not depend on σ^2, a solution that minimizes the MSE needs to satisfy some conditions below.

$$\begin{cases}\text{const:} & (\partial/\partial\lambda_1)\,(\lambda_0\lambda_1+\lambda_1^2/4)=\lambda_1/2+\lambda_0=0\Longrightarrow\lambda_1=-2\lambda_0,\\ \sigma^2\text{:} & (\partial/\partial\lambda_1)\,(2\lambda_0\lambda_2+\lambda_1\lambda_2/2+3b\lambda_0^2\lambda_1/4)=\lambda_2/2+3b\lambda_0^2/4=0\Longrightarrow\lambda_2=-3b\lambda_0^2/2,\\ & (\partial/\partial\lambda_2)\,(2\lambda_0\lambda_2+\lambda_1\lambda_2/2+3b\lambda_0^2\lambda_1/4)=2\lambda_0+\lambda_1/2=0\Longrightarrow\lambda_1=-4\lambda_0,\\ \sigma^4\text{:} & (\partial/\partial\lambda_2)\,(\lambda_2^2/4+3b\lambda_0\lambda_2/4)=\lambda_2/2+3b\lambda_0^2/4=0\Longrightarrow\lambda_2=-3b\lambda_0^2/2.\end{cases}$$

There is no solution that satisfies all four equations above. If the constant term is more important comparing with the σ^2 term, we can choose $(\lambda_1,\lambda_2)=(-2\lambda_0,-3b\lambda_0^2/2)$, and we will have ZG-4 estimator in (26). Another situation is the σ^2 term is more important than the constant term, we then use $(\lambda_1,\lambda_2)=(-4\lambda_0,-3b\lambda_0^2/2)$, and have a new estimator $\widehat{\theta}_\psi$ with parameter ψ which is

$$\psi=1-3a^2d/b-4(1-3a^2d/b)/m-3b(1-3a^2d/b)^2S^2/(2m),$$

where $\lambda_0=1-3a^2d/b$ and $d=O(1)$. □

Derivation of Equation (28). Assume that $\psi=\lambda_0+\lambda_1/m$ where λ_0 and λ_1 do not depend on m, and λ_0 does not depend on σ^2. For any integer k, we have

$$\mathbb{E}\left[\exp\left(ka\widehat{\mu}+kb\psi S^2/2\right)\right]=\exp(ka\mu+k^2a^2d\sigma^2/2)\cdot\left(1-kb(\lambda_0+\lambda_1/m)\sigma^2/m\right)^{-m/2}$$

$$= \exp(ka\mu + k^2a^2d\sigma^2/2)\cdot \exp\big(kb\sigma^2\lambda_0/2 + kb\sigma^2\lambda_1/(2m) + k^2b^2\sigma^4\lambda_0^2/(4m)$$
$$+k^2b^2\sigma^4\lambda_0\lambda_1/(2m^2) + k^3b^3\sigma^6\lambda_0^3/(6m^2) + o(1/m^2)\big) \tag{32}$$

When $d = O(1/n)$, the $O(1)$ term of $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^3]$ is

$$\theta^3\cdot \big(\exp(3b\sigma^2(\lambda_0 - 1)/2) - 3\exp(b\sigma^2(\lambda_0 - 1)) + 3\exp(b\sigma^2(\lambda_0 - 1)/2) - 1\big)\,,$$

which equals zero when $\lambda_0 = 1$. It follows that

$$\mathbb{E}\left[\exp\big(ka\widehat{\mu} + kb\psi S^2/2\big)\right] = \theta^3\cdot \big[1 + b\sigma^2/m\cdot \big(k^2a^2dm/(2b) + k\lambda_1/2\big) + b^2\sigma^4/m\cdot \big(k^2/4\big)$$
$$+b^2\sigma^4/m^2\cdot \big(k^2\lambda_1/2 + k^2\lambda_1^2/8 + k^3a^2dm\lambda_1/(4b) + k^4a^4d^2m^2/(8b^2)\big)$$
$$+b^3\sigma^6/m^2\cdot \big(k^3/6 + k^3\lambda_1/8 + k^4a^2dm/(8b)\big) + b^4\sigma^8/m^2\cdot \big(k^4/32\big) + o(1/m^2)\big]\,. \tag{33}$$

We compute $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^3]$ up to $O(1/m^2)$ terms, which is

$$\mathbb{E}[(\widehat{\theta}_\psi - \theta)^3] = \theta^3\cdot \big[b^2\sigma^4/m^2\cdot \big(9a^4d^2m^2/(2b^2) + 3a^2dm\lambda_1/(2b)\big)$$
$$+b^3\sigma^6/m^2\cdot \big(1 + 9a^2dm/(2b) + 3\lambda_1/4\big) + b^4\sigma^8/m^2\cdot (9/8) + o(1/m^2)\big]\,.$$

Solving $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^3] = 0$ without considering $o(1/m^2)$ terms to find λ_1, we have

$$\lambda_1 = -\frac{9a^4d^2m^2 + 2b^3\sigma^2 + 9a^2dmb^2\sigma^2 + 9b^4\sigma^4/4}{3a^2dmb + 3b^3\sigma^2/2},$$

and an estimator $\widehat{\theta}_\psi$ is followed with

$$\psi = 1 - \frac{3a^4d^2m^2 + 2b^3S^2/3 + 3a^2dmb^2S^2 + 3b^4S^4/4}{m\cdot(a^2dmb + b^3S^2/2)},$$

where σ^2 is replaced by its estimator S^2.

In addition, when σ^2 is small, we can approximate the solution of ψ in the equation above using the form $1 + \lambda_1/m + \lambda_2S^2/m$. This approximate solution is

$$\psi = 1 - 3a^2d/b - bS^2/m\cdot \big(3/2 + 2b/(3a^2dm)\big)\,,$$

where the term involving S^{2k} $(k \geq 2)$ are omitted.

When σ^2 is not very small, we can only keep the terms with high exponents of σ in the expression of $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^3]$, and solve

$$\theta^3\cdot \big[b^3\sigma^6/m^2\cdot \big(1 + 9a^2dm/(2b) + 3\lambda_1/4\big) + b^4\sigma^8/m^2\cdot (9/8)\big] = 0,$$

and the solution is

$$\lambda_1 = -4/3 - 6a^2dm/b - 3b\sigma^2/2,$$

and the corresponding estimator $\widehat{\theta}_\psi$ is

$$\psi = 1 - 6a^2d/b - 4/(3m) - 3bS^2/(2m),$$

using S^2 to replace σ^2. □

Derivation of Equation (29). Let $\psi = \lambda_0 + \lambda_1/m$ where λ_0 and λ_1 are $O(1)$ terms, and λ_0 does not depend on σ^2. When calculating $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^4]$, we need to compute $\mathbb{E}[\widehat{\theta}_\psi^4]$ until $O(1/m^3)$ terms. So we extend formula (32) for any integer k and have

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(ka\widehat{\mu} + kb\psi S^2/2\right)\right] = \exp(ka\mu + k^2a^2d\sigma^2/2)\\
&\quad\cdot\exp\big(kb\sigma^2\lambda_0/2 + kb\sigma^2\lambda_1/(2m) + k^2b^2\sigma^4\lambda_0^2/(4m) + k^2b^2\sigma^4\lambda_0\lambda_1/(2m^2) + k^3b^3\sigma^6\lambda_0^3/(6m^2)\\
&\qquad + k^2b^2\lambda_1^2\sigma^4/(4m^3) + k^3b^3\lambda_0^2\lambda_1\sigma^6/(2m^3) + k^4b^4\lambda_0^4\sigma^8/(8m^3) + o(1/m^3)\big)
\end{aligned}\tag{34}$$

Assuming $d = O(1/n)$, the $O(1)$ term of $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^4]$ reaches its minimum value zero when $\lambda_0 = 1$, and $\mathbb{E}\left[\exp\left(ka\widehat{\mu} + kb\psi S^2/2\right)\right]$ in (34) becomes

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(ka\widehat{\mu} + kb\psi S^2/2\right)\right] = \theta^k\cdot\big[1 + k^2a^2d\sigma^2/2 + b\sigma^2/m\cdot(k\lambda_1/2 + k^2b\sigma^2/4)\\
&\quad + b^2\sigma^4/m^2\cdot(k^2\lambda_1/2 + k^3b\sigma^2/6) + k^4a^4d^2\sigma^4/8 + b^2\sigma^4/(2m^2)\cdot(k\lambda_1/2 + k^2b\sigma^2/4)^2\\
&\quad + k^2a^2db\sigma^4/(2m^2)\cdot(k\lambda_1/2 + k^2b\sigma^2/4) + b^3\sigma^6/m^3\cdot(k^2\lambda_1^2/(4b\sigma^2) + k^3\lambda_1/2 + k^4b^2\sigma^4/8)\\
&\quad + k^2a^2db^2\sigma^6/(2m^2)\cdot(k^2\lambda_1/2 + k^3b\sigma^2/6) + b^3\sigma^6/m^3\cdot(k\lambda_1/2 + k^2b\sigma^2/4)\cdot(k^2\lambda_1/2 + k^3b\sigma^2/6)\\
&\quad + k^6a^6d^3\sigma^6/48 + k^4a^4d^2b\sigma^6/(8m)\cdot(k\lambda_1/2 + k^2b\sigma^2/4)\\
&\quad + k^2a^2db^2\sigma^6/(4m^2)\cdot(k\lambda_1/2 + k^2b\sigma^2/4)^2 + b^3\sigma^6/(6m^3)\cdot(k\lambda_1/2 + k^2b\sigma^2/4)^3 + o(1/m^3)\big]\,.
\end{aligned}$$

Next, we compute $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^4]$ up to $O(1/m^3)$ terms with $d = O(1/n)$ and $\lambda_0 = 1$, which is

$$\begin{aligned}
&\mathbb{E}[(\widehat{\theta}_\psi - \theta)^4] = \theta^4\cdot\big[3a^4d^2\sigma^4 + 3b^4\sigma^8/(4m^2) + 3a^2db^2\sigma^6/(4m) + 13b^5\sigma^{10}/m^3 + 65b^6\sigma^{12}/(16m^3)\\
&\quad + 20a^2db^3\sigma^8/m^2 + 195a^2db^4\sigma^{10}/(8m^2) + 195a^4d^2b^2\sigma^8/(4m) + 65a^6d^3\sigma^6/2\\
&\quad + b^3\sigma^6/m^3\cdot\big(5b\sigma^2\lambda_1 + 15b^2\sigma^4\lambda_1/4 + 6a^2dm\lambda_1/b + 15a^2dm\sigma^2\lambda_1\\
&\qquad + 15a^4d^2m^2\lambda_1/b^2 + 3b\sigma^2\lambda_1^2/4 + 3a^2dm\lambda_1^2/(2b)\big) + o(1/m^3)\big]\,.
\end{aligned}$$

Note that $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^4]$ equals a quadratic function of λ_1 plus $o(1/m^3)$ terms. The sum of all linear and quadratic terms of λ_1 in $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^4]$ is proportional to

$$\begin{aligned}
&5b\sigma^2\lambda_1 + 15b^2\sigma^4\lambda_1/4 + 6a^2dm\lambda_1/b + 15a^2dm\sigma^2\lambda_1 + 15a^4d^2m^2\lambda_1/b^2\\
&\quad + 3b\sigma^2\lambda_1^2/4 + 3a^2dm\lambda_1^2/(2b).
\end{aligned}\tag{35}$$

When σ^2 is small, we can only keep $3a^2dm\lambda_1^2/(2b) + 6a^2dm\lambda_1/b + 15a^2dm\sigma^2\lambda_1$ in (35), and the minimum value of $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^4]$ is approximately achieved at $\lambda_1 = -2 - 5a^2dm/b$, and the corresponding ψ parameter in $\widehat{\theta}_\psi$ is

$$\psi = 1 - 2/m - 5a^2d/b.$$

When σ^2 is large, we keep $3b\sigma^2\lambda_1^2/4 + 5b\sigma^2\lambda_1 + 15b^2\sigma^4\lambda_1/4 + 15a^2dm\sigma^2\lambda_1$ in (35). The approximate solution to achieve the minimum value of $\mathbb{E}[(\widehat{\theta}_\psi - \theta)^4]$ is $\lambda_1 = -10/3 - 5b\sigma^2/2 - 10a^2dm/b$. It is followed that the ψ parameter in $\widehat{\theta}_\psi$ is

$$\psi = 1 - 10/(3m) - 10a^2d/b - 5bS^2/(2m),$$

where σ^2 is substituted by its estimator S^2.

A relatively more complex estimator can be found by directly minimizing the expression in (35), and the solution of λ_1 is

$$\lambda_1 = -\frac{5b\sigma^2/3 + 5b^2\sigma^4/4 + 2a^2dm/b + 5a^2dm\sigma^2 + 5a^4d^2m^2/b^2}{b\sigma^2/2 + a^2dm/b}, \tag{36}$$

and the corresponding estimator $\widehat{\theta}_\psi$ can be found using $\psi = 1 + \lambda_1/m$ and replacing σ^2 with S^2.

The expression of λ_1 in (36) can be simplified. We first assume σ^2 is small. It follows that

$$\begin{aligned}\lambda_1 &= -\frac{5b^2\sigma^2/(3a^2dm) + 5b^3\sigma^4/(4a^2dm) + 2 + 5b\sigma^2 + 5a^2dm/b}{1 + b^2\sigma^2/(2a^2dm)}\\ &\approx -2 - 5a^2dm/b - 5b\sigma^2/2 + 2b^2\sigma^2/(3a^2dm).\end{aligned}$$

Replacing σ^2 with S^2, the corresponding estimator $\widehat{\theta}_\psi$ is

$$\psi = 1 - 2/m - 5a^2d/b - 5bS^2/(2m) - 2b^2S^2/(3a^2dm^2).$$

Similarly, if σ^2 is small, it follows that

$$\begin{aligned}\lambda_1 &= -\frac{10/3 + 5b\sigma^2/2 + 4a^2dm/(b^2\sigma^2) + 10a^2dm/b + 10a^4b^2m^2/(b^3\sigma^2)}{1 + 2a^2dm/(b^2\sigma^2)}\\ &\approx -10/3 - 5b\sigma^2/2 - 5a^2dm/b,\end{aligned}$$

and

$$\psi = 1 - 10/(3m) - 5a^2d/b - 5bS^2/(2m),$$

is the corresponding paramter ψ of estimator $\widehat{\theta}_\psi$. □

References

Aitchison, J. and Brown, J. A. C. (1957). *The Lognormal Distribution*. Cambridge University Press, Cambridge, England.

Bachelier, L. (1900). Théorie de la spéculation. *Annales Scientifiques de l'École Normale Supérieure*, 3(17):21–86.

Basu, D. (1955). On statistics independent of a complete sufficient statistic. *Sankhya*, 15(4):377–380.

Baur, P. (1997). Lognormal distribution of water permeability and organic solute mobility in plant cuticles. *Plant, Cell & Environment*, 20(2):167–177.

Biondini, R. (1976). Cloud motion and rainfall statistics. *Journal of Applied Meteorology*, 15(3):205–224.

Bradu, D. and Mundlak, Y. (1970). Estimation in lognormal linear models. *Journal of the American Statistical Association*, 65(329):198–211.

Cohn, T. A., Delong, L. L., Gilroy, E. J., Hirsch, R. M., and Wells, D. K. (1989). Estimating constituent loads. *Water Resources Research*, 25(5):937–942.

Crow, E. L. and Shimizu, K. (1988). *Lognormal Distribution: Theory and Applications*. Marcel Dekker, New York.

de Valk, C. and Cai, J.-J. (2018). A high quantile estimator based on the log-generalized Weibull tail limit. *Econometrics and Statistics*, 6:107–128.

Dhrymes, P. J. (1962). On devising unbiased estimators for the parameters of the Cobb-Douglas production function. *Econometrica*, 30(2):297–304.

El-Shaarawi, A. H. and Viveros, R. (1997). Inference about the mean in log-regression with environmental applications. *Environmetrics*, 8(5):569–582.

Evans, I. G. and Shaban, S. A. (1974). A note on estimation in lognormal models. *Journal of the American Statistical Association*, 69(347):779–781.

Evans, I. G. and Shaban, S. A. (1976). New estimators (of smaller MSE) for parameters of a lognormal distribution. *Biometrische Zeitschrift*, 18:453–466.

Fabrizi, E. and Trivisano, C. (2012). Bayesian estimation of log-normal means with finite quadratic expected loss. *Bayesian Analysis*, 7(4):975–996.

Finney, D. J. (1941). On the distribution of a variate whose logarithm is normally distributed. *Supplement to the Journal of the Royal Statistical Society*, 7(2):155–161.

Fontenla, J. M., Curdt, W., Avrett, E. H., and Harder, J. (2007). Log-normal intensity distribution of the quiet-Sun FUV continuum observed by SUMER. *Astronomy and Astrophysics*, 468(2):695–699.

Goldberger, A. S. (1968). The interpretation and estimation of Cobb-Douglas functions. *Econometrica*, 368:464–472.

Gou, J. and Tamhane, A. C. (2017). Estimation of a parametric function associated with the lognormal distribution. *Communications in Statistics - Theory and Methods*, 46(16):8134–8154.

Heien, D. M. (1968). A note on log-linear regression. *Journal of the American Statistical Association*, 63(323):1034–1038.

Herdan, G. (1958). The relation between the dictionary distribution and the occurrence distribution of word length and its importance for the study of quantitative linguistics. *Biometrika*, 45(1-2):222–228.

Hong, Y. (2017). *Probability and Statistics for Economists*. World Scientific Publishing Co., New Jersey.

Kennedy, P. (1981). Estimation with correctly interpreted dummy variables in semilogarithmic equations. *American Economic Review*, 71(4):801.

Koch, A. L. (1966). The logarithm in biology 1. mechanisms generating the log-normal distribution exactly. *Journal of Theoretical Biology*, 12(2):276–290.

Kondo, K. (1977). The log-normal distribution of the incubation time of exogenous diseases. *Japanese Journal of Human Genetics*, 21:217–237.

Krige, D. G. (1966). A study of gold and uranium distribution patterns in the Klerksdorp gold field. *Geoexploration*, 4(1):43–53.

Laurent, A. G. (1963). The lognormal distribution and the translation method: Description and estimation problems. *Journal of the American Statistical Association*, 58(301):231–235.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. Wiley, New York.

Limpert, E., Stahel, W. A., and Abbt, M. (2001). Log-normal distributions across the sciences: keys and clues. *BioScience*, 51(5):341–352.

Longford, N. T. (2009). Inference with the lognormal distribution. *Journal of Statistical Planning and Inference*, 139(7):2329–2340.

Loper, J. E., Suslow, T. V., and Schroth, M. N. (1984). Log-normal distribution of bacterial populations in the rhizosphere. *Phytopathology*, 74:1454—-1460.

Marwell, G., Oliver, P. E., and Prahl, R. (1988). Social networks and collective action: A theory of the critical mass. III. *American Journal of Sociology*, 94(3):502–534.

Meulenberg, M. T. G. (1965). On the estimation of an exponential function. *Econometrica*, 33:863–868.

Packard, G. C. (2013). Is logarithmic transformation necessary in allometry? *Biological Journal of the Linnean Society*, 109(2):476–486.

Pirjol, D. (2011). Phase transition in a log-normal Markov functional model. *Journal of Mathematical Physics*, 52(1):013301.

Preston, F. W. (1981). Pseudo-lognormal distributions. *Ecology*, 62(2):355–364.

Rukhin, A. L. (1986). Improved estimation in lognormal models. *Journal of the American Statistical Association*, 81(396):1046–1049.

Rukhin, A. L. (1988). Improved estimation in lognormal regression models. *Journal of Statistical Planning and Inference*, 18(3):291–297.

Serfling, R. (2002). Efficient and robust fitting of lognormal distributions. *North American Actuarial Journal*, 6(4):95–109.

Shen, H., Brown, L. D., and Zhi, H. (2006). Efficient estimation of log-normal means with application to pharmacokinetic data. *Statistics in medicine*, 25(17):3023–3038.

Shen, H. and Zhu, Z. (2008). Efficient mean estimation in log-normal linear models. *Journal of Statistical Planning and Inference*, 138(3):552–567.

Siano, D. B. (1972). The log-normal distribution function. *Journal of Chemical Education*, 49(11):755–757.

Smith, R. J. (1993). Logarithmic transformation bias in allometry. *American Journal of Physical Anthropology*, 90(2):215–228.

Ulrich, R. and Miller, J. (1993). Information processing models generating lognormally distributed reaction times. *Journal of Mathematical Psychology*, 37(4):513–525.

Wilson, E. B. (1923). First and second laws of error. *Journal of the American Statistical Association*, 18(143):841–851.

Zellner, A. (1971). Bayesian and non-Bayesian analysis of the log-normal distribution and log-normal regression. *Journal of the American Statistical Association*, 66(334):327–330.

Zhang, F., Tapera, T. M., and Gou, J. (2018). Application of a new dietary pattern analysis method in nutritional epidemiology. *BMC Medical Research Methodology*, 18(1):119.

Zhou, X.-H. (1998). Estimation of the log-normal mean. *Statistics in Medicine*, 17(19):2251–2264.