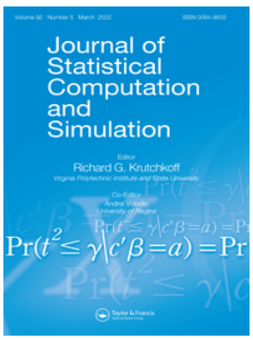


Taylor & Francis
Taylor & Francis Group

Journal of Statistical Computation and Simulation

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/gscs20

On estimating the boundaries of a uniform distribution under additive measurement errors

Ali A. Al-Sharadqah & Alexandre G. Patriota

To cite this article: Ali A. Al-Sharadqah & Alexandre G. Patriota (2022): On estimating the boundaries of a uniform distribution under additive measurement errors, Journal of Statistical Computation and Simulation, DOI: 10.1080/00949655.2021.2022149

To link to this article: https://doi.org/10.1080/00949655.2021.2022149

JOURNAL OF STATISTICAL COMPUTATION AND SIMULATION
https://doi.org/10.1080/00949655.2021.2022149

On estimating the boundaries of a uniform distribution under additive measurement errors

Ali A. Al-Sharadqah[a] and Alexandre G. Patriota[b]

[a]Department of Mathematics, California State University at Northridge, Northridge, CA, USA; [b]Institute of Mathematics and Statistics, The University of São Paulo, São Paulo, Brazil

ABSTRACT

This paper presents the method-of-moments (MM) and the maximum likelihood (ML) estimators for the width of a uniform random variable U when measured with additive errors. We study situations where the support is either the symmetric interval $[-a, a]$ or asymmetric interval $[a, b]$. Both MM and ML estimators for the boundaries are discussed in two cases, namely, when the noise level σ^2 is known and when it is unknown. While the MM estimators have a closed form, the ML estimators need an iterative algorithm to solve a system of nonlinear equations. A reliable algorithm to compute the ML estimators is proposed. We also establish sufficient conditions for the existence of the MM estimators and their asymptotic normality. Numerical experiments on real and synthetic data are presented to validate our results.

ARTICLE HISTORY
Received 3 January 2021
Accepted 19 December 2021

KEYWORDS
Additive-error model; asymptotic distribution; delta method; the method-of-moments estimator; the maximum likelihood estimator

1. Introduction

In image and signal processing, it is common to measure random variables with noise [1,2]. It is widely known that neglecting measurement errors can lead to erroneous conclusions [3,4]. This is known as the measurement-error problem [5]. In this paper, we study the problem of estimating the boundaries of a uniform random variable that is measured imprecisely along a line segment. This inferential problem is of great interest in edge detection [6–9].

Schneeweiss [4] derives the maximum likelihood (ML) and the method-of-moment (MM) estimators for the boundary of a uniform random variable U with support $[0, a]$ when it is measured imprecisely. The additive structure to represent the relation between the unobservable variable U and the observable one X is employed, namely, $X = U + E$, where E is a white noise with zero mean and variance σ^2. Although the ML outperforms the MM, the former is computationally cumbersome even under this simple model. Furthermore, the ML tends to diverge when observations are measured with moderate noise levels σ [4].

CONTACT Ali A. Al-Sharadqah ali.alsharadqah@csun.edu Department of Mathematics, California State University at Northridge, Northridge, CA 91330, USA

Supplemental data for this article can be accessed here. https://doi.org/10.1080/00949655.2021.2022149

Benšić and Sabo [10] consider the uniform distribution on a symmetric line segment $[-a, a]$, where $a > 0$. They derive the ML and MM estimators when the noise level σ^2 is known. They only establish the asymptotic distribution of the ML estimator when σ^2 is known. The asymptotic distribution of the MM estimator was not derived. Later, they extend their results to the situation when the noise level σ^2 is unknown [11] and to the problem of estimating the border of a disc [10,11] but without exploring their asymptotic distributions. Then their work is extended later to estimate the width of uniform distribution when corrupted with Laplace additive error noise [3] supported by their R package [12] which is devoted to estimate borders of uniform distribution on intervals as well as on elliptical domains.

The problem of determining the width of uniform distribution contaminated with additive errors appears in many image processing applications. To recognize the border of elliptical objects in an image, one needs to estimate the width of its $X-$ and $Y-$coordinates. In passing, Benšić *et. al* [10] outline how a one-dimension model can be used to in more practical situations. Later, they [3,12,13] apply their approaches to some real life-applications, chief among them, fluorescent microscope [2], and estimating the size of black fungi colonies on the basis of a monochromatic image [14]. Moreover, Baum and Hanebeck [15] use *Extreme Value Theorem* and Bayesian estimation to estimate the boundary of a rectangle, and then they apply their proposed estimator to the problem of tracking a group of point targets that arise in many air surveillance where aircraft are tracked with a high-resolution radar device.

Benšić and Sabo [10] state that estimating the boundaries of a uniform variable distributed on the line segment $[a, b]$ or of circular or elliptical shapes could be a cornerstone to many important problems. Moreover, they [10,11] mention that if the mean of $U \sim \text{Uniform}[a, b]$ is known, then one can firstly centre the data at its mean $\frac{a+b}{2}$ and then estimate the boundaries of the symmetric line segment using either the ML or the MM methods. However, the mean $\frac{a+b}{2}$ is generally unknown and it shall be estimated first. Therefore this adds extra variability to the estimation process that should be taken into account. Thus, estimating the parameters of uniform random variables distributed on a line segment $[a, b]$ shall be investigated.

In this paper, we study the MM and the ML estimators for both line segments $[-a, a]$ and $[a, b]$ under measurement errors. To our best knowledge, no one has studied yet the asymptotic distribution of the MM estimators. Given the difficulty of computing the asymptotic distribution of the ML estimators [4], we only investigate their asymptotic distributions numerically. The rest of this paper is organized as follows. Section 2 discusses the ML and the MM estimators of the boundaries of a uniform random variable on the symmetric line segment $[-a, a]$ under additive errors. It also discusses the sufficient conditions for the existence and the asymptotic distributions of the MM. Section 3 extends the latter results to the general line segment $[a, b]$. Section 4 summarizes our intensive numerical results and Section 5 concludes our findings. The proofs of our theorems and other technical details are deferred to the Appendix 1.

2. Uniform distribution on a symmetric line segment $[-a, a]$

Benšić and Sabo [10,11] derive the ML and the MM estimators under the additive-error model. Let U be a random variable uniformly distributed on the line segment $[-a, a]$.

The random variable U, however, cannot be observed directly. Instead, a surrogate variable X is observed with additive error, i.e. $X = U + E$. Here E has a normal distribution with zero mean and variance σ^2 and it is statistically independent of U. The probability density function (pdf) of X is given by

$$f_{\boldsymbol{\theta}}^{(1)}(x) := \frac{1}{2a}\left[\Phi\left(\frac{a-x}{\sigma}\right) - \Phi\left(\frac{-a-x}{\sigma}\right)\right], \tag{1}$$

where $\boldsymbol{\theta} = a$ when σ^2 is known and $\boldsymbol{\theta} = (a, \sigma^2)^\top$ when σ^2 is unknown. In this section, we shall discuss the ML and the MM estimators and present their asymptotic distributions.

The ML estimator

The ML estimator of $\boldsymbol{\theta} = (a, \sigma^2)^\top$ is attained by maximizing the likelihood function of $\boldsymbol{\theta}$

$$L_1(\boldsymbol{\theta}) := \prod_{i=1}^{n} f_{\boldsymbol{\theta}}^{(1)}(x_i), \tag{2}$$

where Φ is the cumulative distribution function of the standard normal random variable. Equivalently, one could minimize the negative of the log-likelihood function of $\boldsymbol{\theta}$

$$l_1(\boldsymbol{\theta}) := -\sum_{i=1}^{n} \ln f_{\boldsymbol{\theta}}^{(1)}(x_i). \tag{3}$$

This is a nonlinear optimization problem and there is not an explicit closed-form formula for ML estimators. Moreover, Benšić and Sabo [10] state that the Newton method often fails to find the maximum value of (2) and, for this reason, they implement the *Nelder-Mead downhill simplex (NM) algorithm* by using some built-in MATLAB functions. In our simulations, we experience the same issues with the Newton-Raphson algorithm.

In this paper, instead of using NM algorithm to overcome divergence issues of the Newton method, we use the *Damped-Newton algorithm*. It is not only much simpler and faster than the NM algorithm, but also the convergence rate of the former is similar to the latter one. A generic outline for the adopted algorithm is presented in Table 1.

As Table 1 shows, the algorithm is initialized with the MM estimates and, in case of nonexistence of them, we use the following estimate to initialize the iterative process $\hat{\boldsymbol{\theta}}^{(0)} = (\max|X_i|, \frac{1}{n}\sum_{i=1}^{n} x_i^2)^\top$. Then, the objective function $l_1(\boldsymbol{\theta})$, the Jacobian $\mathbf{J}_1 := \nabla_{\boldsymbol{\theta}} l_1(\boldsymbol{\theta})$, and

Table 1. Damped Newton method.

Let $l(\cdot)$ be the negative of a generic log-likelihood function.
1. Initialize $\boldsymbol{\theta}^{(0)}$, set $\lambda \leftarrow 1$, and compute $l(\boldsymbol{\theta}^{(0)})$.
2. Given $\boldsymbol{\theta}^{(k)}$, compute
$l^{(k)} \leftarrow l(\boldsymbol{\theta}^{(k)})$, $\mathbf{J}^{(k)} \leftarrow \nabla_{\boldsymbol{\theta}} l(\boldsymbol{\theta}^{(k)})$, and $\mathbf{H}_{\lambda}^{(k)} \leftarrow \nabla_{\boldsymbol{\theta}}^2 l(\boldsymbol{\theta}^{(k)}) + \lambda \mathbf{I}_2$.
3. Solve $\mathbf{H}_{\lambda}^{(k)}\, \mathbf{h}^{(k+1)} = -\mathbf{J}^{(k)}$ for $\mathbf{h}^{(k+1)}$ by SV decomposition.
4a. If $\frac{\|\mathbf{h}^{(k+1)}\|}{\max(\|\boldsymbol{\theta}^{(k)}\|,1)} < 10^{-5}$, stop and select $\hat{\boldsymbol{\theta}}_{\text{ml}} = \boldsymbol{\theta}^{(k)}$.
4b. Otherwise, compute $\boldsymbol{\theta}^* \leftarrow \boldsymbol{\theta}^{(k)} + \mathbf{h}^{(k+1)}$.
5a. If $l(\boldsymbol{\theta}^*) \geq l^{(k)}$, reject $\boldsymbol{\theta}^*$, and update $\lambda \leftarrow 2\lambda$, then go to Step 3.
5b. Otherwise $\boldsymbol{\theta}^*$ is accepted and $\boldsymbol{\theta}^{(k+1)} \leftarrow \boldsymbol{\theta}^*$.
6a. Update $k \leftarrow k+1$, $\lambda \leftarrow 0.5\lambda$. If $k <$ N.max, go to Step 2.
6.b Otherwise, the algorithm is divergent.

the augmented Hessian matrix $\mathbf{H}_{\lambda,1} := \nabla^2_{\boldsymbol{\theta}} l_1(\boldsymbol{\theta}) + \lambda \mathbf{I}_2$ are evaluated, where $\mathbf{I}_2$ is the (2×2) identity matrix. The algorithm is iteratively applied until the stopping criteria is satisfied, i.e. see Step (4.a) in Table 1. If the maximum number of iterations N.max $= 200$ is reached before achieving the stopping criterion, the algorithm is considered divergent. It is worth mentioning that the algorithm described here finds the ML estimate of σ not σ^2. To find the ML estimate of σ^2 the invariance property of the maximum likelihood estimators is employed.

The asymptotic distribution of the ML estimator is important to compute approximate confidence intervals for the parameters and to test null hypotheses. Schneeweiss [4] and Benšić and Sabo [10,11] derive the asymptotic variance of the ML estimator $\hat{a}_{\mathrm{ml}}$ when σ^2 is known under both uniform random distributions on $[0, a]$ and $[-a, a]$. The derivation of their asymptotic variances even for those simple models is not straightforward since it requires numerical integrals and log-likelihood function can be unstable [4]. When σ^2 is unknown, one extra layer of complexity is added to the inferential process and the analysis of the asymptotic distributions of the ML estimators becomes more complicated. For this reason, in both cases, we will investigate the asymptotic variance of the ML estimators only numerically.

For both cases, the asymptotic distribution of the ML estimator is $\sqrt{n}(\hat{\boldsymbol{\theta}}_{\mathrm{ml}} - \boldsymbol{\theta}) \xrightarrow{D} \mathcal{N}_p(\mathbf{0}, \boldsymbol{\Gamma}^{-1}_{\mathrm{ml}})$, where the $(ij)^{\mathrm{th}}$ element of $\boldsymbol{\Gamma}_{\mathrm{ml}}$ is

$$\gamma^{(1)}_{ij} = \gamma^{(1)}_{ij}(\boldsymbol{\theta}) = \int_{-\infty}^{\infty} \left(\frac{\partial \ln f^{(1)}_{\boldsymbol{\theta}}(x)}{\partial \theta_i} \right) \left(\frac{\partial \ln f^{(1)}_{\boldsymbol{\theta}}(x)}{\partial \theta_j} \right) f^{(1)}_{\boldsymbol{\theta}}(x)\, \mathrm{d}x, \tag{4}$$

for $i, j = 1, \ldots, p$, where $p = 2$ if σ^2 is known and $p = 3$, if σ^2 is unknown; and $f^{(1)}_{\boldsymbol{\theta}}$ is the pdf given in (1). These quantities are estimated by replacing $\boldsymbol{\theta}$ with its ML estimate $\hat{\boldsymbol{\theta}}$, namely,

$$\widehat{\gamma}^{(1)}_{ij} = \gamma^{(1)}_{ij}(\hat{\boldsymbol{\theta}}), \quad i, j = 1, \ldots, p. \tag{5}$$

In our algorithm, we plug-in the ML estimates in formula (4) and then compute the integral numerically. The algorithm produces stable results as shown by our simulations in Section 4.

The MM estimator

Now we turn our attention to the MM estimator of (a, σ^2). First recall that the non-central k^{th} population moment of U is $\mathbb{E}(U^k) = \frac{a^k}{2}(1 - (-1)^k)$. Therefore, if we denote the k^{th} non-central population moment of X by $\mu_k = \mathbb{E}(X^k)$ and the k^{th} non-central sample-moment by $m_k = \frac{1}{n}\sum_{i=1}^{n} X_i^k$, then the second- and the fourth-moments of X are

$$\mu_2 = \frac{a^2}{3} + \sigma^2 \quad \text{and} \quad \mu_4 = \frac{1}{5}\left(a^4 + 10a^2\sigma^2 + 15\sigma^4\right). \tag{6}$$

These moments can be computed with the aid of $\mathbb{E}(E^j) = \frac{j!\sigma^j}{2^{j/2}(j/2!)}$ for any even positive number j. As the distribution of X is symmetric around zero, its odd moments are zero.

Equating the population moments with their sample moments leads to the solution

$$\hat{a}_{\text{mm}} = \begin{cases} \sqrt{3(m_2 - \sigma^2)} & \text{if } \sigma^2 \quad \text{is known and} \quad m_2 \geq \sigma^2, \\ \sqrt[4]{\frac{15}{2}}\sqrt[4]{3m_2^2 - m_4} & \text{if } \sigma^2 \quad \text{is unknown and} \quad \frac{m_4}{m_2^2} \leq 3. \end{cases} \tag{7}$$

That is, for unknown σ^2, the MM $\widehat{a}_{\text{mm}}$ exists only if $m_4 \leq 3m_2^2$. In the latter case, the MM estimator of σ^2 is

$$\hat{\sigma}^2_{\text{mm}} = m_2 - \sqrt{\frac{5}{6}}\sqrt{3m_2^2 - m_4} \quad \text{if } \frac{m_4}{m_2^2} \in (9/5, 3). \tag{8}$$

The lower bound 9/5 ensures that $\widehat{\sigma}^2_{\text{mm}}$ is positive.

Benšić and Sabo derive the MM estimator of a and σ^2 as in the preceding equations but without exploring their asymptotic properties. To our best knowledge, no one has addressed these properties. In this paper, we present in Theorem 2.1 the asymptotic variances of the MM estimators for both cases: when σ^2 is known and unknown. As expected by the usual asymptotic theory, the theorem requires the existence of the eighth moment of X. Under our statistical assumptions (the sum of uniform and normal random variables), this condition is trivially satisfied.

Theorem 2.1: *Under the additive-error model with $X = U + E$ and $\mathbb{E}(X^8) < \infty$, the MM estimator is asymptotically normally distributed. More specifically,*

(1) *When σ^2 is known, then $\hat{a}_{\text{mm}}$ exists if $m_2 \geq \sigma^2$ and its asymptotic distribution is $\sqrt{n}(\hat{a}_{\text{mm}} - a) \xrightarrow{D} \mathcal{N}_1(0; \Psi^{(1)}_{\text{mm}})$, where*

$$\Psi^{(1)}_{\text{mm}} = 3\left(\frac{a^2}{15} + \sigma^2 + \frac{3\sigma^4}{2a^2}\right). \tag{9}$$

(2) *When σ^2 is unknown, then $\hat{\boldsymbol{\theta}}_{\text{mm}} = (\hat{a}_{\text{mm}}, \hat{\sigma}^2_{\text{mm}})^\top$ exists if $\frac{m_4}{m_2^2} \in (9/5, 3)$ and its asymptotic distribution is*

$$\sqrt{n}\left(\widehat{\boldsymbol{\theta}}_{\text{mm}} - \boldsymbol{\theta}\right) \xrightarrow{D} \mathcal{N}_2\left(\mathbf{0}; \boldsymbol{\Psi}^{(1)}_{\text{mm}}\right), \qquad \boldsymbol{\Psi}^{(1)}_{\text{mm}} = \begin{bmatrix} \psi^{(1)}_{11} & \psi^{(1)}_{12} \\ \psi^{(1)}_{12} & \psi^{(1)}_{22} \end{bmatrix},$$

where

$$\begin{aligned} \psi^{(1)}_{11} &= \frac{3a^2}{7} + \frac{30\sigma^2}{7} + \frac{45\sigma^4}{2a^2} + \frac{225\sigma^6}{2a^4} + \frac{675\sigma^8}{8a^6}, \\ \psi^{(1)}_{12} &= -\left(\frac{2a^3}{21} + \frac{6a\sigma^2}{7} + \frac{15\sigma^4}{a} + \frac{75\sigma^6}{a^3} + \frac{225\sigma^8}{4a^5}\right), \\ \psi^{(1)}_{22} &= \frac{8a^4}{315} + \frac{4a^2\sigma^2}{7} + 12\sigma^4 + \frac{50\sigma^6}{a^2} + \frac{75\sigma^8}{2a^4}. \end{aligned} \tag{10}$$

Proof: See the Appendix. ■

The asymptotic variances are estimated by replacing the unknown parameters by their respective MM estimates. For known σ^2, the asymptotic variance of $\hat{a}_{\mathrm{mm}}$ is estimated by

$$\widehat{\mathrm{Var}_\infty}(\hat{a}_{\mathrm{mm}}) = \frac{1}{n}\widehat{\Psi}^{(1)}_{\mathrm{mm}}$$

and the asymptotic variance-covariance matrix of $\hat{\boldsymbol{\theta}}_{\mathrm{mm}}$ is estimated by

$$\widehat{\mathrm{Var}_\infty}(\hat{\boldsymbol{\theta}}_{\mathrm{mm}}) = \frac{1}{n}\widehat{\boldsymbol{\Psi}}^{(1)}_{\mathrm{mm}},$$

where the 'hat' symbol indicates that unknown parameters are being replaced with their MM estimates. Observe that, as expected, the asymptotic variance of $\hat{a}_{\mathrm{mm}}$ when σ^2 is unknown is larger than when σ^2 is known.

From the asymptotic distributions, one can build approximate confidence intervals such as $\hat{a}_{\mathrm{mm}} \pm z_\gamma\sqrt{\frac{1}{n}\widehat{\Psi}^{(1)}_{\mathrm{mm}}}$, where z_γ is such that $P(-z_\gamma < \mathcal{N}_1(0,1) < z_\gamma) = \gamma$. Also, one can test the null hypothesis $H_0 : a = a_0$ by using the statistic $z(\widehat{a}_{\mathrm{mm}}) = \frac{\sqrt{n}(\widehat{a}_{\mathrm{mm}} - a_0)}{\sqrt{\Psi^{(1)}_{\mathrm{mm}}}}$ which converges in distribution to a standard normal distribution.

3. Uniform distribution on the real line segment [a,b]

We now consider the case where the uniform random variable U is distributed on a general line segment $[a, b]$, where $a < b$. Consider again that U cannot be observed directly and only a surrogate random variable $X = U + E$ is observable, where E has a normal distribution with zero mean and variance σ^2 and it is statistically independent of the original non-observable U. The pdf of X is given by

$$f^{(2)}_{\boldsymbol{\theta}}(x) = \frac{1}{b-a}\left[\Phi\left(\frac{a-x}{\sigma}\right) - \Phi\left(\frac{-b-x}{\sigma}\right)\right], \tag{11}$$

where $\boldsymbol{\theta} = (a, b)^\top$ when σ^2 is known and $\boldsymbol{\theta} = (a, b, \sigma^2)^\top$ when σ^2 is unknown. In this section, we shall discuss the ML and the MM estimators and present their asymptotic distributions for both cases: σ^2 is known and σ^2 is unknown. To the best of our knowledge, no one has considered this errors-in-variables model.

The ML estimator

Here, we find the asymptotic distribution of the ML estimator. In order to compute the ML estimator of the parameter vector under both cases (σ^2 known and unknown), we write $a_i = \frac{x_i - a}{\sigma}$ and $b_i = \frac{x_i - b}{\sigma}$.

We start by considering σ^2 known. In this case, the likelihood function of $\boldsymbol{\theta} = (a, b)^\top$ is

$$L_2(\boldsymbol{\theta}) = \prod_{i=1}^{n} f^{(2)}_{\boldsymbol{\theta}}(x_i) = \prod_{i=1}^{n}\frac{I_i}{b-a}, \quad I_i = \Phi(-b_i) - \Phi(-a_i), \tag{12}$$

and the negative-likelihood function of $\boldsymbol{\theta}$ is

$$l_2(\boldsymbol{\theta}) = -\sum_{i=1}^{n}\ln I_i + n\ln(b-a). \tag{13}$$

For the known σ case, the Jacobian is $\mathbf{J}_2 = \nabla_{(a,b)} l_2(\boldsymbol{\theta})$ with components

$$\begin{aligned}\frac{\partial l_2}{\partial a} &= \sigma^{-1} \sum_{i=1}^{n} \frac{\phi(-a_i)}{I_i} - \frac{n}{b-a},\\ \frac{\partial l_2}{\partial b} &= \sigma^{-1} \sum_{i=1}^{n} \frac{-\phi(-b_i)}{I_i} + \frac{n}{b-a},\end{aligned} \tag{14}$$

To find the ML we need to minimize l_2, so, we adopt the Damped Newton algorithm as described previously in Table 1. The algorithm also requires the evaluation of the Hessian matrix $\nabla^2_{(a,b)} l_2$ which entities are given in the first three formulas in (15).

For the unknown σ^2 case, the first two components of the Jacobian vector $\mathbf{J}_2 = (\frac{\partial l_2}{\partial a}, \frac{\partial l_2}{\partial b}, \frac{\partial l_2}{\partial \sigma})^\top$ shall be evaluated as in (14) and the third component is

$$\frac{\partial l_2}{\partial \sigma} = -\sigma^{-1} \sum_{i=1}^{n} \frac{w_i}{I_i},$$

where $w_i = b_i\phi(-b_i) - a_i\phi(-a_i)$. The entities of the Hessian matrix $\mathbf{H}_2$ are

$$\begin{aligned}\frac{\partial^2 l_2}{\partial a^2} &= -\frac{1}{\sigma^2} \sum_{i=1}^{n} \left[\frac{-a_i\phi(-a_i)}{I_i} - \frac{\phi^2(-a_i)}{I_i^2} \right] - \frac{n}{(b-a)^2},\\ \frac{\partial^2 l_2}{\partial a \partial b} &= -\frac{1}{\sigma^2} \sum_{i=1}^{n} \left[\frac{\phi(-a_i)\phi(-b_i)}{I_i^2} \right] + \frac{n}{(b-a)^2},\\ \frac{\partial^2 l_2}{\partial b^2} &= -\frac{1}{\sigma^2} \sum_{i=1}^{n} \left[\frac{b_i\phi(-b_i) - \phi^2(-b_i)}{I_i^2} \right] - \frac{n}{(b-a)^2},\\ \frac{\partial^2 l_2}{\partial a \partial \sigma} &= -\frac{1}{\sigma^2} \sum_{i=1}^{n} \left[\frac{\phi(-a_i) - a_i^2\phi(-a_i)}{I_i} + \frac{w_i\phi(-a_i)}{I_i^2} \right],\\ \frac{\partial^2 l_2}{\partial a \partial \sigma} &= -\frac{1}{\sigma^2} \sum_{i=1}^{n} \left[\frac{\phi(-a_i) - a_i^2\phi(-a_i)}{I_i} + \frac{w_i\phi(-a_i)}{I_i^2} \right],\\ \frac{\partial^2 l_2}{\partial b \partial \sigma} &= -\frac{1}{\sigma^2} \sum_{i=1}^{n} \left[\frac{-\phi(-b_i) + b_i^2\phi(-b_i)}{I_i} - \frac{w_i\phi(-b_i)}{I_i^2} \right],\\ \frac{\partial^2 l_2}{\partial \sigma^2} &= -\frac{1}{\sigma^2} \sum_{i=1}^{n} \left[\frac{2w_i - b_i^3\phi(-b_i) + a_i^3\phi(-a_i)}{I_i} + \frac{w_i^2}{I_i^2} \right].\end{aligned} \tag{15}$$

Thus, after computing the ML estimate $(\hat{a}_{\text{ml}}, \hat{b}_{\text{ml}}, \hat{\sigma}_{\text{ml}})$, we find the ML estimate of σ^2 by the the invariance property of the ML estimators.

For both cases, the asymptotic distribution of the ML estimators is $\sqrt{n}(\hat{\boldsymbol{\theta}}_{\text{ml}} - \boldsymbol{\theta}) \xrightarrow{D} \mathcal{N}_p(\mathbf{0}, \boldsymbol{\Gamma}_{\text{ml}}^{-1})$, where the $(ij)^{\text{th}}$ element of $\boldsymbol{\Gamma}_{\text{ml}}$ is

$$\gamma_{ij}^{(2)} = \gamma_{ij}^{(2)}(\boldsymbol{\theta}) = \int_{-\infty}^{\infty} \left(\frac{\partial \ln f_{\boldsymbol{\theta}}^{(2)}(x)}{\partial \theta_i} \right) \left(\frac{\partial \ln f_{\boldsymbol{\theta}}^{(2)}(x)}{\partial \theta_j} \right) f_{\boldsymbol{\theta}}^{(2)}(x)\, \mathrm{d}x, \tag{16}$$

for $i, j = 1, \ldots, p$, where $p = 2$ if σ^2 is known and $p = 3$ if σ^2 is unknown. This asymptotic variance is estimated by plugging-in the numerical estimates into Formula (16)

$$\widehat{\gamma_{ij}}^{(2)} = \gamma_{ij}^{(2)}(\hat{\boldsymbol{\theta}}), \quad i, j = 1, \ldots, p. \tag{17}$$

In our algorithm, we plug-in the ML estimates in formula (16) and then compute the integral numerically.

The MM estimator

Now we turn our attention to the derivation of the MM estimators. Some simple calculations yield the first four moments of X, i.e.

$$\begin{aligned} \mu_1 &= \frac{a+b}{2}, \quad \mu_3 = \frac{b^4 - a^4}{4(b-a)} + \frac{3}{2}(b+a)\sigma^2, \\ \mu_2 &= \frac{b^2 + ab + a^2}{3} + \sigma^2, \quad \mu_4 = \frac{b^5 - a^5}{5(b-a)} + \frac{2(b^3 - a^3)}{b-a}\sigma^2 + 3\sigma^4. \end{aligned} \tag{18}$$

Note first that if σ^2 is unknown, then one can estimate σ^2 by its MM estimator:

$$\hat{\sigma}^2_{\text{mm}} = m_2 - m_1^2 - \sqrt{\frac{5}{6}}\sqrt{3m_2^2 - m_4 - 2m_1^4} \tag{19}$$

which exists under the following two conditions:

$$\begin{aligned} &\text{C1}: \quad 3m_2^2 - m_4 - 2m_1^4 \geq 0 \\ &\text{C2}: \quad 6m_2 - 6m_1^2 - \sqrt{5(3m_2^2 - m_4 - 2m_1^4)} \geq 0. \end{aligned} \tag{20}$$

Depending on whether σ^2 is known or not, the MM estimators of (a, b) under the additive-error model can be summarized as follows

$$\hat{a}_{\text{mm}} = \begin{cases} m_1 - \sqrt{3(m_2 - m_1^2 - \sigma^2)} & \text{if } \sigma^2 \text{ is known and } m_2 - m_1^2 \geq \sigma^2, \\ m_1 - \sqrt[4]{\frac{15}{2}}\sqrt[4]{3m_2^2 - m_4 - 2m_1^4} & \text{if } \sigma^2 \text{ is unknown and C1 and C2 are satisfied} \end{cases} \tag{21}$$

and

$$\hat{b}_{\text{mm}} = \begin{cases} m_1 + \sqrt{3(m_2 - m_1^2 - \sigma^2)} & \text{if } \sigma^2 \text{ is known and } \quad m_2 - m_1^2 \geq \sigma^2, \\ m_1 + \sqrt[4]{\frac{15}{2}}\sqrt[4]{3m_2^2 - m_4 - 2m_1^4} & \text{if } \sigma^2 \text{ is unknown and C1 and C2 are satisfied.} \end{cases} \tag{22}$$

The technical details of the derivations are deferred to the Appendix. It is worth mentioning here that if σ^2 is known then $(\hat{a}_{\text{mm}}, \hat{b}_{\text{mm}})^\top$ exists whenever $m_2 - m_1^2 \geq \sigma^2$. Observe also that $\hat{\sigma}^2_{\text{mm}}$ in (19) satisfies $m_2 - m_1^2 \geq \hat{\sigma}^2_{\text{mm}}$.

The following theorem summarizes the limiting distributions of $\hat{\boldsymbol{\theta}}_{\text{mm}}$.

Theorem 3.1: *Under the additive-error model with $X = U + E$ and $\mathbb{E}(X^8) < \infty$, the MM estimator is asymptotically normally distributed. More specifically,*

(1) *When* σ^2 *is known, then* $\hat{\boldsymbol{\theta}}_{\text{mm}} = (\hat{a}_{\text{mm}}, \hat{b}_{\text{mm}})^\top$ *exists if* $m_2 - m_1^2 \geq \sigma^2$ *and*

$$\sqrt{n}(\widehat{\boldsymbol{\theta}}_{\text{mm}} - \boldsymbol{\theta}) \xrightarrow{D} \mathcal{N}_2\left(\mathbf{0}, \boldsymbol{\Psi}_{\text{mm}}^{(2)}\right), \quad \boldsymbol{\Psi}_{\text{mm}} = \begin{bmatrix} \psi_{11}^{(2)} & \psi_{12}^{(2)} \\ \psi_{12}^{(2)} & \psi_{22}^{(2)} \end{bmatrix},$$

where

$$\begin{aligned} \psi_{11}^{(2)} &= \frac{2}{15}(b-a)^2 + 4\sigma^2 + \frac{18}{(b-a)^2}\sigma^4 = \psi_{22}^{(2)}, \\ \psi_{12}^{(2)} &= \frac{1}{30}(b-a)^2 - 2\sigma^2 - \frac{18}{(b-a)^2}\sigma^4. \end{aligned} \tag{23}$$

(2) *When* σ^2 *is unknown, then* $\hat{\boldsymbol{\theta}}_{\text{mm}} = (\hat{a}_{\text{mm}}, \hat{b}_{\text{mm}}, \hat{\sigma}^2_{\text{mm}})^\top$ *exists if both conditions* C1 *and* C2 *given in (20) are satisfied and* $\sqrt{n}(\widehat{\boldsymbol{\theta}}_{\text{mm}} - \boldsymbol{\theta}) \xrightarrow{D} \mathcal{N}_3(\mathbf{0}, \boldsymbol{\Psi}_{\text{mm}}^{(3)})$*, where the entities of* $\boldsymbol{\Psi}_{\text{mm}}^{(3)}$ *are presented in the appendix, see (A7).*

Proof: See the Appendix. ■

4. Numerical experiments

In this section, numerical experiments on actual and synthetic data have been conducted to assess our analysis for both estimators. This section is organized into four subsections: Subsection 4.1 demonstrates the usefulness of our theory to some practical applications in the field of medical imaging and astronomy. Subsection 4.2 studies the asymptotic distributions, confidence intervals' coverages, and the success rates of both estimators for the symmetric case $U \sim [-a, a]$ when σ^2 is unknown, while Subsection 4.3 is devoted to asymmetric case $U \sim [a, b]$ when σ^2 is known and unknown. Moreover, Subsection 4.4 studies the power functions and the p-values for some hypothesis tests.

4.1. Real-life applications:

To corroborate our analytic findings with a set of practical data points, we applied our methods on real images and we assessed their performances.

Using image processing tools, digitized observations (pixels) from two images were extracted. The first image appearing in Figure 1 (top-left) represents the image of Copernicus[1], a lunar crater, which was captured by *Apollo-NASA 12* in 1969. The second image appearing in Figure 2 (top-left) shows the Zoom-in view of cryo-EM image [16] of purified inactivated SARS-CoV-2 virus particles for a 62-year-old male who was admitted to Shenzhen Third People's Hospital on January 15, 2020. Initially, he was diagnosed with pneumonia and was further diagnosed as the first official patient with positive COVID-19[2].

The top-right subfigures in Figures 1 and 2 show the photos of the detected points obtained from both images using *Sobel Edge Detector*[3]. Figures 1 and 2 (bottom) represent the scatterplots of the digitized observations extracted from Figures 1 (top-right) and 2 (top-right) together with the corresponding histogram for each coordinate. After applying the ML and MM methods to each coordinate, the estimates of (a, b, σ^2) and their standard errors (SE) are summarized in Tables 2 and 3. The sample sizes extracted from Figures 1 and 2 are 50,118 and 1676, respectively.

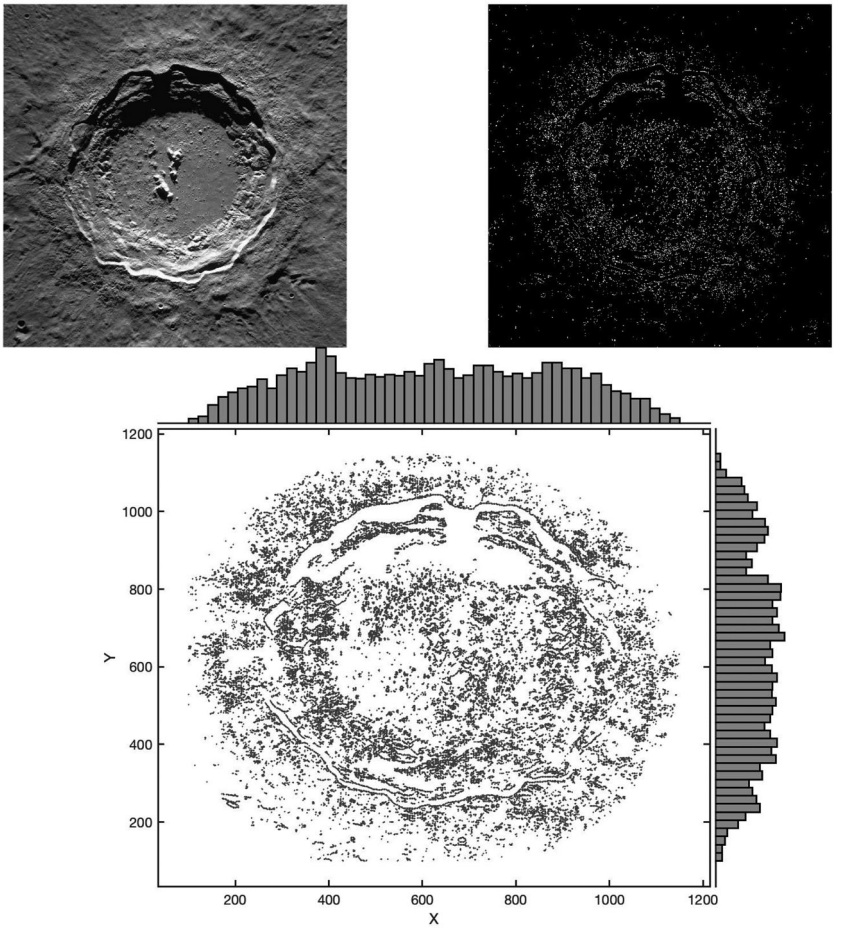


Figure 1. The top-left figure shows the photo of Copernicus. The top-right figure represents the digitized points obtained from the image by using *Sobel Edge Detector*. The lower figure represents the scatterplot of the extracted digitized observations and the corresponding histogram for each coordinate.

Table 2. The ML and the MM estimates of (a, b, σ^2) and their standard errors (SE) for each coordinate of Figure 1.

Coordinate	n	$\hat{a}_{\text{ml}}$	$\text{SE}_{\hat{a}_{\text{ml}}}$	$\hat{b}_{\text{ml}}$	$\text{SE}_{\hat{b}_{\text{ml}}}$	$\hat{\sigma}^2_{\text{ml}}$	$\text{SE}_{\hat{\sigma}^2_{\text{ml}}}$	$\hat{a}_{\text{mm}}$	$\text{SE}_{\hat{a}_{\text{mm}}}$	$\hat{b}_{\text{mm}}$	$\text{SE}_{\hat{b}_{\text{mm}}}$	$\hat{\sigma}^2_{\text{mm}}$	$\text{SE}_{\hat{\sigma}^2_{\text{mm}}}$
***X* (in pixels)**	50,118	188	1.2	1056	1.2	2943	92.4	200	6.5	1041	7.9	6427	1973
***Y* (in pixels)**	50,118	206	1.2	1052	1.2	3374	102	261	10.7	990	11.5	15,984	2669

It can be clearly seen from Tables 2 and 3 that the MM estimates of σ^2 are much larger than that of the ML estimates, and as such, $\hat{a}_{\text{MM}}$ tends to overestimate a while $\hat{b}_{\text{MM}}$ tends to underestimate b. Overall, the MM tends to return smaller estimates for $b-a$ compared to the ML estimates. Moreover, the standard errors of the MM estimates are much larger than that of the ML estimates. Those observations will be demonstrated in our intensive Monte-Carlo simulations presented in Subsection 4.3.

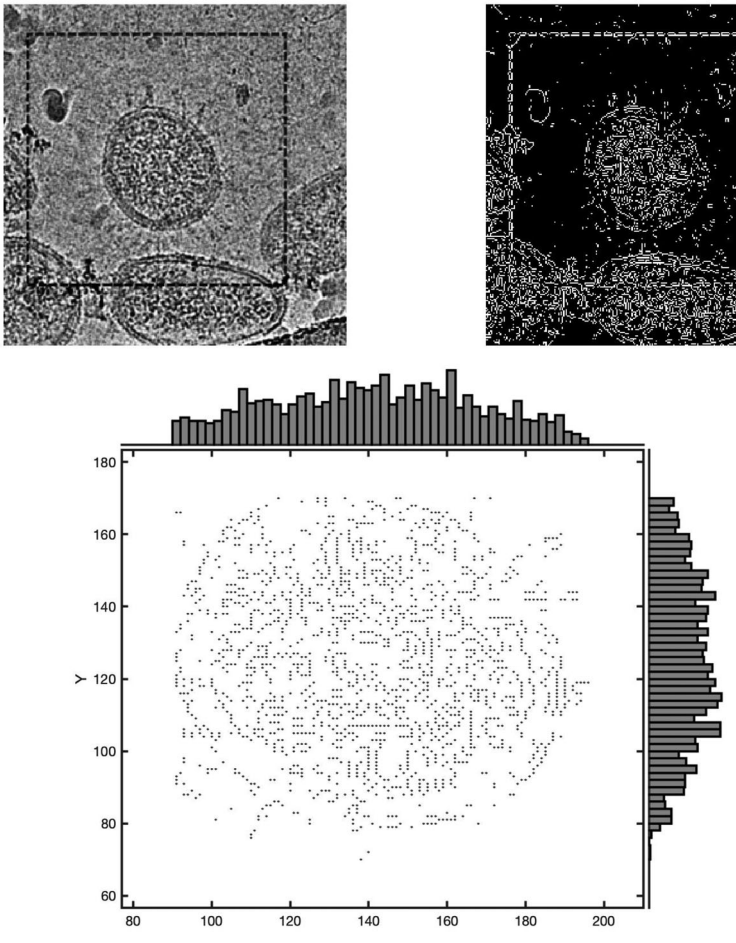


Figure 2. The top-left figure shows the photo of purified inactivated SARS-CoV-2 virus particles. The top-right figure represents the digitized points obtained from the image by using *Sobel Edge Detector*. The lower figure represents the scatterplot of the extracted digitized observations and the corresponding histogram for each coordinate.

Table 3. The ML and the MM estimates of (a, b, σ^2) and their standard errors (SE) for each coordinate of Figure 2.

Coordinate	n	$\hat{a}_{\text{ml}}$	$\text{SE}_{\hat{a}_{\text{ml}}}$	$\hat{b}_{\text{ml}}$	$\text{SE}_{\hat{b}_{\text{ml}}}$	$\hat{\sigma}^2_{\text{ml}}$	$\text{SE}_{\hat{\sigma}^2_{\text{ml}}}$	$\hat{a}_{\text{mm}}$	$\text{SE}_{\hat{a}_{\text{mm}}}$	$\hat{b}_{\text{mm}}$	$\text{SE}_{\hat{b}_{\text{mm}}}$	$\hat{\sigma}^2_{\text{mm}}$	$\text{SE}_{\hat{\sigma}^2_{\text{mm}}}$
***X* (in pixels)**	1676	96	0.8	186	0.8	42	6.7	99	8.9	183	9.7	83	259
***Y* (in pixels)**	1676	87	0.8	162	0.8	60	8.3	89	9.7	159	10.2	103	232

4.2. Monte-Carlo simulations on $[-a, a]$

The experimental results from the real data confirm our theoretical study but it does not reveal the whole picture. To get more insight, we ran more numerical experiments using Monte Carlo simulations. In order to compare between the asymptotic distributions of the MM and the ML estimators and to assess their performances, Monte-Carlo simulations were conducted when σ^2 is unknown.

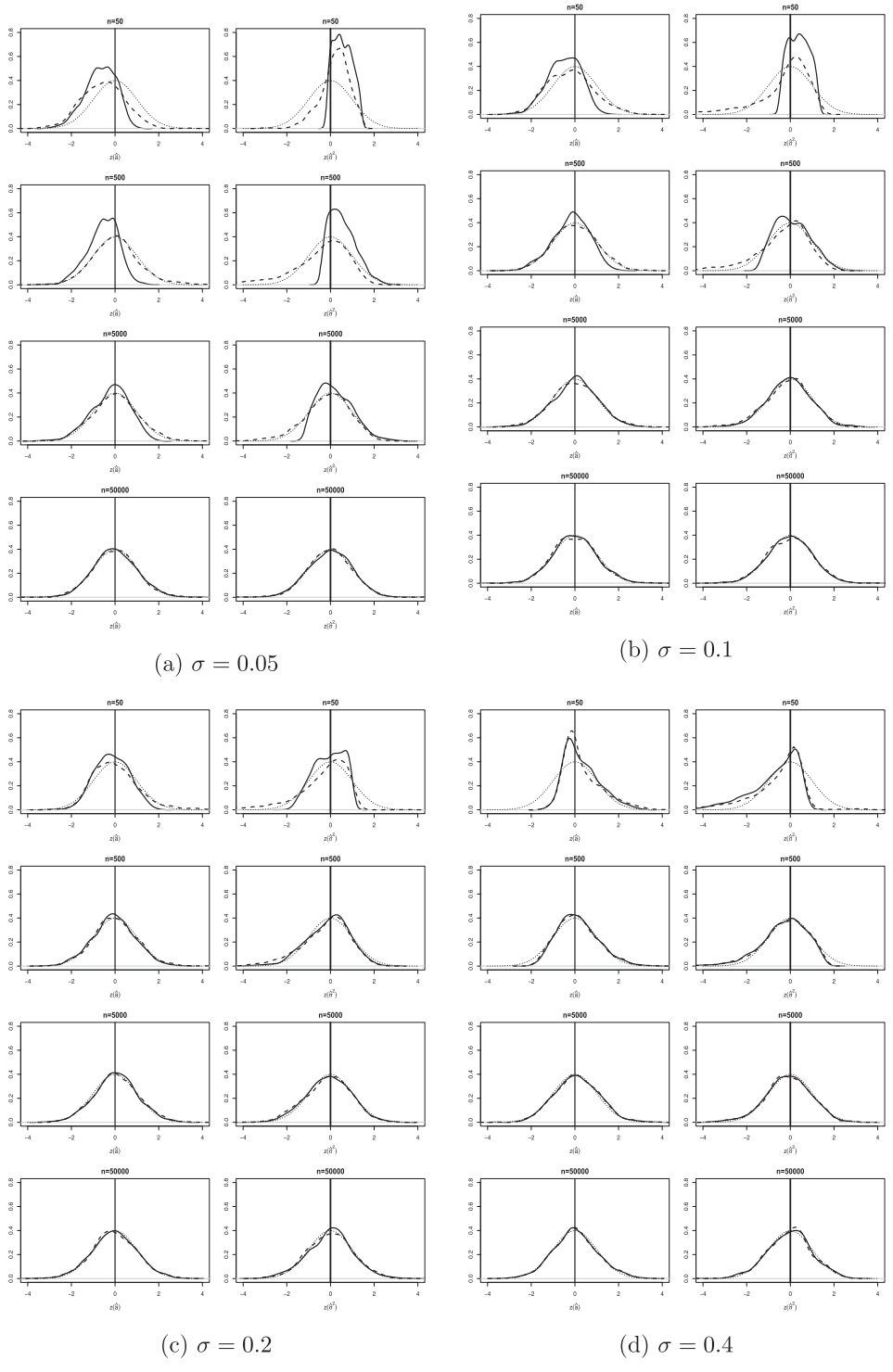
Table 4. MSE, coverage percentages for their 95% CIs, success rates (convergence rates), and average run time (ART) for ML and MM estimates. The data is distributed uniformly on the symmetric line segment $[-1, 1]$ corrupted by additive error with variance σ^2 (considered to be unknown in the estimation process).

		MSE($\times 10^{-5}$)		95%Covrg(%)		95%Covrg(%)		Success(%)			ART($\times 10^{-6}$)	
σ	n	$\hat{\boldsymbol{\theta}}_{\text{ml}}$	$\hat{\boldsymbol{\theta}}_{\text{mm}}$	$\hat{a}_{\text{ml}}$	$\hat{a}_{\text{mm}}$	$\hat{\sigma}^2_{\text{ml}}$	$\hat{\sigma}^2_{\text{mm}}$	$\hat{\boldsymbol{\theta}}_{\text{ml}}$	$\hat{a}_{\text{mm}}$	$\hat{\sigma}^2_{\text{mm}}$	ML	MM
0.05	50	137.0	409.4	93.5	95.8	99.6	99.3	66	100	52	250	0
	250	18.7	67.4	90.8	95.6	83.1	98.3	94	100	58	312	0
	500	7.2	30.5	92.6	96.4	86.2	97.9	98	100	62	560	2
	750	6.2	20.9	93.8	95.2	88.9	97.7	98	100	65	890	2
	1000	4.4	14.6	93.8	97.8	89.4	97.3	98	100	67	1040	5
	5000	0.8	2.6	94.3	97.2	94.2	97.4	99	100	86	33,995	23
	10,000	0.4	1.8	95.2	96.5	95.8	97.2	98	100	92	56,679	34
	50,000	0.08	0.4	94.8	96.1	94.8	96.2	100	100	100	188,620	184
0.1	50	207.0	363.6	94.1	97.8	87.4	99.9	69	100	63	250	0
	250	39.1	70.8	92.4	97.0	86.3	99.1	98	100	78	320	1
	500	18.6	33.8	93.8	96.1	89.7	99.1	99	100	91	460	2
	750	13.8	29.1	93.5	96.4	92.5	98.1	100	100	96	740	2
	1000	10.1	20.7	95.3	97.5	94.4	98.4	100	100	100	902	2
	5000	2.0	3.8	94.5	95.6	93.2	95.4	100	100	100	35,352	27
	10,000	0.9	2.5	94.4	95.1	94.7	94.4	100	100	100	63,112	45
	50,000	0.2	0.5	95.7	95.4	95.2	94.8	100	100	100	212,018	224
0.2	50	508.0	531.0	94.0	99.1	85.6	88	100	100	87	222	1
	250	115.9	150.1	94.4	97.1	90.6	95.8	100	100	99	342	1
	500	51.9	59.9	93.8	95.7	91.8	94.8	100	100	100	638	1
	750	40.2	48.0	95.4	95.2	93.3	94.7	100	100	100	1140	2
	1000	27.2	33.4	95.5	95.4	94.7	94.9	100	100	100	1350	1
	5000	5.4	6.3	95.0	95.4	93.3	94.6	100	100	100	43,308	50
	10,000	2.8	3.5	94.8	95.1	93.8	94.8	100	100	100	79,315	63
	50,000	0.6	0.7	94.7	94.4	93.8	93.8	100	100	100	283,015	210
0.4	50	2524.0	1934.0	92.3	96.0	81.9	85.5	100	99	93	205	1
	250	608.2	564.2	94.1	94.9	91.9	89.1	100	99	99	205	1
	500	283.0	267.2	95.4	95.8	91.6	92.4	100	99	99	534	1
	750	189.5	192.1	96.1	96.4	94.5	93.5	100	100	100	777	3
	1000	121.9	130.1	95.7	96.3	95.8	94.2	100	100	100	1014	7
	5000	32.3	31.8	94.1	94.9	94.1	93.7	100	100	100	39,366	50
	10,000	13.3	14.4	94.7	95.2	94.6	93.9	100	100	100	62,543	65
	50,000	2.7	2.9	94.5	93.8	94.4	94.5	100	100	100	226,800	240

In our experiments, we set $a = 1$ and we used four values for σ, namely, $\sigma = \{0.05, 0.1, 0.2, 0.4\}$. For each σ, different sample sizes n have been used, $n \in \{50, 250, 500, 750, 1000, 5000, 10,000, 50,000\}$, and for each (σ, n), $N = 3000$ independent trials were generated under the additive-error model. Our results are summarized in Table 4 and Figure 3.

Table 4 shows the estimated MSE of $\hat{\boldsymbol{\theta}} = (\hat{a}, \hat{\sigma}^2)$, the coverage (in percentage) for the 95% confidence interval (CI) of each estimator, the success convergence rates (in percentage), and the average run time (ART) for each method. Figure 3 shows the sampling distributions of each (standardized) estimator of $\hat{a}$ (left) and $\hat{\sigma}$ (right). The dashed curves represent the asymptotic distributions of the ML while the solid curves represent the MM. The dotted curves stand for the standard normal density function.

Since the estimates may not exist for some trials, we also computed their convergence rates (success rates). From Equations (7) and (8), one can clearly see that it is not necessarily the case that 'if $\hat{a}_{\text{mm}}$ exists, then $\hat{\sigma}_{\text{mm}}$ exists', so we present the success rates for each



(a) $\sigma = 0.05$ (b) $\sigma = 0.1$

(c) $\sigma = 0.2$ (d) $\sigma = 0.4$

Figure 3. The sampling distributions of standardized $\hat{a}$ and $\hat{\sigma}^2$, namely, $z(\hat{a}) = (\hat{a} - a)/\sqrt{\widehat{\text{Var}_\infty(\hat{a})}}$ and $z(\hat{\sigma}^2) = (\hat{\sigma}^2 - \sigma^2)/\sqrt{\widehat{\text{Var}_\infty(\hat{\sigma}^2)}}$ for both the ML (dashed line) and the MM (solid line), compared with the density of standard normal distribution (dotted-line). The generated data is distributed uniformly along $[-1, 2]$ under the additive-error model with four unknown noise level σ: (a) $\sigma = 0.05$, (b) $\sigma = 0.1$, (c) $\sigma = 0.2$, (d) $\sigma = 0.4$.

parameter. As for the ML, both $\hat{a}_{\text{ml}}$ and $\hat{\sigma}^2_{\text{ml}}$ are estimated jointly so either both exist or both diverge, and hence, we present the success rate for the parameter vector.

Now we summarize our observations as follows:

- As expected, the estimated MSEs for the ML estimators are smaller than the ones for the MM estimators for large samples, but this difference decreases as the noise level increases. In particular, for $\sigma = 0.4$ and $n \leq 150$, the estimated MSEs for the MM estimators are smaller than the ones for ML estimators.
- All the 95% CI coverages get closer to the nominal level (95%) as n increases. In particular,
  - o the coverages under the MM approach are larger than the ones under the ML approach.
  - o the coverages for σ^2 converge slower to the nominal level than the ones for a.
  - o a sample $n \geq 5000$ is required to have at least 90% of coverage. Our other numerical simulations show that the nominal level is attained (with one decimal place accuracy) when $n \geq 100,000$.
- The existence rates of $\hat{a}_{\text{mm}}$ are always very high while $\hat{\sigma}^2_{\text{mm}}$ are suffering from the divergence especially for small σ. The ML estimates diverge more often for small sample size and small σ.
- The MM approach is faster than the ML one. The iterative ML algorithm takes 2–4 iterations to converge under large σ and 7–9 iterations under small σ.

4.3. Monte-Carlo simulations on [a, b]

The setting of the experiments here were designed similarly to the symmetric case but with $a = -1$ and $b = 2$. We tested two cases:

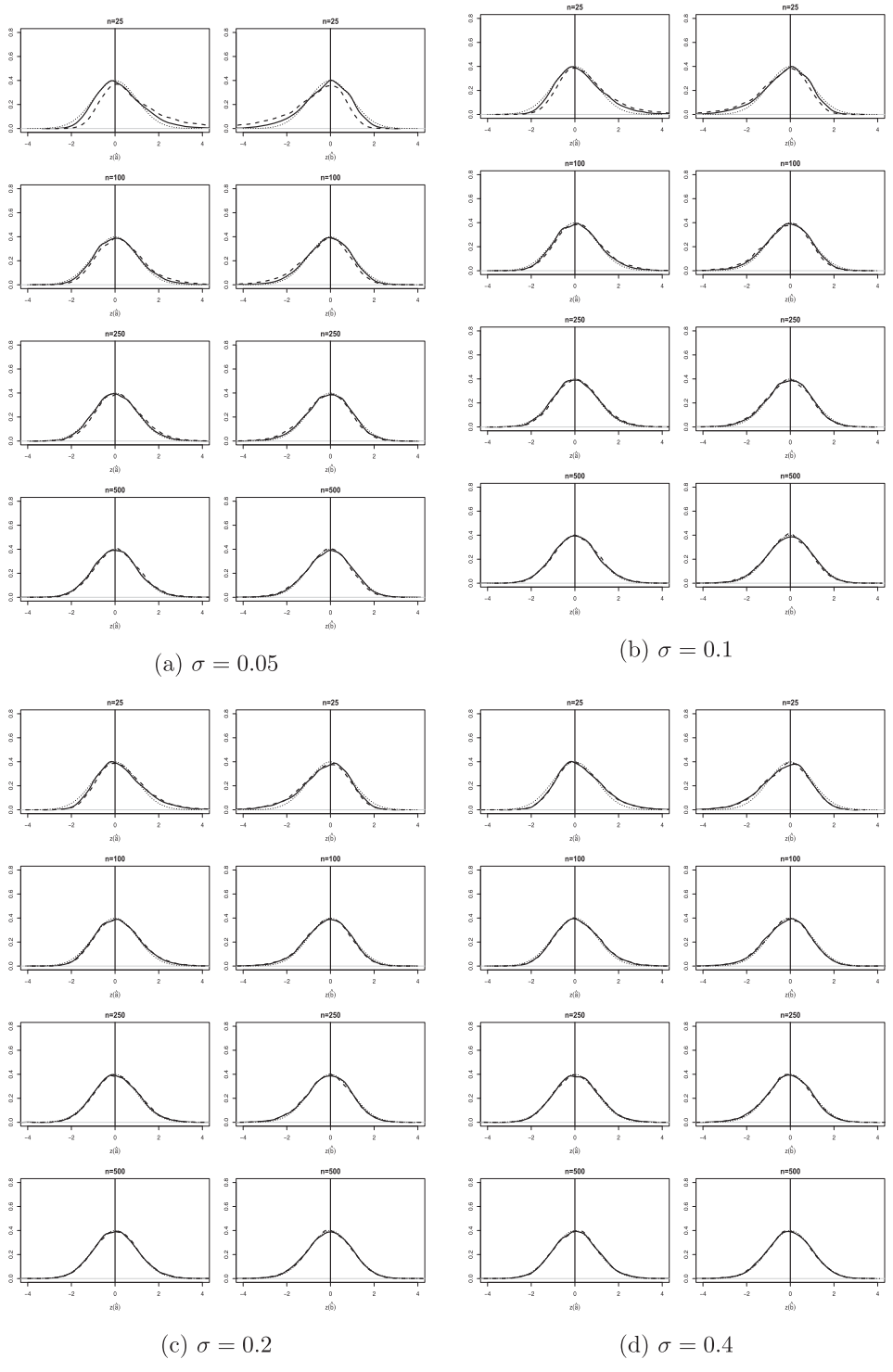
(i) known σ^2 (see Table 5 and Figure 4).
(ii) unknown σ^2 (see Table 6 and Figure 5).

Tables 5 and 6 show the estimated MSE of $\hat{\boldsymbol{\theta}} = (\hat{a}, \hat{b})^\top$ and $\hat{\boldsymbol{\theta}} = (\hat{a}, \hat{b}, \hat{\sigma}^2)^\top$, respectively, and both tables present also the coverage rates (in percentage) for the 95% confidence interval (CI) for each estimator, and the success rates (in percentage) and the average run time (ART) for each method. Figure 4 shows the sampling distributions of each estimate of a (left) and b (right) while Figure 5 shows the sampling distributions of $(\hat{a}, \hat{b})$ and $\hat{\sigma}^2$ (positioned at the rightmost panel.)

(i) Known σ^2. We use smaller sample sizes in this case than the others because the standardized versions of the sampling distributions of $(\hat{a}, \hat{b})$ for both approaches are approaching normal distributions for $n \approx 100$, as can be seen in Columns 5-8 of Table 5 and Figure 4. Here, the sample size was set as $n \in \{50, 100, 250, 500, 750, 1000\}$.

Accordingly, we observe the following noteworthy observations:

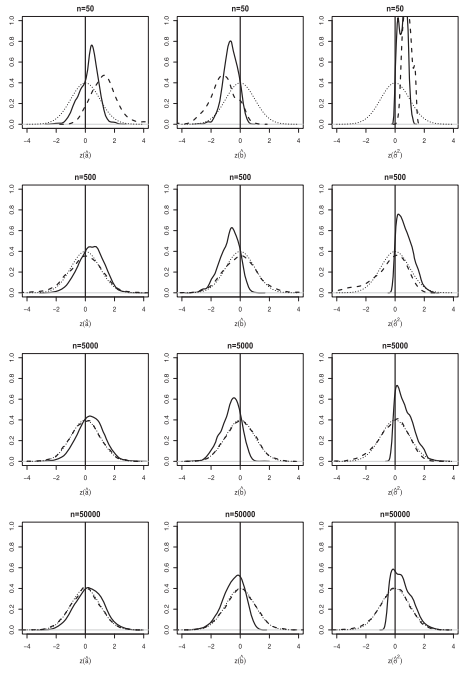
- The MSEs for the ML approach are smaller than the ones for the MM one. However, the difference decreases as n or σ increases.
- The coverages of the 95% CIs of (a, b) achieve 95% (with one decimal place accuracy) when $n \geq 500$.



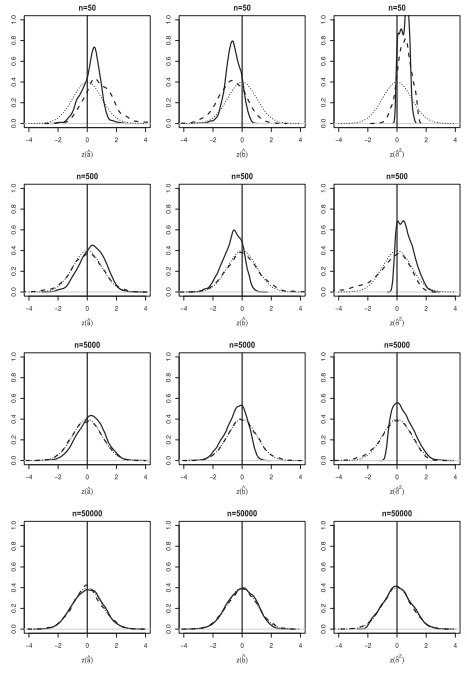
(a) $\sigma = 0.05$ (b) $\sigma = 0.1$

(c) $\sigma = 0.2$ (d) $\sigma = 0.4$

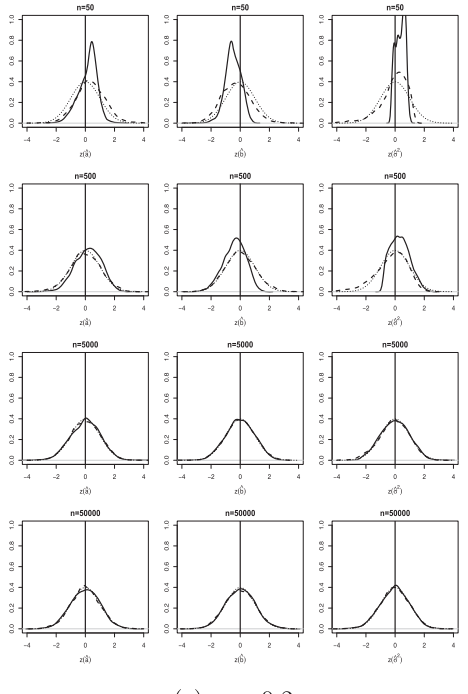
Figure 4. The sampling distributions of standardized $\hat{a}$ and $\hat{b}$, namely, $z(\hat{a}) = (\hat{a} - a)/\sqrt{\widehat{\text{Var}_\infty(\hat{a})}}$ and $z(\hat{b}) = (\hat{b} - b)/\sqrt{\widehat{\text{Var}_\infty(\hat{b})}}$ for both the ML (dashed line) and the MM (solid line), compared with the density of standard normal distribution (dotted-line). The generated data is distributed uniformly along $[-1, 2]$ under the additive-error model with four known noise level σ: (a) $\sigma = 0.05$, (b) $\sigma = 0.1$, (c) $\sigma = 0.2$, (d) $\sigma = 0.4$.



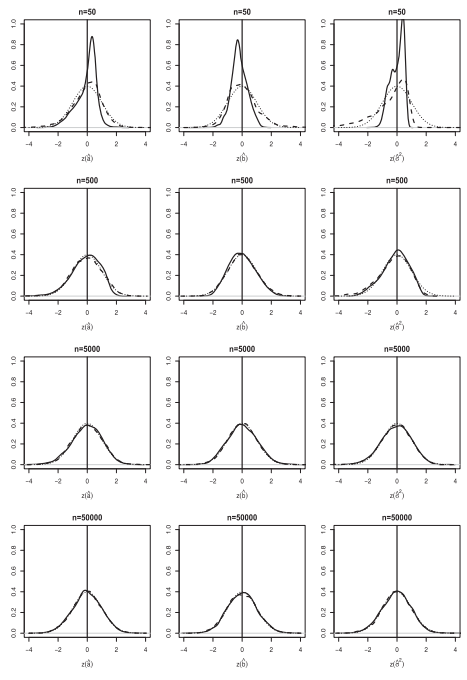
(a) $\sigma = 0.05$



(b) $\sigma = 0.1$



(c) $\sigma = 0.2$



(d) $\sigma = 0.4$

Figure 5. The sampling distributions of standardized $\hat{a}$, $\hat{b}$, and $\hat{\sigma}^2$, namely, $z(\hat{a}) = (\hat{a} - a)/\sqrt{\widehat{\text{Var}_\infty(\hat{a})}}$, $z(\hat{b}) = (\hat{b} - b)/\sqrt{\widehat{\text{Var}_\infty(\hat{b})}}$ and $z(\hat{\sigma}^2) = (\hat{\sigma}^2 - \sigma^2)/\sqrt{\widehat{\text{Var}_\infty(\hat{\sigma}^2)}}$ for both the ML (dashed line) and the MM (solid line), compared with the density of standard normal distribution (dotted-line). The generated data is distributed uniformly along $[-1, 2]$ under the additive-error model with four unknown noise level σ: (a) $\sigma = 0.05$, (b) $\sigma = 0.1$, (c) $\sigma = 0.2$, (d) $\sigma = 0.4$.

Table 5. A comparison between the MM and the ML of $\boldsymbol{\theta} = (a, b)^\top$ for data distributed uniformly on the line segment $[-1, 2]$ corrupted by additive error of known level σ^2. The table includes the MSE of both the ML and the MM estimators, their coverage percentages for their 95% CIs, success (existence) rates, and their average run time (ART).

				Covrg(%)		Covrg(%)		Success(%)			ART($\times 10^{-6}$)	
σ	n	MSE ($\times 10^{-5}$)		$\hat{a}_{\text{ml}}$	$\hat{a}_{\text{mm}}$	$\hat{b}_{\text{ml}}$	$\hat{b}_{\text{mm}}$	$\hat{\boldsymbol{\theta}}_{\text{ml}}$	$\hat{a}_{\text{mm}}$	b_{mm}	ML	MM
0.05	25	1478.0	6853.0	81.3	92.5	81.4	92.3	99.0	100	100	58	1
	50	557.0	3258.0	88.7	93.9	88.4	93.8	99.6	100	100	101	1
	100	264.0	1652.0	91.6	94.4	91.1	94.6	100	100	100	233	1
	250	96.8	654.4	93.7	95.0	93.4	94.6	100	100	100	248	1
	500	46.1	330.1	94.8	94.8	94.5	94.8	100	100	100	483	2
	750	31.7	219.5	94.6	94.9	94.4	94.9	100	100	100	726	1
	1000	23.4	165.2	94.6	95.2	94.6	94.8	100	100	100	945	2
	5000	4.7	33.9	94.9	94.8	94.8	94.7	100	100	100	33,903	47
0.1	25	12,220.0	7001.0	88.1	92.5	88.2	92.2	100	100	100	63	2
	50	1017.0	3387.0	91.7	93.8	91.6	93.9	100	100	100	111	2
	100	505.0	1708.0	93.2	94.5	93.0	94.5	100	100	100	223	2
	250	195.0	670.0	94.5	95.1	94.0	94.5	100	100	100	253	2
	500	94.4	338.3	95.0	94.9	94.7	94.8	100	100	100	490	2
	750	64.4	226.8	94.7	94.9	94.6	94.8	100	100	100	700	2
	1000	47.8	169.1	94.7	95.0	94.7	94.9	100	100	100	940	3
	5000	35.0	34.9	94.9	94.9	95.0	94.7	100	100	100	32,651	44
0.2	25	4310.0	7836.0	91.5	92.6	91.4	92.2	100	100	100	56	1
	50	2063.0	3742.0	93.0	93.8	93.5	93.7	100	100	100	97	1
	100	1019.0	1901.0	93.8	94.5	94.1	94.4	100	100	100	220	2
	250	405.7	743.2	95.1	95.1	94.2	94.4	100	100	100	244	1
	500	199.3	375.7	94.9	95.1	94.8	94.9	100	100	100	510	2
	750	130.7	252.2	95.0	94.8	94.6	94.7	100	100	100	793	2
	1000	100.4	185.2	94.7	95.0	94.6	95.1	100	100	100	991	4
	5000	19.9	38.5	95.1	95.1	94.9	95.0	100	100	100	34,137	47
0.4	25	9245.0	10,993.0	93.1	93.0	92.7	92.6	100	100	100	64.9	1
	50	4452.0	5271.0	93.8	94.0	93.8	93.7	100	100	100	108	1
	100	2223.0	2677.0	94.5	94.5	94.6	94.4	100	100	100	245	2
	250	880.0	1042.0	95.0	95.2	94.6	94.5	100	100	100	270	2
	500	441.9	532.0	95.0	95.0	94.9	95.0	100	100	100	546	2
	750	293.1	350.0	95.2	94.9	94.8	94.8	100	100	100	800	3
	1000	224.2	264.0	94.7	94.9	95.0	94.9	100	100	100	1008	3
	5000	43.8	53.0	95.2	95.2	95.2	95.2	100	100	100	34,440	48

- The success rates for the MM are 100%, while the success rates of the ML vary depending on (n, σ). For example, the ML estimates diverge more often when $\sigma = 0.05$, i.e. its convergence rate is about 50%.

(ii) Unknown σ^2. Among all the scenarios, this is the most challenging one. In our experiments, we used the same values of n as the ones considered in the symmetric case. Our observations for this case are summarized as follows:

- The MSEs for the ML approach are smaller than the ones for the MM one, except when $\sigma = 0.4$ and $n \geq 10,000$.
- The coverages of the 95% CIs are equal or greater than 95% (with one decimal place accuracy) only for $n \geq 50,000$.
- The existence rates of $(\hat{a}_{\text{mm}}, \hat{b}_{\text{mm}})^\top$ are always 100%, but the MM estimates $\hat{\sigma}^2_{\text{mm}}$ diverge more often for small σ, where its divergence rate is $\approx 50\%$. The ML estimates diverge more often for $n \leq 250$ or when σ is small σ, but those rates decrease as σ increases.

Table 6. A Comparison between the MM and the ML of $\boldsymbol{\theta} = (a, b, \sigma^2)^\top$ for data distributed uniformly on the line segment $[-1, 2]$ corrupted by additive error of unknown level σ. The table includes the MSE of both the ML and the MM estimators, their coverage percentages for their 95% CIs, success (existence) rates, and their average run time (ART).

		MSE($\times 10^{-5}$)		95%Covrg(%)		95%Covrg(%)		95%Covrg(%)		Success(%)			ART($\times 10^{-6}$)	
σ	n	$\hat{\boldsymbol{\theta}}_{\text{ml}}$	$\hat{\boldsymbol{\theta}}_{\text{mm}}$	$\hat{a}_{\text{ml}}$	$\hat{a}_{\text{mm}}$	$\hat{b}_{\text{ml}}$	$\hat{b}_{\text{mm}}$	$\hat{\sigma}^2_{\text{ml}}$	$\hat{\sigma}_{\text{mm}}$	$\hat{\boldsymbol{\theta}}_{\text{ml}}$	$\hat{a}_{\text{mm}}$	$\hat{\sigma}^2_{\text{mm}}$	ML	MM
0.05	50	4980.0	7938.0	80.9	99.5	82.7	98.9	100	99.3	23	100	50	388	3
	250	126.8	1463.8	92.5	98.5	91.3	95.4	88.7	99.3	65	100	50	537	2
	500	65.4	770.5	91.3	97.4	91.1	95.5	82.9	98.0	93	100	50	949	1
	750	41.9	521.1	92.2	96.8	91.3	94.7	83.2	96.6	98	100	51	1447	1
	1000	31.0	366.2	91.8	97.2	92.8	94.9	86.3	96.4	99	100	53	1830	1
	5000	60.1	64.3	95.3	96.5	95.1	96.0	94.5	96.5	100	100	52	50,487	27
	10,000	3.1	30.4	95.2	95.6	94.2	95.7	94.2	96.2	100	100	62	87,647	45
	50,000	0.6	5.6	94.8	95.8	94.6	96.7	95.5	96.9	98	100	69	319,826	233
0.1	50	2627.0	7585.4	89.1	99.4	90.7	98.8	100	99.8	27	99	51	310	1
	250	262.9	1339.6	91.4	98.4	91.1	96.1	81.7	99.8	92	100	55	470	2
	500	132.5	694	92.7	97.3	92.4	96.3	86.4	98.3	99	100	59	870	4
	750	85.9	467.3	92.8	97.6	93.2	95.8	88.3	97.4	99	100	58	1310	1
	1000	64.5	315.7	93.9	97.7	94.6	95.1	89.5	97.4	97	100	60	1700	1
	5000	12.6	53.4	95.1	96.4	94.9	97.0	94.4	97.9	97	100	72	46,132	48
	10,000	6.3	29.3	95.2	95.7	94.6	96.9	94.9	97.1	98	100	82	77,565	85
	50,000	1.2	7.9	95.3	95.5	94.6	96.7	95.3	97.5	98	100	97	213,346	290
0.2	50	3015.5	7708.0	93.5	99.5	94.2	99.9	94.8	100	54	100	58	232	2
	250	5,68.4	12,836.0	92.5	98.7	92.9	97.7	85.8	100	99	100	70	408	2
	500	297.9	681.0	93.8	97.8	93.8	97.1	91.1	99.4	99	100	77	778	2
	750	203.0	465.3	93.7	97.2	93.2	96.7	91.4	98.4	99	100	80	1160	2
	1000	144.1	357.1	94.4	97.0	95.6	96.9	91.8	98.5	98	100	85	1450	2
	5000	29.2	91.3	94.5	95.5	94.9	96.2	93.9	96.4	99	100	98	36,323	56
	10,000	14.3	44.3	94.6	95.1	94.6	94.7	94.6	94.8	100	100	100	65,285	75
	50,000	2.9	9.4	94.9	95.4	94.8	95.5	95.8	95.6	100	100	100	230,418	324
0.4	50	7080.1	9298.2	94.5	98.8	94.4	99.2	82.7	100	99	99	66	199	1
	250	1600.3	2653.8	94.1	97.1	93.7	98.9	89.7	99.5	99	99	94	361	1
	500	797.6	1400.5	94.3	96.0	94.8	97.9	93.1	96.8	99	100	98	630	1
	750	525.2	1019.0	94.3	95.3	93.9	96.0	92.4	95.2	100	100	99	1096	1
	1000	381.5	774.5	95.0	95.9	94.8	95.6	93.4	95.4	100	100	100	1096	3
	5000	76.4	160.8	94.8	94.6	94.9	95.2	94.6	94.7	100	100	100	31,131	45
	10,000	39.3	75.1	94.9	94.9	94.7	94.5	94.2	94.9	100	100	100	58,828	75
	50,000	7.6	0.4	94.5	95.4	94.8	95.2	94.6	95.5	100	100	100	230,046	315

- For finite sample sizes, $\hat{a}_{\text{mm}}$ tends to overestimate a and $\hat{b}_{\text{mm}}$ tends to underestimate b. That is, for small n, $\hat{b}_{\text{mm}} - \hat{a}_{\text{mm}}$ underestimates the interval length $b-a$.
- The MM is very fast. Also, the number of iterations required to compute the ML estimates shows the same pattern as the symmetric interval one.

4.4. P-Value and power of tests

We also conducted Monte Carlo simulations to verify if the asymptotic p-values have uniform distribution and if the power of test statistics at 5% significance level increases as the sample size increases. The parameter settings are the same employed in the previous simulations. For the symmetric case, Figure 6 shows the density estimates of the P-values computed based on the ML (dashed line) and the MM (solid line) approaches. The P-values were computed under the null hypothesis $H_0 : a = 1$ by considering the Wald statistics $t_1 = (\hat{a} - 1)^2 / \widehat{\text{Var}}_\infty(\hat{a})$ with its Chi-Squared asymptotic distribution with 1 degree-of-freedom. As expected, under all studied parameter settings, the density estimates of the

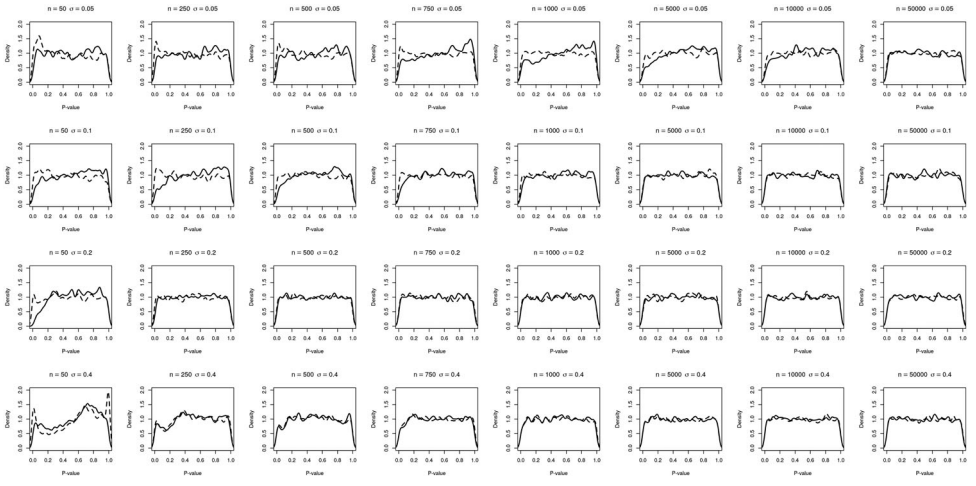


Figure 6. Density estimates of the P-values computed based on the ML (dashed line) and the MM (solid line) approaches under $H_0 : a = 1$. The tests have been conducted under the assumption that the noise level σ is unknown.

P-values computed based on the ML and MM approaches are uniformly distributed over the interval $[0, 1]$ when $n \geq 250$. This demonstrates that the tests are reliable for those sample sizes.

Moreover, we expanded our numerical verification to include a power study as shown in Figure 7. This figure depicts the rejection rates for the ML (dashed line) and the MM (solid line) approaches. The statistics were computed under alternative hypotheses, namely, $H_1 : a = a_1$, where $a_1 = 0.96, 0.97, 0.98, 0.99, 1.01, 1.02, 1.03, 1.04$ by considering the null hypothesis $H_0 : a = 1$ and the same Wald statistics $t_1 = (\hat{a} - 1)^2 / \widehat{\text{Var}}_\infty(\hat{a})$ at 5% significance level. For most parameters settings, the figure shows that the powers converge to 1 as n increases for both approaches. The only exception occurs when $|a_1 - 1| \leq 0.01$, for which the alternative hypotheses are very close to the null making it difficult to detect the effects.. Naturally, the farther a_1 from 1 is, the larger the power is–even for small sample sizes. But if the hypothesized value is very close to the true value of a parameter, then the power of the test is low even with large sample sizes.

Finally it is worth mentioning here that the asymmetrical scenarios with known and unknown σ lead to the same conclusions. For the asymmetric cases only the lower bound was changed under the alternative hypothesis, while the upper bound was kept fixed and equal to 2. Their corresponding figures are provided in the supplementary[4].

5. Conclusion

In this paper, we study the MM and the ML estimators for the width of a uniform random variable U when measured with additive errors. We study situations where the support is either the symmetric interval $[-a, a]$ or asymmetric interval $[a, b]$. Both MM and ML estimators for the boundaries are discussed in two cases, namely, when the noise level σ^2 is known and when it is unknown. While the MM estimators have a closed form, the ML estimators need an iterative algorithm to solve a system of nonlinear equations. A reliable

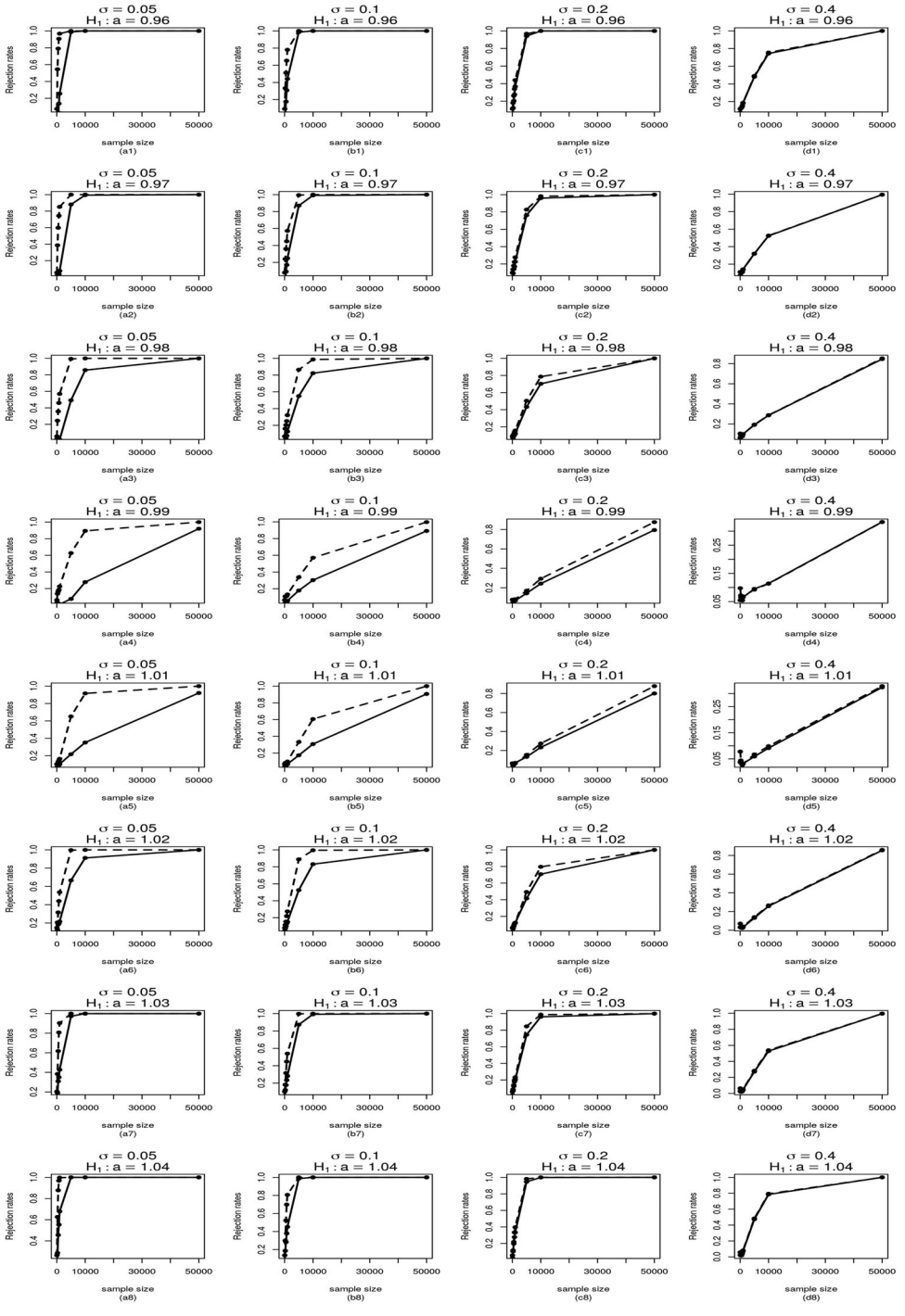


Figure 7. Rejection rates computed for testing $H_0 : a = 1$ under the alternative hypothesis $H_1 : a = a_1$ for each $a_1 \in \{0.96, 0.97, 0.98, 0.99, 1.01, 1.02, 1.03, 1.04\}$ using the MM (dashed line) and the ML (solid line) approaches. The tests have been conducted under the assumption that the noise level σ is unknown.

algorithm to compute the ML estimators is proposed with the aid of the Damped-Newton algorithm. We also establish sufficient conditions for the existence of the MM estimators and their asymptotic normality. The simulations showed that, if σ is known, the asymptotic normality is reliable even for small samples sizes, but when σ is unknown and large, this approximation is not good and large samples sizes are needed.

Notes

1. This image resolution of this image was 1264×1264 pixels, so we used the scale factor $f_0 = 1000$ pixels to normalize the extracted data so that the transformed image stays within a reasonable numerical range. Note that both the ML and the MM are invariant under scaling.
2. This image resolution of this image was 270×270 pixels, so we used the scale factor $f_0 = 100$ pixels to normalize the extracted data so that the transformed image stays within a reasonable numerical range.
3. A technique implemented on Image Processing Toolbox of Matlab.
4. Supplementary Material: The online Supplementary Material includes figures for the density estimates of the P-values and the rejection rates (powers) corresponding to the ML the MM approaches when testing $H_0 : a = -1$ versus $H_1 : a = a_1$, where $U \sim [-1, 2]$. The settings for those numerical experiments are the same as the ones implemented here.

Acknowledgments

We thank the Editor, Associate Editor, and referees for making comments and suggestions that helped to improve the paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by São Paulo Research Foundation (FAPESP) [grant number 2018/21934-5].

ORCID

Alexandre G. Patriota http://orcid.org/0000-0003-1327-3473

References

[1] Bankman I. Handbook of medical imaging: processing and analysis management. 1st Edition. Academic Press; 2000. eBook ISBN: 9780080533100.
[2] Ruzin SE. Plant microtechnique and microscopy. Vol. 198, New York: Oxford University Press; 1999.
[3] Benšić M, Sabo K. Uniform distribution width estimation from data observed with Laplace additive error. J Korean Stat Soc. 2016;45(4):505–517.
[4] Schneeweiss H. Estimating the endpoint of a uniform distribution under measurement errors. CEJOR. 2004;12:221–231.
[5] Wang XF, Wang B. Deconvolution estimation in measurement error models: the R package decon. J Stat Softw. 2011;39(10).
[6] Davidov O. Estimating the slope in measurement error models– a different perspective. Stat Probab Lett. 2005;71:215–223.

[7] Madsen CB, Christensen HI. Modeling and testing the stability of edge segments: length and orientation. Proceeding of the Ninth Scandinavian on Image Analysis; Uppsala, Sweden; 1995.
[8] McAssey M, Hsieh F. Slope estimation in structural line-segment heteroscedastic measurement error models. Stat Med. 2010;29:2631–2642.
[9] Vosselman G, Haralick RM. Performance analysis of line and circle fitting in digital images. Workshop on Performance Characteristics of Vision Algorithm; Cambridge; 1996.
[10] Benšić M, Sabo K. Estimating the width of a uniform distribution when data are measured with additive normal errors with known variance. Comp Stat Data Anal. 2007;51: 4731–4741.
[11] Benšić M, Sabo K. Estimating a uniform distribution when data are measured with a normal additive error with unknown variance. Statistics. 2010;44(3):235–246.
[12] Benšić M, Hamedović S, Sabo K. Taler.LeArEst: Border and area estimation of data measured with additive error, R package version 1.0.0. Available from: https://cran.r-project.org/web/packages/LeArEst/index.html.
[13] Benšić M, Sabo K. Border estimation of a disc observed with random errors solved in two steps. J Comput Appl Math. 2009;229:16–26.
[14] Garlipp T, Müller C. Detection of linear and circular shapes in image analysis. Comput Stat Data Anal. 2006;51:1479–1490.
[15] Baum M, Hanebeck U. Tracking a minimum bounding rectangle based on extreme value theory. 2010 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems; University of Utah, Salt Lake City, UT, USA; Sept. 5–7, 2010.
[16] Liu C, Mendonça L, Yang Y, et al. The architecture of inactivated SARS-CoV-2 with postfusion spikes revealed by Cryo-EM and Cryo-ET. Structure. 2020;28(11):1218–1224.

Appendices

Appendix 1

Proof of Theorem 2.1: From the classical Central Limit Theorem (CLT), $\mathbb{E}(X^4) < \infty$ implies that $\sqrt{n}\,(m_2 - \mu_2) \xrightarrow{D} \mathcal{N}(0, \mathrm{Var}(X^2))$. For the additive-error model, $\mathrm{Var}(X^2) = \frac{4}{45}a^4 + \frac{4}{3}a^2\sigma^2 + 2\sigma^4$. Now if we define $g(x) = \sqrt{3(x - \sigma^2)}$, then $g(\mu_2) = a$. Accordingly, by Delta method, we obtain $\sqrt{n}(\hat{a}_{\mathrm{mm}} - a) \xrightarrow{D} \mathcal{N}(0, \Psi^{(1)}_{\mathrm{mm}})$. The asymptotic variance of $\hat{a}_{\mathrm{mm}}$ for each model can be easily obtained now. Next we consider the case when σ^2 is unknown. Since the support of U is compact and the moments of a Gaussian distribution are finite, then assumption $\mathbb{E}(X^8) < \infty$, and as such

$$\sqrt{n}\,(\boldsymbol{m} - \boldsymbol{\mu}) \xrightarrow{D} \mathcal{N}_2(\mathbf{0}, \boldsymbol{\Pi}), \quad \text{where } \boldsymbol{\Pi} = \begin{bmatrix} \mu_4 - \mu_2^2 & \mu_6 - \mu_2\mu_4 \\ \cdot & \mu_8 - \mu_4^2 \end{bmatrix}, \tag{A1}$$

where $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$ and $\boldsymbol{m} = (m_1, m_2)^\top$. Now using Delta method, one can derive the asymptotic distribution of $\hat{\boldsymbol{\theta}}_{\mathrm{mm}} = (\hat{a}_{\mathrm{mm}}, \hat{\sigma}^2_{\mathrm{mm}})$, which depends on the moments of X. The non-central moment of X can be evaluated according to binomial formula, i.e.

$$\mu_k = \sum_{j=0}^{k} \binom{k}{j} \mathbb{E}(U^{k-j})\mathbb{E}(E^j).$$

For even values of $k-j$ and j, $\mathbb{E}(U^{k-j}) = \frac{a^{k-j}}{k-j+1}$ and $\mathbb{E}(E^j) = \frac{\sigma^j j!}{2^{j/2}(j/2!)}$, respectively; while all other odd powers for U and E have zero expectations. Therefore,

$$\begin{aligned} \mu_6 &= \frac{1}{7}\left(a^6 + 21a^4\sigma^2 + 105a^2\sigma^4 + 105\sigma^6\right) \\ \mu_8 &= \frac{1}{9}\left(a^8 + 36a^6\sigma^2 + 378a^4\sigma^4 + 1260a^2\sigma^6 + 945\sigma^8\right), \end{aligned} \tag{A2}$$

while μ_2 and μ_4 are given in (6). Now, recall that $\hat{\boldsymbol{\theta}}_{\text{mm}} = (\hat{a}_{\text{mm}}, \hat{\sigma}^2_{\text{mm}})^\top$ is given in (7) and (8). Also, define $g(\boldsymbol{m})$ by

$$g(\boldsymbol{m}) = \begin{bmatrix} \sqrt{\frac{15}{2}}\sqrt[4]{3m_2^2 - m_4} \\ m_2 - \sqrt{\frac{5}{6}}\sqrt{3m_2^2 - m_4} \end{bmatrix}.$$

Clearly g generates the MM of $\boldsymbol{\theta}$ under the additive model. Using (6) and (A2), the gradient of g with respect to $\boldsymbol{m}$, when *evaluated* at $\boldsymbol{\mu}$, is

$$g'(\boldsymbol{\mu}) = \left.\frac{\partial g(\boldsymbol{m})}{\partial \boldsymbol{m}^\top}\right|_{\boldsymbol{m}=\boldsymbol{\mu}} = \frac{15}{4a^2}\begin{bmatrix} \frac{a^2+3\sigma^2}{a} & -\frac{1}{2a} \\ -\frac{2(a^2+5\sigma^2)}{5} & \frac{1}{3} \end{bmatrix}.$$

After some algebra, the Delta method yields $\sqrt{n}(\widehat{\boldsymbol{\theta}}_{\text{mm}} - \boldsymbol{\theta}) \xrightarrow{D} \mathcal{N}_2(\mathbf{0}, \boldsymbol{\Psi}^{(1)}_{\text{mm}})$, where $\boldsymbol{\Psi}^{(1)}_{\text{mm}} = g'(\boldsymbol{\mu})\,\boldsymbol{\Pi}\,(g'(\boldsymbol{\mu}))^\top$, and the components of $\boldsymbol{\Psi}^{(1)}_{\text{mm}}$ are presented in Theorem 2.1. This completes the proof of Theorem 2.1. ■

Derivations of (19)–(22). We will first derive the MM estimators for $\boldsymbol{\theta} = (a, b)^\top$ for the case where σ^2 is known. One can now equate the first two sample moments with their corresponding population moments given in (18) to obtain $\hat{a}_{\text{mm}} = 2m_1 - \hat{b}_{\text{mm}}$ and

$$3m_2 = (b - m_1)^2 + 3m_1^2 + 3\sigma^2. \tag{A3}$$

Thus, the MM estimators of $\hat{\boldsymbol{\theta}}_{\text{mm}} = (\hat{a}_{\text{mm}}, \hat{b}_{\text{mm}})^\top$ exist if $\sigma^2 \le m_2 - m_1^2$ and equal to

$$\hat{a}_{\text{mm}} = m_1 - \sqrt{3(m_2 - m_1^2 - \sigma^2)}, \quad \hat{b}_{\text{mm}} = m_1 + \sqrt{3(m_2 - m_1^2 - \sigma^2)} \tag{A4}$$

Now, we will derive the MM estimators when the noise level is unknown. The third moment does not give us helpful information under the additive-error model. We need the fourth moment, which is attained after some algebraic manipulations by

$$m_4 = \frac{1}{5}\left((b - m_1)^4 + 10m_1^2(b - m_1)^2 + 5m_1^4)\right) + 2((b - m_1)^2 + 3m_1^2)\sigma^2 + 3\sigma^4. \tag{A5}$$

By Equation (A3), we eliminate b from Equation (A5)

$$5m_4 = \left(9(m_2 - m_1^2 - \sigma^2)^2 + 30m_1^2(m_2 - m_1^2 - \sigma^2) + 5m_1^4)\right) + 30(m_2 - \sigma^2)\sigma^2 + 15\sigma^4,$$

and, therefore,

$$\sigma^2 = m_2 - m_1^2 \pm \sqrt{\frac{10}{12}}\sqrt{m_2^2 - m_4 + 2(m_2^2 - m_1^4)}.$$

There are two possible solutions for σ^2, but its larger solution when replaced in (A4) would lead to complex-valued solutions for $\hat{a}$ and $\hat{b}$. Therefore, the MM estimator of σ^2 is given by the smaller solution, provided it is positive,

$$\hat{\sigma}^2_{\text{mm}} = m_2 - m_1^2 - \sqrt{\frac{10}{12}}\sqrt{m_2^2 - m_4 + 2(m_2^2 - m_1^4)}.$$

Then, under conditions C1 and C2 in (20), we have the MM estimators

$$\hat{a}_{\text{mm}} = m_1 - \sqrt{\frac{15}{2}}\sqrt[4]{m_2^2 - m_4 + 2(m_2^2 - m_1^4)}$$

$$\hat{b}_{\text{mm}} = m_1 + \sqrt{\frac{15}{2}}\sqrt[4]{m_2^2 - m_4 + 2(m_2^2 - m_1^4)}.$$

This completes the derivations of (19)–(22).

Proof of Theorem 3.1: We start with Part (i) by deriving the asymptotic variance of $\hat{\boldsymbol{\theta}}_{\text{mm}} = (\hat{a}_{\text{mm}}, \hat{b}_{\text{mm}})$. Since $\mathbb{E}(X^8) < \infty$, then μ_1 and μ_2 must be finite and also $\mathbb{E}(3(m_2 - m_1^2 - \sigma^2)) = \frac{(b-a)^2}{4}$. Accordingly, if we define

$$g(\boldsymbol{m}) = \begin{bmatrix} m_1 - \sqrt{(3(m_2 - m_1^2 - \sigma^2)} \\ m_1 + \sqrt{(3(m_2 - m_1^2 - \sigma^2)} \end{bmatrix}, \tag{A6}$$

where $\boldsymbol{m} = (m_1, m_2)^\top$, then one can now easily show that $g(\boldsymbol{\mu}) = \boldsymbol{\theta}$ and also its gradient with respect to $\boldsymbol{m}$, when evaluated at $\boldsymbol{\mu} = (\mu_1, \mu_2)^\top$, is

$$g'(\boldsymbol{\mu}) = \left.\frac{\partial g(\boldsymbol{m})}{\partial \boldsymbol{m}^\top}\right|_{\boldsymbol{m}=\boldsymbol{\mu}} = \begin{bmatrix} 1 + \frac{3(a+b)}{b-a} & \frac{-3}{b-a} \\ 1 - \frac{3(a+b)}{b-a} & \frac{3}{b-a} \end{bmatrix}.$$

Since $\mathbb{E}(X^4) < \infty$, then $\sqrt{n}(\boldsymbol{m} - \boldsymbol{\mu}) \xrightarrow{D} \mathcal{N}_2(\mathbf{0}; \boldsymbol{\Pi})$, where

$$\boldsymbol{\Pi} = \begin{bmatrix} \mu_2 - \mu_1^2 & \mu_3 - \mu_1\mu_2 \\ \cdot & \mu_4 - \mu_2^2 \end{bmatrix}.$$

The asymptotic variance can be established by using the Delta method. That is, $\sqrt{n}(\hat{\boldsymbol{\theta}}_{\text{mm}} - \boldsymbol{\theta}) \xrightarrow{D} \mathcal{N}_2(\mathbf{0}; \boldsymbol{\Psi}^{(2)}_{\text{mm}})$, where $\boldsymbol{\Psi}^{(2)}_{\text{mm}} = g'(\boldsymbol{\mu})\, \boldsymbol{\Pi}\, (g'(\boldsymbol{\mu}))^\top$, and the components of $\boldsymbol{\Psi}^{(2)}_{\text{mm}}$ are presented in (23).

Next we prove Part (ii), let us first define

$$g(x, y, z) = \left(x - \sqrt{\frac{15}{2}}\sqrt[4]{w},\ x + \sqrt{\frac{15}{2}}\sqrt[4]{w},\ y - x^2 - \sqrt{\frac{10}{12}w} \right)^\top,$$

where $w = 3y^2 - 2x^4 - z$. Notice that $3\mu_2^2 - 2\mu_1^4 - \mu_4 = (b-a)^4/120$. After some algebra, this identity can be employed to show that $g(\mu_1, \mu_2, \mu_4) = (a, b, \sigma^2)^\top$.

Hence, by the Delta Method, we have that $\sqrt{n}(\hat{\boldsymbol{\theta}}_{\text{mm}} - \boldsymbol{\theta}) \xrightarrow{D} \mathcal{N}_3(\mathbf{0}; \boldsymbol{\Psi}^{(3)}_{\text{mm}})$, where $\boldsymbol{\Psi}^{(3)}_{\text{mm}} = g'(\boldsymbol{\mu})\, \boldsymbol{\Pi}\, (g'(\boldsymbol{\mu}))^\top$ whose components are attained after a lengthy algebra and by using the following higher moments of $X = U + E$,

$$\begin{aligned}
\mu_5 &= \frac{b^6 - a^6}{6(b-a)} + 2.5\sigma^2 \frac{b^4 - a^4}{b-a} + 7.5\sigma^4(a+b), \\
\mu_6 &= \frac{b^7 - a^7}{7(b-a)} + 3\sigma^2 \frac{b^5 - a^5}{b-a} + 15\sigma^4 \frac{b^3 - a^3}{b-a} + 15\sigma^6, \\
\mu_8 &= \frac{b^9 - a^9}{9(b-a)} + 4\sigma^2 \frac{b^7 - a^7}{b-a} + 42\sigma^4 \frac{b^5 - a^5}{b-a} + 140\sigma^6 \frac{b^3 - a^3}{b-a} + 105\sigma^8.
\end{aligned}$$

Following the same steps of Part (i), one can derive the components of $\boldsymbol{\Psi}^{(3)}_{\text{mm}}$, which are given by

$$\begin{aligned}
\psi^{(3)}_{11} &= \frac{1}{21}\left(37a^2 + 37ab + 16b^2\right) + \chi_1 \\
\psi^{(3)}_{22} &= \frac{1}{21}\left(16a^2 + 37ab + 37b^2\right) + \chi_1 \\
\psi^{(3)}_{33} &= \frac{2}{315}(b-a)^2\left(19a^2 + 37ab + 19b^2\right) + \chi_2 \\
\psi^{(3)}_{12} &= -\frac{1}{21}\left(23a^2 + 44ab + 23b^2\right) + \chi_3 \\
\psi^{(3)}_{13} &= -\frac{1}{42}(b-a)\left(19a^2 + 29ab + 12b^2\right) + \chi_4 \\
\psi^{(3)}_{23} &= \frac{1}{21}\left(16a^2 + 37ab + 37b^2\right) + \chi_5,
\end{aligned} \tag{A7}$$

where

$$\chi_1 = \frac{4}{7}\left(88a^2 + 139ab + 88b^2\right)L_0^2 + 360\left(4a^2 + 7ab + 4b^2\right)L_0^4 + 7200\left(a^2 + ab + b^2\right)L_0^6 + 5400(b-a)^2L_0^8,$$

and

$$\chi_2 = \frac{2\,(b-a)^2}{315}\left[90\left(9a^2 + 17ab + 9b^2\right)L_0^2 + 945\left(27a^2 + 46ab + 27b^2\right)L_0^4 + 126000\left(a^2 + ab + b^2\right)L_0^6 + 94500(b-a)^2L_0^8\right],$$

$$\chi_3 = -\frac{1}{21}\left(6\left(169a^2 + 292ab + 169b^2\right)L_0^2 + 7560\left(4a^2 + 7ab + 4b^2\right)L_0^4 + 151200\left(a^2 + ab + b^2\right)L_0^6 + 113400L_0^6\sigma^2\right),$$

$$\chi_4 = -\frac{b-a}{42}\left(72\left(9a^2 + 17ab + 9b^2\right)L_0^2 + 5040\left(4a^2 + 7ab + 4b^2\right)L_0^4 + 100800\left(a^2 + ab + b^2\right)L_0^6 + 75600(b-a)^2L_0^8\sigma^3\right],$$

$$\chi_5 = \frac{4}{7}\left(88a^2 + 139ab + 88b^2\right)L_0^2 + 360\left(4a^2 + 7ab + 4b^2\right)L_0^4 + 7200\left(a^2 + ab + b^2\right)L_0^6 + 5400(b-a)^2L_0^8.$$

Here $L_0 = \sigma/(b-a)$ represents the ratio of the noise level to the length of the line segment. This completes the proof of Theorem 3.1. ■

Appendix 2. Supplementary material to *On estimating the boundaries of a uniform distribution under additive measurement errors*

This online supplementary material includes figures for the density for P-values and the rejection rates for testing the hypothesis $H_0 : a = -1$ versus $H_1 : a = a_1$ for the asymmetric cases. Figures 7 and 9 present the P-values computed under the null hypothesis and Figures A1 and 10 depict the rejection rates computed under the alternative hypothesis. In those cases, the lower bound was changed under the alternative $a_1 \in \{-1.04, -1.03, -1.02, -1.01, -0.99, -0.98, -0.97, -0.96\}$, while the upper bound was kept fixed and equal to 2. The settings for those numerical experiments are the same as the ones implemented here.

(1) Figure A1 shows the density estimates of the P-values computed based on the ML (dashed line) and the MM (solid line) approaches when σ is known.
(2) Figure A2 shows the rejection rates (power of the test) computed based on the ML (dashed line) and the MM (solid line) approaches when σ is known.
(3) Figure A3 shows the density estimates of the P-values computed based on the ML (dashed line) and the MM (solid line) approaches when σ is unknown.
(4) Figure A4 shows the rejection rates (power of the test) computed based on the ML (dashed line) and the MM (solid line) approaches when σ is unknown.

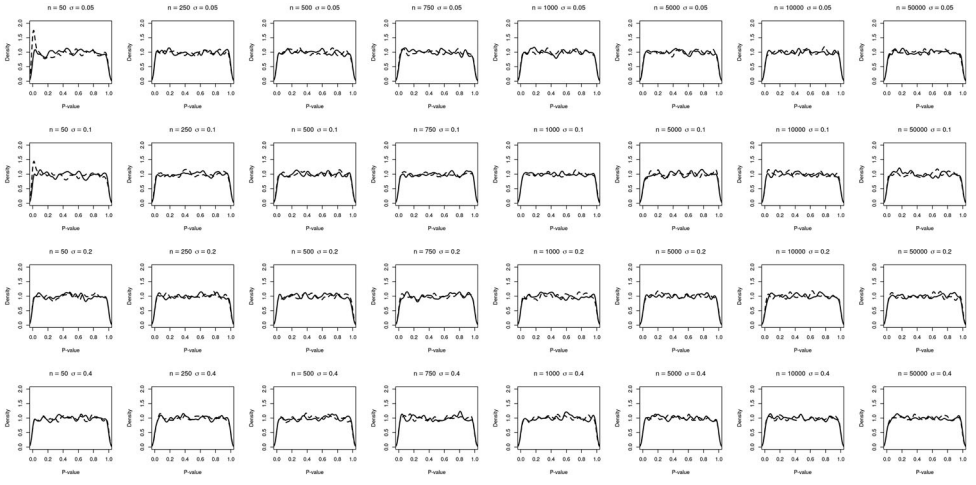


Figure A1. Density estimates of the P-values computed based on the ML (dashed line) and the MM (solid line) approaches under $H_0 : a = 1$. The upper bound of the interval was kept fixed and equal to 2. The tests have been conducted under the assumption that the noise level σ is known.

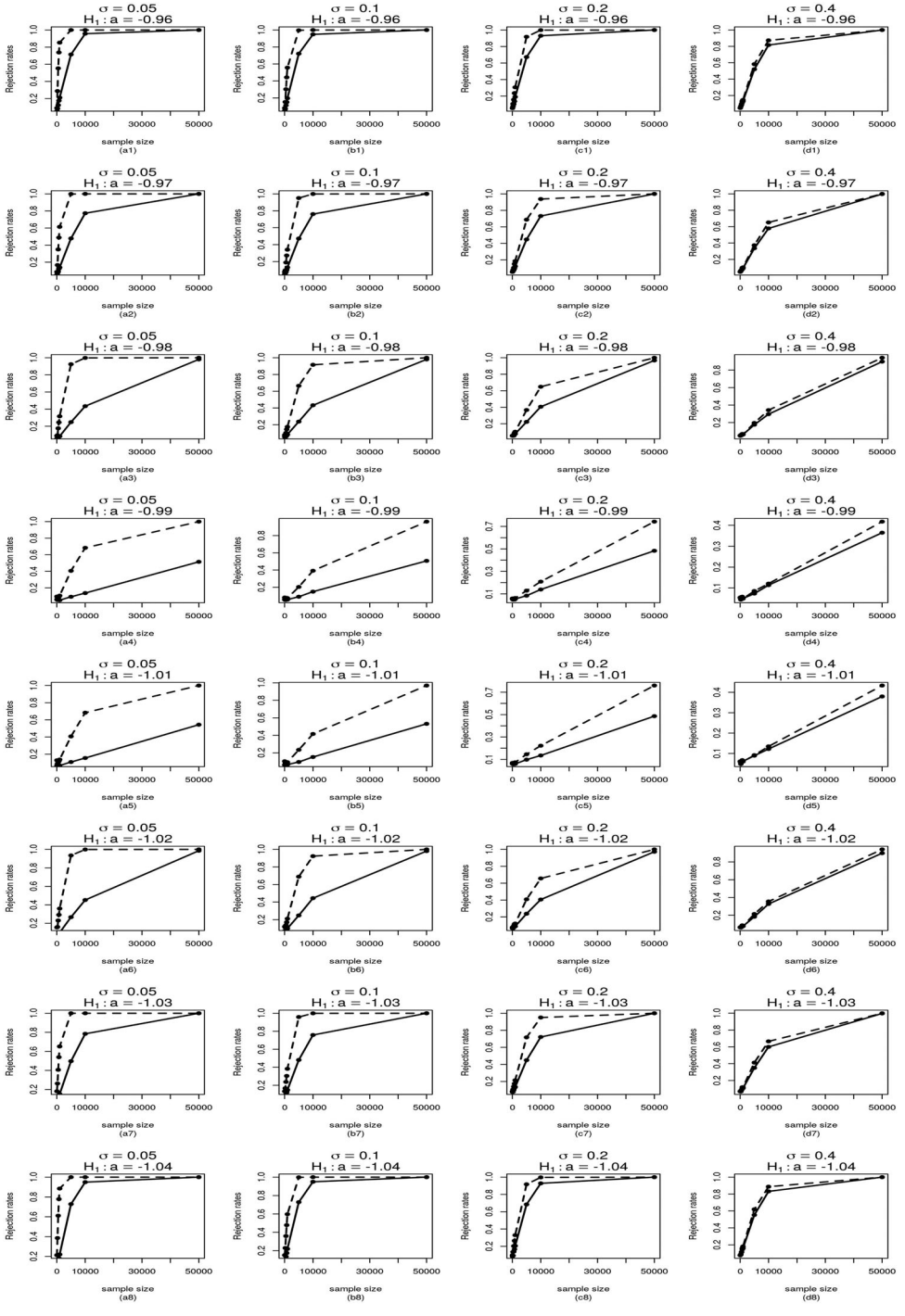


Figure A2. Rejection rates computed for testing $H_0 : a = 1$ under the alternative hypothesis $H_1 : a = a_1$ using the MM (dashed line) and the ML (solid line) approaches. The lower bound was changed under the alternative $a_1 \in \{-1.04, -1.03, -1.02, -1.01, -0.99, -0.98, -0.97, -0.96\}$, while the upper bound of the interval was kept fixed and equal to 2. The tests have been conducted under the assumption that the noise level σ is known.

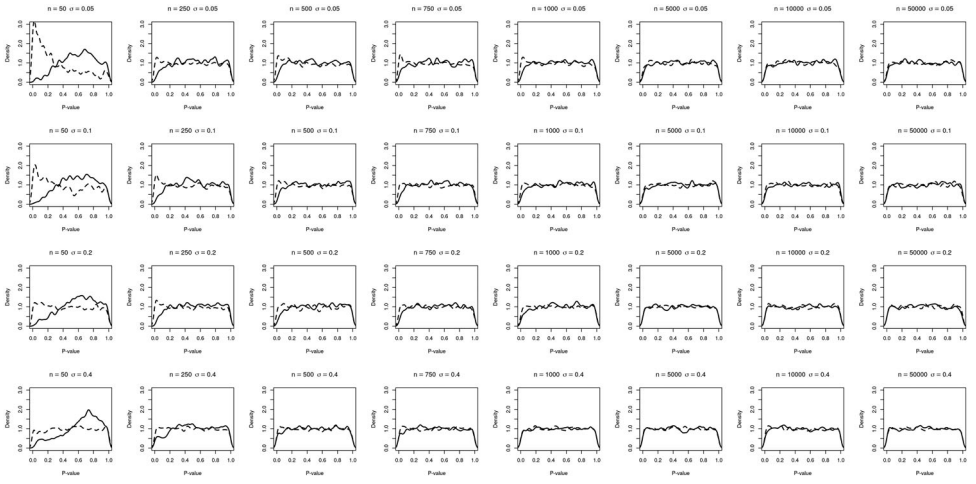


Figure A3. Density estimates of the P-values computed based on the ML (dashed line) and the MM (solid line) approaches under $H_0 : a = -1$ while the upper bound of the interval was kept fixed and equal to 2. The tests have been conducted under the assumption that the noise level σ is unknown.

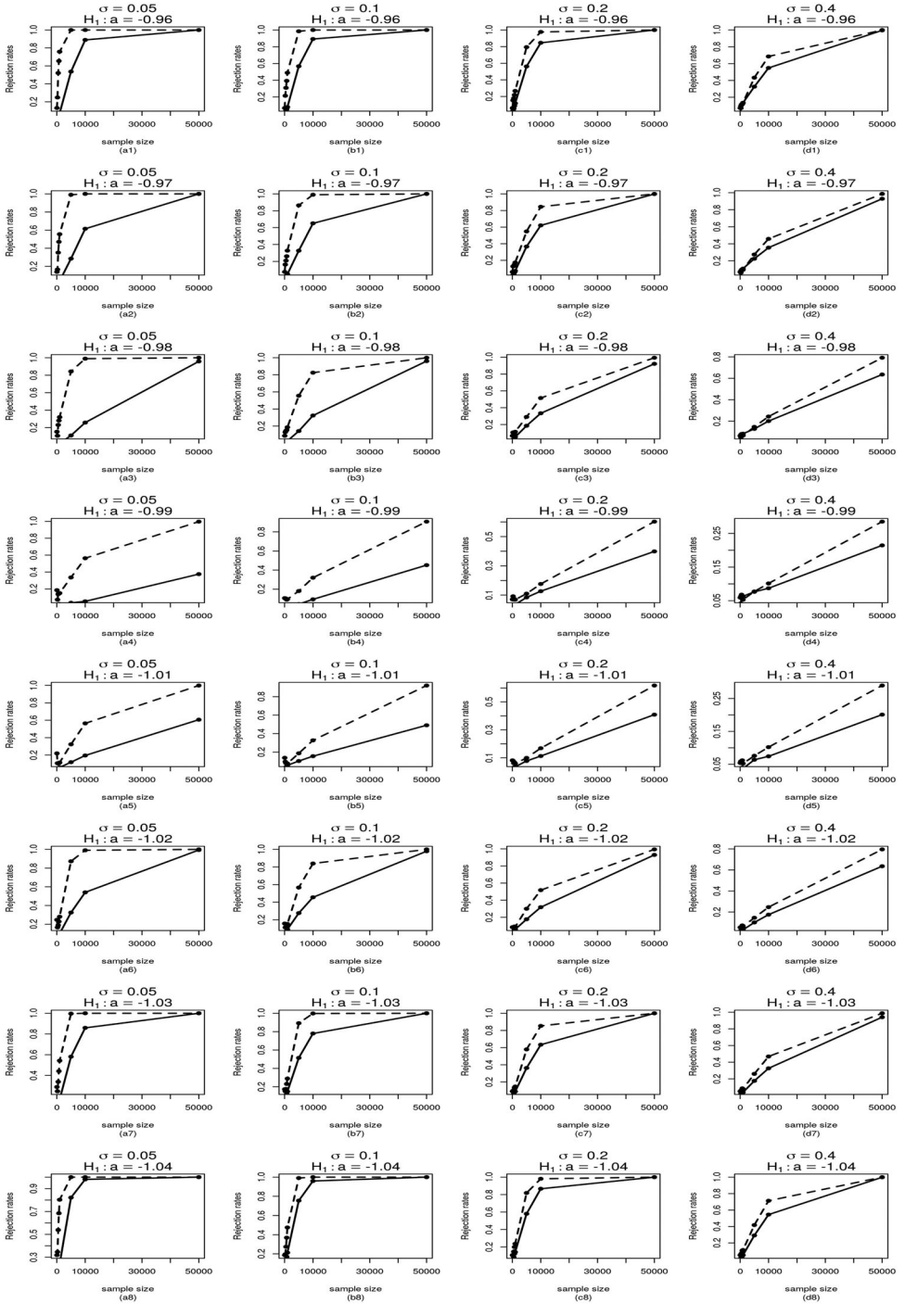


Figure A4. Rejection rates computed for testing $H_0 : a = 1$ under the alternative hypothesis $H_1 : a = a_1$ using the MM (dashed line) and the ML (solid line) approaches. The lower bound was changed under the alternative $a_1 \in \{-1.04, -1.03, -1.02, -1.01, -0.99, -0.98, -0.97, -0.96\}$, while the upper bound of the interval was kept fixed and equal to 2. The tests have been conducted under the assumption that the noise level σ is unknown.