Journal of Statistical Planning and Inference 157–158 (2015) 54–76

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Rigorous error control methods for estimating means of bounded random variables

Zhengjia Chen [a,∗], Xinjia Chen [b]

[a] Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA 30322, United States

[b] Department of Electrical Engineering, Southern University at Baton Rouge, LA 70813, United States

ARTICLE INFO

Article history:
Received 22 April 2013
Received in revised form 27 August 2014
Accepted 28 August 2014
Available online 6 September 2014

Keywords:
Rigorous error control method
Mean estimation
Bounded random variable

ABSTRACT

In this article, we propose rigorous sample size methods for estimating the means of random variables, which require no information of the underlying distributions except that the random variables are known to be bounded in a certain interval. Our sample size methods can be applied without assuming that the samples are identical and independent. Moreover, our sample size methods involve no approximation. We demonstrate that the sample complexity can be significantly reduced by using a mixed error criterion. We derive explicit sample size formulae to ensure the statistical accuracy of estimation.

1. Introduction

Many problems of engineering and sciences boil down to estimating the mean value of a random variable (Mitzenmacher and Upfal, 2005; Motwani and Raghavan, 1995). More formally, let X be a random variable with mean μ. It is a frequent problem to estimate μ based on samples $X_1, X_2, \ldots, X_n$ of X, which are defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P}_\mu)$, where the subscript in the probability measure $\mathbb{P}_\mu$ indicates its association with μ. In many situations, the information on the distribution of X is not available except that X is known to be bounded in some interval $[a, b]$. For example, in clinical trials, many quantities under investigation are bounded random variables, such as biomarker, EGFR, K-Ras, B-Raf, Akt, etc. (see., e.g., Arellano et al., 2012; Janik et al., 2010; Wang et al., 2012, and the references therein). Moreover, the samples $X_1, X_2, \ldots, X_n$ may not be identical and independent (i.i.d). This gives rise to the significance of estimating μ under the assumption that

$$a \leq X_k \leq b \quad \text{almost surely for } k \in \mathbb{N}, \tag{1}$$

$$\mathbb{E}[X_k \mid \mathscr{F}_{k-1}] = \mu \quad \text{almost surely for } k \in \mathbb{N}, \tag{2}$$

where $\mathbb{N}$ denotes the set of positive integers, and $\{\mathscr{F}_k,\ k = 0, 1, \ldots, \infty\}$ is a sequence of σ-subalgebras such that $\{\emptyset, \Omega\} = \mathscr{F}_0 \subset \mathscr{F}_1 \subset \mathscr{F}_2 \subset \ldots \subset \mathscr{F}$, with $\mathscr{F}_k$ being generated by $X_1, \ldots, X_k$. The motivation we propose to consider the estimation of μ under dependency assumption (2) is twofold. First, from a theoretical point of view, we want the results to hold under the most general conditions. Clearly, (2) is satisfied in the special case that $X_1, X_2, \ldots$ are i.i.d. Second, from a practical standpoint, we want to weaken the independency assumption for more applications. For example, in the Monte Carlo estimation technique based on adaptive importance sampling, the samples $X_1, X_2, \ldots$ are not necessarily independent. However, as demonstrated in page 6 of Gajek et al. (2013), it may be shown that the samples satisfy (2). An example of adaptive importance sampling is given in Section 5.8 of Fishman (1996) on the study of catastrophic failure.

∗ Corresponding author.
E-mail addresses: zchen38@emory.edu (Z. Chen), xinjia_chen@subr.edu (X. Chen).

http://dx.doi.org/10.1016/j.jspi.2014.08.007
0378-3758/

An unbiased estimator for μ can be taken as

$$\overline{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}.$$

Let $\varepsilon \in (0, 1)$ and $\delta \in (0, 1)$ be pre-specified margin of absolute error and confidence parameter, respectively. Since the probability distributions of $X_1, X_2, \ldots$ are usually unknown, one would use an absolute error criterion and seek the sample size, n, as small as possible such that for all values of μ,

$$\mathbb{P}_\mu \left\{ |\overline{X}_n - \mu| < \varepsilon \right\} > 1 - \delta \tag{3}$$

holds for all distributions having common mean μ. It should be noted that it is difficult to specify a margin of absolute error ε, without causing undue conservatism, for controlling the accuracy of estimation if the underlying mean value μ can vary in a wide range. To achieve acceptable accuracy, it is necessary to choose small ε for small μ. However, this leads to unnecessarily large sample sizes for large μ.

In addition to the absolute error criterion, a relative error criterion is frequently used for the purpose of error control. Let $\eta \in (0, 1)$ and $\delta \in (0, 1)$ be the pre-specified margin of relative error and confidence parameter, respectively. It is desirable to determine the sample size, n, as small as possible such that for all values of μ,

$$\mathbb{P}_\mu \left\{ \left|\overline{X}_n - \mu\right| < \eta|\mu| \right\} > 1 - \delta \tag{4}$$

holds for all distributions having common mean μ. Unfortunately, the determination of sample size, n, requires a good lower bound for μ, which is usually not available. Otherwise, the sample size n needs to be very large, or infinity.

To overcome the aforementioned difficulties, a mixed criterion may be useful. The reason is that, from a practical point of view, an estimate can be acceptable if either an absolute criterion or a relative criterion is satisfied. More specifically, let $\varepsilon > 0$, $\eta \in (0, 1)$ and $\delta \in (0, 1)$. To control the reliability of estimation, it is crucial that the sample size n is as small as possible, such that for all values of μ,

$$\mathbb{P}_\mu \left\{ |\overline{X}_n - \mu| < \varepsilon \text{ or } |\overline{X}_n - \mu| < \eta|\mu| \right\} > 1 - \delta \tag{5}$$

holds for all distributions having common mean μ.

In the estimation of parameters, a margin of absolute error is usually chosen to be much smaller than the margin of relative error. For instance, in the estimation of a binomial proportion, a margin of relative error $\eta = 0.1$ may be good enough for most situations, while a margin of absolute error may be expected to be $\varepsilon = 0.001$ or even smaller. In many applications, a practitioner accepting a relative error normally expects a much smaller absolute error, i.e., $\varepsilon \ll \eta$. On the other hand, one accepting an absolute error ε typically tolerates a much larger relative error, i.e., $\eta \gg \varepsilon$. It will be demonstrated that the required sample size can be substantially reduced by using a mixed error criterion.

Given that the measure of precision is chosen, the next task is to determine appropriate sample sizes. A conventional method is to determine the sample size by normal approximation derived from the central limit theorem (Chow et al., 2008; Desu and Raghavarao, 1990). Such an approximation method inevitably leads to unknown statistical error due to the fact that the sample size n must be a finite number (Fishman, 1996; Hampel, 1998). This motivates us to explore rigorous methods for determining sample sizes.

In this paper, we consider the problem of estimating the means of bounded random variables based on a mixed error criterion. The remainder of the paper is organized as follows. In Section 2, we introduce some martingale inequalities. In Section 3, we derive explicit sample size formulae by virtue of concentration inequalities and martingale inequalities. In Section 4, we extend the techniques to the problem of estimating the difference of means of two bounded random variables. Illustrative examples are given in Section 5. Section 6 provides our concluding remarks. Most proofs are given in Appendices.

2. Martingale inequalities

Under assumption (2), it can be readily shown that $\{X_k - \mu\}$ is actually a sequence of martingale differences (see, e.g. Doob, 1953; Willams, 1991, and the references therein). In the sequel, we shall introduce some martingale inequalities which are crucial for the determination of sample sizes to guarantee pre-specified statistical accuracy.

Define function

$$\psi(\varepsilon, \mu) = (\mu + \varepsilon) \ln \left(\frac{\mu + \varepsilon}{\mu} \right) + (1 - \mu - \varepsilon) \ln \left(\frac{1 - \mu - \varepsilon}{1 - \mu} \right)$$

for $0 < \varepsilon < 1 - \mu < 1$. Under the assumption that $0 \le X_k \le 1$ almost surely and (2) holds for all $k \in \mathbb{N}$, Hoeffding (1963) established that

$$\mathbb{P}_\mu \{\overline{X}_n \ge \mu + \varepsilon\} < \exp\left(-n\psi(\varepsilon, \mu)\right) \quad \text{for } 0 < \varepsilon < 1 - \mu. \tag{6}$$

To see that such result is due to Hoeffding, see Theorem 1 and the remarks on page 18, the second paragraph, of his paper (Hoeffding, 1963). For bounds tighter than Hoeffding's inequality, see a recent paper (Bentkus, 2004).

To obtain simpler probabilistic inequalities, define bivariate function

$$\varphi(\varepsilon, \mu) = \frac{\varepsilon^2}{2\left(\mu + \frac{\varepsilon}{3}\right)\left(1 - \mu - \frac{\varepsilon}{3}\right)}.$$

It is shown by Massart (1990) that

$$\psi(\varepsilon, \mu) > \varphi(\varepsilon, \mu). \tag{7}$$

By virtue of Hoeffding's inequality and Massart's inequality, the following results can be justified.

Theorem 1. *Assume that* $0 \leq X_k \leq 1$ *almost surely and* (2) *holds for all* $k \in \mathbb{N}$. *Then,*

$$\mathbb{P}_\mu\{\overline{X}_n \geq \mu + \varepsilon\} \leq \exp\left(-n\varphi(\varepsilon, \mu)\right) \quad \textit{for } 0 < \varepsilon < 3(1 - \mu), \tag{8}$$

$$\mathbb{P}_\mu\{\overline{X}_n \leq \mu - \varepsilon\} \leq \exp\left(-n\varphi(-\varepsilon, \mu)\right) \quad \textit{for } 0 < \varepsilon < 3\mu. \tag{9}$$

Proof. To prove Theorem 1, note that

$$\mathbb{P}_\mu\{\overline{X}_n \geq \mu + \varepsilon\} = 0 < \exp\left(-n\varphi(\varepsilon, \mu)\right) \quad \text{for } \varepsilon > 1 - \mu. \tag{10}$$

From Hoeffding's inequality (6) and Massart's inequality (7), we have

$$\mathbb{P}_\mu\{\overline{X}_n \geq \mu + \varepsilon\} < \exp\left(-n\varphi(\varepsilon, \mu)\right) \quad \text{for } 0 < \varepsilon < 1 - \mu. \tag{11}$$

Observe that $\mathbb{P}_\mu\{\overline{X}_n \geq z\}$ is a left-continuous function of z and that $\varphi(\varepsilon, \mu)$ is a continuous function of ε. Making use of this observation and (11), we have

$$\mathbb{P}_\mu\{\overline{X}_n \geq \mu + \varepsilon\} \leq \exp\left(-n\varphi(\varepsilon, \mu)\right) \quad \text{for } \varepsilon = 1 - \mu. \tag{12}$$

Note that

$$\mathbb{P}_\mu\{\overline{X}_n \geq \mu + \varepsilon\} \leq \mathbb{P}_\mu\{\overline{X}_n > 1\} = 0 \leq \exp\left(-n\varphi(\varepsilon, \mu)\right) \quad \text{for } 1 - \mu < \varepsilon < 3(1 - \mu). \tag{13}$$

Combining (10)–(13) yields

$$\mathbb{P}_\mu\{\overline{X}_n \geq \mu + \varepsilon\} \leq \exp\left(-n\varphi(\varepsilon, \mu)\right) \quad \text{for } 0 < \varepsilon < 3(1 - \mu).$$

This proves (8). To show (9), define $Y_i = 1 - X_i$ for $i = 1, \ldots, n$. Define $\overline{Y}_n = 1 - \overline{X}_n$ and $\nu = 1 - \mu$. Then, $\mathbb{E}[\overline{Y}_n] = \nu$. Applying (8), we have

$$\mathbb{P}_\mu\left\{\overline{Y}_n \geq \nu + \varepsilon\right\} \leq \exp\left(-n\varphi(\varepsilon, \nu)\right)$$

for $0 < \varepsilon < 3(1 - \nu)$. By the definitions of ν and $\overline{Y}_n$, we can rewrite the above inequality as

$$\mathbb{P}_\mu\left\{\overline{Y}_n \geq \nu + \varepsilon\right\} = \mathbb{P}_\mu\left\{1 - \overline{Y}_n \leq \mu - \varepsilon\right\} \leq \exp\left(-n\varphi(\varepsilon, 1 - \mu)\right)$$

for $0 < \varepsilon < 3\mu$. Observing that $\overline{X}_n = 1 - \overline{Y}_n$ and that $\varphi(\varepsilon, 1 - \mu) = \varphi(-\varepsilon, \mu)$, we have (9). This completes the proof of Theorem 1. □

It should be noted that Theorem 1 extends Massart's inequality in two aspects. First, the random variables are not required to be i.i.d. Bernoulli random variables. Second, the inequalities hold for wider supports.

3. Explicit sample size formulae

In this section, we shall investigate sample size methods for estimating the mean of bounded random variable X.

If $X_1, \ldots, X_n$ are i.i.d. samples of X bounded in interval $[0, 1]$, it can be shown by Chebyshev's inequality that (3) holds provided that

$$n \geq \frac{1}{4\delta\varepsilon^2}. \tag{14}$$

Under the assumption that $0 \leq X_k \leq 1$ and $\mathbb{E}[X_k \mid \mathscr{F}_{k-1}] = \mu$ almost surely for all $k \in \mathbb{N}$, Azuma–Hoeffding inequality (Azuma, 1967; Hoeffding, 1963) implies that (3) holds for all $\mu \in (0, 1)$ if

$$n > \frac{\ln \frac{2}{\delta}}{2\varepsilon^2}. \tag{15}$$

Clearly, the ratio of the sample size determined by (14) to that of (15) is equal to

$$\frac{\frac{1}{4\delta\varepsilon^2}}{\frac{\ln\frac{2}{\delta}}{2\varepsilon^2}} = \frac{1}{2\delta \ln\frac{2}{\delta}},$$

which is substantially greater than 1 for small $\delta \in (0, 1)$. Despite the significant improvement upon the sample size bound (14), the sample size bound (15) usually leads to a very large sample size, since ε is typically a small number in practice. For example, with $\delta = 0.05$, we have $n = 1{,}844{,}440$ and $n = 184{,}443{,}973$ for $\varepsilon = 0.001$ and 0.0001, respectively.

To the best of our knowledge, the sample size bound (15) is the tightest one discovered so far under the assumption that $0 \leq X_k \leq 1$ and $\mathbb{E}[X_k \mid \mathscr{F}_{k-1}] = \mu$ almost surely for all $k \in \mathbb{N}$. In order to reduce the sample complexity, we propose to use the mixed error criterion, which can be viewed as a relaxation of the absolute error criterion. In this direction, we have exploited the application of Chebyshev's inequality to establish the following result.

Theorem 2. *If $X_1, \ldots, X_n$ are i.i.d. samples of X bounded in interval $[0, 1]$, then (5) holds for all $\mu \in (0, 1)$ provided that $\lambda = \frac{\varepsilon}{\eta} \leq \frac{1}{2}$ and that*

$$n > \frac{1-\lambda}{\delta\varepsilon\eta}. \tag{16}$$

See Appendix A for proof.

The sample size formula (16) may be too conservative. To derive tighter sample size formulae, we need to use the martingale inequalities of exponential form presented in the last section. Throughout the remainder of this section, we make the following assumption:

$X_1, X_2, \ldots$ are random variables such that $a \leq X_k \leq b$ and $\mathbb{E}[X_k \mid \mathscr{F}_{k-1}] = \mu$ almost surely for all $k \in \mathbb{N}$.

In the case that $X_1, X_2, \ldots$ are nonnegative random variables, we have the following general result.

Theorem 3. *Let $0 \leq a < b$. Assume that $0 < \varepsilon < b - a$ and $\eta \in (0, \frac{3}{2})$ and that $a < \frac{\varepsilon}{\eta} < b$. Define*

$$N = \begin{cases} \frac{2}{\varepsilon^2}\left(\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} - a\right)\left(b - \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta} & \text{for } \frac{2ab}{a+b} \leq \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} \leq \frac{a+b}{2}, \\ \frac{(b-a)^2}{2ab}\left(\frac{1}{\eta} + \frac{1}{3}\right)^2 \ln\frac{2}{\delta} & \text{for } \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} < \frac{2ab}{a+b}, \\ \frac{(b-a)^2}{2\varepsilon^2}\ln\frac{2}{\delta} & \text{for } \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} > \frac{a+b}{2} \end{cases}$$

and

$$M = \begin{cases} \frac{2}{\varepsilon^2}\left(\frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} - a\right)\left(b - \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta} & \text{for } \frac{2ab}{a+b} \leq \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} \leq \frac{a+b}{2}, \\ \frac{(b-a)^2}{2ab}\left(\frac{1}{\eta} - \frac{1}{3}\right)^2 \ln\frac{2}{\delta} & \text{for } \frac{\eta}{3} < \frac{b-a}{b+a} \text{ and } \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} < \frac{2ab}{a+b}, \\ \left[\frac{2}{3\eta}\left(1 - \frac{a}{b}\right) - \frac{2}{9}\right]\ln\frac{2}{\delta} & \text{for } \frac{b-a}{b+a} < \frac{\eta}{3} < \frac{b-a}{b} \text{ and } \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} \leq \frac{a+b}{2}, \\ 1 & \text{for } \frac{\eta}{3} \geq \frac{b-a}{b}, \\ \frac{(b-a)^2}{2\varepsilon^2}\ln\frac{2}{\delta} & \text{else}. \end{cases}$$

Then, $\mathbb{P}_\mu\{|\overline{X}_n - \mu| < \varepsilon \text{ or } |\overline{X}_n - \mu| < \eta\mu\} > 1 - \delta$ for any $\mu \in (a, b)$ provided that $n > \max(N, M)$.

See Appendix B for proof. In Theorem 3, our purpose of assuming $\varepsilon < b - a$ and $a < \frac{\varepsilon}{\eta} < b$ is to make sure that the absolute error criterion is active for some $\mu \in (a, b)$ and that the relative error criterion is active for some $\mu \in (a, b)$. In Table 1, we list sample sizes for $b = 1$, $\varepsilon = 0.001$, $\eta = 0.1$, $\delta = 0.01$ and various values of a, where N_{mix} denotes the sample sizes calculated by virtue of Theorem 3 and the mixed error criterion, and N_{abs} denotes the sample sizes obtained from the Chernoff–Hoeffding bound. More precisely,

$$N_{\text{mix}} = \lceil \max(N, M) \rceil$$

and

$$N_{\text{abs}} = \left\lceil \frac{(b-a)^2 \ln\frac{2}{\delta}}{2\varepsilon^2} \right\rceil, \tag{17}$$

Table 1
Table of sample sizes.

a	N_{mix}	N_{abs}	a	N_{mix}	N_{abs}
0.001	97,880	499,001	0.006	40,765	494,019
0.002	87,393	498,002	0.007	34,871	493,025
0.003	76,906	497,005	0.008	30,451	492,032
0.004	66,419	496,008	0.009	27,013	491,041
0.005	55,932	495,013	0.01	24,263	490,050

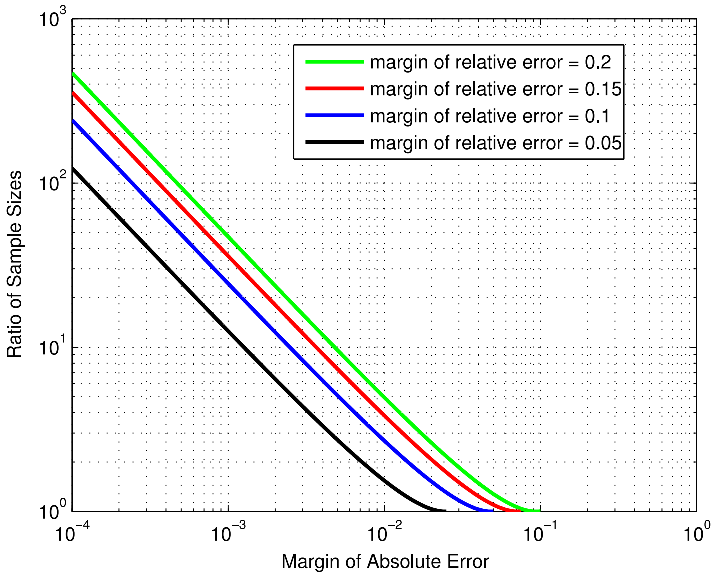

Fig. 1. Comparison of sample sizes.

where $\lceil . \rceil$ denotes the ceiling function. It can be seen from the table that the sample complexity can be significantly reduced by using a mixed error criterion and our sample size formula.

As an immediate application of Theorem 3, we have the following result.

Corollary 1. *Let ε and η be respectively the margins of absolute and relative errors such that*

$$\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} < \frac{1}{2}. \tag{18}$$

Assume that $0 \leq X_k \leq 1$ almost surely for all $k \in \mathbb{N}$. Then, (5) holds for all $\mu \in (0, 1)$ provided that

$$n > 2\left(\frac{1}{\eta} + \frac{1}{3}\right)\left(\frac{1}{\varepsilon} - \frac{1}{\eta} - \frac{1}{3}\right)\ln\frac{2}{\delta}. \tag{19}$$

It should be noted that (18) can be readily satisfied in practice, since $0 < \varepsilon \ll \eta < 1$ is true in most applications.

An appealing feature of formula (19) is that the resultant sample size is much smaller as compared to that of (15) and (16). Moreover, to apply (19), no approximation is involved and no information of μ is needed. Furthermore, the samples need not be i.i.d.

Under the condition that $0 < \varepsilon \ll \eta \ll 1$, the sample size bound of (19) can be approximated as

$$\frac{2\ln\frac{2}{\delta}}{\varepsilon\eta},$$

which indicates that the required sample size is inversely proportional to the product of margins of absolute and relative errors. It can be shown that the ratio of the bound of (16) to the sample size bound of (19) converges to 0 as δ decreases to 0, which implies that the bound of (19) is better for small δ.

The comparison of sample size formulae (15) and (19) is shown in Fig. 1, where it can be seen that the sample size formula (19) leads to a substantial reduction in sample complexity as compared to (15).

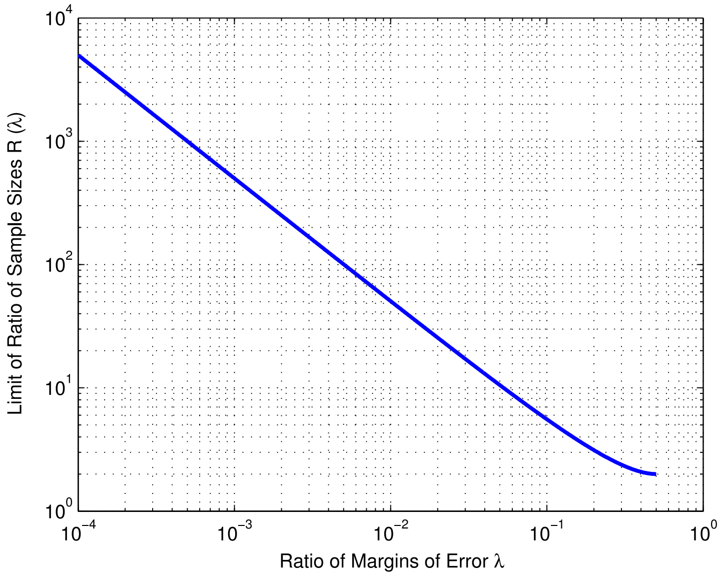

Fig. 2. Limit of ratio of sample sizes.

To obtain more insight into such a reduction of sample size, we shall investigate the ratio of the sample sizes, which is given as

$$\frac{\frac{\ln \frac{2}{\delta}}{2\varepsilon^2}}{2\left(\frac{1}{\eta}+\frac{1}{3}\right)\left(\frac{1}{\varepsilon}-\frac{1}{\eta}-\frac{1}{3}\right)\ln\frac{2}{\delta}} = \frac{1}{4\left(\lambda+\frac{\varepsilon}{3}\right)\left(1-\lambda-\frac{\varepsilon}{3}\right)}.$$

Let $\varepsilon \in (0, 1)$ and $\eta \in (0, 1)$ such that (18) holds. When no information of μ is available except that μ is known to be bounded in $(0, 1)$, the best known sample size bound is given by (15), which asserts that (3) holds for any $\mu \in (0, 1)$ provided that (15) holds. According to Corollary 1, we have that (5) holds for any $\mu \in (0, 1)$ provided that (19) is true. In view of (15) and (19), the ratio of the sample sizes tends to

$$R(\lambda) \stackrel{\text{def}}{=} \frac{1}{4\lambda(1-\lambda)}$$

as $\varepsilon \to 0$ under the restriction that $\lambda = \frac{\eta}{\varepsilon}$ is fixed.

From Fig. 2, it can be seen that the limiting ratio, $R(\lambda)$, of sample sizes is substantially greater than 1 for small $\lambda > 0$. For example, if $\varepsilon = 10^{-5}$ and $\eta = 0.1$, we have $\lambda = \frac{\varepsilon}{\eta} = 10^{-4}$ and $R(\lambda) \approx 2500$. This demonstrates that the required sample size can be significantly reduced by virtue of a mixed error criterion. As mentioned earlier, for small η (e.g. $\eta = 0.1$), the result (5) can be viewed as a slight relaxation of the result (3). Our analysis indicates that such a slight relaxation is well worthy of the significant reduction in the sample complexity.

In Theorem 3, the random variables $X_1, X_2, \ldots$ are assumed to be non-negative. In light of the fact that, in some situations, the random variables may assume positive or negative values, we have derived explicit sample size formula in the following result.

Theorem 4. *Let $a < 0 < b$. Assume that $0 < \varepsilon < b - a$ and $\eta \in (0, \frac{3}{2})$ and that $\varepsilon < \eta \max(|a|, b)$. Define*

$$M = \begin{cases} \frac{2}{\varepsilon^2}\left[|a+b|\left(\frac{\varepsilon}{\eta}+\frac{\varepsilon}{3}\right) - \left(\frac{\varepsilon}{\eta}+\frac{\varepsilon}{3}\right)^2 - ab\right]\ln\frac{2}{\delta} & \text{for } \frac{|a+b|}{2} > \frac{\varepsilon}{\eta}+\frac{\varepsilon}{3}, \\ \frac{(b-a)^2}{2\varepsilon^2}\ln\frac{2}{\delta} & \text{else} \end{cases}$$

Then, $\mathbb{P}_\mu\{|\overline{X}_n - \mu| < \varepsilon$ or $|\overline{X}_n - \mu| < \eta|\mu|\} > 1 - \delta$ for any $\mu \in (a, b)$ provided that $n > M$.

See Appendix C for proof.

It should be noted that the advantage of using the mixed error criterion is more pronounced if the interval $[a, b]$ contains 0 and is more asymmetrical about 0. As an illustration, consider the configuration with $\varepsilon = 0.1$, $\eta = 0.1$ and $\delta = 0.05$. Assume that the lower bound, a, of the interval is fixed as -1 and the upper bound, b, of the interval is a parameter. From formula (15), we know that the sample size required to ensure $\mathbb{P}_\mu\left\{\left|\overline{X}_n - \mu\right| < \varepsilon\right\} \geq 1-\delta$ for any $\mu \in [a, b]$ can be obtained from (17).

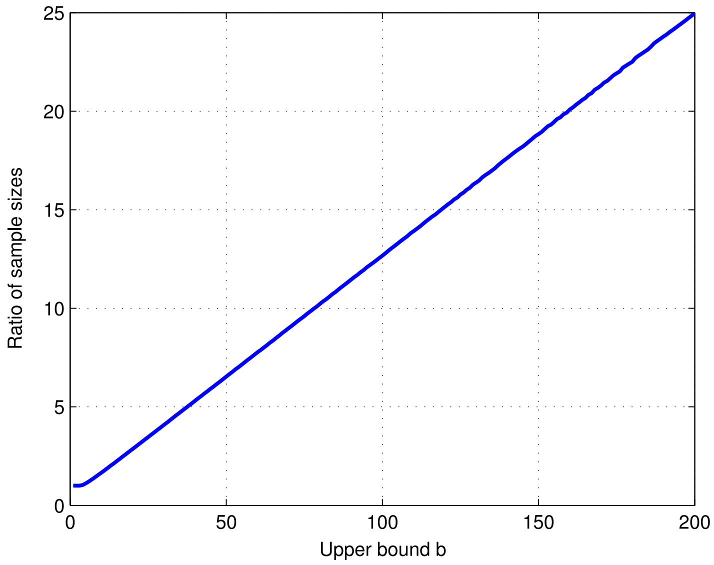

Fig. 3. Ratio of sample sizes ($\varepsilon = 0.1$, $\eta = 0.1$, $\delta = 0.05$ and $a = -1$).

According to Theorem 4, the sample size required to ensure $\mathbb{P}_\mu\{\left|\overline{X}_n - \mu\right| < \varepsilon \text{ or } \left|\overline{X}_n - \mu\right| < \eta|\mu|\} \geq 1 - \delta$ for any $\mu \in [a, b]$ can be calculated as

$$N_{\text{mix}} = \left\lceil \frac{2}{\varepsilon^2}\left[|a+b|\left(\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3}\right) - \left(\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3}\right)^2 - ab\right] \ln\frac{2}{\delta} \right\rceil.$$

Since a is fixed, the ratio, $\frac{N_{\text{abs}}}{N_{\text{mix}}}$, of sample sizes is a function of b. Such a function is shown by Fig. 3, from which it can be seen that the larger b is, the greater the reduction of sample size can be achieved by virtue of a mixed error criterion.

4. Estimating the difference of two population means

Our method can be extended to the estimation of the difference of means of bounded random variables. Let Y and Z be two bounded random variables such that $\mathbb{E}[Y] = \mu_Y$ and $\mathbb{E}[Z] = \mu_Z$. Let $X = Y - Z$ and $\mu = \mu_Y - \mu_Z$. Let $Y_1, \dots Y_n$ be i.i.d. samples of Y. Let $Z_1, \dots Z_n$ be i.i.d. samples of Z. Assume that the samples of Y and Z are independent. Let $X_i = Y_i - Z_i$ for $i = 1, 2, \dots, n$. Then, $X_1, \dots, X_n$ are i.i.d. samples of X. Clearly, X is a bounded random variable. So are $X_1, \dots, X_n$. Define

$$\overline{X}_n = \frac{\sum_{i=1}^n X_i}{n}, \qquad \overline{Y}_n = \frac{\sum_{i=1}^n Y_i}{n}, \qquad \overline{Z}_n = \frac{\sum_{i=1}^n Z_i}{n}.$$

Then, $\overline{X}_n = \overline{Y}_n - \overline{Z}_n$ is an estimator for $\mu = \mu_Y - \mu_Z$. We can apply the sample size methods proposed in Section 3 to determine n such that

$$\mathbb{P}_\mu\left\{|\overline{X}_n - \mu| < \varepsilon \text{ or } |\overline{X}_n - \mu| < \eta|\mu|\right\} > 1 - \delta.$$

To illustrate, consider an example with Y bounded in [0, 10] and Z bounded in [0, 1]. Assume that $\varepsilon = 0.1$, $\eta = 0.1$ and $\delta = 0.05$. Since $X = Y - Z$ is a random variable bounded in the interval $[-1, 10]$, from the discussion in the last section, it can be seen that Theorem 4 can be employed to obtain the minimum sample size as 13,408.

5. Illustrations

In this section, we shall illustrate the applications of our sample size formulae by examples in control and telecommunication engineering.

An extremely important problem of control engineering is to determine the probability that a system will fail to satisfy pre-specified requirements in an uncertain environment. This critical issue has been extensively studied in an area referred to as *probabilistic robustness analysis* (see, e.g. Khargonekar and Tikku, 1996; Lagoa and Barmish, 2002; Tempo et al., 2005, and the references therein). In general, there is no effective deterministic method for computing such failure probability except the Monte Carlo estimation method. To estimate the probability of failure, the uncertain environment is modeled by a random variable $\boldsymbol{\Delta}$, which may be scalar or matrix-valued. Hence, a Bernoulli random variable X can be defined as a function

$\mathscr{X}(.)$ of $\boldsymbol{\Delta}$ such that $X = \mathscr{X}(\boldsymbol{\Delta})$ assumes value 1 if the system associated with $\boldsymbol{\Delta}$ fails to satisfy pre-specified requirements and assumes value 0 otherwise. Clearly, the failure probability p is equal to the mean of X. That is, $p = \mathbb{E}[X] = \mathbb{E}[\mathscr{X}(\boldsymbol{\Delta})]$. For estimating the failure probability p, randomized algorithms have been implemented in a widely used software package RACT (Tremba et al., 2008), in which an absolute error criterion is used for estimating p. Specifically, for *a priori* $\varepsilon, \delta \in (0, 1)$, the objective is to obtain an estimator $\widehat{\boldsymbol{p}}$ such that $\mathbb{P}\{|\widehat{\boldsymbol{p}} - p| < \varepsilon\} > 1 - \delta$ holds regardless of the value of the $p \in (0, 1)$. The estimator is defined as

$$\widehat{\boldsymbol{p}} = \frac{1}{N} \sum_{i=1}^{N} \mathscr{X}(\boldsymbol{\Delta}_i),$$

where N is the sample size and $\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2, \ldots, \boldsymbol{\Delta}_N$ are i.i.d. samples of $\boldsymbol{\Delta}$. In most situations, there is no useful information about the range of the failure probability p due to the complexity of the system. Therefore, the determination of the sample size N should not be dependent on the range of p. It is well-known that, to make $\mathbb{P}\{|\widehat{\boldsymbol{p}} - p| < \varepsilon\} > 1 - \delta$ for any $p \in (0, 1)$, an approximate sample size based on normal approximation is

$$N = \left\lceil \frac{\mathcal{Z}_{\delta/2}^2}{4\varepsilon^2} \right\rceil, \tag{20}$$

where $\mathcal{Z}_{\delta/2}$ is the critical value such that

$$\int_{\mathcal{Z}_{\delta/2}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx = \frac{\delta}{2}.$$

The approximate sample size formula (20) will inevitably lead to unknown statistical error, since the formula (20) is based on the central limit theorem, which is an asymptotic result. In view of this drawback, control theorists and practitioners are reluctant to use the approximate formula (20). To rigorously control the statistical accuracy of the estimation, the Chernoff–Hoeffding bound is most frequently used in control engineering for determination of sample size. To ensure that $\mathbb{P}\{|\widehat{\boldsymbol{p}} - p| < \varepsilon\} > 1 - \delta$ holds for any $p \in (0, 1)$, it suffices to take sample size

$$N = \left\lceil \frac{\ln \frac{2}{\delta}}{2\varepsilon^2} \right\rceil. \tag{21}$$

The ratio of the sample size (21) to the sample size (20) is approximately equal to $\frac{2 \ln \frac{2}{\delta}}{\mathcal{Z}_{\delta/2}^2}$, which tends to 1 as $\delta \to 0$. It can be shown that

$$\frac{2 \ln \frac{2}{\delta}}{\mathcal{Z}_{\delta/2}^2} < \frac{3}{2} \quad \text{for } \delta \in \left(0, \frac{1}{10}\right).$$

This indicates that in most situations, the ratio of the rigorous sample size (21) to the approximate sample size (20) does not exceed $\frac{3}{2}$. From this analysis, it can be seen that it is worthy to obtain a rigorous control of the statistical accuracy by using the sample size (21) at the price of increasing the computational complexity up to 50%. This explains why the sample size (21) is frequently used in control engineering. As a matter of fact, the sample size formula (21) is implemented in RACT to estimate the failure probability.

In control engineering, the absolute error criterion is widely used. Recall that in Section 3, we have shown that a much smaller sample size is sufficient if a mixed error criterion is used. More specifically, the sample size can be significantly reduced by letting $\eta \in (0, 1)$ and relaxing the requirement $\mathbb{P}\{|\widehat{\boldsymbol{p}} - p| < \varepsilon\} > 1 - \delta$ as

$$\mathbb{P}\{|\widehat{\boldsymbol{p}} - p| < \varepsilon \text{ or } |\widehat{\boldsymbol{p}} - p| < \eta p\} > 1 - \delta.$$

In many situations, the margin of absolute error ε needs to be very small (e.g., $\varepsilon \ll 0.1$), since p is usually a very small number. However, the margin of relative error η does not need to be extremely small. For example, $\eta = 0.1$ may be sufficient for most cases.

As a concrete illustrative example, consider an uncertain dynamic system described by the differential equation

$$\frac{d^3 y(t)}{dt^3} + q_1 \frac{d^2 y(t)}{dt^2} + q_2 q_3 \frac{dy(t)}{dt} + q_2 y(t) = u(t),$$

where $u(t)$ is the input, $y(t)$ is the output, and q_1, q_2, q_3 are uncertain parameters. Assume that the tuple (q_1, q_2, q_3) is uniformly distributed over the domain

$$|1 - q_1| \le 1.1, \qquad |1 - q_2| \le 1, \qquad |1 - q_3| \le 0.5.$$

According to control theory, the system is said to be stable if the output is bounded for any bounded input. It can be shown that such a stability criterion is satisfied if and only if all the roots of the polynomial equation

$$s^3 + q_1 s^2 + q_2 q_3 s + q_2 = 0 \tag{22}$$

with respect to s in the field of complex number have negative real parts (see, e.g., Section 3.6 of Franklin et al., 2014, for an explanation of the concept of stability). Since the roots of Eq. (22) are functions of random variables q_1, q_2 and q_3, a Bernoulli random variable X can be defined in terms of q_1, q_2 and q_3 such that X assumes value 0 if all the roots have negative real parts, and otherwise X assumes value 1. For this particular example, we are interested in estimating the probability that the system is unstable. This amounts to the estimation of the probability that the Bernoulli random variable X assumes value 1. Since X is bounded in interval $[0, 1]$, our sample size formula can be useful for the planning of the Monte Carlo experiment. Let $\delta = 10^{-3}$. If the margin of error $\varepsilon = 10^{-3}$, then the sample size is obtained by (21) as 3800,452. If we use a mixed criterion with $\eta = 0.1$ and the same ε and δ, then the sample size can be computed by (19) as 155,463, which is only about 5% of sample size for the absolute criterion. The estimate of the probability of instability is obtained as 0.5403.

In wireless data communications, a frequent problem is to evaluate the bit error rate of a data transmission scheme. The bit error rate is the probability that a bit is transmitted incorrectly. In many situations, due to the complexity of the transmission system, the only tool to obtain the bit error rate is the Monte Carlo simulation method (Proakis, 2000). For example, there is no exact analytical method for computing the bit error rate of a wireless data transmission system employing multiple antennas and space–time block codes. The principle of this transmission system is proposed in Alamouti (1998) (see, e.g., Larsson and Stoica, 2003, and the references therein for a comprehensive discussion). The wireless data transmission process can be modeled by a sequence of Bernoulli random variables $X_1, X_2, \ldots$, where X_i assumes values 0 and 1 in accordance with the correct and incorrect transmissions of the ith bit. If $X_1, X_2, \ldots$ are identically and independently distributed Bernoulli random variables of the same mean $\mu \in (0, 1)$, then the bit error rate is μ and its estimator can be taken as $\frac{\sum_{i=1}^n X_i}{n}$ with n being sufficiently large. However, as a consequence of the application of the space–time block codes, the random variables $X_1, X_2, \ldots$ are not independent. This gives rise to the following question:

Is it possible to estimate the bit error rate without the independence of the random variables $X_1, X_2, \ldots$?

In a wireless data transmission system employing multiple antennas and space–time block codes, the expectation of X_k conditioned upon $X_\ell, \ \ell < k$ is a constant μ with respect to k, since the noise process is stationary and the input data can be treated as a Bernoulli process (Alamouti, 1998; Larsson and Stoica, 2003). This implies that it is reasonable to treat $X_1, X_2, \ldots$ as a martingale process such that condition (2) is satisfied. Hence, despite the lack of independence, the bit error rate can be approximated by $\frac{\sum_{i=1}^n X_i}{n}$. To control the statistical error, the sample size method proposed in the previous section can be applied to determine the appropriate value of n.

6. Concluding remarks

In this paper, we have considered the problem of estimating means of bounded random variables. We have illustrated that in many applications, it may be more appropriate to use a mixed error criterion for quantifying the reliability of estimation. We demonstrated that as a consequence of using the mixed error criterion, the sample complexity can be substantially reduced. By virtue of probabilistic inequalities, we have developed explicit sample size formulae for the purpose of controlling the statistical error of estimation. We have attempted to make our results generally applicable by eliminating the need of i.i.d. assumptions of the samples and the form of the underlying distributions.

Acknowledgments

The authors would like to thank the Associated Editor and referees for their time, effort and comments in reviewing this paper.

This research is supported in part by NIH/NCI Grant Nos. 1 P01 CA116676, P30 CA138292-01, and 5 P50 CA128613.

Appendix A. Proof of Theorem 2

Note that

$$\mathbb{P}_\mu\{|\overline{X}_n - \mu| \geq \varepsilon, \ |\overline{X}_n - \mu| \geq \eta\mu\} = \mathbb{P}_\mu\{|\overline{X}_n - \mu| \geq \max(\varepsilon, \eta\mu)\}. \tag{23}$$

Since $X_1, \ldots, X_n$ are i.i.d. samples of X, it follows from (23) and Chebyshev's inequality that

$$\mathbb{P}_\mu\{|\overline{X}_n - \mu| \geq \varepsilon, \ |\overline{X}_n - \mu| \geq \eta\mu\} \leq \frac{\mathbb{V}(X)}{n[\max(\varepsilon, \eta\mu)]^2}, \tag{24}$$

where $\mathbb{V}(X)$ denotes the variance of X. Since $0 \leq X \leq 1$ almost surely and $\mathbb{E}[X] = \mu$, it must be true that

$$\mathbb{V}(X) \leq \mu(1 - \mu). \tag{25}$$

Combining (24) and (25) yields

$$\mathbb{P}_\mu\{|\overline{X}_n - \mu| \geq \varepsilon, \ |\overline{X}_n - \mu| \geq \eta\mu\} \leq \frac{Q(\mu)}{n}, \tag{26}$$

where

$$Q(\mu) = \frac{\mu(1-\mu)}{[\max(\varepsilon, \eta\mu)]^2}$$

for $\mu \in (0, 1)$. Now we investigate the maximum of $Q(\mu)$ for $\mu \in (0, 1)$ by considering two cases as follows.

Case (i): $0 \le \mu \le \lambda$.
Case (ii): $\lambda < \mu \le 1$.

In Case (i), we have $0 \le \mu \le \lambda = \frac{\varepsilon}{\eta} \le \frac{1}{2}$ and

$$Q(\mu) = \frac{\mu(1-\mu)}{\varepsilon^2} \le \frac{\lambda(1-\lambda)}{\varepsilon^2} = \frac{1-\lambda}{\varepsilon\eta}, \tag{27}$$

where we have used the fact that $\mu(1-\mu)$ is increasing with respect to $\mu \in (0, \frac{1}{2})$. In Case (ii), we have $\lambda \le \mu \le 1$ and

$$Q(\mu) = \frac{\mu(1-\mu)}{(\eta\mu)^2} = \frac{1-\mu}{\eta^2\mu} \le \frac{1-\lambda}{\eta^2\lambda} = \frac{1-\lambda}{\varepsilon\eta}. \tag{28}$$

In view of (27) and (28), we have

$$Q(\mu) \le \frac{1-\lambda}{\varepsilon\eta}, \qquad \forall \mu \in [0, 1]. \tag{29}$$

Making use of (26) and (29), we have

$$\mathbb{P}_\mu\{|\overline{X}_n - \mu| \ge \varepsilon,\ |\overline{X}_n - \mu| \ge \eta\mu\} \le \frac{1-\lambda}{n\varepsilon\eta}, \qquad \forall \mu \in [0, 1],$$

from which the theorem immediately follows. This completes the proof of Theorem 2.

Throughout the proofs of Theorems 3 and 4, we shall use the following definitions. Let

$$\theta = \frac{\mu - a}{b - a}.$$

Let $\mathbb{P}_\theta$ denote the probability measure associated with θ. Define

$$Y_k = \frac{X_k - a}{b - a}, \qquad \overline{Y}_n = \frac{\sum_{i=1}^{k} Y_i}{k} \qquad k \in \mathbb{N},$$

where $X_1, X_2, \ldots$ are random variables such that $a \le X_k \le b$ and $\mathbb{E}[X_k \mid \mathscr{F}_{k-1}] = \mu$ almost surely for all $k \in \mathbb{N}$.

Appendix B. Proof of Theorem 3

To prove the theorem, we need some preliminary results.

Lemma 1. *Let $\zeta \in (0, 1)$. Define*

$$\mathscr{Q}_1(\theta) = \begin{cases} \exp(-n\ \varphi(\zeta, \theta)) & \text{for } \theta \in \left(0, 1 - \dfrac{\zeta}{3}\right), \\ 0 & \text{for } \theta \in \left(1 - \dfrac{\zeta}{3}, 1\right). \end{cases}$$

Then, $\mathbb{P}_\theta\{\overline{Y}_n \ge \theta + \zeta\} \le \mathscr{Q}_1(\theta)$ for $\theta \in (0, 1)$. Moreover, $\mathscr{Q}_1(\theta)$ is increasing with respect to $\theta \in (0, \frac{1}{2} - \frac{\zeta}{3})$ and non-increasing with respect to $\theta \in (\frac{1}{2} - \frac{\zeta}{3}, 1)$.

Proof. For $\theta \in (0, 1 - \frac{\zeta}{3})$, we have $0 < \zeta < 3(1-\theta)$, it follows from (8) that $\mathbb{P}_\theta\{\overline{Y}_n \ge \theta + \zeta\} \le \exp(-n\ \varphi(\zeta, \theta))$ for $(0, 1 - \frac{\zeta}{3})$. For $\theta \in (1 - \frac{\zeta}{3}, 1)$, we have $\theta + \zeta > 1$ and consequently,

$$\mathbb{P}_\theta\{\overline{Y}_n \ge \theta + \zeta\} \le \mathbb{P}_\theta\{\overline{Y}_n > 1\} = 0.$$

Thus, we have shown that $\mathbb{P}_\theta\{\overline{Y}_n \ge \theta + \zeta\} \le \mathscr{Q}_1(\theta)$ for $\theta \in (0, 1)$. To establish the monotonicity of $\mathscr{Q}_1(\theta)$, it is sufficient to observe that

$$\frac{\partial \varphi(\zeta, \theta)}{\partial \theta} = \frac{\zeta^2}{\left(\theta + \frac{\zeta}{3}\right)^2 \left(1 - \theta - \frac{\zeta}{3}\right)^2} \left[2\left(\theta + \frac{\zeta}{3}\right) - 1\right],$$

which is negative for any $\theta \in \left(0,\ \frac{1}{2} - \frac{\zeta}{3}\right)$ and positive for any $\theta \in \left(\frac{1}{2} - \frac{\zeta}{3}, 1\right)$. □

Lemma 2. *Let* $\zeta \in (0, 1)$. *Define*

$$\mathcal{Q}_2(\theta) = \begin{cases} \exp(-n\,\varphi(-\zeta,\theta)) & \text{for } \theta \in \left(\frac{\zeta}{3}, 1\right), \\ 0 & \text{for } \theta \in \left(0, \frac{\zeta}{3}\right). \end{cases}$$

Then, $\mathbb{P}_\theta\{\overline{Y}_n \le \theta - \zeta\} \le \mathcal{Q}_2(\theta)$ *for* $\theta \in (0, 1)$. *Moreover,* $\mathcal{Q}_2(\theta)$ *is non-decreasing with respect to* $\theta \in (0, \frac{1}{2} + \frac{\zeta}{3})$ *and decreasing with respect to* $\theta \in (\frac{1}{2} + \frac{\zeta}{3}, 1)$.

Proof. For $\theta \in (\frac{\zeta}{3}, 1)$, we have $0 < \zeta < 3\theta$, it follows from (9) that $\mathbb{P}_\theta\{\overline{Y}_n \le \theta - \zeta\} \le \exp(-n\,\varphi(-\zeta,\theta))$ for $(\frac{\zeta}{3}, 1)$. For $\theta \in (\frac{\zeta}{3}, 1)$, we have $\theta - \zeta < 0$ and consequently,

$$\mathbb{P}_\theta\{\overline{Y}_n \le \theta - \zeta\} \le \mathbb{P}_\theta\{\overline{Y}_n < 0\} = 0.$$

Thus, we have shown that $\mathbb{P}_\theta\{\overline{Y}_n \le \theta - \zeta\} \le \mathcal{Q}_2(\theta)$ for $\theta \in (0, 1)$. To establish the monotonicity of $\mathcal{Q}_2(\theta)$, it is sufficient to observe that

$$\frac{\partial\varphi(-\zeta,\theta)}{\partial\theta} = \frac{\zeta^2}{(\theta - \frac{\zeta}{3})^2(1 - \theta + \frac{\zeta}{3})^2}\left[2\left(\theta - \frac{\zeta}{3}\right) - 1\right],$$

which is negative for $\theta \in (\frac{\zeta}{3}, \frac{1}{2} + \frac{\zeta}{3})$ and positive for $\theta \in (\frac{1}{2} + \frac{\zeta}{3}, 1)$. □

Lemma 3. *Let* $-\frac{3}{\eta} < c \le 0$. *Define*

$$r^* = \frac{3 + \eta c}{3 + \eta}, \qquad \nu^* = \frac{1}{3 + \eta}\left(\eta c + \frac{3c}{2c - 1}\right)$$

and

$$\mathcal{Q}_3(\theta) = \begin{cases} \exp\left(-n\,\varphi\left(\eta(\theta - c), \theta\right)\right) & \text{for } \theta \in (0, r^*), \\ 0 & \text{for } \theta \in [r^*, 1). \end{cases}$$

Then, the following assertions hold.

(I) : $\mathbb{P}_\theta\{\overline{Y}_n \ge \theta + \eta(\theta - c)\} \le \mathcal{Q}_3(\theta)$ *for* $\theta \in (0, 1)$.
(II) : *If* $\nu^* > 0$, *then* $\mathcal{Q}_3(\theta)$ *is increasing with respect to* $\theta \in (0, \nu^*)$ *and non-increasing with respect to* $\theta \in (\nu^*, 1)$.
(III) : *If* $\nu^* \le 0$, *then* $\mathcal{Q}_3(\theta)$ *is non-increasing with respect to* $\theta \in (0, 1)$.

Proof. To show assertion (I), note that $\theta + \eta(\theta - c) > 1$ for $\theta \in [r^*, 1)$. Consequently,

$$\mathbb{P}_\theta\left\{\overline{Y}_n \ge \theta + \eta\,(\theta - c)\right\} \le \mathbb{P}_\theta\left\{\overline{Y}_n > 1\right\} = 0$$

for $\theta \in [r^*, 1)$. On the other hand, $0 < \eta\,(\theta - c) < 3(1 - \theta)$ for $\theta \in (0, r^*)$. Hence, it follows from inequality (8) that

$$\mathbb{P}_\theta\left\{\overline{Y}_n \ge \theta + \eta\,(\theta - c)\right\} \le \exp\left(-n\,\varphi\left(\eta(\theta - c), \theta\right)\right) = \exp\left(-\frac{n\eta^2}{2}g(\theta)\right) \quad \text{for } \theta \in (0, r^*),$$

where

$$g(\theta) = \frac{(\theta - c)^2}{\rho(\theta)[1 - \rho(\theta)]}$$

with $\rho(\theta) = \theta + \frac{\eta}{3}\,(\theta - c)$. Clearly, $0 < r^* < 1$ and $0 < \rho(\theta) < 1$ for $\theta \in (0, r^*)$. This proves assertion (I).

To show assertions (II) and (III), consider the derivative of $g(\theta)$ with respect to θ. Let $x = (1 + \frac{\eta}{3})\theta$ and $\alpha = \frac{c\eta}{3}$. Then, $\rho(\theta) = x - \alpha$ and

$$\begin{aligned} g'(\theta) &= \frac{2\,(\theta - c)}{\rho(\theta)[1 - \rho(\theta)]} - \frac{(\theta - c)^2\,\{\rho'(\theta)[1 - \rho(\theta)] - \rho(\theta)\rho'(\theta)\}}{[\rho(\theta)]^2[1 - \rho(\theta)]^2} \\ &= \frac{2(\theta - c)\{\rho(\theta)[1 - \rho(\theta)] - (1 + \frac{\eta}{3})\,(\theta - c)\left[\frac{1}{2} - \rho(\theta)\right]\}}{[\rho(\theta)]^2[1 - \rho(\theta)]^2} \\ &= \frac{2(\theta - c)}{[\rho(\theta)]^2[1 - \rho(\theta)]^2}\left[(x - \alpha)(1 - x + \alpha) - (x - \gamma)\left(\frac{1}{2} - x + \alpha\right)\right] \\ &= \frac{2(\theta - c)}{[\rho(\theta)]^2[1 - \rho(\theta)]^2}\left[\left(\frac{1}{2} + \alpha - \gamma\right)x - \alpha(1 + \alpha) + \gamma\left(\frac{1}{2} + \alpha\right)\right] \end{aligned}$$

$$= \frac{2(\theta - c)}{[\rho(\theta)]^2[1-\rho(\theta)]^2}\left[\left(\frac{1}{2} - c\right)x - \alpha(1+\alpha) + (c+\alpha)\left(\frac{1}{2}+\alpha\right)\right]$$
$$= \frac{2(\theta - c)}{[\rho(\theta)]^2[1-\rho(\theta)]^2}\left[\left(\frac{1}{2} - c\right)\rho(\theta) + \frac{c}{2}\right]. \tag{30}$$

Since $\theta - c > 0$ for $\theta \in (0, 1)$, it follows from (30) that $g'(\theta) \geq 0$ if and only if

$$\left(\frac{1}{2} - c\right)\rho(\theta) + \frac{c}{2} \geq 0,$$

which is equivalent to $\theta \geq \nu^*$. As a consequence of $c < 1$, we have $\nu^* < r^*$. It follows that assertions (II) and (III) hold. □

Lemma 4. *Define*

$$N = \begin{cases} \frac{2}{\varepsilon^2}\left(\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} - a\right)\left(b - \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta} & \text{for } \frac{2ab}{a+b} < \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} < \frac{a+b}{2}, \\ \frac{(b-a)^2}{2ab}\left(\frac{1}{\eta} + \frac{1}{3}\right)^2 \ln\frac{2}{\delta} & \text{for } \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} < \frac{2ab}{a+b}, \\ \frac{(b-a)^2}{2\varepsilon^2}\ln\frac{2}{\delta} & \text{for } \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} > \frac{a+b}{2}. \end{cases}$$

Then, $\mathbb{P}_\mu\{\overline{X}_n \geq \mu + \max(\varepsilon, \eta\mu)\} \leq \delta$ *for all* $\mu \in (a, b)$ *provided that* $n > N$.

Proof. For simplicity of notations, define $\zeta = \frac{\varepsilon}{b-a}$, $\lambda = \frac{\varepsilon}{\eta}$, $c = \frac{a}{a-b}$,

$$p^* = \frac{1}{2} - \frac{\zeta}{3}, \qquad \theta^* = \frac{\lambda}{b-a} + c, \qquad \nu^* = \frac{1}{3+\eta}\left(\eta c + \frac{3c}{2c-1}\right)$$

and $\mathrm{Q}^+(\theta) = \mathbb{P}_\theta\{\overline{Y}_n \geq \theta + \max(\zeta, \eta(\theta - c))\}$ for $\theta \in (0, 1)$. Then, $\mathbb{P}_\mu\{\overline{X}_n \geq \mu + \max(\varepsilon, \eta\mu)\} = \mathrm{Q}^+(\theta)$. It suffices to show the lemma for the following three cases.

Case (1): $\frac{2ab}{a+b} \leq \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} \leq \frac{a+b}{2}$.

Case (2): $\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} < \frac{2ab}{a+b}$.

Case (3): $\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} > \frac{a+b}{2}$.

First, consider Case (1). Clearly, as a consequence of $\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} \leq \frac{a+b}{2}$, we have $p^* \geq \theta^*$. As a consequence of $\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} \geq \frac{2ab}{a+b}$, we have $\theta^* \geq \nu^*$. Therefore, it follows from $\frac{2ab}{a+b} \leq \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} \leq \frac{a+b}{2}$ that $p^* \geq \theta^* \geq \nu^*$. Since $\frac{1}{2} > p^* \geq \theta^* > 0$, it follows from Lemma 1 that $\mathscr{Q}_1(\theta)$ is increasing for $\theta \in (0, \theta^*]$. Hence,

$$\mathrm{Q}^+(\theta) \leq \mathscr{Q}_1(\theta) \leq \mathscr{Q}_1(\theta^*) = \mathscr{Q}_3(\theta^*)$$

for $\theta \in (0, \theta^*]$. Since $\theta^* \geq \nu^*$, it follows from Lemma 3 that $\mathscr{Q}_3(\theta)$ is decreasing for $\theta \in [\theta^*, 1)$. Hence,

$$\mathrm{Q}^+(\theta) \leq \mathscr{Q}_3(\theta) \leq \mathscr{Q}_3(\theta^*)$$

for $\theta \in [\theta^*, 1)$. Therefore, $\mathrm{Q}^+(\theta) \leq \frac{\delta}{2}$ for $\theta \in (0, 1)$ provided that $\mathscr{Q}_3(\theta^*) \leq \frac{\delta}{2}$. Observing that

$$\mathscr{Q}_3(\theta^*) = \exp\left(-n\varphi(\zeta, \theta^*)\right),$$

we have that $\mathrm{Q}^+(\theta) \leq \frac{\delta}{2}$ for $\theta \in (0, 1)$ provided that

$$n > \frac{\ln\frac{2}{\delta}}{\varphi(\zeta, \theta^*)} = \frac{2}{\zeta^2}\left(\theta^* + \frac{\zeta}{3}\right)\left(1 - \theta^* - \frac{\zeta}{3}\right)\ln\frac{2}{\delta}$$
$$= \frac{2}{\zeta^2}\left(\frac{\lambda}{b-a} + c + \frac{\zeta}{3}\right)\left(1 - \frac{\lambda}{b-a} - c - \frac{\zeta}{3}\right)\ln\frac{2}{\delta}$$
$$= \frac{2}{\varepsilon^2}\left(\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} - a\right)\left(b - \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta}.$$

Next, consider Case (2). As a consequence of $\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} < \frac{2ab}{a+b}$, we have $\theta^* < \min(p^*, \nu^*)$. Since $\frac{1}{2} > p^* > \theta^* > 0$, it follows from Lemma 1 that $\mathscr{Q}_1(\theta)$ is increasing for $\theta \in (0, \theta^*]$. Hence,

$$\mathrm{Q}^+(\theta) \leq \mathscr{Q}_1(\theta) \leq \mathscr{Q}_1(\theta^*) = \mathscr{Q}_3(\theta^*) \leq \mathscr{Q}_3(\nu^*)$$

for $\theta \in (0, \theta^*]$. Since $\theta^* < \nu^*$, it follows from Lemma 3 that $\mathscr{Q}_3(\theta)$ is increasing for $\theta \in [\theta^*, \nu^*)$ and is decreasing for $\theta \in [\nu^*, 1)$. Hence,

$$Q^+(\theta) \leq \mathscr{Q}_3(\theta) \leq \mathscr{Q}_3(\nu^*)$$

for $\theta \in [\theta^*, 1)$. Therefore, $Q^+(\theta) \leq \frac{\delta}{2}$ for $\theta \in (0, 1)$ provided that $\mathscr{Q}_3(\nu^*) \leq \frac{\delta}{2}$. Since $\mathscr{Q}_3(\nu^*) = \exp(-n\varphi(\eta(\nu^* - c), \nu^*))$, it follows that $Q^+(\theta) \leq \frac{\delta}{2}$ for $\theta \in (0, 1)$ provided that

$$n > \frac{\ln \frac{2}{\delta}}{\varphi(\eta(\nu^* - c), \nu^*)} = \frac{(b-a)^2}{2ab} \left(\frac{1}{\eta} + \frac{1}{3}\right)^2 \ln \frac{2}{\delta},$$

where we have used the definitions of ν^* and c.

Finally, consider Case (3). In this case, we have $Q^+(\theta) \leq \mathbb{P}_\theta\{\overline{Y}_n \geq \theta + \zeta\} \leq \exp(-2n\zeta^2)$. Therefore, $Q^+(\theta) \leq \frac{\delta}{2}$ provided that

$$n > \frac{\ln \frac{2}{\delta}}{2\zeta^2} = \frac{(b-a)^2}{2\varepsilon^2} \ln \frac{2}{\delta}.$$

This completes the proof of the lemma. □

Lemma 5. *Let $c \leq 0$. Define*

$$r^* = -\frac{\eta c}{3 - \eta}, \qquad \nu^* = \frac{1}{3-\eta}\left(\frac{3c}{2c-1} - \eta c\right)$$

and

$$\mathscr{Q}_4(\theta) = \begin{cases} \exp\left(-n\,\varphi\left(\eta(c-\theta), \theta\right)\right) & \text{for } \theta \in (r^*, 1), \\ 0 & \text{for } \theta \in (0, r^*). \end{cases}$$

Then, the following assertions hold.

(I) *: $\mathbb{P}_\theta\{\overline{Y}_n \leq \theta - \eta(\theta - c)\} \leq \mathscr{Q}_4(\theta)$ for $\theta \in (0, 1)$.*

(II) *: If $\nu^* \leq 1$, then $\mathscr{Q}_4(\theta)$ is non-decreasing with respect to $\theta \in (0, \nu^*)$ and decreasing with respect to $\theta \in (\nu^*, 1)$.*

(III) *: If $\nu^* > 1$, then $\mathscr{Q}_4(\theta)$ is non-decreasing with respect to $\theta \in (0, 1)$.*

Proof. Clearly, $r^* \geq 0$. To show assertion (I), note that $\theta - \eta(\theta - c) < 0$ for $\theta \in (0, r^*)$. It follows that

$$\mathbb{P}_\theta\{\overline{Y}_n \leq \theta - \eta(\theta - c)\} \leq \mathbb{P}_\theta\{\overline{Y}_n < 0\} = 0$$

for $\theta \in (0, r^*)$. On the other hand, $0 < \eta\,(\theta - c) < 3\theta$ for $\theta \in (r^*, 1)$, it follows from (9) that

$$\mathbb{P}_\theta\left\{\overline{Y}_n \leq \theta - \eta\,(\theta - c)\right\} \leq \exp\left(-n\,\varphi\left(\eta(c-\theta), \theta\right)\right) = \exp\left(-\frac{n\eta^2}{2} g(\theta)\right),$$

where

$$g(\theta) = \frac{(\theta - c)^2}{\rho(\theta)[1 - \rho(\theta)]}$$

with $\rho(\theta) = \theta - \frac{\eta}{3}(\theta - c)$. Clearly, $\rho(\theta) < \theta < 1$. Since $\theta > r^*$, we have $\rho(\theta) > 0$. Hence, $0 < \rho(\theta) < 1$ for $\theta \in (r^*, 1)$. This establishes assertion (I).

To show assertions (II) and (III), consider the derivative of $g(\theta)$ with respective to θ. Tedious computation shows that

$$g'(\theta) = \frac{2(\theta - c)}{[\rho(\theta)]^2[1-\rho(\theta)]^2}\left[\left(\frac{1}{2} - c\right)\rho(\theta) + \frac{c}{2}\right]. \tag{31}$$

Since $\theta - c > 0$ for $\theta \in (0, 1)$, it follows from (31) that $g'(\theta) \geq 0$ if and only if $\left(\frac{1}{2} - c\right)\rho(\theta) + \frac{c}{2} \geq 0$, which is equivalent to $\theta \geq \nu^*$. Direct computation shows that $0 < r^* < \nu^*$. It follows that assertions (II) and (III) hold. □

Lemma 6. *Define*

$$M = \begin{cases} \frac{2}{\varepsilon^2}\left(\frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} - a\right)\left(b - \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta} & \text{for } \frac{2ab}{a+b} < \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} < \frac{a+b}{2}, \\ \frac{(b-a)^2}{2ab}\left(\frac{1}{\eta} - \frac{1}{3}\right)^2 \ln\frac{2}{\delta} & \text{for } \frac{\eta}{3} < \frac{b-a}{b+a} \text{ and } \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} < \frac{2ab}{a+b}, \\ \left[\frac{2}{3\eta}\left(1 - \frac{a}{b}\right) - \frac{2}{9}\right]\ln\frac{2}{\delta} & \text{for } \frac{b-a}{b+a} < \frac{\eta}{3} < \frac{b-a}{b} \text{ and } \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} < \frac{a+b}{2}, \\ 1 & \text{for } \frac{\eta}{3} > \frac{b-a}{b}, \\ \frac{(b-a)^2}{2\varepsilon^2}\ln\frac{2}{\delta} & \text{else.} \end{cases}$$

Then, $\mathbb{P}_\mu\{\overline{X}_n \le \mu - \max(\varepsilon, \eta\mu)\} \le \delta$ *for all* $\mu \in (a, b)$ *provided that* $n > M$.

Proof. For simplicity of notations, define $\zeta = \frac{\varepsilon}{b-a}$, $\lambda = \frac{\varepsilon}{\eta}$, $c = \frac{a}{a-b}$,

$$q^* = \frac{1}{2} + \frac{\zeta}{3}, \qquad \theta^* = \frac{\lambda}{b-a} + c, \qquad \nu^* = \frac{1}{3-\eta}\left(\frac{3c}{2c-1} - \eta c\right), \qquad r^* = -\frac{\eta c}{3-\eta}.$$

It can be checked that $0 < \theta^* < 1$. Define $Q^-(\theta) = \mathbb{P}_\theta\{\overline{Y}_n \le \theta - \max(\zeta, \eta(\theta - c))\}$ for $\theta \in (0, 1)$. Then, $\mathbb{P}_\mu\{\overline{X}_n \le \mu - \max(\varepsilon, \eta\mu)\} = Q^-(\theta)$. We need to show the lemma for the following six cases.

Case (i): $\frac{2ab}{a+b} \le \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} \le \frac{a+b}{2}$.

Case (ii): $\frac{\eta}{3} < \frac{b-a}{b+a}$ and $\frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} < \frac{2ab}{a+b}$.

Case (iii): $\frac{b-a}{b+a} < \frac{\eta}{3} < \frac{b-a}{b}$ and $\frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} \le \frac{a+b}{2}$.

Case (iv): $\frac{\eta}{3} \ge \frac{b-a}{b}$.

Case (v): Else.

First, consider Case (i). Clearly, as a consequence of $\frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} \le \frac{a+b}{2}$, we have $q^* \ge \theta^*$. As a consequence of $\frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} \ge \frac{2ab}{a+b}$, we have $\theta^* \ge \nu^*$. It follows from $\frac{2ab}{a+b} \le \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} \le \frac{a+b}{2}$ that $q^* \ge \theta^* \ge \nu^*$. Since $q^* \ge \theta^*$, it follows from Lemma 2 that $\mathscr{Q}_2(\theta)$ is non-decreasing for $\theta \in (0, \theta^*]$. Hence, for $\theta \in (0, \theta^*]$,

$$Q^-(\theta) \le \mathscr{Q}_2(\theta) \le \mathscr{Q}_2(\theta^*).$$

Since $\theta^* \ge \nu^*$, it follows from Lemma 5 that $\mathscr{Q}_4(\theta)$ is decreasing for $\theta \in [\theta^*, 1)$. Hence,

$$Q^-(\theta) \le \mathscr{Q}_4(\theta) \le \mathscr{Q}_4(\theta^*) = \mathscr{Q}_2(\theta^*)$$

for $\theta \in [\theta^*, 1)$. Hence, $Q^-(\theta) \le \frac{\delta}{2}$ for $\theta \in (0, 1)$ provided that $\mathscr{Q}_2(\theta^*) \le \frac{\delta}{2}$. Observing that $\mathscr{Q}_2(\theta^*) = \exp(-n\varphi(-\zeta, \theta^*))$, we have that $Q^-(\theta) \le \frac{\delta}{2}$ for $\theta \in (0, 1)$ provided that

$$\begin{aligned} n > \frac{\ln\frac{2}{\delta}}{\varphi(-\zeta, \theta^*)} &= \frac{2}{\zeta^2}\left(\theta^* - \frac{\zeta}{3}\right)\left(1 - \theta^* + \frac{\zeta}{3}\right)\ln\frac{2}{\delta} \\ &= \frac{2}{\zeta^2}\left(\frac{\lambda}{b-a} + c - \frac{\zeta}{3}\right)\left(1 - \frac{\lambda}{b-a} - c + \frac{\zeta}{3}\right)\ln\frac{2}{\delta} \\ &= \frac{2}{\varepsilon^2}\left(\frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} - a\right)\left(b - \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta}. \end{aligned}$$

Second, consider Case (ii). As a consequence of $\frac{\eta}{3} < \frac{b-a}{b+a}$, we have $\nu^* < 1$. Making use of $\frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} < \frac{2ab}{a+b}$, we have $\theta^* < \min(q^*, \nu^*)$. Since $q^* > \theta^*$, it follows from Lemma 2 that $\mathscr{Q}_2(\theta)$ is non-decreasing for $\theta \in (0, \theta^*]$. Hence, for $\theta \in (0, \theta^*]$,

$$Q^-(\theta) \le \mathscr{Q}_2(\theta) \le \mathscr{Q}_2(\theta^*). \tag{32}$$

Since $\theta^* < \nu^*$, it follows from Lemma 5 that $\mathscr{Q}_4(\theta)$ is increasing for $\theta \in [\theta^*, \nu^*)$ and is decreasing for $\theta \in [\nu^*, 1)$. Hence,

$$Q^-(\theta) \le \mathscr{Q}_4(\theta) \le \mathscr{Q}_4(\nu^*) \tag{33}$$

for $\theta \in [\theta^*, 1)$. Note that

$$\mathscr{Q}_4(\nu^*) \ge \mathscr{Q}_4(\theta^*) = \mathscr{Q}_2(\theta^*). \tag{34}$$

In view of (32)–(34), we have that $Q^-(\theta) \leq \frac{\delta}{2}$ for $\theta \in (0, 1)$ provided that $\mathscr{Q}_4(\nu^*) \leq \frac{\delta}{2}$. Observing that

$$\mathscr{Q}_4(\nu^*) = \exp\left(-n\varphi(-\eta(\nu^* - c), \nu^*)\right),$$

we have that $Q^-(\theta) \leq \frac{\delta}{2}$ for $\theta \in (0, 1)$ provided that the corresponding sample size

$$n > \frac{\ln \frac{2}{\delta}}{\varphi(-\eta(\nu^* - c), \nu^*)} = \frac{(b-a)^2}{2ab}\left(\frac{1}{\eta} - \frac{1}{3}\right)^2 \ln \frac{2}{\delta},$$

where we have used the definitions of ν and c.

Third, consider Case (iii). As a consequence of

$$\frac{b-a}{b+a} < \frac{\eta}{3} < \frac{b-a}{b}, \qquad \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} \leq \frac{a+b}{2},$$

we have $r^* < 1 < \nu^*$ and $q^* \geq \theta^*$. Since $q^* \geq \theta^*$, it follows from Lemma 2 that $\mathscr{Q}_2(\theta)$ is non-decreasing for $\theta \in (0, \theta^*]$. Hence, for $\theta \in (0, \theta^*]$,

$$Q^-(\theta) \leq \mathscr{Q}_2(\theta) \leq \mathscr{Q}_2(\theta^*).$$

Since $\nu^* > 1$, it follows from Lemma 5 that $\mathscr{Q}_4(\theta)$ is non-decreasing for $\theta \in [\theta^*, 1)$. Hence,

$$Q^-(\theta) \leq \mathscr{Q}_4(\theta) \leq \mathscr{Q}_4(1)$$

for $\theta \in [\theta^*, 1)$. Note that $\mathscr{Q}_4(1) \geq \mathscr{Q}_4(\theta^*) = \mathscr{Q}_2(\theta^*)$. Hence, $Q^-(\theta) \leq \frac{\delta}{2}$ for $\theta \in (0, 1)$ provided that $\mathscr{Q}_4(1) \leq \frac{\delta}{2}$. Since $\mathscr{Q}_4(1) = \exp(-n\varphi(-\eta(1-c), 1))$, it follows that $Q^-(\theta) \leq \frac{\delta}{2}$ for $\theta \in (0, 1)$ provided that the corresponding sample size

$$n > \frac{\ln \frac{2}{\delta}}{\varphi(-\eta(1-c), 1)} = \frac{2(1 - \frac{\eta(1-c)}{3})(\frac{\eta(1-c)}{3})}{[\eta(1-c)]^2} \ln \frac{2}{\delta} = \left[\frac{2}{3\eta}\left(1 - \frac{a}{b}\right) - \frac{2}{9}\right] \ln \frac{2}{\delta}.$$

Now, consider Case (iv). As a consequence of $\frac{\eta}{3} \geq \frac{b-a}{b}$, we have $r^* \geq 1$, which implies that $Q^-(\theta) = 0$ for $\theta \in (0, 1)$. Hence, $Q^-(\theta) \leq \frac{\delta}{2}$ for $\theta \in (0, 1)$ for any sample size $n \geq 1$.

Finally, consider Case (v). In this case, we have $Q^-(\theta) \leq \mathbb{P}_\theta\{\overline{Y}_n \leq \theta - \zeta\} \leq \exp(-2n\zeta^2)$. Therefore, $Q^-(\theta) \leq \frac{\delta}{2}$ provided that

$$n > \frac{\ln \frac{2}{\delta}}{2\zeta^2} = \frac{(b-a)^2}{2\varepsilon^2} \ln \frac{2}{\delta}.$$

This completes the proof of the lemma. □

Finally, Theorem 3 can be established by making use of Lemmas 4 and 6.

Appendix C. Proof of Theorem 4

To prove the theorem, we need some preliminary results.

Lemma 7. *Let $c \in (0, 1)$. Define $r^* = \frac{3+\eta c}{3+\eta}$ and*

$$\mathscr{H}_1(\theta) = \begin{cases} \exp\left(-n\,\varphi\left(\eta(\theta - c), \theta\right)\right) & \text{for } \theta \in (c, r^*), \\ 0 \quad \text{for } \theta \in (r^*, 1). \end{cases}$$

Then, $\mathbb{P}_\theta\{\overline{Y}_n \geq \theta + \eta(\theta - c)\} \leq \mathscr{H}_1(\theta)$ for $\theta \in (c, 1)$. Moreover, $\mathscr{H}_1(\theta)$ is non-increasing with respect to $\theta \in (c, 1)$.

Proof. By the definition of r^*, it can be checked that $c < r^* < 1$. Note that $\theta + \eta(\theta - c) > 1$ for $\theta \in (r^*, 1)$. Hence, $\mathbb{P}_\theta\{\overline{Y}_n \geq \theta + \eta(\theta - c)\} \leq \mathbb{P}_\theta\{\overline{Y}_n > 1\} = 0$ for $\theta \in (r^*, 1)$. On the other hand, $0 < \eta\,(\theta - c) < 3(1 - \theta)$ for $\theta \in (c, r^*)$. Thus, it follows from inequality (8) that

$$\mathbb{P}_\theta\left\{\overline{Y}_n \geq \theta + \eta\,(\theta - c)\right\} \leq \exp\left(-n\,\varphi\left(\eta(\theta - c), \theta\right)\right) = \exp\left(-\frac{n\eta^2}{2} g(\theta)\right),$$

where

$$g(\theta) = \frac{(\theta - c)^2}{\rho(\theta)[1 - \rho(\theta)]}$$

with $\rho(\theta) = \theta + \frac{\eta}{3}(\theta - c)$. It can be verified that $0 < \rho(\theta) < 1$ for $\theta \in (c, r^*)$. This shows that $\mathbb{P}_\theta\{\overline{Y}_n \geq \theta + \eta(\theta - c)\} \leq \mathscr{H}_1(\theta)$ for $\theta \in (c, 1)$.

To show that $\mathscr{H}_1(\theta)$ is non-increasing with respect to $\theta \in (c, 1)$, consider the derivative of $g(\theta)$ with respective to θ. Tedious computation shows that the derivative of $g(\theta)$ is given as

$$g'(\theta) = \frac{2(\theta - c)}{[\rho(\theta)]^2[1-\rho(\theta)]^2}\left[\left(\frac{1}{2} - c\right)\rho(\theta) + \frac{c}{2}\right]. \tag{35}$$

We claim that the derivative $g'(\theta)$ is positive for $\theta \in (c, 1)$. In view of (35) and the fact that $\theta - c > 0$ for $\theta \in (c, 1)$, it is sufficient to show that $\left(\frac{1}{2} - c\right)\rho(\theta) + \frac{c}{2} \geq 0$ for $\theta \in (c, 1)$ in the case that $\frac{1}{2} - c \geq 0$ and the case that $\frac{1}{2} - c < 0$. By the definition of $\rho(\theta)$, we have

$$c \leq \rho(\theta) \leq 1 + \frac{\eta}{3}(1 - c) \tag{36}$$

for $\theta \in (c, 1)$. In the case of $\frac{1}{2} - c \geq 0$, using the lower bound of $\rho(\theta)$ given by (36), we have $\left(\frac{1}{2} - c\right)\rho(\theta) + \frac{c}{2} \geq \left(\frac{1}{2} - c\right)c + \frac{c}{2} > 0$ for $\theta \in (c, 1)$. In the case of $\frac{1}{2} - c < 0$, using the upper bound of $\rho(\theta)$ given by (36), we have

$$\begin{aligned}\left(\frac{1}{2} - c\right)\rho(\theta) + \frac{c}{2} &\geq \left(\frac{1}{2} - c\right)\left[1 + \frac{\eta}{3}(1 - c)\right] + \frac{c}{2}\\ &= \frac{1}{2} + \frac{\eta}{6} - \frac{1+\eta}{2}c + c^2\frac{\eta}{3}\\ &= \frac{\eta}{3}(c - 1)\left(c - \frac{3}{2\eta} - \frac{1}{2}\right) > 0\end{aligned}$$

for $\theta \in (c, 1)$ and $\eta \in (0, 1)$. Thus, we have shown the claim that $g'(\theta) > 0$ in all cases. This implies that $\mathscr{H}_1(\theta)$ is non-increasing with respect to $\theta \in (c, 1)$. □

Lemma 8. *Let $c \in (0, 1)$. Define $\mathscr{H}_2(\theta) = \exp(-n\,\varphi(\eta(c - \theta), \theta))$ for $\theta \in (c, 1)$. Then, $\mathbb{P}_\theta\{\overline{Y}_n \leq \theta - \eta(\theta - c)\} \leq \mathscr{H}_2(\theta)$ for $\theta \in (c, 1)$. Moreover, $\mathscr{H}_2(\theta)$ is decreasing with respect to $\theta \in (c, 1)$.*

Proof. Clearly, $0 < \eta(\theta - c) < 3\theta$ for $\theta \in (c, 1)$. It follows from inequality (9) that

$$\mathbb{P}_\theta\left\{\overline{Y}_n \leq \theta - \eta(\theta - c)\right\} \leq \exp(-n\,\varphi(\eta(c - \theta), \theta)) = \exp\left(-\frac{n\eta^2}{2}g(\theta)\right),$$

where

$$g(\theta) = \frac{(\theta - c)^2}{\rho(\theta)[1 - \rho(\theta)]}$$

with $\rho(\theta) = \theta - \frac{\eta}{3}(\theta - c)$. Clearly, $0 < \rho(\theta) < 1$ for $\theta \in (c, 1)$. This shows that $\mathbb{P}_\theta\{\overline{Y}_n \leq \theta - \eta(\theta - c)\} \leq \mathscr{H}_2(\theta)$ for $\theta \in (c, 1)$.

To show that $\mathscr{H}_2(\theta)$ is decreasing with respect to $\theta \in (c, 1)$, consider the derivative of $g(\theta)$ with respect to θ. Tedious computation shows that the derivative of $g(\theta)$ is given as

$$g'(\theta) = \frac{2(\theta - c)}{[\rho(\theta)]^2[1-\rho(\theta)]^2}\left[\left(\frac{1}{2} - c\right)\rho(\theta) + \frac{c}{2}\right]. \tag{37}$$

We claim that the derivative $g'(\theta)$ is positive for $\theta \in (c, 1)$. In view of (37) and the fact that $\theta - c > 0$ for $\theta \in (c, 1)$, it is sufficient to show that $\left(\frac{1}{2} - c\right)\rho(\theta) + \frac{c}{2} \geq 0$ for $\theta \in (c, 1)$ in the case that $\frac{1}{2} - c \geq 0$ and the case that $\frac{1}{2} - c < 0$. By the definition of $\rho(\theta)$, we have

$$0 < \rho(\theta) < 1 - \frac{\eta}{3}(1 - c) \tag{38}$$

for $\theta \in (c, 1)$. In the case of $\frac{1}{2} - c \geq 0$, using the lower bound of $\rho(\theta)$ given by (38), we have $\left(\frac{1}{2} - c\right)\rho(\theta) + \frac{c}{2} \geq \frac{c}{2} > 0$ for $\theta \in (c, 1)$. In the case of $\frac{1}{2} - c < 0$, using the upper bound of $\rho(\theta)$ given by (38), we have

$$\begin{aligned}\left(\frac{1}{2} - c\right)\rho(\theta) + \frac{c}{2} &\geq \left(\frac{1}{2} - c\right)\left[1 - \frac{\eta}{3}(1 - c)\right] + \frac{c}{2} = \frac{1}{2} - \frac{\eta}{6} - \frac{1-\eta}{2}c - c^2\frac{\eta}{3}\\ &> \frac{1}{2} - \frac{\eta}{6} - \frac{1-\eta}{2}c - c\frac{\eta}{3} = \frac{1}{6}(3 - \eta)(1 - c) > 0.\end{aligned}$$

Thus, we have established the claim that $g'(\theta) > 0$ for $\theta \in (c, 1)$. It follows that $\mathscr{H}_2(\theta)$ is decreasing with respect to $\theta \in (c, 1)$. This completes the proof of the lemma. □

Lemma 9. *Let $c \in (0,1)$. Define $\mathscr{H}_3(\theta) = \exp\left(-n\,\varphi\left(\eta(c-\theta),\theta\right)\right)$ for $\theta \in (0,c)$. Then, $\mathbb{P}_\theta\{\overline{Y}_n \geq \theta + \eta(c-\theta)\} \leq \mathscr{H}_3(\theta)$ for $\theta \in (0,c)$. Moreover, $\mathscr{H}_3(\theta)$ is increasing with respect to $\theta \in (0,c)$.*

Proof. Note that $0 < \eta\,(c-\theta) < 3(1-\theta)$ for $\theta \in (0,c)$. It follows from inequality (8) that

$$\mathbb{P}_\theta\left\{\overline{Y}_n \geq \theta + \eta\,(c-\theta)\right\} \leq \exp\left(-n\,\varphi\left(\eta(c-\theta),\theta\right)\right) = \exp\left(-\frac{n\eta^2}{2}g(\theta)\right),$$

where

$$g(\theta) = \frac{(c-\theta)^2}{\rho(\theta)[1-\rho(\theta)]}$$

with $\rho(\theta) = \theta + \frac{\eta}{3}\,(c-\theta)$. Clearly, $\rho(\theta) > 0$ for $\theta \in (0,c)$. Since $c \in (0,1)$ and $\eta \in (0,3)$, we have $\rho(\theta) < \theta + \frac{\eta}{3}\,(1-\theta) < \theta + (1-\theta) = 1$ for $\theta \in (0,c)$. Hence, we have established that $\mathbb{P}_\theta\{\overline{Y}_n \geq \theta + \eta(c-\theta)\} \leq \mathscr{H}_3(\theta)$ for $\theta \in (0,c)$.

To show that $\mathscr{H}_3(\theta)$ is increasing with respect to $\theta \in (0,c)$, consider the derivative of $g(\theta)$ with respect to θ. Tedious computation shows that the derivative of $g(\theta)$ is given as

$$g'(\theta) = \frac{2(\theta-c)}{[\rho(\theta)]^2[1-\rho(\theta)]^2}\left[\left(\frac{1}{2}-c\right)\rho(\theta) + \frac{c}{2}\right]. \tag{39}$$

We claim that $g'(\theta)$ is negative for $\theta \in (0,c)$. In view of (39) and the fact that $\theta - c < 0$ for $\theta \in (0,c)$, it suffices to show $\left(\frac{1}{2}-c\right)\rho(\theta) + \frac{c}{2} > 0$ for the case that $\frac{1}{2} - c \geq 0$ and the case that $\frac{1}{2} - c < 0$. Note that $\frac{\eta}{3}c < \rho(\theta) < c$ for $\theta \in (0,c)$. In the case of $\frac{1}{2} - c \geq 0$, we have

$$\left(\frac{1}{2}-c\right)\rho(\theta) + \frac{c}{2} \geq \left(\frac{1}{2}-c\right)\frac{\eta}{3}c + \frac{c}{2} > \frac{c}{2} > 0.$$

In the case of $\frac{1}{2} - c < 0$, we have

$$\left(\frac{1}{2}-c\right)\rho(\theta) + \frac{c}{2} \geq \left(\frac{1}{2}-c\right)c + \frac{c}{2} > 0.$$

Therefore, we have established the claim that $g'(\theta) < 0$ for $\theta \in (0,c)$. This implies that $\mathscr{H}_3(\theta)$ is increasing with respect to $\theta \in (0,c)$. The proof of the lemma is thus completed. □

Lemma 10. *Let $c \in (0,1)$. Define $r^* = \frac{\eta c}{3+\eta}$ and*

$$\mathscr{H}_4(\theta) = \begin{cases} \exp\left(-n\,\varphi\left(\eta(\theta-c),\theta\right)\right) & \text{for } \theta \in (r^*, c), \\ 0 & \text{for } \theta \in (0, r^*). \end{cases}$$

Then, $\mathbb{P}_\theta\{\overline{Y}_n \leq \theta + \eta(\theta-c)\} \leq \mathscr{H}_4(\theta)$ for $\theta \in (0,c)$. Moreover, $\mathscr{H}_4(\theta)$ is non-decreasing with respect to $\theta \in (0,c)$.

Proof. Clearly, $0 < r^* < c$. Note that $\theta + \eta\,(\theta - c) < 0$ for $\theta \in (0, r^*)$. Hence, $\mathbb{P}_\theta\{\overline{Y}_n \leq \theta + \eta(\theta-c)\} \leq \mathbb{P}_\theta\{\overline{Y}_n < 0\}$ for $\theta \in (0,r^*)$. On the other hand, it can be checked that $0 < \eta\,(c-\theta) < 3\theta$ for $\theta \in (r^*, c)$. It follows from inequality (9) that

$$\mathbb{P}_\theta\left\{\overline{Y}_n \leq \theta + \eta\,(\theta-c)\right\} \leq \exp\left(-n\,\varphi\left(\eta(\theta-c),\theta\right)\right) = \exp\left(-\frac{n\eta^2}{2}g(\theta)\right)$$

for $\theta \in (r^*, c)$, where

$$g(\theta) = \frac{(c-\theta)^2}{\rho(\theta)[1-\rho(\theta)]}$$

with $\rho(\theta) = \theta + \frac{\eta}{3}\,(\theta - c)$. It can be verified that $\rho(\theta) > 0$ for $\theta \in (r^*, c)$. Since $c \in (0,1)$ and $\eta \in (0,3)$, we have $\rho(\theta) < \theta + \frac{\eta}{3}\,(1-\theta) < \theta + (1-\theta) = 1$ for $\theta \in (r^*, c)$. Thus, we have shown that $\mathbb{P}_\theta\{\overline{Y}_n \leq \theta + \eta(\theta-c)\} \leq \mathscr{H}_4(\theta)$ for $\theta \in (0,c)$.

To show that $\mathscr{H}_4(\theta)$ is non-decreasing with respect to $\theta \in (0,c)$, consider the derivative of $g(\theta)$ with respective to θ. Tedious computation shows that the derivative of $g(\theta)$ is given by

$$g'(\theta) = \frac{2(\theta-c)}{[\rho(\theta)]^2[1-\rho(\theta)]^2}\left[\left(\frac{1}{2}-c\right)\rho(\theta) + \frac{c}{2}\right]. \tag{40}$$

Note that $-\frac{\eta}{3}c < \rho(\theta) < c$ for $\theta \in (0,c)$. It follows that

$$\left(\frac{1}{2}-c\right)\rho(\theta) + \frac{c}{2} \geq \left(\frac{1}{2}-c\right)c + \frac{c}{2} \geq 0 \tag{41}$$

for $c \in (\frac{1}{2}, 1)$ and $\theta \in (0, c)$. Moreover,

$$\left(\frac{1}{2} - c\right)\rho(\theta) + \frac{c}{2} \geq -\left(\frac{1}{2} - c\right)\frac{\eta}{3}c + \frac{c}{2} > -\frac{\eta}{6}c + \frac{c}{2} > 0 \tag{42}$$

for $c \in (0, \frac{1}{2}]$ and $\theta \in (0, c)$. Making use of (41), (41) and (42), we have $g'(\theta) < 0$ for $\theta \in (r^*, c)$. So, we have established that $\mathscr{H}_4(\theta)$ is non-decreasing with respect to $\theta \in (0, c)$. This completes the proof of the lemma. □

Lemma 11. *Assume that $a < 0 < b$. Define*

$$N = \begin{cases} \frac{2}{\varepsilon^2}\left(\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} - a\right)\left(b - \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta} & \text{for } \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} < \frac{a+b}{2}, \\ \frac{2}{\varepsilon^2}\left(-\frac{\varepsilon}{3} - a\right)\left(b + \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta} & \text{for } \frac{\varepsilon}{3} + \frac{a+b}{2} < 0, \\ \frac{(b-a)^2}{2\varepsilon^2}\ln\frac{2}{\delta} & \text{else.} \end{cases} \tag{43}$$

Then, $\mathbb{P}_\mu\{|\overline{X}_n - \mu| \geq \max(\varepsilon,\ \eta|\mu|)\} \leq \delta$ for any $\mu \in (0, b)$ provided that $n > N$.

Proof. For simplicity of notations, define $\zeta = \frac{\varepsilon}{b-a}$,

$$\lambda = \frac{\varepsilon}{\eta}, \qquad c = \frac{a}{a-b}, \qquad \theta^* = \frac{\lambda}{b-a} + c, \qquad p^* = \frac{1}{2} - \frac{\varepsilon}{3(b-a)}, \qquad q^* = \frac{1}{2} + \frac{\varepsilon}{3(b-a)}.$$

Define functions

$$Q^+(\theta) = \mathbb{P}_\theta\{\overline{Y}_n \geq \theta + \max(\zeta, \eta(\theta - c))\},$$
$$Q^-(\theta) = \mathbb{P}_\theta\{\overline{Y}_n \leq \theta - \max(\zeta, \eta(\theta - c))\}$$

for $\theta \in (c, 1)$. For $\mu \in (0, b)$, putting $\theta = \frac{\mu}{b-a} + c$, we have $c < \theta < 1$ and

$$\mathbb{P}_\mu\{\overline{X}_n \geq \mu + \max(\varepsilon, \eta\mu)\} = Q^+(\theta),$$
$$\mathbb{P}_\mu\{\overline{X}_n \leq \mu - \max(\varepsilon, \eta\mu)\} = Q^-(\theta).$$

To prove the lemma, it suffices to consider the following three cases.

Case (I): $\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} < \frac{a+b}{2}$.
Case (II): $\frac{\varepsilon}{3} + \frac{a+b}{2} < 0$.
Case (III): $-\frac{\varepsilon}{3} \leq \frac{a+b}{2} \leq \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3}$.

First, consider Case (I). As a consequence of $\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} < \frac{a+b}{2}$, we have $\theta^* < p^*$. Since $p^* > \theta^*$, it follows from Lemma 1 that $\mathscr{Q}_1(\theta)$ is increasing for $\theta \in (0, \theta^*]$. Moreover, according to Lemma 7, we have that $\mathscr{H}_1(\theta)$ is non-increasing for $\theta \in [\theta^*, 1)$. It follows that

$$Q^+(\theta) \leq \mathscr{Q}_1(\theta) \leq \mathscr{Q}_1(\theta^*) \quad \text{for } \theta \in [c, \theta^*] \tag{44}$$

and that

$$Q^+(\theta) \leq \mathscr{H}_1(\theta) \leq \mathscr{H}_1(\theta^*) = \mathscr{Q}_1(\theta^*) \quad \text{for } \theta \in [\theta^*, 1). \tag{45}$$

Observing that $q^* > \theta^*$ and making use of Lemma 2, we have that $\mathscr{Q}_2(\theta)$ is non-decreasing for $\theta \in (c, \theta^*]$. According to Lemma 8, we have that $\mathscr{H}_2(\theta)$ is decreasing for $\theta \in [\theta^*, 1)$. It follows that

$$Q^-(\theta) \leq \mathscr{Q}_2(\theta) \leq \mathscr{Q}_2(\theta^*) \quad \text{for } \theta \in [c, \theta^*] \tag{46}$$

and that

$$Q^-(\theta) \leq \mathscr{H}_2(\theta) \leq \mathscr{H}_2(\theta^*) = \mathscr{Q}_2(\theta^*) \quad \text{for } \theta \in [\theta^*, 1). \tag{47}$$

Combining (44)–(47), we have

$$Q^+(\theta) + Q^-(\theta) \leq \mathscr{Q}_1(\theta^*) + \mathscr{Q}_2(\theta^*)$$

for $\theta \in (c, 1)$. Observing that

$$\mathscr{Q}_1(\theta^*) = \exp\left(-\frac{n\zeta^2}{2(\theta^* + \frac{\zeta}{3})(1 - \theta^* - \frac{\zeta}{3})}\right), \qquad \mathscr{Q}_2(\theta^*) = \exp\left(-\frac{n\zeta^2}{2(\theta^* - \frac{\zeta}{3})(1 - \theta^* + \frac{\zeta}{3})}\right)$$

and that

$$\left(\theta^* + \frac{\zeta}{3}\right)\left(1 - \theta^* - \frac{\zeta}{3}\right) - \left(\theta^* - \frac{\zeta}{3}\right)\left(1 - \theta^* + \frac{\zeta}{3}\right) = \frac{2\zeta}{3}(1 - 2\theta^*) > \frac{2\zeta}{3}(1 - 2p^*) > 0,$$

we have $\mathscr{Q}_1(\theta^*) > \mathscr{Q}_2(\theta^*)$ and consequently,

$$\mathrm{Q}^+(\theta) + \mathrm{Q}^-(\theta) \le 2\mathscr{Q}_1(\theta^*)$$

for $\theta \in (c, 1)$. It follows that

$$\mathbb{P}_\mu\{|\overline{X}_n - \mu| \ge \max(\varepsilon,\ \eta|\mu|)\} \le \mathrm{Q}^+(\theta) + \mathrm{Q}^-(\theta) \le 2\mathscr{Q}_1(\theta^*) = 2\exp\left(-n\,\varphi(\zeta, \theta^*)\right)$$

for $\mu \in (0, b)$. This implies that $\mathbb{P}_\mu\{|\overline{X}_n - \mu| \ge \max(\varepsilon,\ \eta|\mu|)\} < \delta$ provided that the corresponding sample size

$$\begin{aligned} n &> \frac{\ln\frac{2}{\delta}}{\varphi(\zeta, \theta^*)} = \frac{2\left(\theta^* + \frac{\zeta}{3}\right)\left(1 - \theta^* - \frac{\zeta}{3}\right)}{\zeta^2}\ln\frac{2}{\delta} \\ &= \frac{2\left(\frac{\lambda}{b-a} + c + \frac{\zeta}{3}\right)\left(1 - \frac{\lambda}{b-a} - c - \frac{\zeta}{3}\right)}{\zeta^2}\ln\frac{2}{\delta} \\ &= \frac{2}{\varepsilon^2}\left(\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} - a\right)\left(b - \frac{\varepsilon}{\eta} - \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta}. \end{aligned}$$

Next, consider Case (II). As a consequence of $\frac{\varepsilon}{3} + \frac{a+b}{2} < 0$, we have $q^* < c$. Clearly, $p^* < q^* < c < \theta^*$. By Lemma 1, $\mathscr{Q}_1(\theta)$ is non-increasing for $\theta \in (c, \theta^*]$. By Lemma 7, $\mathscr{H}_1(\theta)$ is non-increasing for $\theta \in [\theta^*, 1)$. It follows that

$$\mathrm{Q}^+(\theta) \le \mathscr{Q}_1(\theta) \le \mathscr{Q}_1(c) \quad \text{for } \theta \in [c, \theta^*]. \tag{48}$$

Moreover,

$$\mathrm{Q}^+(\theta) \le \mathscr{H}_1(\theta) \le \mathscr{H}_1(\theta^*) = \mathscr{Q}_1(\theta^*) \le \mathscr{Q}_1(c) \quad \text{for } \theta \in [\theta^*, 1). \tag{49}$$

Similarly, according to Lemma 2, $\mathscr{Q}_2(\theta)$ is decreasing for $\theta \in (c, \theta^*]$. By Lemma 8, $\mathscr{H}_2(\theta)$ is decreasing for $\theta \in [\theta^*, 1)$. It follows that

$$\mathrm{Q}^-(\theta) \le \mathscr{Q}_2(\theta) \le \mathscr{Q}_2(c) \quad \text{for } \theta \in [c, \theta^*]. \tag{50}$$

Moreover,

$$\mathrm{Q}^-(\theta) \le \mathscr{H}_2(\theta) \le \mathscr{H}_2(\theta^*) = \mathscr{Q}_2(\theta^*) \le \mathscr{Q}_2(c) \quad \text{for } \theta \in [\theta^*, 1). \tag{51}$$

Making use of (48)–(51), we have

$$\mathrm{Q}^+(\theta) + \mathrm{Q}^-(\theta) \le \mathscr{Q}_1(c) + \mathscr{Q}_2(c)$$

for $\theta \in (c, 1)$. Observing that

$$\mathscr{Q}_1(c) = \exp\left(-\frac{n\zeta^2}{2(c + \frac{\zeta}{3})(1 - c - \frac{\zeta}{3})}\right), \qquad \mathscr{Q}_2(c) = \exp\left(-\frac{n\zeta^2}{2(c - \frac{\zeta}{3})(1 - c + \frac{\zeta}{3})}\right)$$

and that

$$\left(c + \frac{\zeta}{3}\right)\left(1 - c - \frac{\zeta}{3}\right) - \left(c - \frac{\zeta}{3}\right)\left(1 - c + \frac{\zeta}{3}\right) = \frac{2\zeta}{3}(1 - 2c) < \frac{2\zeta}{3}(1 - 2q^*) < 0,$$

we have $\mathscr{Q}_1(c) < \mathscr{Q}_2(c)$ and consequently,

$$\mathrm{Q}^+(\theta) + \mathrm{Q}^-(\theta) \le 2\mathscr{Q}_2(c)$$

for $\theta \in (c, 1)$. It follows that

$$\mathbb{P}_\mu\{|\overline{X}_n - \mu| \ge \max(\varepsilon,\ \eta|\mu|)\} \le \mathrm{Q}^+(\theta) + \mathrm{Q}^-(\theta) \le 2\mathscr{Q}_2(c) = 2\exp\left(-n\,\varphi(-\zeta, c)\right)$$

for $\mu \in (0, b)$. This implies that $\mathbb{P}_\mu\{|\overline{X}_n - \mu| \ge \max(\varepsilon,\ \eta|\mu|)\} < \delta$ provided that the corresponding sample size

$$\begin{aligned} n &> \frac{\ln\frac{2}{\delta}}{\varphi(-\zeta, c)} = \frac{2\left(c - \frac{\zeta}{3}\right)\left(1 - c + \frac{\zeta}{3}\right)}{\zeta^2}\ln\frac{2}{\delta} \\ &= \frac{2\left(-a - \frac{\varepsilon}{3}\right)\left(b - a + a + \frac{\varepsilon}{3}\right)}{\varepsilon^2}\ln\frac{2}{\delta} \\ &= \frac{2}{\varepsilon^2}\left(-\frac{\varepsilon}{3} - a\right)\left(b + \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta}. \end{aligned}$$

Finally, consider Case (III). In this case, we have $\mathbb{P}_\mu\{|\overline{X}_n - \mu| \geq \max(\varepsilon,\ \eta|\mu|)\} \leq Q^+(\theta) + Q^-(\theta) \leq \mathbb{P}_\theta\{|\overline{Y}_n - \theta| \geq \zeta\} \leq 2\exp(-2n\zeta^2)$. Therefore, $\mathbb{P}_\mu\{|\overline{X}_n - \mu| \geq \max(\varepsilon,\ \eta|\mu|)\} < \delta$ provided that

$$n > \frac{\ln\frac{2}{\delta}}{2\zeta^2} = \frac{(b-a)^2}{2\varepsilon^2}\ln\frac{2}{\delta}.$$

This completes the proof of the lemma. □

Lemma 12. *Assume that $a < 0 < b$. Define*

$$M = \begin{cases} \frac{2}{\varepsilon^2}\left(-\frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} - a\right)\left(b + \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta} & \text{for } \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} < -\frac{a+b}{2}, \\ \frac{2}{\varepsilon^2}\left(\frac{\varepsilon}{3} - a\right)\left(b - \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta} & \text{for } \frac{\varepsilon}{3} < \frac{a+b}{2}, \\ \frac{(b-a)^2}{2\varepsilon^2}\ln\frac{2}{\delta} & \text{else.} \end{cases} \tag{52}$$

Then, $\mathbb{P}_\mu\{|\overline{X}_n - \mu| \geq \max(\varepsilon,\ \eta|\mu|)\} \leq \delta$ for any $\mu \in (a, 0)$ provided that $n > M$.

Proof. For simplicity of notations, define $\zeta = \frac{\varepsilon}{b-a}$,

$$\lambda = \frac{\varepsilon}{\eta}, \qquad c = \frac{a}{a-b}, \qquad \vartheta^* = c - \frac{\lambda}{b-a}, \qquad p^* = \frac{1}{2} - \frac{\varepsilon}{3(b-a)}, \qquad q^* = \frac{1}{2} + \frac{\varepsilon}{3(b-a)}.$$

Define functions

$$Q^+(\theta) = \mathbb{P}_\theta\{\overline{Y}_n \geq \theta + \max(\zeta, \eta(c-\theta))\},$$
$$Q^-(\theta) = \mathbb{P}_\theta\{\overline{Y}_n \leq \theta - \max(\zeta, \eta(c-\theta))\}$$

for $\theta \in (0, c)$. For $\mu \in (a, 0)$, putting $\theta = \frac{\mu}{b-a} + c$, we have $0 < \theta < c$ and

$$\mathbb{P}_\mu\{\overline{X}_n \geq \mu + \max(\varepsilon, \eta|\mu|)\} = Q^+(\theta),$$
$$\mathbb{P}_\mu\{\overline{X}_n \leq \mu - \max(\varepsilon, \eta|\mu|)\} = Q^-(\theta).$$

To prove the lemma, it suffices to consider the following three cases.

Case (I): $\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} < -\frac{a+b}{2}$.
Case (II): $\frac{\varepsilon}{3} < \frac{a+b}{2}$.
Case (III): $-\left(\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3}\right) \leq \frac{a+b}{2} \leq \frac{\varepsilon}{3}$.

First, consider Case (I). As a consequence of $\frac{\varepsilon}{\eta} + \frac{\varepsilon}{3} < -\frac{a+b}{2}$, we have $q^* < \vartheta^*$. Since $p^* < q^* < \vartheta^*$, it follows from Lemma 9 that $\mathscr{H}_3(\theta)$ is increasing for $\theta \in (0, \vartheta^*]$. According to Lemma 1, $\mathscr{Q}_1(\theta)$ is non-increasing for $\theta \in [\vartheta^*, c)$. Hence,

$$Q^+(\theta) \leq \mathscr{H}_3(\theta) \leq \mathscr{H}_3(\vartheta^*) = \mathscr{Q}_1(\vartheta^*) \quad \text{for } \theta \in (0, \vartheta^*]. \tag{53}$$

Moreover,

$$Q^+(\theta) \leq \mathscr{Q}_1(\theta) \leq \mathscr{Q}_1(\vartheta^*) \quad \text{for } \theta \in (\vartheta^*, c]. \tag{54}$$

Since $q^* < \vartheta^*$, it follows from Lemma 10 that $\mathscr{H}_4(\theta)$ is non-decreasing for $\theta \in (0, \vartheta^*]$. From Lemma 2, $\mathscr{Q}_2(\theta)$ is decreasing for $\theta \in [\vartheta^*, c)$. Hence,

$$Q^-(\theta) \leq \mathscr{H}_4(\theta) \leq \mathscr{H}_4(\vartheta^*) = \mathscr{Q}_2(\vartheta^*) \quad \text{for } \theta \in (0, \vartheta^*]. \tag{55}$$

Moreover,

$$Q^-(\theta) \leq \mathscr{Q}_2(\theta) \leq \mathscr{Q}_2(\vartheta^*) \quad \text{for } \theta \in (\vartheta^*, c]. \tag{56}$$

Making use of (53)–(56), we have

$$Q^+(\theta) + Q^-(\theta) \leq \mathscr{Q}_1(\vartheta^*) + \mathscr{Q}_2(\vartheta^*)$$

for $\theta \in (0, c)$. Observing that

$$\mathscr{Q}_1(\vartheta^*) = \exp\left(-\frac{n\zeta^2}{2(\vartheta^* + \frac{\zeta}{3})(1 - \vartheta^* - \frac{\zeta}{3})}\right), \qquad \mathscr{Q}_2(\vartheta^*) = \exp\left(-\frac{n\zeta^2}{2(\vartheta^* - \frac{\zeta}{3})(1 - \vartheta^* + \frac{\zeta}{3})}\right)$$

and that

$$\left(\vartheta^* + \frac{\zeta}{3}\right)\left(1 - \vartheta^* - \frac{\zeta}{3}\right) - \left(\vartheta^* - \frac{\zeta}{3}\right)\left(1 - \vartheta^* + \frac{\zeta}{3}\right) = \frac{2\zeta}{3}(1 - 2\vartheta^*) < \frac{2\zeta}{3}(1 - 2q^*) < 0,$$

we have $\mathscr{Q}_1(\vartheta^*) < \mathscr{Q}_2(\vartheta^*)$ and consequently,

$$\mathrm{Q}^+(\theta) + \mathrm{Q}^-(\theta) \le 2\mathscr{Q}_2(\vartheta^*)$$

for $\theta \in (0, c)$. It follows that

$$\mathbb{P}_\mu\{|\overline{X}_n - \mu| \ge \max(\varepsilon,\ \eta|\mu|)\} \le \mathrm{Q}^+(\theta) + \mathrm{Q}^-(\theta) \le 2\mathscr{Q}_2(\theta^*) = 2\exp\left(-n\,\varphi(-\zeta, \theta^*)\right)$$

for $\mu \in (a, 0)$. This implies that $\mathbb{P}_\mu\{|\overline{X}_n - \mu| \ge \max(\varepsilon,\ \eta|\mu|)\} < \delta$ provided that the corresponding sample size

$$\begin{aligned} n &> \frac{\ln\frac{2}{\delta}}{\varphi(-\zeta, \vartheta^*)} = \frac{2\left(\vartheta^* - \frac{\zeta}{3}\right)\left(1 - \vartheta^* + \frac{\zeta}{3}\right)}{\zeta^2}\ln\frac{2}{\delta} \\ &= \frac{2\left(-\frac{\lambda}{b-a} + c - \frac{\zeta}{3}\right)\left(1 + \frac{\lambda}{b-a} - c + \frac{\zeta}{3}\right)}{\zeta^2}\ln\frac{2}{\delta} \\ &= \frac{2}{\varepsilon^2}\left(-\frac{\varepsilon}{\eta} - \frac{\varepsilon}{3} - a\right)\left(b + \frac{\varepsilon}{\eta} + \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta}. \end{aligned}$$

Case (II). As a consequence of $\frac{\varepsilon}{3} < \frac{a+b}{2}$, we have $p^* > c$. Clearly, $q^* > p^* > c > \vartheta^*$. By Lemma 9, $\mathscr{H}_3(\theta)$ is increasing for $\theta \in (0, \vartheta^*]$. By Lemma 1, $\mathscr{Q}_1(\theta)$ is increasing for $\theta \in [\vartheta^*, c)$. Hence,

$$\mathrm{Q}^+(\theta) \le \mathscr{H}_3(\theta) \le \mathscr{H}_3(\vartheta^*) = \mathscr{Q}_1(\vartheta^*) \le \mathscr{Q}_1(c) \quad \text{for } \theta \in (0, \vartheta^*]. \tag{57}$$

Moreover,

$$\mathrm{Q}^+(\theta) \le \mathscr{Q}_1(\theta) \le \mathscr{Q}_1(c) \quad \text{for } \theta \in (\vartheta^*, c]. \tag{58}$$

Similarly, by Lemma 10, $\mathscr{H}_4(\theta)$ is non-decreasing for $\theta \in (0, \vartheta^*]$. By Lemma 2, $\mathscr{Q}_2(\theta)$ is non-decreasing for $\theta \in [\vartheta^*, c)$. Hence,

$$\mathrm{Q}^-(\theta) \le \mathscr{H}_4(\theta) \le \mathscr{H}_4(\vartheta^*) = \mathscr{Q}_2(\vartheta^*) \le \mathscr{Q}_2(c) \quad \text{for } \theta \in (0, \vartheta^*]. \tag{59}$$

Moreover,

$$\mathrm{Q}^-(\theta) \le \mathscr{Q}_2(\theta) \le \mathscr{Q}_2(c) \quad \text{for } \theta \in (\vartheta^*, c]. \tag{60}$$

Making use of (57)–(60), we have

$$\mathrm{Q}^+(\theta) + \mathrm{Q}^-(\theta) \le \mathscr{Q}_1(c) + \mathscr{Q}_2(c)$$

for $\theta \in (0, c)$. Observing that

$$\mathscr{Q}_1(c) = \exp\left(-\frac{n\zeta^2}{2(c + \frac{\zeta}{3})(1 - c - \frac{\zeta}{3})}\right), \qquad \mathscr{Q}_2(c) = \exp\left(-\frac{n\zeta^2}{2(c - \frac{\zeta}{3})(1 - c + \frac{\zeta}{3})}\right)$$

and that

$$\left(c + \frac{\zeta}{3}\right)\left(1 - c - \frac{\zeta}{3}\right) - \left(c - \frac{\zeta}{3}\right)\left(1 - c + \frac{\zeta}{3}\right) = \frac{2\zeta}{3}(1 - 2c) > \frac{2\zeta}{3}(1 - 2p^*) > 0,$$

we have $\mathscr{Q}_1(c) > \mathscr{Q}_2(c)$ and consequently,

$$\mathrm{Q}^+(\theta) + \mathrm{Q}^-(\theta) \le 2\mathscr{Q}_1(c)$$

for $\theta \in (0, c)$. It follows that

$$\mathbb{P}_\mu\{|\overline{X}_n - \mu| \ge \max(\varepsilon,\ \eta|\mu|)\} \le \mathrm{Q}^+(\theta) + \mathrm{Q}^-(\theta) \le 2\mathscr{Q}_1(c) = 2\exp\left(-n\,\varphi(\zeta, c)\right)$$

for $\mu \in (a, 0)$. This implies that $\mathbb{P}_\mu\{|\overline{X}_n - \mu| \ge \max(\varepsilon,\ \eta|\mu|)\} < \delta$ provided that the corresponding sample size

$$\begin{aligned} n &> \frac{\ln\frac{2}{\delta}}{\varphi(\zeta, c)} = \frac{2(c + \frac{\zeta}{3})(1 - c - \frac{\zeta}{3})}{\zeta^2}\ln\frac{2}{\delta} \\ &= \frac{2(-a + \frac{\varepsilon}{3})(b - a + a - \frac{\varepsilon}{3})}{\varepsilon^2}\ln\frac{2}{\delta} \\ &= \frac{2}{\varepsilon^2}\left(\frac{\varepsilon}{3} - a\right)\left(b - \frac{\varepsilon}{3}\right)\ln\frac{2}{\delta}. \end{aligned}$$

Finally, consider Case (III). In this case, we have $\mathbb{P}_\mu\{|\overline{X}_n - \mu| \geq \max(\varepsilon, \ \eta|\mu|)\} \leq Q^+(\theta) + Q^-(\theta) \leq \mathbb{P}_\theta\{|\overline{Y}_n - \theta| \geq \zeta\} \leq 2\exp(-2n\zeta^2)$. Therefore, $\mathbb{P}_\mu\{|\overline{X}_n - \mu| \geq \max(\varepsilon, \ \eta|\mu|)\} < \delta$ provided that

$$n > \frac{\ln \frac{2}{\delta}}{2\zeta^2} = \frac{(b-a)^2}{2\varepsilon^2} \ln \frac{2}{\delta}.$$

This completes the proof of the lemma. □

Lemma 13. *Let N and M be defined by (43) and (52), respectively. Then,*

$$\max(N, M) = \begin{cases} \dfrac{2}{\varepsilon^2}\left[|a+b|\left(\dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3}\right) - \left(\dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3}\right)^2 - ab\right] \ln \dfrac{2}{\delta} & \textit{for } \dfrac{|a+b|}{2} > \dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3}, \\ \dfrac{(b-a)^2}{2\varepsilon^2} \ln \dfrac{2}{\delta} & \textit{else.} \end{cases} \tag{61}$$

Proof. To prove the lemma, it suffices to consider two cases as follows.

Case (A): $\frac{a+b}{2} \geq 0$.

Case (B): $\frac{a+b}{2} < 0$.

In Case (A), as a consequence of $\frac{a+b}{2} \geq 0$, we have

$$N = \begin{cases} \dfrac{2}{\varepsilon^2}\left(\dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3} - a\right)\left(b - \dfrac{\varepsilon}{\eta} - \dfrac{\varepsilon}{3}\right) \ln \dfrac{2}{\delta} & \text{for } \dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3} < \dfrac{a+b}{2}, \\ \dfrac{(b-a)^2}{2\varepsilon^2} \ln \dfrac{2}{\delta} & \text{else} \end{cases}$$

and

$$M = \begin{cases} \dfrac{2}{\varepsilon^2}\left(\dfrac{\varepsilon}{3} - a\right)\left(b - \dfrac{\varepsilon}{3}\right) \ln \dfrac{2}{\delta} & \text{for } \dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3} < \dfrac{a+b}{2}, \\ \dfrac{2}{\varepsilon^2}\left(\dfrac{\varepsilon}{3} - a\right)\left(b - \dfrac{\varepsilon}{3}\right) \ln \dfrac{2}{\delta} & \text{for } \dfrac{\varepsilon}{3} < \dfrac{a+b}{2}, \ \dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3} > \dfrac{a+b}{2}, \\ \dfrac{(b-a)^2}{2\varepsilon^2} \ln \dfrac{2}{\delta} & \text{else.} \end{cases}$$

Therefore, in Case (A), we have

$$\max(N, M) = \begin{cases} \dfrac{2}{\varepsilon^2}\left(\dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3} - a\right)\left(b - \dfrac{\varepsilon}{\eta} - \dfrac{\varepsilon}{3}\right) \ln \dfrac{2}{\delta} & \text{for } \dfrac{a+b}{2} > \dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3}, \\ \dfrac{(b-a)^2}{2\varepsilon^2} \ln \dfrac{2}{\delta} & \text{for } 0 < \dfrac{a+b}{2} < \dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3}. \end{cases} \tag{62}$$

In Case (B), as a consequence of $\frac{a+b}{2} < 0$, we have

$$N = \begin{cases} \dfrac{2}{\varepsilon^2}\left(-\dfrac{\varepsilon}{3} - a\right)\left(b + \dfrac{\varepsilon}{3}\right) \ln \dfrac{2}{\delta} & \text{for } \dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3} + \dfrac{a+b}{2} < 0, \\ \dfrac{2}{\varepsilon^2}\left(-\dfrac{\varepsilon}{3} - a\right)\left(b + \dfrac{\varepsilon}{3}\right) \ln \dfrac{2}{\delta} & \text{for } \dfrac{\varepsilon}{3} + \dfrac{a+b}{2} < 0, \ \dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3} + \dfrac{a+b}{2} > 0, \\ \dfrac{(b-a)^2}{2\varepsilon^2} \ln \dfrac{2}{\delta} & \text{else} \end{cases}$$

and

$$M = \begin{cases} \dfrac{2}{\varepsilon^2}\left(-\dfrac{\varepsilon}{\eta} - \dfrac{\varepsilon}{3} - a\right)\left(b + \dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3}\right) \ln \dfrac{2}{\delta} & \text{for } \dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3} < -\dfrac{a+b}{2}, \\ \dfrac{(b-a)^2}{2\varepsilon^2} \ln \dfrac{2}{\delta} & \text{else.} \end{cases}$$

Therefore, in Case (B), we have

$$\max(N, M) = \begin{cases} \dfrac{2}{\varepsilon^2}\left(-\dfrac{\varepsilon}{\eta} - \dfrac{\varepsilon}{3} - a\right)\left(b + \dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3}\right) \ln \dfrac{2}{\delta} & \text{for } \dfrac{a+b}{2} < -\left(\dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3}\right), \\ \dfrac{(b-a)^2}{2\varepsilon^2} \ln \dfrac{2}{\delta} & \text{for } -\left(\dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3}\right) < \dfrac{a+b}{2} < 0. \end{cases} \tag{63}$$

Combining (62) and (63), we have

$$\max(N, M) = \begin{cases} \dfrac{2}{\varepsilon^2}\left(\dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3} - a\right)\left(b - \dfrac{\varepsilon}{\eta} - \dfrac{\varepsilon}{3}\right)\ln\dfrac{2}{\delta} & \text{for } \dfrac{a+b}{2} > \dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3}, \\ \dfrac{2}{\varepsilon^2}\left(-\dfrac{\varepsilon}{\eta} - \dfrac{\varepsilon}{3} - a\right)\left(b + \dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3}\right)\ln\dfrac{2}{\delta} & \text{for } \dfrac{a+b}{2} < -\left(\dfrac{\varepsilon}{\eta} + \dfrac{\varepsilon}{3}\right), \\ \dfrac{(b-a)^2}{2\varepsilon^2}\ln\dfrac{2}{\delta} & \text{else} \end{cases}$$

which implies (61). This completes the proof of the lemma. □

Finally, Theorem 4 can be established by making use of Lemmas 11–13.

References

Alamouti, S.M., 1998. A simple transmit diversity technique for wireless communications. IEEE J. Sel. Areas Commun. 16 (8), 1451–1458.

Arellano, M., Pakkala, S., Langston, A., Tighiouart, M., Pan, L., Chen, Z., Heffner, L.T., Lonial, S., Winton, E., Khoury, H.J., 2012. Early clearance of peripheral blood blasts predicts response to induction chemotherapy in acute myeloid leukemia. Cancer 118 (21), 5278–5282.

Azuma, K., 1967. Weighted sums of certain dependent random variables. Tôkuku Math. J. 19 (3), 357–367.

Bentkus, V., 2004. On Hoeffdings inequalities. Ann. Probab. 32 (2), 1650–1673.

Chow, S.C., Shao, J., Wang, H., 2008. Sample Size Calculations in Clinical Trials, second ed. Chapman & Hall.

Desu, M.M., Raghavarao, D., 1990. Sample Size Methodology. Academic Press.

Doob, J., 1953. Stochastic Processes. Wiley.

Fishman, G.S., 1996. Monte Carlo – Concepts, Algorithms and Applications. Spring-Verlag.

Franklin, G.F., Powell, J.D., Emami-Naeini, A., 2014. Feedback Control of Dynamic Systems. Pearson Higher Education, Inc.

Gajek, L., Niemiro, W., Pokarowski, P., 2013. Optimal Monte Carlo integration with fixed relative precision. J. Complexity 29, 4–26.

Hampel, F., 1998. Is statistics too difficult? The Canadian Journal of Statistics 26, 497–513.

Hoeffding, W., 1963. Probability inequalities for sums of bounded variables. J. Amer. Statist. Assoc. 58, 13–29.

Janik, M., Hartlage, G., Alexopoulos, N., Mirzoyev, Z., McLean, D.S., Arepalli, C.D., Chen, Z., Stillman, A.E., Raggi, P., 2010. Epicardial adipose tissue volume and coronary artery calcium to predict myocardial ischemia on positron emission tomography-computed tomography studies. J. Nuclear Cardiology 17 (5), 841–847.

Khargonekar, P., Tikku, A., Randomized algorithms for robust control analysis and synthesis have polynomial complexity, in: Proceedings of the IEEE Conference on Decision and Control, 1996.

Lagoa, C.M., Barmish, B.R., Distributionally robust Monte Carlo simulation: a tutorial survey, in: Proceedings of the IFAC World Congress, 2002.

Larsson, E., Stoica, P., 2003. Space-Time Block Coding For Wireless Communications. Cambridge University Press, UK.

Massart, P., 1990. The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. Ann. Probab. 18, 1269–1283.

Mitzenmacher, M., Upfal, E., 2005. Probability and Computing. Cambridge University Press.

Motwani, R., Raghavan, P., 1995. Randomized Algorithms. Cambridge University Press.

Proakis, J.G., 2000. Digital Communications. Mcgraw-Hill.

Tempo, R., Calafiore, G., Dabbene, F., 2005. Randomized Algorithms for Analysis and Control of Uncertain Systems. Springer.

Tremba, G. Calafiore, Dabbene, F., Gryazina, E., Polyak, B.T., Shcherbakov, P.S., Tempo, R., RACT: Randomized Algorithms Control Toolbox for MATLAB, in: Proc. of the IFAC World Congress, Seoul, Korea, July 2008.

Wang, D., Müller, S., Amin, A.R., Huang, D., Su, L., Hu, Z., Rahman, M.A., Nannapaneni, S., Koenig, L., Chen, Z., Tighiouart, M., Shin, D.M., Chen, Z.G., 2012. The pivotal role of integrin $\beta 1$ in metastasis of head and neck squamous cell carcinoma. Clinical Cancer Res. 18 (17), 4589–4599.

Willams, D., 1991. Probability with Martingales. Cambriage University Press.