THE CONVERGENCE RATE AND ASYMPTOTIC DISTRIBUTION OF BOOTSTRAP QUANTILE VARIANCE ESTIMATOR FOR IMPORTANCE SAMPLING

JINGCHEN LIU AND XUAN YANG,
∗ Columbia University

Abstract

Importance sampling is a widely used variance reduction technique to compute sample quantiles such as value-at-risk. The variance of the weighted sample quantile estimator is usually a difficult quantity to compute. In this paper, we present the exact convergence rate and asymptotic distributions of the bootstrap variance estimators for quantiles of weighted empirical distributions. Under regularity conditions, we show that the bootstrap variance estimator is asymptotically normal and has relative standard deviation of order $O(n^{-1/4})$.

1. Introduction

In this paper, we derive the asymptotic distributions of the bootstrap quantile variance estimators for weighted samples. Let $F$ be a cumulative distribution function (c.d.f.), $f$ be its density function, and $\alpha_p = \inf\{x : F(x) \geq p\}$ be its $p$-th quantile. It is well known that the asymptotic variance of the $p$-th sample quantile is inversely proportional to $f(\alpha_p)$ (c.f. [6]). When $f(\alpha_p)$ is close to zero (e.g. $p$ is close to zero or one), the sample quantile becomes very unstable since the “effective samples” size is small. In the scenario of Monte Carlo, one solution is using importance sampling for variance reduction by distributing more samples around a neighborhood of the interesting quantile $\alpha_p$. Such a technique has been widely employed in multiple disciplines. In portfolio risk management, the $p$-th quantile of a portfolio’s total asset price is an important risk measure. This quantile is also known as the value-at-risk.

∗ Postal address: 1255 Amsterdam Ave, Department of Statistics, New York, NY 10027, U.S.A.
Typically, the probability $p$ in this context is very close to zero (or one). A partial list of literature of using importance sampling to compute the value-at-risk includes [25, 26, 28, 42, 43, 44, 14, 23]. A recent work by [30] discussed efficient importance sampling for risk measure computation for heavy-tailed distributions. In the system stability assessment of engineering, the extreme quantile evaluation is of interest. In this context, the interesting probabilities are typically of a smaller order than those of the portfolio risk analysis.

Upon considering $p$ being close to zero or one, the computation of $\alpha_p$ can be viewed as the inverse problem of rare-event simulation. The task of the latter topic is computing the tail probabilities $1 - F(b)$ when $b$ tends to infinity. Similar to the usage in the quantile estimation, importance sampling is also a standard variance reduction technique for rare-event simulation. The first work on this topic is given by [41], which not only presents an efficient importance sampling estimator but also defines a second-moment-based efficiency measure. We will later see that such a measure is also closely related to the asymptotic variance of the weighted quantiles. Such a connection allows people to adapt the efficient algorithms designed for rare-event simulations to the computation of quantiles (c.f. [24, 30]). More recent works of rare-event simulations for light-tailed distributions include [19, 39, 17] and for heavy-tailed distributions include [2, 4, 18, 31, 7, 8, 10, 11]. There are also standard textbooks such as [13, 3].

Another related field of this line of work is survey sampling where unequal probability sampling and weighted samples are prevailing (c.f. [32, 36]). The weights are typically defined as the inverse of the inclusion probabilities.

The estimation of distribution quantile is a classic topic. The almost sure result of sample quantile is established by [6]. The asymptotic distribution of (unweighted) sample quantile can be found in standard textbook such as [15]. Estimation of the (unweighted) sample quantile variance via bootstrap was proposed by [37, 38, 40, 5, 22]. There are also other kernel based estimators
Bootstrap for Weighted Quantile Variances

(to estimate $f(\alpha_p)$) for such variances (c.f. [21]).

There are several pieces of works immediately related to the current one. The first one is [29], which derived the asymptotic distribution of the bootstrap quantile variance estimator for unweighted i.i.d. samples. Another one is given by [27] who derived the asymptotic distribution of weighted quantile estimators; see also [14] for a confidence interval construction. A more detailed discussion of these results is given in Section 2.2.

The asymptotic variance of weighted sample quantile, as reported in [27], contains the density function $f(\alpha_p)$, whose evaluation typically consists of computation of high dimensional convolutions and therefore is usually not straightforward. In this paper, we propose to use bootstrap method to compute/estimate the variance of such a weighted quantile. Bootstrap is a generic method that is easy to implement and does not consist of tuning parameters in contrast to the kernel based methods for estimating $f(\alpha_p)$. This paper derives the convergence rate and asymptotic distribution of the bootstrap variance estimator for weighted quantiles. More specifically, the main contributions are to first provide conditions under which the quantiles of weighted samples have finite variances and develop their asymptotic approximations. Second, we derive the asymptotic distribution of the bootstrap estimators for such variances. Let $n$ denote the sample size. Under regularity conditions (for instance, moment conditions and continuity conditions for the density functions), we show that the bootstrap variance estimator is asymptotically normal with a convergence rate of order $O(n^{-5/4})$. Given that the quantile variance decays at the rate of $O(n^{-1})$, the relative standard deviation of a bootstrap estimator is $O(n^{-1/4})$. Lastly, we present the asymptotic distribution of the bootstrap estimator for one particular case where $p \to 0$.

The technical challenge lies in that many classic results of order statistics are not applicable. This is mainly caused by the variations introduced by the weights, which in the current context is the Radon-Nikodym derivative, and
the weighted sample quantile does not map directly to the ordered statistics. In this paper, we employed Edgeworth expansion combined with the strong approximation of empirical processes ([33]) to derive the results.

This paper is organized as follows. In Section 2, we present our main results and summarize the related results in literature. A numerical implementation is given in Section 3 to illustrate the performance of the bootstrap estimator. The proofs of the theorems are provided in Sections 4, 5, and 6.

2. Main results

2.1. Problem setting

Consider a probability space $(\Omega, \mathcal{F}, P)$ and a random variable $X$ admitting cumulative distribution function (c.d.f.) $F(x) = P(X \leq x)$ and density function

$$f(x) = F'(x)$$

for all $x \in \mathbb{R}$. Let $\alpha_p$ be its $p$-th quantile, that is,

$$\alpha_p = \inf \{x : F(x) \geq p \}.$$ 

Consider a change of measure $Q$, under which $X$ admits a cumulative distribution function $G(x) = Q(X \leq x)$ and density

$$g(x) = G'(x).$$

Let

$$L(x) = \frac{f(x)}{g(x)},$$

and $X_1, \ldots, X_n$ be i.i.d. copies of $X$ under $Q$. Assume that $P$ and $Q$ are absolutely continuous with respect to each other, then $E^Q L(X_i) = 1$. The corresponding weighted empirical c.d.f. is

$$\hat{F}_X(x) = \frac{\sum_{i=1}^n L(X_i) I(X_i \leq x)}{\sum_{i=1}^n L(X_i)}. \quad (1)$$
A natural estimator of $\alpha_p$ is

$$\hat{\alpha}_p(X) = \inf\{x \in \mathbb{R} : \hat{F}_X(x) \geq p\}. \quad (2)$$

Of interest in this paper is the variance of $\hat{\alpha}_p(X)$ under the sampling distribution of $X_i$, that is

$$\sigma_n^2 = Var^Q(\hat{\alpha}_p(X)). \quad (3)$$

The notations $E^Q(\cdot)$ and $Var^Q(\cdot)$ are used to denote the expectation and variance under measure $Q$.

Let $Y_1, \ldots, Y_n$ be i.i.d. bootstrap samples from the empirical distribution

$$\hat{G}(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x).$$

The bootstrap estimator for $\sigma_n^2$ in (3) is defined as

$$\hat{\sigma}_n^2 = \sum_{i=1}^{n} \hat{Q}(\hat{\alpha}_p(Y) = X_i) (X_i - \hat{\alpha}_p(X))^2, \quad (4)$$

where $Y = (Y_1, \ldots, Y_n)$ and $\hat{Q}$ is the measure induced by $\hat{G}$, that is, under $\hat{Q}$ $Y_1, \ldots, Y_n$ are i.i.d. following empirical distribution $\hat{G}$. Note that both $\hat{G}$ and $\hat{Q}$ depend on $X$. To simplify the notations, we do not include the index of $X$ in the notations of $\hat{Q}$ and $\hat{G}$.

**Remark 1.** There are multiple ways to form an estimate of $F$. One alternative to (1) is

$$\tilde{F}_X(x) = \frac{1}{n} \sum_{i=1}^{n} L(X_i) I(X_i \leq x). \quad (5)$$

The analysis of $\tilde{F}_X$ is analogous to and simpler than that of (1). This is because the denominator is a constant. The weighted sample c.d.f. in (5) only depends on samples below $x$. This is an important feature for variance reduction of extreme quantile estimation when the change of measure $Q$ is designed to be concentrated on the region below $F^{-1}(p)$. We will provide more detailed discussions later.
2.2. Related results

In this section, we present two related results in the literature. First, [29] established asymptotic distribution of the the bootstrap variance estimators for the (unweighted) sample quantiles. In particular, it showed that if the density function $f(x)$ is Hölder continuous with index $\frac{1}{2} + \delta_0$ then

$$n^{5/4}(\hat{\sigma}_n^2 - \sigma_n^2) \Rightarrow N(0, 2\pi^{-1/2}[p(1 - p)]^{3/2}f(\alpha_p)^{-4})$$ (6)

as $n \to \infty$. This is consistent with the results in Theorem 2 by setting $L(x) \equiv 1$. This paper can be viewed as a natural extension of [29], though the proof techniques are different.

In the context of importance sampling, as shown by [27], if $E^Q|L(x)|^3 < \infty$, the asymptotic distribution of a weighted quantile is

$$\sqrt{n}(\hat{\alpha}_p(X) - \alpha_p) \Rightarrow N\left(0, \frac{\text{Var}^Q(W_p)}{f(\alpha_p)^2}\right)$$ (7)

as $n \to \infty$, where $W_p = L(X)(I(X < \alpha_p) - p)$. More general results in terms of weighted empirical processes are given by [30]

We now provide a brief discussion on the efficient quantile computation via importance sampling. The sample quantile admits a large variance when $f(\alpha_p)$ is small. One typical situation is that $p$ is very close to zero or one. To fix ideas, we consider the case where $p$ tends to zero. The asymptotic variance of the $p$-th quantile of $n$ i.i.d. samples is

$$\frac{1 - p}{np} \frac{p^2}{f(\alpha_p)^2}.$$ 

Then, in order to obtain an estimate of an $\varepsilon$ error with at least $1 - \delta$ probability, the necessary number of i.i.d. samples is proportional to $p^{-1} \frac{p^2}{f(\alpha_p)}$, which grows to infinity as $p \to 0$. Typically, the inverse of the hazard function, $p/f(\alpha_p)$, varies slowly as $p$ tends to zero. For instance, $p/f(\alpha_p)$ is bounded if $X$ is a light-tailed random variable and grows at the most linearly in $\alpha_p$ for most heavy-tailed distributions (e.g., regularly varying distribution, log-normal distribution).
The asymptotic variance of the quantiles of \( \hat{F}_X \) defined in (5) is

\[
\frac{\text{Var}^Q(L(X)I(X \leq \alpha_p))}{np^2} p^2 f(\alpha_p)^2.
\]

There is a wealth of literature on the design of importance sampling algorithms particularly adapted to the context in which \( p \) is close to zero. A well accepted efficiency measure is precisely based on the relative variance \( p^{-2}\text{Var}^Q(L(X)I(X \leq \alpha_p)) \) as \( p \to 0 \). More precisely, the change of measure is called strongly efficient, if \( p^{-2}\text{Var}^Q(L(X)I(X \leq \alpha_p)) \) is bounded for arbitrarily small \( p \). A partial list of recent developments of importance sampling algorithms in the rare event setting includes [1, 9, 10, 12, 19]. Therefore, the change of measure designed to estimate \( p \) can be adapted without much additional effort to the quantile estimation problem. For a more thorough discussion, see [30, 14]. We will provide the analysis of one special in Theorem 3.

2.3. The results for regular quantiles

In this subsection, we provide an asymptotic approximation of \( \sigma_n^2 \) and the asymptotic distribution of \( \hat{\sigma}_n^2 \). We first list a set of conditions which we will refer to in the statements of our theorems.

A1 There exists an \( \alpha > 4 \) such that

\[
E^Q|L(X)|^\alpha < \infty.
\]

A2 There exists a \( \beta > 3 \) such that

\[
E^Q|X|^\beta < \infty.
\]

A3 Assume that

\[
\frac{\alpha}{3} > \frac{\beta + 2}{\beta - 3}.
\]

A4 There exists a \( \delta_0 > 0 \) such that the density functions \( f(x) \) and \( g(x) \) are Hölder continuous with index \( \frac{1}{2} + \delta_0 \) in a neighborhood of \( \alpha_p \), that is, there
exists a constant $c$ such that

$$|f(x) - f(y)| \leq c|x - y|^\frac{1}{2} + \delta_0, \quad |g(x) - g(y)| \leq c|x - y|^\frac{1}{2} + \delta_0,$$

for all $x$ and $y$ in a neighborhood of $\alpha_p$.

**A5** The measures $P$ and $Q$ are absolutely continuous with respect to each other. The likelihood ratio $L(x) \in (0, \infty)$ is Lipschitz continuous in a neighborhood of $\alpha_p$.

**A6** Assume $f(\alpha_p) > 0$.

**Theorem 1.** Let $F$ and $G$ be the cumulative distribution functions of a random variable $X$ under probability measures $P$ and $Q$ respectively. The distributions $F$ and $G$ have density functions $f(x) = F'(x)$ and $g(x) = G'(x)$. We assume that conditions A1 - A6 hold. Let

$$W_p = L(X)I(X \leq \alpha_p) - pL(X),$$

and $\hat{\alpha}_p(X)$ be as defined in (2). Then,

$$\sigma_n^2 \triangleq \text{Var}_Q(\hat{\alpha}_p(X)) = \frac{\text{Var}_Q(W_p)}{n f(\alpha_p)^2} + o(n^{-5/4}), \quad E^Q(\hat{\alpha}_p(X)) = \alpha_p + o(n^{-3/4})$$

as $n \to \infty$.

**Theorem 2.** Suppose that the conditions in Theorem 1 hold and $L(X)$ has density under $Q$. Let $\hat{\sigma}_n^2$ be defined as in (4). Then, under $Q$

$$n^{5/4}(\hat{\sigma}_n^2 - \sigma_n^2) \Rightarrow N(0, \tau_p^2)$$

as $n \to \infty$, where "$\Rightarrow$" denotes weak convergence and

$$\tau_p^2 = 2\pi^{-1/2}L(\alpha_p)f(\alpha_p)^{-1}(\text{Var}_Q(W_p))^{3/2}.$$

**Remark 2.** In Theorem 1, we provide bounds on the errors of the asymptotic approximations for $E^Q(\hat{\alpha}_p(X))$ and $\sigma_n^2$ in order to assist the analysis of the bootstrap estimator. In particular, in order to approximate $\sigma_n^2$ with an accuracy of order $o(n^{-5/4})$, it is sufficient to approximate $E^Q(\hat{\alpha}_p(X))$ with an
accuracy of order $o(n^{-5/8})$. Thus, Theorem 1 indicates that $\hat{\alpha}_p$ can be viewed as asymptotically unbiased. In addition, given that the bootstrap estimator has a convergence rate of $O_p(n^{-5/4})$, Theorem 1 suggests that when computing the distribution of the bootstrap estimator, we can use the approximation of $\sigma_n^2$ to replace the true variance.

In Theorem 2, if we let $L(x) \equiv 1$, that is, $P = Q$, then $\hat{\alpha}_p$ is the regular quantile and the asymptotic distribution (8) recovers the result of [29] given in (6).

**Remark 3.** Note that the weak convergence in (7) requires weaker conditions than those in Theorems 1 and 2. The weak convergence does not require $\hat{\alpha}_p(X)$ to have a finite variance. In contrast, in order to apply the bootstrap variance estimator, one needs to have the estimand well defined, that is, $Var_Q(\hat{\alpha}_p(X)) < \infty$. Conditions A1-3 are imposed to insure that $\hat{\alpha}_p(X)$ has a finite variance under $Q$.

The continuity assumptions on the density function $f$ and the likelihood ratio function $L$ (conditions A4 and A5) are typically satisfied in practice. Condition A6 is necessary for the quantile to have a variance of order $O(n^{-1})$.

### 2.4. Results for extreme quantile estimation

In this subsection, we consider one particular scenario when $p$ tends to zero. The analysis in the context of extreme quantile is sensitive to the underlying distribution and the choice of change of measure. Here, we only consider a stylized case, one of the first two cases considered in the rare-event simulation literature. Let $X = \sum_{j=1}^{m} Z_i$ where $Z_i$’s are i.i.d. random variables with mean zero and density function $h(z)$. The random variable $X$ has density function $f(x)$ that is the $m$-th convolution of $h(z)$. Note that both $X$ and $f(x)$ depend on $m$. To simplify notation, we omit the index $m$ when there is no ambiguity.

We further consider the exponential change of measure

$$Q(X \in dx) = e^{\theta x - m\varphi(\theta)} f(x) dx,$$
where \( \varphi(\theta) = \log \int e^{\theta x} h(x) dx \). We say \( \varphi \) is steep if, for every \( a \), \( \varphi(\theta) = a \) has a solution. For \( \varepsilon > 0 \), let \( \alpha_p = -m\varepsilon \) be in the large deviations regime. We let \( \theta \) be the solution to \( \sup_{\theta'}(-\theta' \varepsilon - \varphi(\theta')) \) and \( I = \sup_{\theta'}(-\theta' \varepsilon - \varphi(\theta')) \). Then, a well known approximation of the tail probability is given by

\[
P(X < -m\varepsilon) = \frac{c(\theta) + o(1)}{\sqrt{m}} e^{-mI},
\]
as \( m \to \infty \). The likelihood ratio is given by

\[
LR(x) = e^{-\theta x + m\varphi(\theta)}.
\]

We use the notation “\( LR(x) \)” to make a difference from the previous likelihood ratio denoted by “\( L(x) \)”.

**Theorem 3.** Suppose that \( X = \sum_{j=1}^{m} Z_i \), where \( Z_i \)’s are i.i.d mean-zero random variable’s with Lipschitz continuous density function. The log-MGF \( \varphi(\theta) \) is steep. For \( \varepsilon > 0 \), equation

\[
\varphi'(\theta) = -\varepsilon
\]
has one solution denoted by \( \theta \). Let \( \alpha_p = -m\varepsilon \) and \( X_1, ..., X_m \) be i.i.d. samples generated from exponential change of measure

\[
Q(X \in dx) = e^{\theta x - m\varphi(\theta)} f(x).
\]

Let

\[
\tilde{F}_{X}(x) = \frac{1}{n} \sum_{i=1}^{n} LR(X_i) I(X_i \leq x), \quad \hat{\alpha}_p(X) = \inf(x \in \mathbb{R} : \tilde{F}_{X}(x) \geq p).
\]

Let \( Y_1, ..., Y_n \) be i.i.d. samples from the empirical measure \( \hat{Q} \) and \( \hat{\sigma}_n^2 \) be as defined in (4). If \( m \) (growing as a function of \( n \) and denoted by \( m_n \)) admits the limit \( m_n^3/n \to c_* \in [0, +\infty) \) as \( n \to \infty \), then

\[
\frac{n^{5/4}}{\hat{\tau}_p}(\hat{\sigma}_n^2 - \sigma_n^2) \Rightarrow N(0, 1)
\]
as \( n \to \infty \), where \( \sigma_n^2 = \text{Var}^Q(\hat{\alpha}_p(X)) \)

\[
\hat{\tau}_p^2 = 2\pi^{-1/2} LR(\alpha_p) f(\alpha_p)^{-4} (\text{Var}^Q(\hat{\alpha}_p))^3/2, \quad \hat{W}_p = LR(X) I(X \leq \alpha_p).
\]
Remark 4. The above theorem establishes the asymptotic distribution of the bootstrap variance estimator for a very stylized rare-event simulation problem. For more general situations, further investigations are necessary, such as the heavy-tailed cases where the likelihood ratios do not behave as well as those of the lighted cases even for strongly efficient estimators.

Remark 5. To simplify the notations, we drop the subscript of $n$ in the notation of $m_n$ and write $m$ whenever there is no ambiguity.

As $m$ tends to infinity, the term $\tilde{\tau}_p^2$ is no longer a constant. With the standard large deviations results (e.g. Lemma 6 stated later in the proof), we have that $\tilde{\tau}_p^2$ is of order $O(m^{5/4})$. Therefore, the convergence rate of $\hat{\sigma}_n^2$ is $O(m^{5/8}n^{-5/4})$. In addition, $\sigma_n^2$ is of order $O(m^{1/2}n^{-1})$. Thus, the relative convergence rate of $\hat{\sigma}_n^2$ is $O(m^{1/8}n^{-1/4})$. Choosing $n$ so that $m^3 = O(n)$ is sufficient to estimate $\alpha_p$ with $\epsilon$ accuracy and $\sigma_n^2$ with $\epsilon$ relative accuracy.

The empirical c.d.f. in Theorem 3 is different from that in Theorems 1 and 2. We emphasize that it is necessary to use (9) to obtain the asymptotic results. This is mainly because the variance of $LR(X)$ grows exponentially fast as $m$ tends to infinity. Then, the normalizing constant of the empirical c.d.f. in (1) is very unstable. In contrast, the empirical c.d.f. in (9) only depends on the samples below $x$. Note that the change of measure is designed to reduce the variance of $LR(X)I(X \leq \alpha_p)$. Thus, the asymptotic results hold when $n$ grows on the order of $m^3$ or faster.

3. A numerical example

In this section, we provide one numerical example to illustrate the performance of the bootstrap variance estimator. In order to compare the bootstrap estimator with the asymptotic approximation in Theorem 1, we choose an example for which the marginal density $f(x)$ is in a closed form and $\alpha_p$ can be
\[ p_1^{\alpha} - p_1^{\hat{\alpha}} - p_1^{\sigma_n^2} - p_1^{\tilde{\sigma}_n^2} - p_1^{\hat{\sigma}_n^2} \]

\begin{center}
\begin{tabular}{ccccccc}
\hline
\( p \) & \( p_1^{\alpha} \) & \( p_1^{\hat{\alpha}} \) & \( p_1^{\sigma_n^2} \) & \( p_1^{\tilde{\sigma}_n^2} \) & \( p_1^{\hat{\sigma}_n^2} \) \\
\hline
0.05 & 15.70 & 15.67 & 0.0032 & 0.0031 & 0.0027 \\
0.04 & 16.16 & 16.13 & 0.0034 & 0.0033 & 0.0029 \\
0.03 & 16.73 & 16.70 & 0.0037 & 0.0036 & 0.0032 \\
0.02 & 17.51 & 17.47 & 0.0042 & 0.0041 & 0.0037 \\
0.01 & 18.78 & 18.74 & 0.0054 & 0.0052 & 0.0047 \\
\hline
\end{tabular}
\end{center}

\( \sigma_n^2 \): the quantile variance computed by crude Monte Carlo.

\( \tilde{\sigma}_n^2 \): the asymptotic approximation of \( \sigma_n^2 \) in Theorem 1.

\( \hat{\sigma}_n^2 \): the bootstrap estimate of \( \sigma_n^2 \).

**Table 1:** Comparison of variance estimators for fixed \( m = 10 \) and \( n = 10,000 \).

\begin{center}
\begin{tabular}{cccccccc}
\hline
\( m \) & \( p \) & \( \alpha_{1-p} \) & \( \bar{\alpha}_{1-p} \) & \( \sigma_n^2 \) & \( \tilde{\sigma}_n^2 \) & \( \hat{\sigma}_n^2 \) \\
\hline
10 & 7.0e-02 & 15 & 14.95 & 1.1e-03 & 1.0e-03 & 1.2e-03 \\
30 & 7.3e-03 & 45 & 44.95 & 2.2e-03 & 2.2e-03 & 1.7e-03 \\
50 & 9.0e-04 & 75 & 74.98 & 3.0e-03 & 3.2e-03 & 2.4e-03 \\
100 & 5.9e-06 & 150 & 149.99 & 4.8e-03 & 4.9e-03 & 5.0e-03 \\
\hline
\end{tabular}
\end{center}

**Table 2:** Comparison of variance estimators as \( m \to \infty \), \( \alpha_p = 1.5m \), and \( n = 10,000 \).

computed numerically. Consider a partial sum

\[ X = \sum_{i=1}^{m} Z_i \]

where \( Z_i \)'s are i.i.d. exponential random variables with rate one. Then, the density function of \( X \) is

\[ f(x) = \frac{x^{m-1}}{(m-1)!} e^{-x}. \]

We are interested in computing \( X \)'s \((1-p)\)-th quantile via exponential change of measure that is

\[ \frac{dQ_\theta}{dP} = \prod_{i=1}^{m} e^{\theta Z_i - \varphi(\theta)}, \quad (11) \]

where \( \varphi(\theta) = -\log(1 - \theta) \) for \( \theta < 1 \). We further choose \( \theta = \arg \sup_{\theta'}(\theta' \alpha_p - m\varphi(\theta')) \).
We generate \( n \) i.i.d. replicates of \((Z_1, ..., Z_m)\) from \( Q_\theta \), that is, \((Z_1^{(k)}, ..., Z_m^{(k)})\) for \( k = 1, ..., n \); then, use \( X_k = \sum_{i=1}^m Z_i^{(k)}, k = 1, ..., n \) and the associated weights to form an empirical distribution and further \( \hat{\alpha}_{1-p}(X) \). Let \( \sigma_n^2 = \text{Var}_{Q}(\hat{\alpha}_{1-p}(X)) \), \( \bar{\sigma}_n^2 \) be the asymptotic approximation of \( \sigma_n^2 \), and \( \hat{\sigma}_n^2 \) be the bootstrap estimator of \( \sigma_n^2 \). We use Monte Carlo to compute both \( \sigma_n^2 \) and \( \hat{\sigma}_n^2 \) by generating independent replicates of \( \hat{\alpha}_{1-p}(X) \) under \( Q \) and bootstrap samples under \( \hat{Q} \) respectively.

We first consider the situation when \( m = 10 \). Table 1 shows the results using the estimators based on the empirical c.d.f. in Theorem 2 with \( n = 10,000 \). In the table, the column \( \sigma_n^2 \) shows the variances of \( \hat{\alpha}_{1-p} \) estimated using 1000 Monte Carlo simulations. In addition, we consider the case that \( \alpha_p = 1.5m \) and send \( m \) to infinity. Table 2 shows the numerical results using the estimators in Theorem 3 based on \( n = 10,000 \) simulations.

4. Proof of Theorem 1

Throughout our discussion we use the following notations for asymptotic behavior. We say that \( 0 \leq g(b) = O(h(b)) \) if \( g(b) \leq ch(b) \) for some constant \( c \in (0, \infty) \) and all \( b \geq b_0 > 0 \). Similarly, \( g(b) = \Omega(h(b)) \) if \( g(b) \geq ch(b) \) for all \( b \geq b_0 > 0 \). We also write \( g(b) = \Theta(h(b)) \) if \( g(b) = O(h(b)) \) and \( g(b) = \Omega(h(b)) \). Finally, \( g(b) = o(h(b)) \) as \( b \to \infty \) if \( g(b)/h(b) \to 0 \) as \( b \to \infty \).

Before the proof of Theorem 1, we first present a few useful lemmas.

**Lemma 1.** \( X \) is random variable with finite second moment, then

\[
EX = \int_{x>0} P(X > x)dx - \int_{x<0} P(X < x)dx
\]

\[
EX^2 = \int_{x>0} 2xP(X > x)dx - \int_{x<0} 2xP(X < x)dx
\]

**Lemma 2.** Let \( X_1, ..., X_n \) be i.i.d. random variables with \( EX_i = 0 \) and \( E|X_i|^\alpha < \infty \) for some \( \alpha > 2 \). For each \( \varepsilon > 0 \), there exists a constant \( \kappa \) depending on \( \varepsilon \),
$EX_i^2$ and $E|X_i|^\alpha$ such that

$$E \left| \sum_{i=1}^{n} \frac{X_i}{\sqrt{n}} \right|^{\alpha - \varepsilon} \leq \kappa$$

for all $n > 0$.

The proofs of Lemmas 1 and 2 are elementary. Therefore, we omit them.

**Lemma 3.** Let $h(x)$ be a non-negative function. There exists $\zeta_0 > 0$ such that $h(x) \leq x^{\zeta_0}$ for all $x$ sufficiently large. Then, for all $\zeta_1, \zeta_2, \lambda > 0$ such that $(\zeta_1 - 1)\lambda < \zeta_2$, we obtain

$$\int_0^{n^\lambda} h(x)\Phi(-x + o(x^{\zeta_1}n^{-\zeta_2}))dx = \int_0^{n^\lambda} h(x)\Phi(-x)dx + o(n^{-\zeta_2}),$$

as $n \to \infty$, where $\Phi$ is the c.d.f. of a standard Gaussian distribution. In addition, we write $a_n(x) = o(x^{\zeta_1}n^{-\zeta_2})$ if $a_n(x)x^{-\zeta_1}n^{\zeta_2} \to 0$ as $n \to \infty$ uniformly for $x \in (\varepsilon, n^\lambda)$.

**Proof of Lemma 3.** We first split the integral into

$$\int_0^{n^\lambda} h(x)\Phi(-x + o(x^{\zeta_1}n^{-\zeta_2}))dx$$

$$= \int_0^{(\log n)^2} h(x)\Phi(-x + o(x^{\zeta_1}n^{-\zeta_2}))dx + \int_{(\log n)^2}^{n^\lambda} h(x)\Phi(-x + o(x^{\zeta_1}n^{-\zeta_2}))dx.$$

Note that the second term

$$\int_{(\log n)^2}^{n^\lambda} h(x)\Phi(-x + o(x^{\zeta_1}n^{-\zeta_2}))dx \leq n^{\zeta_0+1}\Phi(-(\log n)^2/2) = o(n^{-\zeta_2}).$$

For the first term, note that for all $0 \leq x \leq (\log n)^2$,

$$\Phi(-x + o(x^{\zeta_1}n^{-\zeta_2})) = (1 + o(x^{\zeta_1+1}n^{-\zeta_2}))\Phi(-x).$$

Then

$$\int_0^{(\log n)^2} h(x)\Phi(-x + o(x^{\zeta_1}n^{-\zeta_2}))dx = \int_0^{(\log n)^2} h(x)\Phi(-x)dx + o(n^{-\zeta_2}).$$

Therefore, the conclusion follows immediately.
Proof of Theorem 1. Let \( \hat{\alpha}_p(X) \) be defined in (2). To simplify the notation, we omit the index \( X \) and write \( \hat{\alpha}_p(X) \) as \( \hat{\alpha}_p \). We use Lemma 1 to compute the moments. In particular, we need to approximate the following probability,

\[
Q\left(n^{1/2}(\hat{\alpha}_p - \alpha_p) > x\right) = Q\left(\hat{F}_X(\alpha_p + x n^{-1/2}) < p\right)
\]

For some \( \lambda \in \left(\frac{1}{4(\alpha - 2)}, \frac{1}{8}\right) \), we provide approximations for (12) of the following three cases: \( 0 < x \leq n^\lambda \), \( n^\lambda \leq x \leq c \sqrt{n} \), and \( x > \sqrt{n} \). The development for

\[
Q\left(n^{1/2}(\hat{\alpha}_p - \alpha_p) < x\right)
\]

on the region that \( x \leq 0 \) is the same as that of the positive side.

Case 1: \( 0 < x \leq n^\lambda \).

Let

\[
W_{x,n,i} = L(X_i) \left(I(X_i \leq \alpha_p + x n^{-1/2}) - p\right) - F(\alpha_p + x n^{-1/2}) + p.
\]

According to Berry-Esseen bound (c.f. [20]),

\[
Q\left(\sum_{i=1}^n L(X_i) \left(I(X_i \leq \alpha_p + x n^{-1/2}) - p\right) < 0\right)
\]

\[
= Q\left(\frac{\frac{1}{n} \sum_{i=1}^n W_{x,n,i}}{\sqrt{\frac{1}{n} Var^Q W_{x,n,1}}} < -\frac{F(\alpha_p + x n^{-1/2}) - p}{\sqrt{\frac{1}{n} Var^Q W_{x,n,1}}}\right)
\]

\[
= \Phi\left(-\frac{(F(\alpha_p + x n^{-1/2}) - p)}{\sqrt{\frac{1}{n} Var^Q W_{x,n,1}}}\right) + D_1(x).
\]

There exists a constant \( \kappa_1 \) such that

\[
|D_1(x)| \leq \frac{\kappa_1}{(Var^Q W_{x,n,1})^{3/2} n^{-1/2}}.
\]
Case 2: \( n^{\lambda} \leq x \leq c\sqrt{n} \).

Thanks to Lemma 2, for each \( \varepsilon > 0 \), the \((\alpha - \varepsilon)\)-th moment of \( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} W_{x,n,i} \) is bounded. By Chebyshev’s inequality, we obtain that

\[
Q \left( \sum_{i=1}^{n} L(X_i)(I(X_i \leq \alpha + x^{-1/2}) - p) < 0 \right) = Q \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} W_{x,n,i} \leq \sqrt{n} \left( p - F(\alpha + x/\sqrt{n}) \right) \right) \leq \kappa_1 \left( \frac{1}{\sqrt{n}} \left[ F(\alpha + x^{-1/2}) - p \right] \right)^{\alpha-\varepsilon} \leq \kappa_2 x^{-\alpha+\varepsilon}.
\]

Since \( \lambda > \frac{1}{4(\alpha-2)} \), we choose \( \varepsilon \) small enough such that

\[
\int_{n^{\lambda}}^{c\sqrt{n}} x Q(\hat{\alpha}_p - \alpha > x^{-1/2}) dx = O(n^{-\lambda(\alpha-2-\varepsilon)}) = o(n^{-1/4}). \tag{14}
\]

Case 3: \( x > c\sqrt{n} \).

Note that

\[
Q \left( \sum_{i=1}^{n} L(X_i)(I(X_i \leq \alpha + x^{-1/2}) - p) < 0 \right)
\]

is a non-increasing function of \( x \). Therefore, for all \( x > c\sqrt{n} \), from Case 2, we obtain that

\[
Q \left( \sum_{i=1}^{n} L(X_i)(I(X_i \leq \alpha + x^{-1/2}) - p) < 0 \right) \leq \kappa_2 (c\sqrt{n})^{-\alpha+\varepsilon} = \kappa_3 n^{-\alpha/2+\varepsilon/2}.
\]

For \( c\sqrt{n} < x \leq n^{\alpha/6-\varepsilon/6} \), we have that

\[
Q \left( \sum_{i=1}^{n} L(X_i)(I(X_i \leq \alpha + x^{-1/2}) - p) < 0 \right) \leq \kappa_3 n^{-\alpha/2+\varepsilon/2} \leq \kappa_3 x^{-3}.
\]

In addition, note that for all \( x^{\beta-3} > n^{1+\beta/2} \),

\[
Q \left( \hat{\alpha}_p > \alpha + x/\sqrt{n} \right) \leq Q(\sup_i X_i > \alpha + x^{-1/2}) = 1 - G^n(\alpha + x^{-1/2}) \leq O(1)n^{1+\beta/2}x^{-\beta} = O(x^{-3}).
\]

Therefore, \( Q(\hat{\alpha}_p > \alpha + x/\sqrt{n}) = O(x^{-3}) \) on the region \( \{c\sqrt{n} < x \leq n^{\alpha/6-\varepsilon/3}\} \cup \{x > n^{\frac{\beta+2}{\beta-3}}\} \), since \( \frac{\alpha}{3} > \frac{\beta + 2}{\beta - 3} \), one can choose \( \varepsilon \) small enough such that
$x > n^{\alpha/6 - \varepsilon/6}$ implies $x^{\beta-3} > n^{1+\beta/2}$. Therefore, for all $x > c\sqrt{n}$, we obtain that

\[ Q\left(\hat{\alpha}_p > \alpha_p + x/\sqrt{n}\right) \leq x^{-3}, \]

and

\[ \int_{c\sqrt{n}}^{\infty} x Q\left(\hat{\alpha}_p > \alpha_p + x/\sqrt{n}\right) = O(n^{-1/2}). \tag{15} \]

**A summary of Cases 1, 2, and 3.**

Summarizing the Cases 2 and 3, more specifically (14) and (15), we obtain that

\[ \int_{\lambda n}^{\infty} x Q\left(\hat{\alpha}_p > \alpha_p + x/\sqrt{n}\right) \, dx = o(n^{-1/4}). \]

Using the result in (13), we obtain that

\[
\begin{align*}
\int_0^{\infty} x Q\left(\hat{\alpha}_p > \alpha_p + x/\sqrt{n}\right) \, dx &= \int_0^{n^\lambda} x \left[ \Phi\left( -\frac{F(\alpha_p + xn^{-1/2} - p)}{\sqrt{\frac{1}{n} \text{Var}^Q_{W_{x,n,1}}} + O(n^{-1/2})} \right) + O(n^{-1/2}) \right] \, dx + o(n^{-1/4}) \\
&= \int_0^{n^\lambda} x \Phi\left( -\frac{F(\alpha_p + xn^{-1/2} - p)}{\sqrt{\frac{1}{n} \text{Var}^Q_{W_{x,n,1}}}} \right) \, dx + O(n^{2\lambda-1/2}) + o(n^{-1/4}). \tag{16}
\end{align*}
\]

Given that $\lambda < \frac{1}{8}$, we have that $O(n^{2\lambda-1/2}) = o(n^{-1/4})$. Thanks to condition A4 and the fact that $\text{Var}^Q(W_{x,n,1}) = (1 + O(xn^{-1/2}))\text{Var}^Q(W_{0,n,1})$, we have

\[ -\frac{F(\alpha_p + xn^{-1/2} - p)}{\sqrt{\frac{1}{n} \text{Var}^Q_{W_{x,n,1}}}} = -\frac{xf(\alpha_p)}{\sqrt{\text{Var}^Q_{W_{0,n,1}}}} + O(x^{3/2 + \delta_0}n^{-1/4 - \delta_0/2}). \]

Insert this approximation to (16). Together with the results from Lemma 3, we obtain that

\[ \int_0^{n^\lambda} x Q\left(\hat{\alpha}_p > \alpha_p + x/\sqrt{n}\right) \, dx = \int_0^{n^\lambda} x \Phi\left( -\frac{xf(\alpha_p)}{\sqrt{\text{Var}^Q_{W_{0,n,1}}}} \right) \, dx + o(n^{-1/4}). \]
Therefore
\[ \int_0^\infty xQ(\hat{\alpha}_p - \alpha_p > x) \, dx = \frac{1}{n} \int_0^\infty xQ(\hat{\alpha}_p > \alpha_p + x/\sqrt{n}) \, dx \]
\[ = \frac{1}{n} \int_0^\infty x\Phi \left( -\frac{xf(\alpha_p)}{\sqrt{VarQW_{0,n,1}}} \right) \, dx + o(n^{-5/4}) \]
\[ = \frac{VarQW_{0,n,1}}{nf^2(\alpha_p)} \int_0^\infty x\Phi(-x) \, dx + o(n^{-5/4}) \]
\[ = \frac{VarQW_{0,n,1}}{2nf^2(\alpha_p)} + o(n^{-5/4}). \]

Similarly,
\[ \int_0^\infty Q(\hat{\alpha}_p > \alpha_p + x) \, dx = \frac{1}{\sqrt{n}} \int_0^\infty Q(\hat{\alpha}_p > \alpha_p + x/\sqrt{n}) \, dx \]
\[ = \frac{1}{\sqrt{n}} \int_0^\infty \Phi \left( -\frac{xf(\alpha_p)}{\sqrt{VarQW_{0,n,1}}} \right) \, dx + o(n^{-3/4}). \]

For \( Q(\hat{\alpha}_p < \alpha_p - x) \) and \( x > 0 \), the approximations are completely the same and therefore are omitted. We summarize the results of \( x > 0 \) and \( x \leq 0 \) and obtain that
\[ E^Q(\hat{\alpha}_p - \alpha_p)^2 = \int_0^\infty xQ(\hat{\alpha}_p > \alpha_p + x) \, dx + \int_0^\infty xQ(\hat{\alpha}_p < \alpha_p - x) \, dx \]
\[ = n^{-1} \left[ VarQW_{0,n,1} + o(n^{-1/4}) \right] \]
\[ E^Q(\hat{\alpha}_p - \alpha_p) = \int_0^\infty Q(\hat{\alpha}_p > \alpha_p + x) \, dx - \int_0^\infty Q(\hat{\alpha}_p < \alpha_p - x) \, dx \]
\[ = o(n^{-3/4}). \]

5. Proof of Theorem 2

We first present a lemma that localizes the event. This lemma can be proven straightforwardly by standard results of empirical processes (c.f. [33, 34, 35]) along with the strong law of large numbers and the central limit theorem. Therefore, we omit it. Let \( Y_1, ..., Y_n \) be i.i.d. bootstrap samples and \( Y \) be a generic random variable equal in distribution to \( Y_i \). Let \( \hat{Q} \) be the probability measure associated with the empirical distribution \( \hat{G}(x) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x) \).
Lemma 4. Let $C_n$ be the set in which the following events occur

$E1$ $E^Q[L(Y)]^\zeta < 2E^Q[L(X)]^\zeta$ for $\zeta = 2, 3, \alpha$; $E^Q[L(Y)]^2 > \frac{1}{2}E^Q[L(X)]^2$; $E^Q[X]^\beta \leq 2E^Q[X]^\beta$.

$E2$ Suppose that $\hat{\alpha}_p = X(r)$. Then, assume that $|r/n - G(\alpha_p)| < n^{-1/2} \log n$ and $|\hat{\alpha}_p - \alpha_p| < n^{-1/2} \log n$.

$E3$ There exists $\delta \in (0, 1)$ such that for all $1 < x < \sqrt{n}$

$$\delta \leq \frac{\sum_{i=1}^{n} I(X(i) \in (\hat{\alpha}_p, \hat{\alpha}_p + xn^{-1/2})]}{nQ(\alpha_p < X \leq \alpha_p + n^{-1/2}x)} \leq \delta^{-1},$$

and

$$\delta \leq \frac{\sum_{i=1}^{n} I(X(i) \in (\hat{\alpha}_p - xn^{-1/2}, \hat{\alpha}_p)]}{nQ(\alpha_p - n^{-1/2}x < X \leq \alpha_p)} \leq \delta^{-1}.$$  

Then,

$$\lim_{n \to \infty} Q(C_n) = 1.$$  

Lemma 5. Under conditions A1 and A5, let $Y$ be a random variable following c.d.f. $\hat{G}$. Then, for each $\lambda \in (0, 1/2)$

$$\sup_{|x| \leq cn^{-1/2+\lambda}} |Var^Q[L(Y)(I(Y \leq \hat{\alpha}_p + x) - p)] - Var^Q[L(X)(I(\alpha_p + x) - p)]| = O_p(n^{-1/2+\lambda}).$$

Proof of Lemma 5. Note that

$$E^Q L^2(Y)(I(Y \leq \hat{\alpha}_p + x) - p)^2 = \frac{1}{n} \sum_{i=1}^{n} L^2(X_i)(I(X_i \leq \hat{\alpha}_p + x) - p)^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} L^2(X_i)(I(X_i \leq \alpha_p + x) - p)^2$$

$$+ L^2(\alpha_p) O_p \left( \frac{1}{n} \sum_{i=1}^{n} I(\min(\alpha_p, \hat{\alpha}_p) \leq X_i - x \leq \max(\alpha_p, \hat{\alpha}_p)) \right).$$

For the first term, by central limit theorem, continuity of $L(x)$, and Taylor’s expansion, we obtain that

$$\sup_{|x| \leq cn^{-1/2+\lambda}} \left| \frac{1}{n} \sum_{i=1}^{n} L^2(X_i)(I(X_i \leq \alpha_p + x) - p)^2 - E^Q(L^2(X)(I(X \leq \alpha_p + x) - p)^2) \right| = O_p(n^{-1/2+\lambda}).$$
Thanks to the weak convergence of empirical measure and \( \hat{\alpha}_p - \alpha_p = O(n^{-1/2}) \), we have that the second term
\[
L^2(\alpha_p) O_p \left( \frac{1}{n} \sum_{i=1}^{n} I(\min(\alpha_p, \hat{\alpha}_p) \leq X_i - x \leq \max(\alpha_p, \hat{\alpha}_p)) \right) = O_p(n^{-1/2}).
\]
Therefore,
\[
\sup_{|x| \leq cn^{-1/2+\lambda}} \left| E^\hat{Q} L^2(Y)(I(Y \leq \hat{\alpha}_p + x) - p)^2 - E^Q (L^2(X)(I(X \leq \alpha_p + x) - p)^2) \right| = O_p(n^{-1/2+\lambda}).
\]
With a very similar argument, we have that
\[
\sup_{-cn^{-1/2} \leq x \leq cn^{-1/2}} \left| E^\hat{Q} L(Y)(I(Y \leq \hat{\alpha}_p + x) - p) - E^Q (L(X)(I(X \leq \alpha_p + x) - p)) \right| = O_p(n^{-1/2+\lambda}).
\]
Thereby, we conclude the proof.

**Proof of Theorem 2.** Let \( X(1), \ldots, X(n) \) be the order statistics of \( X_1, \ldots, X_n \) in an ascending order. Since we aim at proving weak convergence, it is sufficient to consider the case that \( X \in C_n \) as in Lemma 4. Throughout the proof, we assume that \( X \in C_n \).

Similar to the notations in the proof of Theorem 1, we write \( \hat{\alpha}_p(X) \) as \( \hat{\alpha}_p \) and keep the notation \( \hat{\alpha}_p(Y) \) to differentiate them. We use Lemma 1 to compute the second moment of \( \hat{\alpha}_p(Y) - \hat{\alpha}_p \) under \( \hat{Q} \), that is,
\[
\hat{\sigma}_n^2 = \int_0^\infty x \hat{Q}(\hat{\alpha}_p(Y) > \hat{\alpha}_p + x) dx + \int_0^\infty x \hat{Q}(\hat{\alpha}_p(Y) < \hat{\alpha}_p - x) dx.
\]
We first consider the case that \( x > 0 \) and proceed to a similar derivation as that of Theorem 1. Choose \( \lambda \in \left( \frac{1}{4(\alpha-2)}, \frac{1}{8} \right) \).

**Case 1: \( 0 < x \leq n^\lambda. \)**

Similar to the proof of Theorem 1 by Berry-Esseen bound, for all \( x \in R \)
\[
\hat{Q} \left( n^{1/2}(\hat{\alpha}_p(Y) - \hat{\alpha}_p) > x \right) = \Phi \left( \frac{-\sum_{i=1}^{n} L(X_i) (I(X_i \leq \hat{\alpha}_p + xn^{-1/2} - p)}{\sqrt{n \text{Var} \hat{Q} \tilde{W}_{x,n}}} \right) + D_2,
\]
where
\[ \tilde{W}_{x,n} = L(Y) \left( I(Y \leq \hat{\alpha}_p + xn^{-1/2}) - p \right) - \frac{1}{n} \sum_{i=1}^{n} L(X_i) \left( I(X_i \leq \hat{\alpha}_p + x/\sqrt{n}) - p \right), \]

and (thanks to E1 in Lemma 4)
\[ |D_2| \leq \frac{3E^Q \left| \tilde{W}_{x,n} \right|^3}{\sqrt{n} \left( \text{Var}^Q \tilde{W}_{x,n} \right)^{3/2}} = O(n^{-1/2}). \]

In what follows, we further consider the cases that \( x > n^\lambda \). We will essentially follow the Cases 2 and 3 in the proof of Theorem 1.

Case 2: \( n^\lambda \leq x \leq c\sqrt{n} \).

Note that
\[ \sum_{i=1}^{n} L(X_i) \left( I(X_i \leq \hat{\alpha}_p) - p \right) = O(1). \]

With exactly the same argument as in Case 2 of Theorem 1 and thanks to E1 in Lemma 4, we obtain that for each \( \varepsilon > 0 \)
\[ \hat{Q} \left( \hat{\alpha}_p(Y) - \hat{\alpha}_p > xn^{-1/2} \right) \]
\[ \leq \kappa \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} L(X_i) \left( I(X_i \leq \hat{\alpha}_p + x/\sqrt{n}) - p \right) \right)^{-\alpha+\varepsilon} \]
\[ = \kappa \left( \frac{1}{\sqrt{n}} \sum_{i=1}^{n} L(X_i) I(\hat{\alpha}_p < X_i \leq \hat{\alpha}_p + x/\sqrt{n}) + O(1/\sqrt{n}) \right)^{-\alpha+\varepsilon} \]

Further, thanks to E3 in Lemma 4, we have
\[ \hat{Q} \left( \hat{\alpha}_p(Y) - \hat{\alpha}_p > xn^{-1/2} \right) = O(x^{-\alpha+\varepsilon}). \]

With \( \varepsilon \) sufficiently small, we have
\[ \int_{n^\lambda}^{\sqrt{n}} x \hat{Q} \left( \hat{\alpha}_p(Y) - \hat{\alpha}_p > xn^{-1/2} \right) dx = O(n^{-\lambda(\alpha-\varepsilon-2)}) = o(n^{-1/4}). \]

Case 3: \( x > c\sqrt{n} \).

Note that
\[ \hat{Q} \left( \sum_{i=1}^{n} L(Y_i) \left( I(Y_i \leq \hat{\alpha}_p + xn^{-1/2}) - p \right) < 0 \right), \]
is a monotone non-increasing function of $x$. Therefore, for all $x > c\sqrt{n}$, from Case 2, we obtain that

$$\hat{Q}\left(\sum_{i=1}^{n} L(Y_i) \left( I(Y_i \leq \hat{\alpha}_p + xn^{-1/2}) - p \right) < 0 \right) \leq \kappa_3 n^{-\alpha/2+\epsilon/2}.$$

For $x \leq n^{\alpha/6-\epsilon/6}$, we obtain that

$$\hat{Q}\left(\sum_{i=1}^{n} L(Y_i) \left( I(Y_i \leq \hat{\alpha}_p + xn^{-1/2}) - p \right) < 0 \right) \leq \kappa_3 n^{-\alpha/2+\epsilon/2} \leq \kappa_3 x^{-3}.$$

Thanks to condition A3, with $\epsilon$ sufficiently small, we have that $x > n^{\alpha/6-\epsilon/6}$ implies that $x^{\beta-3} > n^{1+\beta/2}$. Therefore, because of E1 in Lemma 4, for all $x > n^{\alpha/6-\epsilon/6}$ (therefore $x^{\beta-3} > n^{1+\beta/2}$)

$$\hat{Q}(\hat{\alpha}(Y) > \hat{\alpha}_p + xn^{-1/2}) \leq \hat{Q}(\sup_i Y_i > \hat{\alpha}_p + xn^{-1/2}) = O(1)n^{1+\beta/2}x^{-\beta} = O(x^{-3})$$

Therefore, we have that

$$\int_{c\sqrt{n}}^{\infty} x\hat{Q}\left(\hat{\alpha}_p(Y) - \hat{\alpha}_p > xn^{-1/2}\right) dx = O(n^{-1/2}).$$

Summary of Cases 2 and 3.

From the results of Cases 2 and 3, we obtain that for $X \in C_n$

$$\int_{n^\lambda}^{\infty} x\hat{Q}(\hat{\alpha}_p(Y) > \hat{\alpha}_p + x/\sqrt{n})dx = o(n^{-1/4}). \quad (17)$$

With exactly the same proof, we can show that

$$\int_{n^\lambda}^{\infty} x\hat{Q}(\hat{\alpha}_p(Y) < \hat{\alpha}_p - x/\sqrt{n})dx = o(n^{-1/4}). \quad (18)$$
Case 1 revisit.

Cases 2 and 3 imply that the integral in the region where \(|x| > n^\lambda\) can be ignored. In the region \(0 \leq x \leq n^\lambda\), on the set \(C_n\), for \(\lambda < 1/8\), we obtain that

\[
\begin{align*}
\int_0^{n^\lambda} x \hat{Q}(\hat{\alpha}_p(Y) > \hat{\alpha}_p + x/\sqrt{n})dx &= \int_0^{n^\lambda} x \Phi \left(-\sum_{i=1}^n L(X_i) \left(I(X_i \leq \hat{\alpha}_p + xn^{-1/2}) - p\right)\right) + D_2 dx \\
&= \int_0^{n^\lambda} x \Phi \left(-\sum_{i=1}^n L(X_i) \left(I(X_i \leq \hat{\alpha}_p + xn^{-1/2}) - p\right)\right) dx + o(n^{-1/4}).
\end{align*}
\]

We now take a closer look at the integrand. Note that

\[
\sum_{i=1}^n L(X_i) \left(I(X_i \leq \hat{\alpha}_p + xn^{-1/2}) - p\right) = \sum_{i=1}^n L(X_i) (I(X_i \leq \hat{\alpha}_p) - p) + \sum_{i=1}^n L(X_i) I(\hat{\alpha}_p < X_i \leq \hat{\alpha}_p + xn^{-1/2}).
\]

Suppose that \(\hat{\alpha}_p = X_{(r)}\). Then,

\[
\sum_{i=1}^r L(X_{(i)}) \geq p \sum_{i=1}^n L(X_i), \quad \text{ and } \quad \sum_{i=1}^{r-1} L(X_{(i)}) < p \sum_{i=1}^n L(X_i).
\]

Therefore,

\[
p \sum_{i=1}^n L(X_i) \leq \sum_{i=1}^n L(X_i) I(X_i \leq \hat{\alpha}_p) < L(\hat{\alpha}_p) + p \sum_{i=1}^n L(X_i).
\]

We plug this back to (20) and obtain that

\[
\sum_{i=1}^n L(X_i) \left(I(X_i \leq \hat{\alpha}_p + xn^{-1/2}) - p\right) = O(L(\hat{\alpha}_p)) + \sum_{i=1}^n L(X_i) I(\hat{\alpha}_p < X_i \leq \hat{\alpha}_p + xn^{-1/2}).
\]

In what follows, we study the dominating term in (19) via (22). For all
we further simplify the denominator and the display (24) equals to
\[ \alpha \]
the conditional distribution of (23) given \( \hat{\alpha} \) being, we proceed by conditioning only on \( X \).

Note that the above display is a functional of \((X_1, ..., X_n)\) and is also a stochastic process indexed by \( x \). In what follows we show that it is asymptotically a Gaussian process. The distribution of (23) is not straightforward to obtain. The strategy is to first consider a slightly different quantity and then connect it to (23). For each \((x_r), r\) such that \(|x_r - \alpha_p| \leq n^{-1/2} \log n\) and \(|r/n - G(\alpha_p)| \leq n^{-1/2} \log n\), conditional on \( X_r = x_r \), \( X_{r+1}, ..., X_n \) are equal in distribution to the order statistics of \((n-r)\) i.i.d. samples from \( Q(X \in \cdot | X > x_r) \). Thanks to the fact that \( L(x) \) is locally Lipschitz continuous and E3 in Lemma 4, we obtain
\[
\Phi \left( -\sum_{i=1}^n L(X_i)(I(X_i \leq \hat{\alpha} + x/n - 1/2) - p) \right) + O(n^{-1/2})
\]
\[
= \Phi \left( -\sum_{i=1}^n L(X_i)I(\hat{\alpha} < X_i \leq \hat{\alpha} + x/n - 1/2) \right) + O(n^{-1/2}) \tag{23}
\]

In the above inequality, we replace \( L(X_i) \) by \( L(X_r) \). The error term is
\[
O(1) \frac{L'(X_r)x/n - 1/2 \sum_{i=r+1}^n I(x_r < X_i \leq x_r + x/n - 1/2)}{\sqrt{n\text{Var}^2W_{x,n}}} = O(x^2n^{-1/2}).
\]

Note that the display (24) equals to (23) if \( \hat{\alpha} = X_r = x_r \). For the time being, we proceed by conditioning only on \( X_r = x_r \) and then further derive the conditional distribution of (23) given \( \hat{\alpha} = X_r = x_r \). Due to Lemma 5, we further simplify the denominator and the display (24) equals to
\[
\Phi \left( -\frac{L(X_r)}{\sqrt{n\text{Var}^2W_{0,n}}} \sum_{i=r+1}^n I(X_i \in (x_r, x_r + x/n - 1/2]) + O(x^2n^{-1/2 + \lambda}) \right) \tag{25}
\]
Let
\[ G_{x(r)}(x) = \frac{G(x(r) + x) - G(x(r))}{1 - G(x(r))} = Q(X \leq x(r) + x | X > x(r)). \]

Thanks to the result of strong approximation ([33, 34, 35]), given \( X(r) = x(r) \), there exists a Brownian bridge \( \{B(t) : t \in [0, 1]\} \), such that
\[
\sum_{i=r+1}^{n} I(X_{(i)} \in (x(r), x(r) + xn^{-1/2}]) = (n - r)G_{x(r)}(xn^{-1/2}) + \sqrt{n - r}B(G_{x(r)}(xn^{-1/2})) + O_p(\log(n - r)) \tag{26}
\]
where the \( O_p(\log(n - r)) \) is uniform in \( x \). Again, we can localize the event by considering a set in which the error term in the above display is \( O(\log(n - r))^2 \).

We plug this strong approximation back to (25) and obtain
\[
\Phi \left( -\frac{L(x(r))}{\sqrt{nVar^{QW}_{0,n}}} (n - r)G_{x(r)}(xn^{-1/2}) + O_p(x^2n^{-1/2+\lambda}(\log n)^2) \right) \tag{27}
\]
\[
-\varphi \left( -\frac{L(x(r))}{\sqrt{nVar^{QW}_{0,n}}} (n - r)G_{x(r)}(xn^{-1/2}) + O_p(x^2n^{-1/4}(\log n)^2) \right)
\times \frac{(n - r)^{1/2}L(x(r))}{\sqrt{nVar^{QW}_{0,n}}} B(G_{x(r)}(xn^{-1/2})).
\]

In addition, thanks to condition A4,
\[
\frac{L(x(r))}{\sqrt{nVar^{QW}_{0,n}}} (n - r)G_{x(r)}(xn^{-1/2}) = \frac{f(x(r))}{\sqrt{Var^{QW}_{0,n}}} x + O(x^{\delta_3+3/2}n^{-1/4-\delta_2/2}). \tag{28}
\]

Let
\[
\xi(x) = \frac{(n - r)^{1/2}L(x(r))}{\sqrt{nVar^{QW}_{0,n}}} B(G_{x(r)}(xn^{-1/2})) \tag{29}
\]
which is a Gaussian process with mean zero and covariance function
\[
Cov(\xi(x), \xi(y)) = \frac{(n - r)L^2(x(r))}{nVar^{QW}_{0,n}} G_{x(r)}(x/\sqrt{n})(1 - G_{x(r)}(y/\sqrt{n}))
\]
\[
= (1 + O(n^{-1/4+\lambda/2})) \frac{L(x(r))f(x(r))}{Var^{QW}_{0,n}} \frac{x}{\sqrt{n}} \tag{30}
\]
for \( 0 \leq x \leq y \leq n^{\lambda} \). Insert (28) and (29) back to (27) and obtain that given
\[ X(r) = x(r) \]

\[
\int_0^{n^\lambda} 2x \Phi \left( -\frac{\sum_{i=r+1}^n L(X(i))I(x(r) < X(i) \leq x(r) + x^{-1/2})}{\sqrt{n \text{Var}W_{x,n}}} + O(n^{-1/2}) \right) dx
\]

\[
= \int_0^{n^\lambda} 2x \Phi \left( -\frac{f(x(r))}{\sqrt{\text{Var}QW_{0,n}}} x + o(x^2 n^{-1/4}) \right) dx
\]

\[
- \int_0^{n^\lambda} 2x \varphi \left( -\frac{f(x(r))}{\sqrt{\text{Var}QW_{0,n}}} x + o(x^2 n^{-1/4}) \right) \xi(x) dx + o(n^{-1/4}),
\]

where \( \varphi(x) \) is the standard Gaussian density function. Due to Lemma 3 and \( |x(r) - \alpha_p| \leq n^{-1/2} \log n \), the first term on the right side of (31) is

\[
\int_0^{n^\lambda} 2x \Phi \left( -\frac{f(x(r))}{\sqrt{\text{Var}QW_{0,n}}} x + o(x^2 n^{-1/4}) \right) dx
\]

\[
= (1 + o(n^{-1/4})) \int_0^{\infty} 2x \Phi \left( -\frac{f(\alpha_p)}{\sqrt{\text{Var}QW_{0,n}}} x \right) dx + o_p(n^{-1/4}).
\]

The second term on the right side of (31) multiplied by \( n^{1/4} \) converges weakly to a Gaussian distribution with mean zero and variance

\[
\sqrt{n} \int_0^{n^\lambda} \int_0^{n^\lambda} 4xy \varphi \left( -\frac{f(x(r))}{\sqrt{\text{Var}QW_{0,n}}} x + o(x^2 n^{-1/4}) \right) \varphi \left( -\frac{f(x(r))}{\sqrt{\text{Var}QW_{0,n}}} y + o(x^2 n^{-1/4}) \right) \text{Cov}(\xi(x), \xi(y)) dx dy
\]

\[
= (1 + o(1)) \int_0^{n^\lambda} \int_0^{n^\lambda} 4xy \varphi \left( -\frac{f(x(r))}{\sqrt{\text{Var}QW_{0,n}}} x \right) \varphi \left( -\frac{f(x(r))}{\sqrt{\text{Var}QW_{0,n}}} y \right) \times \frac{L(x(r))f(x(r))}{\text{Var}QW_{0,n}} \min(x, y) dx dy
\]

\[
= (1 + o(1)) \frac{L(\alpha_p) (\text{Var}QW_{0,n})^{3/2}}{f^4(\alpha_p) \sqrt{\pi}}.
\]

(33)
To obtain the last step of the above display, we need the following calculation

\[
\text{Var} \left( \int_0^\infty 2x\varphi(x)B(x)dx \right) = \int_0^\infty \int_0^\infty 4xy\varphi(x)\varphi(y) \min(x, y)dx dy
\]

\[
= \int_0^\infty \int_0^\infty 4r^3 \cos \theta \sin \theta \min(\cos \theta, \sin \theta) \frac{1}{2\pi} e^{-r^2/2} dr d\theta
\]

\[
= 8 \int_0^{\pi/4} \int_0^\infty r^4 \cos \theta \sin^2 \theta \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{2\pi}} e^{-r^2/2} dr d\theta
\]

\[
= 8 \frac{1}{3} \frac{1}{\sqrt{2\pi}} \frac{1}{2} \frac{3}{2} \frac{1}{\sqrt{\pi}}
\]

\[
= \frac{1}{\sqrt{\pi}}.
\]

We insert the estimates in (32) and (33) back to (31) and obtain that conditional on \( X(r) = x(r) \),

\[
n^{1/4} \left\{ \int_0^n 2x\Phi \left( -\frac{\sum_{i=r+1}^n L(X(i))I(X(i) < x(r) \leq x(r) + xn^{-1/2})}{\sqrt{n\text{Var}^Q_{x,n}}} + O(n^{-1/2}) \right) dx 
\]

\[
- \int_0^\infty 2x\Phi \left( -\frac{f(\alpha_p)}{\sqrt{\text{Var}^Q_{0,n}}} x \right) dx \}
\]

\[
\implies N(0, \tau_p^2/2)
\]

as \( n-r, r \to \infty \) subject to the constraint that \(|r/n - G^{-1}(\alpha_p)| \leq n^{-1/2} \log n\), where \( \tau_p^2 \) is defined in the statement of the theorem. One may consider that the left-hand-side of (34) is indexed by \( r \) and \( n-r \). The limit is in the sense that both \( r \) and \( n-r \) tend to infinity in the region that \(|r/n - G^{-1}(\alpha_p)| \leq n^{-1/2} \log n\).

The limiting distribution of (34) conditional on \( \hat{\alpha}_p = X(r) = x(r) \).

We now consider the limiting distribution of the left-hand-side of (34) further conditional on \( \hat{\alpha}_p = X(r) = x(r) \). To simplify the notation, let

\[
V_n = -n^{1/4} \int_0^n 2x\varphi \left( -\frac{f(\alpha_p)}{\sqrt{n\text{Var}^Q_{x,n}}} x + o(x^2n^{-1/4}) \right) \]

\[
\times L(\alpha_p) \sum_{i=r+1}^n I(x(r) < X(i) \leq x(r) + xn^{-1/2}) \frac{1}{\sqrt{n\text{Var}^Q_{x,n}}} dx.
\]
Then,
\[
\int_0^{n^2} \Phi \left( -\frac{\sum_{i=r+1}^n L(X(i)) I(x(r) < X(i) \leq x(r) + x n^{-1/2})}{\sqrt{n \text{Var} Q_{x,n}}} + O(n^{-1/2}) \right) dx
\]
\[
- \int_0^\infty 2x \Phi \left( -\frac{f(\alpha_p)}{\sqrt{\text{Var} Q_{0,n}}} x \right) dx
\]
\[
= n^{-1/4} V_n + o(n^{-1/4}).
\]

The weak convergence result in (34) says that for each compact set \(A\),
\[
Q(V_n \in A|X(r) = x(r)) \rightarrow P(Z \in A),
\]
as \(n - r, r \to \infty\) subject to the constraint that \(|r/n - G^{-1}(\alpha_p)| \leq n^{-1/2} \log n\),
where \(Z\) is a Gaussian random variable with mean zero and variance \(\tau^2_p/2\). Note that \(\hat{\alpha}_p = X(r) = x(r)\) is equivalent to
\[
0 \leq H = \sum_{i=1}^r L(X(i))(1 - p) - p \sum_{i=r+1}^n L(X(i)) \leq L(x(r)).
\]

Let
\[
U_n = \sum_{i=r+1}^n L(X(i)) I(x(r) < X(i) \leq x(r) + n\lambda^{-1/2}) - n \times P(x(r) < X \leq x(r) + n\lambda^{-1/2})
\]
and
\[
B_n = \left\{ \left| U_n \right| \leq n^{\lambda/2 + 1/4} \log n \right\}.
\]

Note that, given the partial sum \(U_n\), \(H\) is independent of the \(X_i\)'s in the interval \((x(r), x(r) + n\lambda^{-1/2})\) and therefore is independent of \(V_n\). For each compact set \(A\) and \(A_n = \{V_n \in A\} \cap B_n\), we have
\[
Q(A_n|\hat{\alpha}_p = X(r) = x(r)) \quad (35)
\]
\[
= Q(0 \leq H \leq L(x(r))|X(r) = x(r), A_n) Q(A_n|X(r) = x)
\]
\[
= P^{Q} \left[ Q(0 \leq H \leq L(x(r))|X(r) = x(r), U_n) \left| X(r) = x(r), A_n \right] Q(A_n|X(r) = x(r)) \right]
\]
The second step of the above equation uses the fact that on the set $B_n$

$$Q \left( 0 \leq H \leq L(x_{(r)})|X_{(r)} = x_{(r)}, U_n \right) = Q \left( 0 \leq H \leq L(x_{(r)})|X_{(r)} = x_{(r)}, U_n, A_n \right).$$

Note that $U_n$ only depends on the $X_i$'s in $(x_{(r)}, x_{(r)} + n^{\lambda - 1/2})$, while $H$ is the weighted sum of all the samples. Therefore, on the set $B_n = \{|U_n| \leq n^{\lambda/2+1/4} \log n\}$

$$\frac{Q \left( 0 \leq H \leq L(x_{(r)})|X_{(r)} = x_{(r)}, U_n \right)}{Q \left( 0 \leq H \leq L(x_{(r)})|X_{(r)} = x_{(r)} \right)} = 1 + o(1),$$

and the $o(1)$ is uniform in $B_n$. The rigorous proof of the above approximation can be developed using the Edgeworth expansion of density functions straightforwardly, but is tedious. Therefore, we omit it. We plug (36) back to (35).

Note that $Q(B_n|X_{(r)} = x_{(r)}) \to 1$ and we obtain that for each $A$

$$Q \left( V_n \in A|\hat{\alpha}_p = X_{(r)} = x_{(r)} \right) - Q(V_n \in A|X_{(r)} = x_{(r)}) \to 0.$$

Therefore, we obtain that conditional on $\hat{\alpha}_p = X_{(r)}$, $|\hat{\alpha}_p - \alpha_p| \leq n^{-1/2} \log n$, and $|r/n - G^{-1}(\alpha_p)| \leq n^{-1/2} \log n$, as $n \to \infty$

$$n^{1/4} \left[ \int_0^{n^{\lambda}} \hat{Q} (\hat{\alpha}_p(Y) > \hat{\alpha}_p + x/\sqrt{n}) \, dx - \int_0^{\infty} 2x \Phi \left( - \frac{f(\alpha_p)}{\sqrt{\text{Var}^q W_{0,n}}} x \right) \, dx \right]$$

$$= n^{1/4} \left\{ \int_0^{n^{\lambda}} \Phi \left( - \sum_{i=r+1}^{n} L(X_{(i)}) I(\hat{\alpha}_p < X_{(i)} \leq \hat{\alpha}_p + x n^{-1/2}) \frac{\sqrt{n} \text{Var}^q \tilde{W}_{x,n}}{\text{Var}^q \tilde{W}_{0,n}} \right) + O(n^{-1/2}) \right\} \, dx$$

$$- \int_0^{\infty} 2x \Phi \left( - \frac{f(\alpha_p)}{\sqrt{\text{Var}^q W_{0,n}}} x \right) \, dx \right) + o_p(1)$$

$$= V_n + o_p(1) \implies N(0, \tau_p^2/2).$$

Together with E2 in Lemma 4, this convergence indicates that asymptotically the bootstrap variance estimator is independent of $\hat{\alpha}_p$. Therefore, the unconditional asymptotic distribution is

$$n^{1/4} \left[ \int_0^{n^{\lambda}} \hat{Q} (\hat{\alpha}_p(Y) > \hat{\alpha}_p + x/\sqrt{n}) \, dx - \int_0^{\infty} 2x \Phi \left( - \frac{f(\alpha_p)}{\sqrt{\text{Var}^q W_{0,n}}} x \right) \, dx \right] \implies N(0, \tau_p^2/2).$$

(37)
With exactly the same argument, we have the asymptotic distribution of the negative part of the integral
\[
\frac{n^{1/4}}{\sqrt{\text{Var} Q(W_0, n)}} \left( \int_0^\infty Q(\hat{\alpha}_p(Y) < \hat{\alpha}_p - x/\sqrt{n}) \, dx - \int_0^\infty 2x \Phi \left( -\frac{f(\alpha_p)}{\sqrt{\text{Var} Q(W_0, n)}} x \right) \, dx \right) \rightsquigarrow N(0, \tau_p^2/2).
\]
(38)

Using a conditional independence argument, we obtain that the negative part and the positive part of the integral are asymptotically independent. Putting together the results in Theorem 1, (17), (18), (37), (38), and the moment calculations of Gaussian distributions, we conclude that
\[
\hat{\sigma}^2_n = \int_0^\infty 2x \left[ Q(\hat{\alpha}_p(Y) < \hat{\alpha}_p - x) + Q(\hat{\alpha}_p(Y) > \hat{\alpha}_p + x) \right] \, dx
\]
\[
= \frac{1}{n} \int_0^\infty 2x \left[ \hat{Q}(\hat{\alpha}_p(Y) < \hat{\alpha}_p - x/\sqrt{n}) + \hat{Q}(\hat{\alpha}_p(Y) > \hat{\alpha}_p + x/\sqrt{n}) \right] \, dx
\]
\[
\leq \frac{\text{Var} Q(W_p)}{nf(\alpha_p)^2} + Zn^{-5/4} + o(n^{-5/4})
\]
\[
= \sigma^2_n + Zn^{-5/4} + o(n^{-5/4}).
\]
where \( Z \sim N(0, \tau_p^2) \).

6. Proof of Theorem 3

Lemma 6. Under the conditions of Theorem 3, we have that
\[
E^Q(LR^\gamma(X); X < \alpha_p) = c(\theta) + o(1) \gamma e^{-m\gamma l}.
\]
We clarify that \( LR^\gamma(X) = (LR(X))^\gamma \) is the \( \gamma \)-th moment of the likelihood ratio. With this result, if we choose \( m^3 = O(n) \), it is sufficient to guarantee that with probability tending to one the ratio between empirical moments and the theoretical moments are within \( \varepsilon \) distance from one. Thus, the localization results (Lemma 4) are in place.

Lemma 7. Under the conditions of Theorem 3, for each \( \gamma > 0 \) there exist constants \( \delta \) (sufficiently small), \( u_\gamma \), and \( l_\gamma \), such that
\[
l_\gamma \delta \leq \frac{E^Q(LR^\gamma(x); \alpha_p < X \leq \alpha_p + \delta)}{E^Q(LR^\gamma(x); X < \alpha_p)} \leq u_\gamma \delta
\]
for all \( m \) sufficiently large.

The proof of the above two Lemmas are standard via exponential change of measure and Edgeworth expansion. Therefore, we omit them.

**Proof of Theorem 3.** The proof of this theorem is very similar to that of Theorems 1 and 2. The only difference is that we need to keep in mind that there is another parameter \( m \) that tends to infinity. Therefore, the main task of this proof is to provide a careful analysis and establish a sufficiently large \( n \) so that similar asymptotic results hold as \( m \) tends to infinity in a slower manner than \( n \).

From a technical point of view, the main reason why we need to choose \( m^3 = O(n) \) is that we use Barry-Esseen bound in the region \([0, n^\lambda]\) to approximate the distribution of \( \sqrt{n}(\hat{\alpha}_p - \alpha_p) \). In order to have the approximation hold (Case 2 of Part 1), it is necessary to have \( m^{1/4} = o(n^\lambda) \). On the other hand, the error term of the Barry-Esseen bound requires that \( n^\lambda \) cannot to too large. The order \( m^3 = O(n) \) is sufficient to guarantee both.

The proof consists of two main parts. In part 1, we establish similar results as those in Theorem 1; in part 2, we establish the corresponding results given in Theorem 2.

**Part 1.**

We now proceed to establish the asymptotic mean and variance of the weighted quantile estimator. Recall that in the proof of Theorem 1 we developed the approximations of the tail probabilities of the quantile estimator and use Lemma 1 to conclude the proof. In particular, we approximate the right tail of \( \hat{\alpha}_p \) in three different regions. We go through the three cases carefully for some \( \max(\frac{1}{4(\alpha-2)}, \frac{1}{10}) < \lambda < \frac{1}{8} \) (recall that \( L(X) \) has at least \( \alpha \)-th moment under \( Q \)).
Case 1: $0 < x \leq n^\lambda$. We approximate the tail probability using Berry-Esseen bound

$$Q(\sqrt{n}(\hat{\alpha}_p - \alpha_p) > x) = \Phi\left(\frac{-(F(\alpha_p + xn^{-1/2}) - p)}{\sqrt{\frac{1}{n} \text{Var} Q_{W_{x,n}}}}\right) + D_1(x)$$  \hspace{1cm} (39)

where

$$W_{x,n,i} = LR(X_i)I(X_i \leq \alpha_p + xn^{-1/2}) - F(\alpha_p + xn^{-1/2})$$

and

$$|D_1(x)| \leq \frac{cE|W_{x,n,1}|^3}{(\text{Var} Q_{W_{x,n,1}})^{3/2} n^{-1/2}} = O(1 \frac{m^{1/4}}{n^{1/2}}).$$

The last step approximation of $D_1(x)$ is from Lemma 6. In the current case, $W_{x,n,i}$ has all the moments, that is, $\alpha$ can be chosen arbitrarily large.

Case 2: $n^\lambda < x \leq c\sqrt{n}$. Applying Theorem 2.18 in [16], for each $\delta > 0$ there exists $\kappa_1(\delta)$ and $\kappa_2(\delta)$ so that

$$Q\left(\sqrt{n}(\hat{\alpha}_p - \alpha_p) > x\right) = Q\left(\sum_{i=1}^{n} W_{x,n,i} < n(p - F(\alpha_p + xn^{-1/2}))\right)$$

$$\leq \left(\frac{3}{1 + \delta n \text{Var}(W_{x,n,1})^{-1}(p - F(\alpha_p + xn^{-1/2}))^2}\right)^{1/\delta}$$

$$+ Q\left(\min_{i=1}^{n} W_{x,n,1} < \delta n(p - F(\alpha_p + xn^{-1/2}))\right)$$

$$\leq \kappa_1(\delta)(m^{-1/4}x)^{-2/\delta} + \frac{\kappa_2(\delta)n^{-\alpha/2+1}}{\sqrt{m}}(m^{-1/2}x)^{-\alpha}$$

The last inequality uses Lemma 7 with $\gamma = 1$. Given that we can choose $\delta$ arbitrarily small and $\alpha$ arbitrarily large (thanks to the steepness) and $m^3 = O(n)$, we obtain that

$$Q(\sqrt{n}(\hat{\alpha}_p - \alpha_p) > x) = O(x^{-\alpha+\epsilon}),$$

where $\epsilon$ can be chosen arbitrarily small. Thus,

$$\int_{n^\lambda}^{c\sqrt{n}} xQ(\sqrt{n}(\hat{\alpha}_p - \alpha_p) > x)\,dx = o(n^{-1/4})$$
Case 3: $x > c \sqrt{n}$. Similar to the proof in Theorem 1, for $c \sqrt{n} < x \leq n^{\alpha/6-\varepsilon/6}$
\[ Q(\sqrt{n}(\hat{\alpha}_p - \alpha_p) > x) \leq \kappa_3 x^{-3}. \]

For $x^{\beta-3} > n^{1+2\beta/3}$ (recall that $X$ has at least $\beta$-th moment under $Q$),
\[ Q(\sqrt{n}(\hat{\alpha}_p - \alpha_p) > x) \leq Q(\sup_i X_i > \alpha_p + xn^{-1/2}) \leq O(1)n^{1+\beta/2}m^{\beta/2}x^{-\beta} = O(x^{-3}). \]

Given the steepness of $\varphi(\theta)$, $\alpha$ and $\beta$ can be chosen arbitrarily large. We then have $1 = \frac{\alpha}{6} \geq \frac{1+2\beta/3}{\beta-3}$. Thus, we conclude that
\[ Q(\sqrt{n}(\hat{\alpha}_p - \alpha_p) > x) = O(x^{-3}) \]
for all $x > c \sqrt{n}$ and
\[ \int_{c \sqrt{n}}^{\infty} xQ(\sqrt{n}(\hat{\alpha}_p - \alpha_p) > x)dx = O(n^{-1/2}). \]

Summarizing the three cases, we have that
\[
\int_{0}^{\infty} xQ(\sqrt{n}(\hat{\alpha}_p - \alpha_p) > x)dx = \int_{0}^{n^4} x\Phi \left( \frac{-(F(\alpha_p + xn^{-1/2}) - p)}{\sqrt{nVar^QW_{x,n,1}}} \right) dx + o(m^{1/4}n^{-1/4})
\]
\[ = \int_{0}^{n^4} x\Phi \left( -\frac{f(\alpha_p)x}{\sqrt{Var^QW_{0,n,1}}} \right) dx + o(m^{1/4}n^{-1/4}). \]

For the last step, we use the fact that $Var^QW_{x,n,1} = (1+O(xn^{-1/2}))Var^QW_{0,n,1}$, which is an application of Lemma 7. Using Lemma 1, we obtain that
\[ E^Q(\hat{\alpha}_p) = \alpha_p + o(m^{1/4}n^{-3/4}), \quad \sigma_n^2 = Var^Q(\hat{\alpha}_p) = \frac{Var^Q(W_p)}{nf(\alpha_p)^2} + o(m^{1/4}n^{-5/4}), \]
where
\[ W_p = LR(X)I(X \leq \alpha_p) \]
and the convergence is uniform in $m$ when $m^3 = O(n)$. Notice that we aim at showing that the bootstrap variance estimator converges to $\sigma_n^2$ at a rate of $O(m^{5/8}n^{-5/4})$. Then, we can basically treat the above approximations as true values in the derivations for the bootstrap variance estimator.
Part 2.

We now proceed to the discussion of the distribution of the bootstrap variance estimator. The proof needs to go through three similar cases as in Part 1 of this proof where we derive the approximation of the mean and variance of \( \hat{\alpha}_p \) under \( Q \). The difference is that we need to handle the empirical measure \( \hat{Q} \) instead of \( Q \). As we explained after the statement of Lemma 6, the localization conditions (Lemma 4) are satisfied when \( m^3 = O(n) \). Let the set \( C_n \) be as defined in Lemma 4. The analyses of these three cases are identical to those in part 1. We obtain similar results as in Part 1 that

\[
\int_{n^3}^{\infty} x \hat{Q}(\hat{\alpha}_p(Y) > \hat{\alpha}_p + x/\sqrt{n}) \, dx = o(n^{-1/4}).
\]  

(40)

We omit the detailed analysis. In what follows, we focus on the “revisit of case 1” in the proof of Theorem 2, which is the leading term of the asymptotic result. Then, we continue the derivation from (19) in the proof of Theorem 2. On the set \( C_n \) and with a similar result (for the empirical measure \( \hat{Q} \)) as in (39), we have that

\[
\int_{0}^{n^3} x \hat{Q}(\hat{\alpha}_p(Y) > \hat{\alpha}_p + x/\sqrt{n}) \, dx
\]

(41)

\[
= \int_{0}^{n^3} x \Phi \left( -\frac{\sum_{i=1}^{n} LR(X_i) I(X_i \leq \hat{\alpha}_p + x/\sqrt{n}) - np}{\sqrt{n \text{Var}^Q W_{x,n}}} \right) \, dx + o(m^{1/4}n^{-1/4}).
\]

We take a closer look at the above Gaussian probability. Using the same argument as in (20) and (21), we obtain that

\[
\Phi \left( -\frac{\sum_{i=1}^{n} LR(X_i) I(X_i \leq \hat{\alpha}_p + x/\sqrt{n}) - np}{\sqrt{n \text{Var}^Q W_{x,n}}} \right)
\]

\[
= \Phi \left( -\frac{\sum_{i=1}^{n} LR(X_i) I(\hat{\alpha}_p < X_i \leq \hat{\alpha}_p + x/\sqrt{n})}{\sqrt{n \text{Var}^Q W_{x,n}}} + O(m^{1/4}n^{-1/2}) \right)
\]

We replace \( LR(X_i) \) by \( LR(\hat{\alpha}_p) \) and obtain

\[
= \Phi \left( -LR(\hat{\alpha}_p) \frac{\sum_{i=1}^{n} I(\hat{\alpha}_p < X_i \leq \hat{\alpha}_p + x/\sqrt{n})}{\sqrt{n \text{Var}^Q W_{x,n}}} + O(x^2m^{-1/4}n^{-1/2} + m^{1/4}n^{-1/2}) \right)
\]
Lastly, we replace the empirical variance in the denominator by the theoretical variance. Similar to the development in Lemma 5, we can obtain the following estimates

\[ |\text{Var}^{Q} W_{0,n} - \text{Var}^{\hat{Q}} W_{x,n}| = (1 + O_{p}(m^{1/4}n^{-1/2} + xn^{-1/2}))\text{Var}^{Q} W_{0,n}, \]

for \(|x| < n^{\lambda}\). Since that we are deriving weak convergence results, we can always localize the events so that \(O_{p}(\cdot)\) can be replace by \(O(\cdot)\). Then, the Gaussian probability is approximated by

\[
\Phi \left( \frac{-LR(\hat{\alpha}_p) \sum_{i=1}^{n} I(\hat{\alpha}_p < X_i \leq \hat{\alpha}_p + x/\sqrt{n})}{\sqrt{n\text{Var}^{Q} W_{0,n}}} + \zeta(x, m, n) \right),
\]

where \(\zeta(x, m, n) = O(x^{2}m^{-1/4}n^{-1/2} + xn^{-1/2} + m^{1/4}n^{-1/2})\). Using the strong approximation of empirical processes in (26) and the Lipschitz continuity of the density function, one can further approximate the above probability by

\[
\Phi \left( \frac{f(\hat{\alpha}_p)}{\sqrt{\text{Var}^{Q} W_{0,n}}} x + \zeta(x, m, n) \right) - \varphi \left( \frac{f(\hat{\alpha}_p)}{\sqrt{\text{Var}^{Q} W_{0,n}}} x + \zeta(x, m, n) \right) \frac{(n - r)^{1/2}L(\hat{\alpha}_p)}{\sqrt{n\text{Var}^{Q} W_{p}}} B(G_{\hat{\alpha}_p}(x/\sqrt{n})) \]

where \(B(t)\) is a standard Brownian bridge and

\[ G_{y}(x) = Q(X \leq y + x|X > y). \]

Note that

\[
\int_{0}^{n^{\lambda}} 2x\Phi \left( \frac{f(\hat{\alpha}_p)}{\sqrt{\text{Var}^{Q} W_{0,n}}} x + \zeta(x, m, n) \right) dx = \int_{0}^{n^{\lambda}} 2x\Phi \left( \frac{f(\hat{\alpha}_p)}{\sqrt{\text{Var}^{Q} W_{0,n}}} x \right) dx + O(m^{3/4}n^{-1/2}).
\]

We write

\[ Z_n = -n^{1/4} \int_{0}^{n^{\lambda}} 2x\varphi \left( \frac{f(\hat{\alpha}_p)}{\sqrt{\text{Var}^{Q} W_{0,n}}} x + \zeta(x, m, n) \right) \frac{(n - r)^{1/2}L(\hat{\alpha}_p)}{\sqrt{n\text{Var}^{Q} W_{p}}} B(G_{\hat{\alpha}_p}(x/\sqrt{n})) dx. \]

The calculation of the asymptotic distribution of \(Z_n\) is completely analogous to (33) in the proof of Theorem 2. Therefore, we omit it. Putting all these results together, the integral in (41) can be written as

\[
\int_{0}^{n^{\lambda}} 2x\bar{Q} (\hat{\alpha}_p(Y) > \hat{\alpha}_p + x/\sqrt{n}) dx = \int_{0}^{\infty} 2x\Phi \left( \frac{f(\hat{\alpha}_p)}{\sqrt{\text{Var}^{Q} W_{0,n}}} x \right) dx + Z_n \frac{n^{1/4}}{n} + o(m^{5/8}n^{-1/4}),
\]
where

\[ \frac{Z_n}{\sqrt{\tau_p/2}} \Rightarrow N(0, 1), \quad \tau_p^2 = 2\pi^{-1/2}L(\alpha_p)f(\alpha_p)^{-4}(\text{Var}^Q(W_p))^{3/2} = O(m^{5/4}) \]

as \( n \to \infty \) uniformly when \( m^3 = O(n) \). The derivation for \( \int_0^{\hat{\alpha}_p} x\hat{Q}(\hat{\alpha}_p(Y) < \hat{\alpha}_p - x/\sqrt{n}) \, dx \)

is analogous to the positive part. Together with (40), we conclude the proof.

References


