On a new method for constructing bootstrap confidence bounds

C Pretorius
20104480

Thesis submitted for the degree *Philosophiae Doctor* in Statistics at the Potchefstroom Campus of the North-West University

Promoter: Prof JWH Swanepoel

May 2017

NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT

It all starts here ™

Acknowledgements

I wish to express special appreciation and gratitude to the following individuals:

- Prof. Jan Swanepoel for his guidance and support throughout this study. I will be forever grateful for the opportunity to have shared in his knowledge of and passion for mathematics and statistics.
- Other colleagues in the Department, especially Prof. James Allison and Dr. Jaco Visagie for insightful discussions and encouragement.
- The National Research Foundation of South Africa and the South African Statistical Association for financial support.
- My parents, Leon and Janettie Pretorius, for providing continued support and the privilege of an excellent education.

Abstract

It is well-known that the standard methods for constructing bootstrap confidence bounds or intervals are in many situations not sufficiently accurate, that is, coverage probabilities converge to the nominal level at unsatisfactory rates. We propose a new method, based on sample splitting, for constructing higher-order accurate bootstrap confidence bounds for a parameter appearing in the regular *smooth function model* introduced by Bhattacharya and Ghosh (1978).

It has been demonstrated by Hall (1986, 1988, 1992) that the well-known *percentile-t* bootstrap confidence bound typically incurs a coverage error of order $O(n^{-1})$, with n being the sample size. Our version of the percentile-t bound reduces this coverage error to order $O(n^{-3/2})$ and in some cases to $O(n^{-2})$. Furthermore, whereas the standard *percentile* bounds typically incur coverage error of $O(n^{-1/2})$, the new percentile bounds have reduced error of $O(n^{-1})$. We show that equal-tailed confidence intervals with coverage error at most $O(n^{-2})$ may be obtained from the newly proposed bounds, as opposed to the typical error $O(n^{-1})$ of the standard intervals.

In the case where the parameter of interest is the population mean we derive, for each confidence bound, the exact coefficient of the leading term in an asymptotic expansion of the coverage error, although similar results may be obtained for other parameters such as the variance, the correlation coefficient, and the ratio of two means. We also derive similar results for the case where the slope parameter in a linear regression model is of interest, showing that the good properties of the new percentile-t method carry over to regression problems.

Results of independent interest are derived, such as a generalisation of a delta method by Cramér (1946) and Hurt (1976), as well as an expression for an Edgeworth polynomial arising in the linear regression setup.

The study is concluded with a modest simulation study, which illustrates the behaviour of the new confidence bounds for small to moderate sample sizes.

Keywords: Bootstrap; confidence bounds; sample splitting; coverage error; smooth function model; Edgeworth polynomials; Cornish-Fisher; delta method; regression.

Uittreksel

Ofskoon daar in die literatuur talle skoenlusmetodes bestaan wat gebruik kan word om vertrouensgrense te konstrueer, is die akkuraatheid van die standaard metodes in sommige gevalle ontoereikend. Dit wil sê, die oordekkingswaarskynlikhede van hierdie grense konvergeer na die voorgeskrewe oordekkingswaarskynlikheid teen onbevredigende tempo's. In hierdie proefskrif stel ons 'n nuwe metode voor wat aangewend kan word om meer akkurate vertrouensgrense te verkry vir 'n parameter wat in die *gladdefunksie-model* van Bhattacharya en Ghosh (1978) voorkom. Die nuwe metode berus deels daarop dat die steekproef in twee onderling uitsluitende versamelings verdeel word.

Hall (1986, 1988, 1992) toon aan dat die welbekende *persentiel-t*-vertrouensgrens tipies tot 'n oordekkingsfout van orde $O(n^{-1})$ lei, waar n die steekproefgrootte aandui. Ons weergawe van die persentiel-t-grens verminder hierdie oordekkingsfout na $O(n^{-3/2})$ en in sommige gevalle na $O(n^{-2})$. Verder, terwyl die standaard *persentiel*-grense tipies lei tot 'n oordekkingsfout van $O(n^{-1/2})$, lewer die nuwe persentiel-grense 'n verminderde fout van $O(n^{-1})$. Ons toon aan dat gelykkantige vertrouensintervalle met 'n oordekkingsfout van hoogstens $O(n^{-2})$ uit die nuwe vertrouensgrense verkry kan word, in teenstelling met die tipiese fout van $O(n^{-1})$ van die standaard intervalle.

In die geval waar die populasiegemiddelde van belang is, verkry ons vir elke nuwe vertrouensgrens die presiese koëffisiënt van die leidende term in 'n asimptotiese ontwikkeling van die oordekkingsfout. Aangesien ons resultate in die algemeen onder die gladdefunksiemodel geld, kan soortgelyke resultate vir ander parameters soos die variansie, die korrelasiekoëffisiënt en die kwosiënt van twee gemiddeldes afgelei word. Ons lei ook soortgelyke uitdrukkings af vir die geval waar die helling in 'n lineêre regressiemodel beraam moet word en toon sodoende aan dat die goeie eienskappe van die nuwe persentiel-t-grens oordra na regressieprobleme.

Daarbenewens vekry ons nuwe resultate van algemene belang in die statistiese literatuur, soos 'n veralgemening van 'n delta-metode van Cramér (1946) en Hurt (1976), asook 'n uitdrukking vir 'n Edgeworth-polinoom wat in 'n regressie-opset voorkom.

Die proefskrif sluit af met 'n simulasiestudie wat die gedrag van die nuwe vertrouensgrense vir klein tot matige steekproewe illustreer.

Sleutelwoorde: Skoenlusmetode; vertrouensgrense; steekproefopdeling; oordekkingsfout; gladdefunksie-model; Edgeworth-polinome; Cornish-Fisher; delta-metode; regressie.

Contents

Chapter 1

Introduction

1.1 Overview

Over the past 37 years the bootstrap has been established as a powerful tool in many problems of statistical inference, including error estimation, bias correction, density estimation, hypothesis testing, prediction, and model selection, to name a few. However, the construction of confidence intervals in particular, as well as their performance in different situations, has enjoyed much attention in the statistical literature. Carpenter and Bithell (2000) even remarked that "since the early 1980s, a bewildering array of methods for constructing bootstrap confidence intervals have been proposed."

Despite this it has been shown that, in some situations, the existing methods yield confidence bounds (or intervals) which are not sufficiently accurate, that is, coverage probabilities converge to the nominal level at unsatisfactory rates. In particular, Shao and Tu (1995, p. 155) make the following remark: "In view of the existing empirical results showing that second-order accurate confidence sets are sometimes still not accurate enough, one might ask whether we can obtain a confidence set with an even higher order of accuracy. The answer is affirmative *if one has the resources for more intensive computations*." Shao and Tu refer specifically to techniques such as bootstrap iteration and bootstrap calibration, which will be discussed briefly in Chapter 2. In this study we show, however, that it is possible to obtain higher-order accurate confidence bounds (and intervals) without having to deal with the computational demands of the above-mentioned methods. This is achieved by carefully splitting the sample into two disjoint parts, using the one part to calculate the point estimator for the parameter of interest and using the other part to approximate the appropriate quantile of the distribution of the standardised/Studentised point estimator. The latter approximation employs the m/n bootstrap, where the choice of the resample size m is crucial to achieve the improved levels of accuracy.

To assess the asymptotic coverage probability of each new confidence bound, we make use of the methodology established by Hall (1988, 1992), which involves finding an asymptotic series expansion in $n^{-1/2}$ (where n is the sample size) of the coverage probability by means of Edgeworth and Cornish-Fisher expansions. These expansions are then used to establish

(i) whether the coverage probability converges to the nominal level (i.e., whether the bound is consistent) and (ii) the rate of this convergence. The results are then used to compare the asymptotic behaviour of the new bounds to that of the existing bounds, the latter results being readily available in the literature.

Specific applications of the new confidence bounds will be of interest, particularly the two cases where the parameter of interest is the population mean and the slope parameter in a simple linear regression model. For these examples asymptotic results have been obtained by Hall (1992), among others. In this study we derive similar results for the new confidence bounds, which require results of independent interest. In particular, we obtain a generalised form of the delta method and derive an expression for a polynomial appearing in an Edgeworth expansion of the distribution of the Studentised estimator for slope in a simple linear regression model.

The study is concluded with a Monte Carlo study to compare the finite sample behaviour of the new confidence bounds to that of existing bootstrap confidence bounds.

1.2 Objectives

The main objectives of this study can be summarised as follows:

- Provide a comprehensive overview of the current literature on existing methods for constructing bootstrap confidence bounds and intervals.
- Review the smooth function model introduced by Bhattacharya and Ghosh (1978) and its connection with Edgeworth and Cornish-Fisher expansions of asymptotically pivotal statistics.
- Derive certain Edgeworth polynomials not available in the existing literature, which are required for some of the asymptotic theory in this study.
- Propose two new percentile-type methods and two new percentile-t-type methods for constructing bootstrap confidence bounds.
- Assess the consistency and accuracy of the newly proposed confidence bounds by deriving the asymptotic converge probability of each new bound by means of Edgeworth and Cornish-Fisher expansions.
- Apply these asymptotic results to the case where the parameter of interest is the population mean.
- Demonstrate that the good properties of the new bounds carry over to simple linear regression problems where the slope parameter is of interest.
- Assess the finite-sample performance of the new confidence bounds by means of a Monte Carlo study.

1.3 Outline

A review of traditional methods for constructing bootstrap confidence bounds and intervals is given in Chapter 2, along with some more refined methods.

Chapter 3 is devoted to describe the regular *smooth function model*, which forms the basis of our asymptotic assessment of the new confidence bounds. In this chapter we also show how expressions for Edgeworth and Cornish-Fisher expansions can be obtained and derive a new result that does not appear in the existing literature.

Two new percentile-type confidence bounds are proposed in Chapter 4 and two new percentile-t-type confidence bounds in Chapter 5. For each new bound, an expression for the asymptotic coverage probability is derived. A concise summary of all these results is given in Chapter 6.

In Chapter 7, for illustrative purposes, we apply the new methods to the case where the parameter of interest is the population mean and, for each new method, derive the exact coefficient of the leading term in an asymptotic expansion of coverage probability. This chapter also contains a useful and interesting generalisation of the well known delta method.

Chapter 8 demonstrates how the good properties of the new percentile-t bound carry over to regression problems, where the slope parameter is of interest. Similar asymptotic results are derived in this chapter.

The main part of the thesis is concluded in Chapter 9 with a limited Monte Carlo study, which illustrates the behaviour of the new confidence bounds for small to moderate sample sizes.

Supporting theoretical results required in the main text are stated and derived in the appendices appearing after Chapter 9.

1.4 Frequently used notation

For easy reference we provide the following list of notation occurring throughout the text.

- $\mathbb{R}^d$ is the set of all d-dimensional real numbers; $\mathbb{R} := \mathbb{R}^1$.
- $\mathbb{N} = \{1,2,\ldots\}$, the set of natural numbers, and $\mathbb{N}_0 = \{0,1,2,\ldots\}$, the set of nonnegative integers.
- $a \wedge b := \min(a,b)$ and $a \vee b := \max(a,b)$ for any $a,b \in \mathbb{R}$.
- x_i, $i = 1,2,\ldots,d$, denotes the ith element of $\mathbf{x} \in \mathbb{R}^d$, i.e.,

$$\mathbf{x} = (x_1, x_2, \ldots, x_d).$$

  We say that $\mathbf{x}$ is a *d-vector*.

- $||\cdot||$ denotes the Euclidean norm of a vector. That is, if $\mathbf{x} \in \mathbb{R}^d$, then

$$||\mathbf{x}|| = \sqrt{x_1^2 + x_2^2 + \cdots + x_d^2}.$$

- For any $x \in \mathbb{R}$, $\lfloor x \rfloor$ denotes the largest integer less than or equal to x and $\lceil x \rceil$ denotes the smallest integer greater than or equal to x.
- $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ and $\Phi(x) = \int_{-\infty}^{x} \phi(u)du$ for $x \in \mathbb{R}$.
- $\xrightarrow{d}$ and $\xrightarrow{p}$ denote convergence *in distribution* and *in probability*, respectively.

To avoid ambiguity, we explicitly define the asymptotic notation $O(\cdot)$, $o(\cdot)$, $O_p(\cdot)$ and $o_p(\cdot)$, which will be used throughout to describe error terms in asymptotic expressions. For a very informative review of related asymptotic notation, the reader is referred to Janson (2009), from where the following definitions have been adapted. In these definitions, $\{a_n\}$ and $\{b_n\}$ are two non-random sequences of real numbers, with $a_n > 0$ for all n. Throughout the text the limit is $n \to \infty$ unless stated otherwise.

Definition 1.1. We say that $b_n = O(a_n)$ as $n \to \infty$ if there exist constants C and n_0 such that $a_n^{-1}|b_n| \leq C$ for all $n \geq n_0$, that is,

$$b_n = O(a_n) \iff \limsup_{n\to\infty} \frac{|b_n|}{a_n} < \infty.$$

Definition 1.2. We say that $b_n = o(a_n)$ as $n \to \infty$ if for every $\varepsilon > 0$ there exists and integer n_ε such that $a_n^{-1}|b_n| \leq \varepsilon$ for all $n \geq n_\varepsilon$, that is,

$$b_n = o(a_n) \iff \lim_{n\to\infty} \frac{b_n}{a_n} = 0.$$

Definition 1.3. Let $\{X_n\}$ be any sequence of random variables. We say that $X_n = O_p(a_n)$ as $n \to \infty$ if for every $\varepsilon > 0$ there exist a constant C_ε and an integer n_ε such that $\mathrm{P}\left(a_n^{-1}|X_n| > C_\varepsilon\right) < \varepsilon$ for all $n \geq n_\varepsilon$, that is,

$$X_n = O_p(a_n) \iff \lim_{C\to\infty} \limsup_{n\to\infty} \mathrm{P}\left(\frac{|X_n|}{a_n} > C\right) = 0.$$

Definition 1.4. Let $\{X_n\}$ be any sequence of random variables. We say that $X_n = o_p(a_n)$ as $n \to \infty$ if for every $\varepsilon > 0$ there exists an integer n_ε such that $\mathrm{P}\left(a_n^{-1}|X_n| > \varepsilon\right) < \varepsilon$ for all $n \geq n_\varepsilon$.

Chapter 2

Bootstrap confidence bounds

The bootstrap is an automated, computerised resampling technique that, since its introduction by Bradley Efron in the 1970s, has proved to be successful in many problems of inference too complex to address adequately by means of traditional analytical methods. In fact, apart from being straightforward to implement, in many situations the bootstrap has been shown to produce results that are superior to results obtained by traditional methods, especially when the sample size is small or when underlying model assumptions cannot be verified. Efron and Tibshirani (1993, p. 45) state that *"The bootstrap [...] enjoys the advantage of being completely automatic. [It] requires no theoretical calculations, and is available no matter how mathematically complicated the estimator may be."*

Throughout this chapter we are concerned solely with the construction of confidence bounds from a random sample. After introducing some notation, we revisit the well-known pivotal method for constructing confidence bounds and intervals. We also define some measures to gauge the asymptotic performance of confidence bounds. The existing, standard confidence bounds are reviewed in Sections 2.4 and 2.5, with some improved methods discussed in Sections 2.6 and 2.7. In Section 2.8 we describe how bootstrap iteration may be used to improve the accuracy of confidence bounds and intervals. We end the chapter by discussing an interesting study by Cheung and Lee (2005) in which the authors investigate the asymptotic performance of m/n bootstrap confidence bounds in a nonregular case.

2.1 Notation

Let $\mathscr{X}_n = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ denote a random sample from an unknown d-variate distribution function F. The objective is to construct a $100(1-\alpha)\%$ upper confidence bound for a parameter $\theta = \theta(F)$, where $\theta(\cdot)$ is some known functional. The *plug-in estimator* for θ is given by $\widehat{\theta}_n = \theta(F_n)$, where $F_n(\mathbf{x}) = n^{-1}\sum_{i=1}^n \mathrm{I}(\mathbf{X}_i \le \mathbf{x})$. Let $\mathscr{X}_n^* = \{\mathbf{X}_1^*, \mathbf{X}_2^*, \ldots, \mathbf{X}_n^*\}$ be a random sample from F_n and define the bootstrap version of θ by $\widehat{\theta}_n^* = \theta(F_n^*)$, where $F_n^*(\mathbf{x}) = n^{-1}\sum_{i=1}^n \mathrm{I}\left(\mathbf{X}_i^* \le \mathbf{x}\right)$.

Denote the asymptotic standard error of $n^{1/2}\widehat{\theta}_n$ by σ and let $\widehat{\sigma}_n$ be an estimator for σ.

Define the *standardised* and *Studentised* forms of $\widehat{\theta}_n$ by

$$S_n := \frac{n^{1/2}(\widehat{\theta}_n - \theta)}{\sigma} \quad \text{and} \quad T_n := \frac{n^{1/2}(\widehat{\theta}_n - \theta)}{\widehat{\sigma}_n}, \tag{2.1}$$

respectively. Let ξ_β and η_β denote the theoretical β-level quantiles of S_n and T_n, i.e.,

$$\mathrm{P}\left(S_n \le \xi_\beta\right) = \mathrm{P}\left(T_n \le \eta_\beta\right) = \beta.$$

In order to present the bootstrap methods, we require some additional notation. As before, denote the *standardised* and *Studentised* forms of $\widehat{\theta}_n^*$ by S_n^* and T_n^* respectively, that is,

$$S_n^* := \frac{n^{1/2}(\widehat{\theta}_n^* - \widehat{\theta}_n)}{\widehat{\sigma}_n} \quad \text{and} \quad T_n^* := \frac{n^{1/2}(\widehat{\theta}_n^* - \widehat{\theta}_n)}{\widehat{\sigma}_n^*},$$

where $\widehat{\sigma}_n^*$ is the same function of $\mathscr{X}_n^*$ as $\widehat{\sigma}_n$ is of $\mathscr{X}_n$. Define $\widehat{\xi}_\beta$ and $\widehat{\eta}_\beta$ such that

$$\mathrm{P}^*\left(S_n^* \le \widehat{\xi}_\beta\right) = \mathrm{P}^*\left(T_n^* \le \widehat{\eta}_\beta\right) = \beta, \tag{2.2}$$

i.e., $\widehat{\xi}_\beta$ and $\widehat{\eta}_\beta$ denote the β-level quantiles of the bootstrap distributions of S_n^* and T_n^* respectively. Note that the quantile $\widehat{\xi}_\beta$ can also be rewritten in an explicit form. Let

$$\widehat{G}_n(t) = \mathrm{P}^*\left(\widehat{\theta}_n^* \le t\right) = \mathrm{P}\left(\widehat{\theta}_n^* \le t \,|\, F_n\right),$$

where P^* is short-hand notation for the conditional law of $\mathscr{X}_n^*$ given $\mathscr{X}_n$. We have from (2.2) that

$$\beta = \mathrm{P}^*\left(\frac{n^{1/2}(\widehat{\theta}_n^* - \widehat{\theta}_n)}{\widehat{\sigma}_n} \le \widehat{\xi}_\beta\right) = \mathrm{P}^*\left(\widehat{\theta}_n^* \le \widehat{\theta}_n + n^{-1/2}\widehat{\sigma}_n\widehat{\xi}_\beta\right) = \widehat{G}_n\left(\widehat{\theta}_n + n^{-1/2}\widehat{\sigma}_n\widehat{\xi}_\beta\right),$$

that is,

$$\widehat{G}_n^{-1}(\beta) = \widehat{\theta}_n + n^{-1/2}\widehat{\sigma}_n\widehat{\xi}_\beta. \tag{2.3}$$

Consequently we can express the quantile $\widehat{\xi}_\beta$ as

$$\widehat{\xi}_\beta = \frac{n^{1/2}\left(\widehat{G}_n^{-1}(\beta) - \widehat{\theta}_n\right)}{\widehat{\sigma}_n}.$$

The quantile $\widehat{\eta}_\beta$ cannot be written explicitly in this way.

2.2 Construction of confidence bounds

To shed some light on some of the properties of the existing bootstrap confidence bounds (and intervals), it is worth reviewing a method commonly used to construct confidence bounds (or intervals) for θ. Noting that S_n and T_n are (at least asymptotically) pivotal, we may use them to obtain exact confidence bounds for θ. For example, note that

$$1 - \alpha = \mathrm{P}(S_n \ge \xi_\alpha) = \mathrm{P}\left(\theta \le \widehat{\theta}_n - n^{-1/2}\sigma\xi_\alpha\right).$$

Therefore, we obtain the following $100(1-\alpha)\%$ *percentile* upper confidence bound for θ:

$$\mathscr{I}(\alpha) := \left(-\infty, \widehat{\theta}_n - n^{-1/2}\sigma\xi_\alpha\right]. \tag{2.4}$$

Analogously, based on T_n, a $100(1-\alpha)\%$ *percentile-t* upper confidence bound for θ is given by

$$\mathscr{J}(\alpha) := \left(-\infty, \widehat{\theta}_n - n^{-1/2}\widehat{\sigma}_n\eta_\alpha\right]. \tag{2.5}$$

Clearly, $\mathscr{I}$ and $\mathscr{J}$ are *exact* confidence bounds as they achieve the *nominal* coverage probability $1-\alpha$, i.e.,

$$\mathrm{P}(\theta \in \mathscr{I}(\alpha)) = \mathrm{P}\left(\theta \in \mathscr{J}(\alpha)\right) = 1-\alpha.$$

Note that $\mathscr{I}$ and $\mathscr{J}$ are ideal confidence bounds that typically cannot be calculated as they contain unknown quantities. The parameter σ and the quantiles ξ_β and η_β appearing in (2.4) and (2.5) are usually unknown and need to be estimated. Moreover, in order to determine the quantiles ξ_β and η_β exactly requires knowledge of the distributions of S_n and T_n, which are unknown in most practical settings. However, the bootstrap provides a standard procedure for approximating ξ_β and η_β.

Lower confidence bounds and equal-tailed confidence intervals

Although only *upper* confidence bounds are studied in this thesis, the results immediately hold also for *lower* confidence bounds. For instance, if $\mathscr{J}(\alpha)$ is an upper $(1-\alpha)$-level confidence bound for θ, then

$$\mathbb{R}\backslash\mathscr{J}(1-\alpha) = \left(\widehat{\theta}_n - n^{-1/2}\widehat{\sigma}_n\eta_{1-\alpha}, \infty\right)$$

is a lower $(1-\alpha)$-level confidence bound for θ. Likewise, a two-sided *equal-tailed* $(1-2\alpha)$-*level confidence interval* for θ is given by

$$\mathscr{J}(\alpha)\backslash\mathscr{J}(1-\alpha) = \left(\widehat{\theta}_n - n^{-1/2}\widehat{\sigma}_n\eta_{1-\alpha}, \widehat{\theta}_n - n^{-1/2}\widehat{\sigma}_n\eta_\alpha\right].$$

2.3 Consistency and accuracy

In our study we would like to assess and compare the performance of the various existing confidence bounds in the literature with the new bounds that will be proposed in Chapters 4 and 5. From an asymptotic perspective, we require that all considered confidence bounds at least be consistent in the following sense.

Definition 2.1. A confidence set $\mathscr{H}_n(\alpha)$ for θ with nominal level $1-\alpha$ is *consistent* if

$$\mathrm{P}(\theta \in \mathscr{H}_n(\alpha)) \to 1-\alpha \quad \text{as} \quad n \to \infty.$$

In addition to consistency, we will mainly compare the asymptotic performance of the various confidence bounds in terms of their asymptotic accuracy, which can be defined as follows (also see Shao and Tu, 1995, p. 144):

Definition 2.2. A confidence set $\mathcal{H}_n(\alpha)$ for θ with nominal level $1-\alpha$ is said to be *kth-order accurate* if

$$\mathrm{P}(\theta \in \mathcal{H}_n(\alpha)) = 1-\alpha + O(n^{-k/2}).$$

Note that if a confidence set is at least first order accurate, it is also consistent.

The reader should note that these definitions relate to asymptotic considerations only. Simulation studies can also provide valuable additional information, especially in the case of small to moderate samples. Indeed, Cheung and Lee (2005) warn against over-reliance on asymptotic implications in practical application. For this reason we include a small simulation study in Chapter 9 to investigate the finite-sample behaviour of the existing and new bounds.

2.4 Percentile-type bounds

The first two types of bounds are based on the standardised quantity S_n defined in (2.1). Bounds based on this quantity are commonly referred to as *percentile* bounds.

2.4.1 Backwards bounds

Probably the most widely used $100(1-\alpha)\%$ bootstrap upper confidence bound for the parameter θ is

$$\widehat{\mathcal{I}}_B(\alpha) := \left(-\infty,\, \widehat{G}_n^{-1}(1-\alpha)\right]. \tag{2.6}$$

This method of constructing a bootstrap confidence bound is known as as the *basic* or *backwards percentile* method (Efron, 1981, p. 146). The corresponding *equal-tailed* $100(1-\alpha)\%$ confidence interval for θ is given by

$$\left[\widehat{G}_n^{-1}\left(\tfrac{\alpha}{2}\right),\, \widehat{G}_n^{-1}\left(1-\tfrac{\alpha}{2}\right)\right].$$

The bound $\widehat{\mathcal{I}}_B(\alpha)$ in (2.6) can be approximated by the following simple Monte Carlo algorithm suggested by Efron (1979).

Algorithm 2.1.

1. Draw B independent bootstrap random samples of size n from $\mathcal{X}_n$.
2. For each bootstrap sample, calculate the bootstrap replications $\widehat{\theta}_1^*, \widehat{\theta}_2^*, \ldots, \widehat{\theta}_B^*$ and the corresponding order statistics $\widehat{\theta}_{(1)}^* \le \widehat{\theta}_{(2)}^* \le \cdots \le \widehat{\theta}_{(B)}^*$.
3. Approximate $\widehat{\mathcal{I}}_B(\alpha)$ by
$$\left(-\infty,\, \widehat{\theta}_{(s)}^*\right],$$
where $s = \lfloor B(1-\alpha) \rfloor$.

Intuitively $\widehat{\mathscr{I}}_B$ seems like an obvious and good choice for a $100(1-\alpha)\%$ upper confidence bound. However, as Hall (1988) points out, this bound is a result of some serious logical flaws. Substituting (2.3) in the expression for $\widehat{\mathscr{I}}_B$ shows that it can also be expressed as

$$\widehat{\mathscr{I}}_B(\alpha) = \left(-\infty, \widehat{\theta}_n + n^{-1/2}\widehat{\sigma}_n\widehat{\xi}_{1-\alpha}\right].$$

There are now two very important observations one has to make. Firstly, comparing this alternative expression of $\widehat{\mathscr{I}}_B$ to the ideal percentile bound in (2.4), we see that instead of simply replacing $\xi_{1-\alpha}$ with its appropriate estimator $\widehat{\xi}_{1-\alpha}$, the upper bound $\mathscr{I}$ is approximated by using a quantile from the wrong tail of the distribution of S_n. Consequently, the bound $\widehat{\mathscr{I}}_B$ has been criticised of being *backwards*, whence its name. Furthermore, the unknown parameter σ has carelessly been replaced by its estimate $\widehat{\sigma}_n$ without substituting the standardised quantile $\xi_{1-\alpha}$ with the Studentised quantile $\eta_{1-\alpha}$. This is akin to using the standard normal table to look up quantiles instead of the t-table when constructing standard traditional confidence bounds (or intervals) for the mean. Thus, as Hall (1988, p. 933) notes, *the backwards percentile method is a result of two errors: looking up the wrong tables, backwards*.

Hall (1988) shows that, typically, $\widehat{\mathscr{I}}_B(\alpha)$ is first-order accurate, i.e.,

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_B(\alpha)\right) = 1-\alpha + O(n^{-1/2}).$$

He also demonstrates that, in the case where θ is the mean of a univariate *symmetrically* distributed population, the coverage error of $\widehat{\mathscr{I}}_B$ reduces to $O(n^{-1})$.

Properties of the backwards percentile bound

The backwards percentile bound $\widehat{\mathscr{I}}_B$ (and the corresponding equal-tailed confidence interval) has the following important properties:

- It is *range-preserving*. This means that if there are any restrictions on the values the parameter can assume (e.g., a correlation coefficient must lie in the interval $[0,1]$), $\widehat{\mathscr{I}}_B$ will produce a confidence bound that falls within the allowable range. According to Efron and Tibshirani (1993, p. 176), range-preserving intervals are sometimes more accurate.

- It is *transformation-respecting* (Efron and Tibshirani, 1993, p. 175). Suppose that, instead of θ, we are interested in constructing a $100(1-\alpha)\%$ upper confidence bound for $\phi = m(\theta)$, where m is any monotone function. For the bootstrap replication $\widehat{\phi}_n^* = m(\widehat{\theta}_n^*)$, define $\widehat{H}_n(x) = \mathrm{P}^*(\widehat{\phi}_n^* \leq x)$. Noting that, for any $\beta \in [0,1]$,

$$\beta = \mathrm{P}^*\left(\widehat{\theta}_n^* \leq \widehat{G}_n^{-1}(\beta)\right) = \mathrm{P}^*\left(\widehat{\phi}_n^* \leq \widehat{H}_n^{-1}(\beta)\right) = \mathrm{P}^*\left(\widehat{\theta}_n^* \leq m^{-1}\left(\widehat{H}_n^{-1}(\beta)\right)\right),$$

  we obtain the relationship

$$\widehat{H}_n^{-1}(\beta) = m\left(\widehat{G}_n^{-1}(\beta)\right).$$

Now if we apply the transformation m directly to the confidence bound given in (2.6), we obtain

$$\left(-\infty,\, m\left(\widehat{G}_n^{-1}(1-\alpha)\right)\right] = \left(-\infty,\, \widehat{H}_n^{-1}(1-\alpha)\right],$$

which is the standard backwards percentile $100(1-\alpha)\%$ confidence bound for ϕ.

2.4.2 Hybrid bounds

An alternative percentile bound commonly found in the literature (see, e.g., Hall, 1988; Davison and Hinkley, 1997) is

$$\widehat{\mathscr{I}}_H(\alpha) := \left(-\infty,\, 2\widehat{\theta}_n - \widehat{G}_n^{-1}(\alpha)\right]. \tag{2.7}$$

We will refer to this bound as the standard *hybrid percentile* bound. The corresponding *equal-tailed* $100(1-\alpha)\%$ confidence interval for θ is given by

$$\left[2\widehat{\theta}_n - \widehat{G}_n^{-1}\left(1-\tfrac{\alpha}{2}\right),\, 2\widehat{\theta}_n - \widehat{G}_n^{-1}\left(\tfrac{\alpha}{2}\right)\right].$$

The bound $\widehat{\mathscr{I}}_H(\alpha)$ in (2.7) may be approximated by the following simple Monte Carlo algorithm.

Algorithm 2.2.

1. Draw B independent bootstrap random samples of size n from $\mathscr{X}_n$.
2. For each bootstrap sample, calculate the bootstrap replications $\widehat{\theta}_1^*, \widehat{\theta}_2^*, \ldots, \widehat{\theta}_B^*$ and the corresponding order statistics $\widehat{\theta}_{(1)}^* \le \widehat{\theta}_{(2)}^* \le \cdots \le \widehat{\theta}_{(B)}^*$.
3. Approximate $\widehat{\mathscr{I}}_H(\alpha)$ by
$$\left(-\infty,\, 2\widehat{\theta}_n - \widehat{\theta}_{(r)}^*\right],$$
where $r = \lfloor B\frac{\alpha}{2} \rfloor$.

Substituting $\widehat{G}_n^{-1}$ by its expression in (2.3) renders $\widehat{\mathscr{I}}_H$ in the following form:

$$\widehat{\mathscr{I}}_H(\alpha) = \left(-\infty,\, \widehat{\theta}_n - n^{-1/2}\widehat{\sigma}_n\widehat{\xi}_\alpha\right].$$

Comparing $\widehat{\mathscr{I}}_H$ to the ideal bound $\mathscr{I}$ in (2.4) we see that the two bounds are almost in agreement, with the quantile ξ_α having been replaced by its appropriate estimator from the correct tail of the distribution. However, as is the case of the backwards percentile bound $\widehat{\mathscr{I}}_B$, the unknown parameter σ has been replaced by its estimate $\widehat{\sigma}_n$ without substituting the standardised quantile ξ_α with the Studentised quantile η_α. The hybrid percentile bound is therefore a result of one error: looking up the wrong tables (Hall, 1988, p. 928).

As shown in Hall (1988), the coverage probability of $\widehat{\mathscr{I}}_H(\alpha)$ is typically

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H(\alpha)\right) = 1 - \alpha + O(n^{-1/2}).$$

Like $\widehat{\mathscr{I}}_B$, the hybrid bound $\widehat{\mathscr{I}}_H$ is second-order accurate if θ is the mean of a symmetric population. However, unlike the backwards percentile bound, the hybrid percentile bound is not range-preserving and clearly need not be transformation-respecting.

2.5 Percentile-t-type bound

Efron (1981, p. 152) defines the *percentile-t* (or *bootstrap-t*) $100(1-\alpha)\%$ upper confidence bound for θ by

$$\widehat{\mathscr{J}}(\alpha) := \left(-\infty,\, \widehat{\theta}_n - n^{-1/2}\widehat{\sigma}_n\widehat{\eta}_\alpha\right].$$

The corresponding *equal-tailed* $100(1-\alpha)\%$ confidence interval for θ is given by

$$\left[\widehat{\theta}_n - n^{-1/2}\widehat{\sigma}_n\widehat{\eta}_{1-\alpha/2},\, \widehat{\theta}_n - n^{-1/2}\widehat{\sigma}_n\widehat{\eta}_{\alpha/2}\right].$$

Comparing $\widehat{\mathscr{J}}$ to the ideal bound in (2.5), we see that $\widehat{\mathscr{J}}$ is obtained by simply replacing the theoretical quantile η_α by its estimator $\widehat{\eta}_\alpha$. This bound therefore does not suffer from the same logical/philosophical flaws as $\widehat{\mathscr{I}}_B$ and $\widehat{\mathscr{I}}_H$. It is for this reason that the percentile-t bounds (and intervals) are favoured by Hall (1988). Beran (1987) also argues that, from the perspective of pivotal statistics, the percentile-t bound is "better" than the percentile hybrid bound. In fact, Hall (1988) shows that the percentile-t bound is typically second-order accurate, i.e.,

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{J}}(\alpha)\right) = 1-\alpha+O(n^{-1}).$$

Despite these good qualities, however, the percentile-t bound $\widehat{\mathscr{J}}$ suffers from one major drawback. Since $\widehat{\sigma}_n^*$ appears in the denominator of T_n^*, the efficacy of $\widehat{\mathscr{J}}$ relies heavily on how well σ can be estimated (Hall, 1992, p. 18). For example, if θ is the correlation coefficient or a ratio of two means, the percentile-t bound (or interval) performs very poorly.

Unlike the backwards percentile bound, the percentile-t bound is neither transformation-respecting nor range-preserving.

2.6 Bias-corrected percentile bounds

Efron (1981, p. 146) introduced a *bias-corrected* (BC) version of (2.6) as

$$\widehat{\mathscr{I}}_{BC}(\alpha) := \left(-\infty,\, \widehat{G}_n^{-1}(\Phi(z_{1-\alpha}+2\widehat{m}))\right],$$

where $z_{1-\alpha} = \Phi^{-1}(1-\alpha)$ and $\widehat{m} = \Phi^{-1}(\widehat{G}_n(\widehat{\theta}_n))$. The quantity $\widehat{m}$ represents the discrepancy between the median of $\widehat{\theta}_n^*$ and the median of $\widehat{\theta}_n$. Note that $\widehat{\mathscr{I}}_{BC}(\alpha) = \widehat{\mathscr{I}}_B(1-\Phi(z_{1-\alpha}+2\widehat{m}))$. Furthermore, if $\widehat{\theta}_n$ is median unbiased for θ, i.e., if $\mathrm{P}(\widehat{\theta}_n \le \theta) = 0.5$, then $\widehat{m} \approx 0$ so that the bias-corrected bound $\widehat{\mathscr{I}}_{BC}(\alpha)$ reduces to the usual backwards percentile bound $\widehat{\mathscr{I}}_B(\alpha)$. To approximate $\widehat{\mathscr{I}}_{BC}(\alpha)$ one may use Algorithm 2.1 with s replaced by

$$s = \lfloor B\,\Phi(z_\alpha + 2\widehat{m}_B)\rfloor, \qquad \text{where} \qquad \widehat{m}_B = \Phi^{-1}\left\{\frac{1}{B}\sum_{b=1}^{B}\mathrm{I}\left(\widehat{\theta}^*_{(b)} \le \widehat{\theta}_n\right)\right\}.$$

The derivation of the bias-corrected method is based on the assumption that there exists a monotone increasing function g such that $\tau^{-1}(g(\widehat{\theta}) - g(\theta)) \sim N(-m, 1)$, where τ denotes the standard error of $g(\widehat{\theta})$ and $m = \Phi^{-1}(G(\theta))$. Schenker (1985) shows that in the simple case where the objective is to construct a confidence interval for the variance of a normal population, such a function g does not exist. This results in the poor performance of the backwards percentile and bias-corrected bounds.

In response to Schenker's counterexample, Efron (1987) proposed an *accelerated bias-corrected* (ABC) bound which relaxes the assumption on the transformation g (also see Chapter 23 of Efron and Tibshirani, 1993). The bound is defined by

$$\widehat{\mathscr{I}}_{ABC}(\alpha) = \left(-\infty,\, \widehat{G}_n^{-1}\left\{\Phi\left(\widehat{m} + \frac{\widehat{m} + z_{1-\alpha}}{1 - a(\widehat{m} + z_{1-\alpha})}\right)\right\}\right],$$

where a is known as the *acceleration constant*. Although Efron (1987) provides a more general framework for calculating the acceleration constant, Efron and Tibshirani (1993) suggest the following simple jackknife estimator:

$$\widehat{a} = \frac{\sum_{i=1}^n \left(\widehat{\theta}_{(\cdot)} - \widehat{\theta}_{(i)}\right)^3}{6\left\{\sum_{i=1}^n \left(\widehat{\theta}_{(\cdot)} - \widehat{\theta}_{(i)}\right)^2\right\}^{3/2}},$$

where $\widehat{\theta}_{(i)}$ is the same function of the sample as $\widehat{\theta}_n$, but with the ith observation deleted, and $\widehat{\theta}_{(\cdot)} = n^{-1}\sum_{i=1}^n \widehat{\theta}_{(i)}$.

The bias-corrected bounds are essentially percentile-type bounds. As a consequence, $\widehat{\mathscr{I}}_{BC}(\alpha)$ typically has coverage probability

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_{BC}(\alpha)\right) = 1 - \alpha + O(n^{-1/2}).$$

However, $\widehat{\mathscr{I}}_{ABC}(\alpha)$ is second-order accurate, i.e.,

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_{ABC}(\alpha)\right) = 1 - \alpha + O(n^{-1}).$$

Since the bias-corrected bounds are versions of the backwards percentile bound $\widehat{\mathscr{I}}_B$ evaluated at adjusted levels, both $\widehat{\mathscr{I}}_{BC}$ and $\widehat{\mathscr{I}}_{ABC}$ are transformation-respecting. See Section 2.4.1.

2.7 The Chung-Lee bound

Chung and Lee (2001) study the construction of bootstrap confidence bounds by means of the m/n bootstrap (see, e.g., Bickel and Freedman, 1981; Swanepoel, 1986). Chung and Lee show that the typical coverage error of $O(n^{-1/2})$ incurred by the *percentile* bounds $\widehat{\mathscr{I}}_B$ and $\widehat{\mathscr{I}}_H$ may be reduced to $O(n^{-1})$ with an optimal choice of m.

Let $\mathscr{X}_m^* = \{\mathbf{X}_1^*, \ldots, \mathbf{X}_m^*\}$ denote a bootstrap sample of size m taken from $\mathscr{X}_n$ and let $\widehat{\theta}_m^*$ be the same function of $\mathscr{X}_m^*$ as $\widehat{\theta}_n$ is of $\mathscr{X}_n$. Also, define $\widehat{G}_m(t) := \mathrm{P}^*\left(\widehat{\theta}_m^* \le t\right)$. Chung and Lee propose the following $100(1-\alpha)\%$ bootstrap upper confidence bound for θ:

$$\widehat{\mathscr{C}}(m, \delta, \alpha) = \begin{cases} \left(-\infty,\, \widehat{\theta} + (m/n)^{1/2}\left(\widehat{G}_m^{-1}(1-\alpha) - \widehat{\theta}\right)\right] & \text{if } \delta = +1, \\ \left(-\infty,\, \widehat{\theta} - (m/n)^{1/2}\left(\widehat{G}_m^{-1}(\alpha) - \widehat{\theta}\right)\right] & \text{if } \delta = -1, \end{cases}$$

where δ is a parameter. They show that choosing

$$m = \left\lfloor n\left(\frac{P_1(z_{1-\alpha})}{Q_1(z_{1-\alpha})}\right)^2\right\rfloor \quad \text{and} \quad \delta = -\operatorname{sgn}\left(\frac{P_1(z_{1-\alpha})}{Q_1(z_{1-\alpha})}\right),$$

where P_1 and Q_1 are Edgeworth polynomials appearing in Edgeworth expansions[1] of the distributions of S_n and T_n, respectively, guarantees (under some regularity and moment conditions) that $\widehat{\mathscr{C}}$ has second-order accuracy, i.e.,

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{C}}(m, \delta, \alpha)\right) = 1 - \alpha + O(n^{-1}). \tag{2.8}$$

As will be seen in Chapter 3, the polynomials P_1 and Q_1 depend on the parameter of interest. For example, in the case where θ is the population mean, they are given by

$$P_1(z_{1-\alpha}) = -\tfrac{1}{6}\kappa_3'(z_{1-\alpha}^2 - 1) \quad \text{and} \quad Q_1(z_{1-\alpha}) = \tfrac{1}{6}\kappa_3'(2z_{1-\alpha}^2 + 1),$$

where $\kappa_3' = \mathrm{E}\{(X_1 - \theta)^3\}/\sigma^3$. In this case $P_1(z_{1-\alpha})/Q_1(z_{1-\alpha}) = -(z_{1-\alpha}^2 - 1)/(2z_{1-\alpha}^2 + 1)$, which does not depend on the underlying distribution F. Hence, for a 95% upper confidence bound for θ we have that $m \approx \lfloor 0.0708n \rfloor$ and $\delta = 1$. Chung and Lee further show that the result in (2.8) is still valid if m and δ are replaced by their *empirical* counterparts (i.e., if population moments are substituted for corresponding sample moments). However, since m and δ do not depend on F in the case where θ is the mean, the estimators $\widehat{m}$ and $\widehat{\delta}$ will be independent of the sample $\mathscr{X}_n$ and hence are *not* data-based choices.

The reader should note that, given the above strategy, the improvement in coverage relies on knowledge of the polynomials P_1 and Q_1, which might be difficult to obtain for complicated statistics. However, Chung and Lee alternatively propose a double bootstrap procedure for approximating $\widehat{m}$ and $\widehat{\delta}$ which avoids explicit evaluation of the expressions for P_1 and Q_1.

Another important contribution by Chung and Lee (2001) is that they established valid Edgeworth expansions for

$$\mathrm{P}^*\left(\frac{\sqrt{m}\left(\widehat{\theta}_m^* - \widehat{\theta}_n\right)}{\widehat{\sigma}_n} \le x\right) \quad \text{and} \quad \mathrm{P}^*\left(\frac{\sqrt{m}\left(\widehat{\theta}_m^* - \widehat{\theta}_n\right)}{\widehat{\sigma}_m^*} \le x\right),$$

where $\widehat{\sigma}_m^*$ is the same function of $\mathscr{X}_m^*$ as $\widehat{\sigma}_n$ is of $\mathscr{X}_n$. Consequently, Cornish-Fisher expansions of the quantiles of the above bootstrap distributions may also be obtained. These results will for our purposes become important in Chapters 4 and 5.

2.8 Bootstrap iteration and calibration

Beran (1987) argues that the percentile-t bound $\widehat{\mathscr{I}}$ is more accurate than the two percentile bounds $\widehat{\mathscr{I}}_B$ and $\widehat{\mathscr{I}}_H$ because T_n is a better pivot than S_n in the sense that T_n is less dependent on the underlying distribution F. Consequently he proposes an iterative bootstrap method known as *bootstrap prepivoting* to obtain "improved" pivots and hence confidence bounds or intervals with improved accuracy. The method proceeds as follows: start with a pivot $R_n^{(0)}$ with distribution function $H_n^{(0)}$ and obtain the improved pivot

$$R_n^{(1)} = \widehat{H}_n^{(0)}\left(R_n^{(0)}\right),$$

[1]Edgeworth expansion will be discussed in detail in Chapter 3.

where $\widehat{H}_n^{(0)}$ denotes the bootstrap estimator for $H_n^{(0)}$. This procedure can be repeated to obtain an even more improved pivot $R_n^{(2)}$. In general, define

$$R_n^{(j+1)} = \widehat{H}_n^{(j)}\left(R_n^{(j)}\right), \qquad j = 0, 1, 2, \ldots,$$

where $\widehat{H}_n^{(j)}$ denotes the bootstrap estimator for the distribution of $R_n^{(j)}$. An upper confidence bound for θ may then be defined as

$$\widehat{\mathscr{L}}^{(j)}(\alpha) := \left\{\theta : R_n^{(j)} \leq \left(\widehat{H}_n^{(j)}\right)^{-1}(1-\alpha)\right\}.$$

Beran (1987) shows that if $R_n^{(0)} = S_n$ and S_n admits a two-term Edgeworth expansion, then $\widehat{\mathscr{L}}^{(1)}(\alpha)$ will be second-order accurate, which is comparable to the percentile-t bound $\widehat{\mathscr{J}}$. Another iteration of this procedure improves coverage even further, yielding a bound $\widehat{\mathscr{L}}^{(2)}(\alpha)$ which is third-order accurate. Using this method, confidence bounds of any desired order of accuracy may be obtained, at least *in theory*. However, iterative methods such as these are computationally intensive and quickly become infeasible. Furthermore, despite the reduction in computational demands offered by the warp-speed method of Giacomini, Politis and White (2013), Chang and Hall (2015) found that using the warp-speed method as a substitute for bootstrap iteration is not effective in offering the above levels of improved accuracy of confidence bounds (or intervals).

Beran's prepivoting is a special case of a more general idea known as *bootstrap iteration* which was later proposed and studied by Hall and Martin (1988). A related concept, termed *bootstrap calibration*, was introduced by Loh (1987, 1988, 1991). DiCiccio and Romano (1988) showed that bootstrap calibration and prepivoting are closely related, and in some cases equivalent.

2.9 The Cheung-Lee-Young bound

The consistency and accuracy results reported for each of the confidence bounds discussed thus far have been established by various authors from an asymptotically pivotal statistic of the form

$$W_n = \sqrt{n}\left(g(\bar{X}_n) - g(\mu)\right),$$

where g is a sufficiently smooth function, $\bar{X}_n = \sum_{i=1}^n X_i$ and $\mu = \mathrm{E}(X_1)$. For simplicity, we assume that $X_1, \ldots, X_n$ are iid random variables from a univariate distribution. It is well known that if X_1 has finite second-order moments and g has continuous derivatives in a neighbourhood of μ, then W_n has a limiting normal distribution (Cramér, 1946). Cases in which the pivotal statistic can be expressed thus are often referred to as *regular* cases. This is in fact the smooth function model studied by Bhattacharya and Ghosh (1978), which will be discussed thoroughly in Chapter 3.

There exist many *nonregular* cases (see, e.g., Bickel and Freedman, 1981) in which the bootstrap fails. Another type of nonregular case is, e.g., when $g'(\mu) = 0$. In this case the limiting distribution of W_n is not normal, but degenerate (see Shao, 1994). However, $\sqrt{n}W_n$

converges weakly to a nongenerate random variable which is a linear combination of χ_1^2 variables. Note that in this case the correct normalising constant of $g(\bar{X}_n) - g(\mu)$ is n, not $\sqrt{n}$. It turns out that the traditional bootstrap is inconsistent in this nonregular case, but Shao (1994) demonstrates that the m/n bootstrap rectifies the problem.

In this specific nonregular case, Cheung and Lee (2005) studied the asymptotic behaviour of the following percentile-type m/n bootstrap confidence bound:

$$\widehat{\mathscr{K}}(\alpha, m) := \left(-\infty, \widehat{\theta}_n - n^{-1}\widetilde{G}_m^{-1}(1-\alpha)\right],$$

where $\widetilde{G}_m(t) = \mathrm{P}^*(m(\widehat{\theta}_m^* - \widehat{\theta}) \leq t)$. They show that in this case, under certain conditions, $\widehat{\mathscr{K}}$ is first-order accurate, i.e., has coverage error $O(n^{-1/2})$. They also found that bootstrap iteration discussed in the previous section is ineffective in improving the accuracy of $\widehat{\mathscr{K}}$. However, they propose a new iterative scheme which is successful in reducing the coverage error to $O(n^{-2/3})$ in this nonregular case.

Chapter 3

Edgeworth expansion and the smooth function model

The idea of Edgeworth expansion was conceived in the wake of the nineteenth century by Edgeworth (1905) and contemporaries. These series expansions afforded a major improvement over the then already existing Gram-Charlier expansions in that the error of approximation (as a function of n) could be controlled (Cramér, 1946). Rigorous theory guaranteeing the validity of such expansions, however, was only developed gradually for specific cases over the course of the twentieth century (Cramér, 1928, 1946; Chibishov, 1972, 1973; Sargan, 1975, 1976). In a review article by Wallace (1958) it was conjectured that Edgeworth expansions could be established formally for more general cases if proper assumptions were made. Finally, in 1978, Bhattacharya and Ghosh were able to show rigorously that a broad class of statistics indeed admit asymptotic expansions which are identical to the known Edgeworth expansions. Their main result holds for statistics that may be expressed in terms of a so-called *smooth function model*, which is discussed in Section 3.1.

Particularly interesting is the vast amount of research devoted to the development of formal expansions for the distribution function of Student's statistic. Perhaps the greatest motivation behind this lies in uncovering details of the well-known fact that, under certain conditions,

$$T_n := \frac{\sqrt{n}\left(\bar{X}_n - \mu\right)}{\widehat{\sigma}_n} \xrightarrow{d} N(0,1),$$

where $\bar{X}_n = n^{-1}\sum_{i=1}^n X_i$ and $\widehat{\sigma}_n^2 = (n-1)^{-1}\sum_{i=1}^n (X_i - \bar{X}_n)^2$ for a random sample $X_1, X_2, \ldots, X_n$ from an unknown distribution with mean μ and variance $0 < \sigma^2 < \infty$. Edgeworth expansion of the distribution function of T_n yields

$$\mathrm{P}(T_n \leq x) = \Phi(x) + \frac{\kappa_3'(2x^2+1)\phi(x)}{6\sqrt{n}} + o(n^{-1/2}), \tag{3.1}$$

uniformly in x, where $\kappa_3' = \mathrm{E}\{(X_1-\mu)^3\}/\sigma^3$, giving a concise account of the *rate of convergence*.

Naturally the question arises: under which conditions is such an expansion valid? Using a method devised by Hsu (1945), Chung (1946) was able to show that the above expansion is valid if the distribution of X_1 is nonsingular and has at least 8 finite moments. Improving on this, the later results of the smooth function model of Bhattacharya and Ghosh show that one need only assume that the 6th moment of X_1 exists. However, looking at (3.1) one notices that the terms up to order $n^{-1/2}$ depend on moments of X_1 only up to the 3rd order. Indeed, Hall (1987) showed that the expansion remains valid, even under the minimal assumption $\mathrm{E}(|X_1|^3) < \infty$, together with the usual assumption of nonsingularity of X_1.

As far as the convergence of T_n to the standard normal distribution is concerned, the most general result we are aware of is that by Giné, Götze and Mason (1997). They show that one has $T_n \xrightarrow{d} N(0,1)$ *if and only if* the distribution of X_1 lies in the domain of attraction of the normal distribution. Hall and Wang (2004) derive the leading term of an Edgeworth expansion of $\mathrm{P}(T_n \le x)$ under this general condition (i.e., without any moment conditions).

In this chapter we discuss Edgeworth expansion in the general smooth function model framework of Bhattacharya and Ghosh (1978). We also describe the related Cornish-Fisher expansions of distribution quantiles, which may be obtained by inverting the Edgeworth expansions. For the two quantities $S_n = n^{1/2}(\widehat{\theta}_n - \theta)/\sigma$ and $T_n = n^{1/2}(\widehat{\theta}_n - \theta)/\widehat{\sigma}_n$ considered in Chapter 2, we provide explicit expansions for the cases where the parameter θ is the population mean and the population variance, respectively. In the case where θ is the slope parameter in a simple linear regression model, one may also obtain appropriate Edgeworth and Cornish-Fisher expansions. For this case we derive some new results which do not appear in the literature.

3.1 The smooth function model

Consider a random sample $\mathscr{X}_n = \{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ from an unknown p-dimensional distribution depending on a scalar parameter θ. For $k = 1, \ldots, n$, set $\mathbf{W}_k = (f_1(\mathbf{X}_k), \ldots, f_d(\mathbf{X}_k))$, where $f_1, \ldots, f_d$ are real-valued Borel-measurable functions on $\mathbb{R}^p$. Define $\boldsymbol{\nu} = \mathrm{E}(\mathbf{W}_1)$ and $\bar{\mathbf{W}}_n = n^{-1}\sum_{i=1}^n \mathbf{W}_i$. Assume that the parameter of interest is of the form $\theta = g(\boldsymbol{\nu})$, where $g : \mathbb{R}^d \to \mathbb{R}$ is a known Borel-measurable function which is sufficiently *smooth*[1] in a neighbourhood of $\boldsymbol{\nu}$. We will estimate θ by $\widehat{\theta}_n := g(\bar{\mathbf{W}}_n)$, which we assume has asymptotic variance of the form $n^{-1}\beta^2 = n^{-1}h^2(\boldsymbol{\nu})$, with $h : \mathbb{R}^d \to \mathbb{R}$ any known, Borel-measurable function, sufficiently *smooth* in a neighbourhood of $\boldsymbol{\nu}$. The sample estimate of β^2 is $\widehat{\beta}^2 := h^2(\bar{\mathbf{W}}_n)$.

Throughout the chapter, the following two particular smooth functions are of interest:

$$A_S(\mathbf{w}) := \frac{g(\mathbf{w}) - g(\boldsymbol{\nu})}{h(\boldsymbol{\nu})},$$

to which we refer as the *standardised* form, and

$$A_T(\mathbf{w}) := \frac{g(\mathbf{w}) - g(\boldsymbol{\nu})}{h(\mathbf{w})},$$

to which we refer as the *Studentised* form. Note that for both of these functions it holds that $A_S(\boldsymbol{\nu}) = A_T(\boldsymbol{\nu}) = 0$. Hence, assuming that $\mathbf{X}_1$ has sufficiently many finite moments, we

[1] A function is said to be *smooth* in a neighbourhood of $\boldsymbol{\nu}$ if it is continuously differentiable up to a sufficiently large order in a neighbourhood of $\boldsymbol{\nu}$. See Theorem 3.1.

know from the results of Chapter 28 of Cramér (1946) that $n^{1/2}A(\bar{\mathbf{W}}_n)$ has a limiting normal distribution with mean zero and variance

$$\sum_{i=1}^{k}\sum_{j=1}^{k}\operatorname{Cov}\left(f_i(\mathbf{X}_1),f_j(\mathbf{X}_1)\right)\left[\frac{\partial A(\mathbf{w})}{\partial w_i}\right]_{\mathbf{w}=\boldsymbol{\nu}}\left[\frac{\partial A(\mathbf{w})}{\partial w_j}\right]_{\mathbf{w}=\boldsymbol{\nu}},$$

where A stands for either A_S or A_T. In fact, Bhattacharya and Ghosh (1978) demonstrate that probabilities of the form $\mathrm{P}\left(n^{1/2}A(\bar{\mathbf{W}}_n)\le x\right)$ admit rigorously-established Edgeworth expansions. We formally present their result in the following theorem in the same form as given by Hall (1992).

Theorem 3.1 (Theorem 2.2 of Hall, 1992)**.** *Assume that the function A has $j+3$ continuous derivatives in a neighbourhood of $\boldsymbol{\nu}=\mathrm{E}(\mathbf{W}_1)$, that $\mathrm{E}(\|\mathbf{W}_1\|^{j+3})<\infty$, and that the characteristic function χ of $\mathbf{W}_1$ satisfies Cramér's condition, that is,*

$$\limsup_{\|\mathbf{t}\|\to\infty}|\chi(\mathbf{t})|<1. \tag{3.2}$$

Assuming that the asymptotic variance of $n^{1/2}A(\bar{\mathbf{W}}_n)$ is 1, it holds for any integer $j\ge 1$ that

$$\begin{aligned}\mathrm{P}\left(n^{1/2}A(\bar{\mathbf{W}}_n)\le x\right)&=\Phi(x)+n^{-1/2}P_1(x)\phi(x)+n^{-1}P_2(x)\phi(x)+\cdots\\&\quad+n^{-j/2}P_j(x)\phi(x)+O(n^{-(j+1)/2}),\end{aligned} \tag{3.3}$$

uniformly in x, where P_j is a polynomial of degree at most $3j-1$, odd for even j and even for odd j, with coefficients depending on moments of $\mathbf{W}_1$ up to order $j+2$.

Proof. See Bhattacharya and Ghosh (1978).

3.2 Deriving the Edgeworth polynomials

The Edgeworth polynomials P_j in Theorem 3.1 above depend on the choice of the function $A(\cdot)$ and hence on the choice of the parameter θ. In this section we show heuristically how these polynomials can be obtained. For more detail on the mathematics behind the derivation of the polynomials, the reader may consult Hall (1992).

Let $\mathbf{X}_1,\mathbf{X}_2,\ldots,\mathbf{X}_n$ be a random sample from an unknown distribution with sufficiently many finite moments. Suppose that $S_n=S_n(\mathbf{X}_1,\mathbf{X}_2,\ldots,\mathbf{X}_n)$ is a statistic with a limiting standard normal distribution. If we denote the jth cumulant of S_n by $\kappa_{j,n}$, we may by definition express the characteristic function of S_n as

$$\chi_n(t)=\exp\left\{\kappa_{1,n}it+\tfrac{1}{2}\kappa_{2,n}(it)^2+\tfrac{1}{3!}\kappa_{3,n}(it)^3+\cdots\right\}. \tag{3.4}$$

Now, if S_n is a smooth statistic, we may, as a consequence of Theorem 2.1 of Hall (1992), express the cumulants of S_n as a power series in n^{-1} as follows:

$$\kappa_{j,n}=n^{-(j-2)/2}\left(k_{j,1}+n^{-1}k_{j,2}+n^{-2}k_{j,3}+\cdots\right),\qquad j=1,2,\ldots, \tag{3.5}$$

where the $k_{j,l}$ are real constants depending on the moments of $\mathbf{X}_1$. In particular, since S_n has an asymptotic mean of zero and asymptotic variance of one, $k_{1,1} = 0$ and $k_{2,1} = 1$. Hence, for example,

$$\kappa_{1,n} = n^{-1/2}k_{1,2} + n^{-3/2}k_{1,3} + O(n^{-5/2}),$$
$$\kappa_{2,n} = 1 + n^{-1}k_{2,2} + n^{-2}k_{2,3} + O(n^{-3}),$$
$$\kappa_{3,n} = n^{-1/2}k_{3,1} + n^{-3/2}k_{3,2} + O(n^{-5/2}),$$

etc. Substituting these expressions in (3.4) we may write (by Taylor expansion)

$$\chi_n(t) = e^{-t^2/2}\left\{1 + n^{-1/2}r_1(it) + n^{-1}r_2(it) + \cdots + n^{-j/2}r_j(it) + \cdots\right\}, \tag{3.6}$$

where the r_j, $j = 1,2,\ldots$, are polynomials with real coefficients depending on the $k_{j,l}$. It can be shown (see Chapter 2 of Hall, 1992) that the polynomials P_j appearing in Theorem 3.1 are then given by the relation

$$P_j(x)\phi(x) = r_j\left(-\frac{d}{dx}\right)\Phi(x). \tag{3.7}$$

For our purposes we will require the polynomials r_1, r_2 and r_3. The first two, which are

$$\begin{aligned} r_1(u) &= k_{1,2}u + \tfrac{1}{6}k_{3,1}u^3, \\ r_2(u) &= \tfrac{1}{2}\left(k_{2,2} + k_{1,2}^2\right)u^2 + \tfrac{1}{24}\left(k_{4,1} + 4k_{3,1}k_{1,2}\right)u^4 + \tfrac{1}{72}k_{3,1}^2u^6, \end{aligned} \tag{3.8}$$

are given in Hall (1992, p. 45). In the following lemma we state and derive an expression for r_3, which we will require later in this chapter. To the best of our knowledge this results does not appear in the literature.

Lemma 3.1. *Suppose the statistic S_n has a limiting standard normal distribution. If the cumulants of S_n allow expansion as in* (3.5)*, then the polynomial r_3 in the expansion in* (3.6) *is given by*

$$\begin{aligned} r_3(u) = k_{1,3}u &+ \tfrac{1}{6}\left(k_{3,2} + 3k_{2,2}k_{1,2} + k_{1,2}^3\right)u^3 + \tfrac{1}{120}\left(k_{5,1} + 5k_{4,1}k_{1,2} + 10k_{3,1}k_{2,2} + 10k_{3,1}k_{1,2}^2\right)u^5 \\ &+ \tfrac{1}{144}k_{3,1}\left(k_{4,1} + 2k_{3,1}k_{1,2}\right)u^7 + \tfrac{1}{1296}k_{3,1}^3u^9. \end{aligned}$$

Proof. We follow here the same method of derivation as found in Section 2.3 in Hall (1992). Substituting (3.5) in (3.4) and collecting terms we obtain

$$\begin{aligned} \chi_n(t) = \exp\Big\{ -\tfrac{1}{2}t^2 &+ n^{-1/2}\left(k_{1,2}it + \tfrac{1}{3!}k_{3,1}(it)^3\right) + n^{-1}\left(\tfrac{1}{2}k_{2,2}(it)^2 + \tfrac{1}{4!}k_{4,1}(it)^4\right) \\ &+ n^{-3/2}\left(k_{1,3}it + \tfrac{1}{3!}k_{3,2}(it)^3 + \tfrac{1}{5!}k_{3,1}(it)^5\right) + O(n^{-2})\Big\}. \end{aligned}$$

From the identity $e^x = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \cdots$ we have

$$\begin{aligned} \chi_n(t) = \exp\left(-\tfrac{1}{2}t^2\right)\Big(1 &+ n^{-1/2}\left\{k_{1,2}it + \tfrac{1}{3!}k_{3,1}(it)^3\right\} \\ &+ n^{-1}\left\{\tfrac{1}{2}k_{2,2}(it)^2 + \tfrac{1}{4!}k_{4,1}(it)^4 + \tfrac{1}{2}\left(k_{1,2}it + \tfrac{1}{3!}k_{3,1}(it)^3\right)^2\right\} \end{aligned}$$

$$+n^{-3/2}\Big\{k_{1,3}it+\tfrac{1}{3!}k_{3,2}(it)^3+\tfrac{1}{5!}k_{3,1}(it)^5$$
$$+\left(k_{1,2}it+\tfrac{1}{3!}k_{3,1}(it)^3\right)\left(k_{2,2}(it)^2+\tfrac{1}{4!}k_{4,1}(it)^4\right)$$
$$+\tfrac{1}{3!}\left(k_{1,2}it+\tfrac{1}{3!}k_{3,1}(it)^3\right)^3\Big\}\Big)+O(n^{-2})$$
$$=\exp\left(-\tfrac{1}{2}t^2\right)\Big(1+n^{-1/2}r_1(it)+n^{-1}r_2(it)+n^{-3/2}r_3(it)\Big)+O(n^{-2}),$$

whence we conclude that $r_1(u)$ and $r_2(u)$ are as in (3.8) and

$$r_3(u)=k_{1,3}u+\tfrac{1}{6}\left(k_{3,2}+3k_{2,2}k_{1,2}+k_{1,2}^3\right)u^3+\tfrac{1}{120}\left(k_{5,1}+5k_{4,1}k_{1,2}+10k_{3,1}k_{2,2}+10k_{3,1}k_{1,2}^2\right)u^5$$
$$+\tfrac{1}{144}k_{3,1}\left(k_{4,1}+2k_{3,1}k_{1,2}\right)u^7+\tfrac{1}{1296}k_{3,1}^3u^9. \qquad \square$$

3.3 Cornish-Fisher expansion

Edgeworth expansions may be inverted to obtain expansions of distribution quantiles. Such expansions are known as Cornish-Fisher expansions. The following theorem, taken from Hall (1992, p. 70), provides the general form of a Cornish-Fisher expansion corresponding to the expansion in Theorem 3.1.

Theorem 3.2 (Theorem 2.4 of Hall, 1992)**.** *Assume the conditions of Theorem 3.1 on the function A and the distribution of $\mathbf{W}_1$ and define*

$$w_{n,\alpha}=\inf\Big\{x:\mathrm{P}\Big(n^{1/2}A(\bar{\mathbf{W}}_n)\le x\Big)\ge\alpha\Big\}.$$

Then there exist polynomials $P_1^{cf},\dots,P_j^{cf}$, completely determined by $P_1,\dots,P_j$ appearing in (3.3)*, such that*

$$w_{n,\alpha}=z_\alpha+n^{-1/2}P_1^{cf}(z_\alpha)+\cdots+n^{-j/2}P_j^{cf}(z_\alpha)+O\left(n^{-(j+1)/2}\right),$$

uniformly in $\alpha\in(\varepsilon,1-\varepsilon)$ for each $\varepsilon>0$.

Proof. See Hall (1992).

The polynomials $P_1^{cf},\dots,P_j^{cf}$ are completely defined by the polynomials $P_1,\dots,P_j$ through the following formal implicit relation, which is easy to see:

$$\Phi\Big(z_\alpha+\sum_{j\ge1}n^{-j/2}P_j^{cf}(z_\alpha)\Big)+\sum_{i\ge1}n^{-i/2}P_i\Big(z_\alpha+\sum_{j\ge1}n^{-j/2}P_j^{cf}(z_\alpha)\Big)\phi\Big(z_\alpha+\sum_{j\ge1}n^{-j/2}P_j^{cf}(z_\alpha)\Big)=\alpha,$$

for $0<\alpha<1$ (cf. Hall, 1992, p. 60). Hall provides explicit expressions for P_1^{cf} and P_2^{cf} in terms of P_1 and P_2. They are,

$$P_1^{cf}(x)=-P_1(x) \tag{3.9}$$

and

$$P_2^{cf}(x)=P_1(x)P_1'(x)-\tfrac{1}{2}xP_1^2(x)-P_2(x), \tag{3.10}$$

for all $x\in\mathbb{R}$. In our work we will also require the third polynomial P_3^{cf}, which we state and prove in the following lemma.

Lemma 3.2. *The third Cornish-Fisher polynomial P_3^{cf} appearing in Theorem 3.2 is given by*

$$P_3^{cf}(x) = -\tfrac{1}{3}(x^2-1)P_1^3(x) + \tfrac{3}{2}xP_1^2(x)P_1'(x) - \tfrac{1}{2}P_1^2(x)P_1''(x) - xP_1(x)P_2(x) + P_1(x)P_2'(x)$$
$$-P_1(x)\big(P_1'(x)\big)^2 + P_1'(x)P_2(x) - P_3(x), \tag{3.11}$$

for all $x \in \mathbb{R}$.

Proof. By Taylor expansion about z_α we have that

$$\begin{aligned}
&\Phi\Big(z_\alpha + \sum_{j\ge1} n^{-j/2}P_j^{cf}(z_\alpha)\Big)\\
&= \Phi\Big(z_\alpha + n^{-1/2}P_1^{cf}(z_\alpha) + n^{-1}P_2^{cf}(z_\alpha) + n^{-3/2}P_3^{cf}(z_\alpha) + O(n^{-2})\Big)\\
&= \Phi(z_\alpha) + \phi(z_\alpha)\Big(n^{-1/2}P_1^{cf}(z_\alpha) + n^{-1}P_2^{cf}(z_\alpha) + n^{-3/2}P_3^{cf}(z_\alpha) + O(n^{-2})\Big)\\
&\quad - \tfrac{1}{2}z_\alpha\phi(z_\alpha)\Big(n^{-1/2}P_1^{cf}(z_\alpha) + n^{-1}P_2^{cf}(z_\alpha) + n^{-3/2}P_3^{cf}(z_\alpha) + O(n^{-2})\Big)^2\\
&\quad + \tfrac{1}{6}(z_\alpha^2-1)\phi(z_\alpha)\Big(n^{-1/2}P_1^{cf}(z_\alpha) + n^{-1}P_2^{cf}(z_\alpha) + n^{-3/2}P_3^{cf}(z_\alpha) + O(n^{-2})\Big)^3 + O(n^{-2})\\
&= \alpha + n^{-1/2}\phi(z_\alpha)P_1^{cf}(z_\alpha) + n^{-1}\phi(z_\alpha)\Big(P_2^{cf}(z_\alpha) - \tfrac{1}{2}z_\alpha\Big(P_1^{cf}(z_\alpha)\Big)^2\Big)\\
&\quad + n^{-3/2}\phi(z_\alpha)\Big(P_3^{cf}(z_\alpha) - z_\alpha P_1^{cf}(z_\alpha)P_2^{cf}(z_\alpha)\tfrac{1}{6}\Big(P_1^{cf}(z_\alpha)\Big)^3\Big) + O(n^{-2}).
\end{aligned} \tag{3.12}$$

Also by Taylor expansion about z_α,

$$\begin{aligned}
&\sum_{i\ge1} n^{-i/2}P_i\Big(z_\alpha + \sum_{j\ge1} n^{-j/2}P_j^{cf}(z_\alpha)\Big)\\
&= n^{-1/2}P_1\Big(z_\alpha + n^{-1/2}P_1^{cf}(z_\alpha) + n^{-1}P_2^{cf}(z_\alpha) + O(n^{-3/2})\Big)\\
&\quad + n^{-1}P_2\Big(z_\alpha + n^{-1/2}P_1^{cf}(z_\alpha) + n^{-1}P_2^{cf}(z_\alpha) + O(n^{-3/2})\Big)\\
&\quad + n^{-3/2}P_3\Big(z_\alpha + n^{-1/2}P_1^{cf}(z_\alpha) + n^{-1}P_2^{cf}(z_\alpha) + O(n^{-3/2})\Big) + O(n^{-2})\\
&= n^{-1/2}\Big[P_1(z_\alpha) + P_1'(z_\alpha)\Big(n^{-1/2}P_1^{cf}(z_\alpha) + n^{-1}P_2^{cf}(z_\alpha) + O(n^{-3/2})\Big)\\
&\qquad + \tfrac{1}{2}P_1''(z_\alpha)\Big(n^{-1/2}P_1^{cf}(z_\alpha) + n^{-1}P_2^{cf}(z_\alpha) + O(n^{-3/2})\Big)^2\Big]\\
&\quad + n^{-1}\Big[P_2(z_\alpha) + P_2'(z_\alpha)\Big(n^{-1/2}P_1^{cf}(z_\alpha) + n^{-1}P_2^{cf}(z_\alpha) + O(n^{-3/2})\Big) + O(n^{-1})\Big]\\
&\quad + n^{-3/2}\Big[P_3(z_\alpha) + O(n^{-1/2})\Big]\\
&= n^{-1/2}P_1(z_\alpha) + n^{-1}\Big(P_1^{cf}(z_\alpha)P_1'(z_\alpha) + P_2(z_\alpha)\Big)\\
&\quad + n^{-3/2}\Big(P_2^{cf}(z_\alpha)P_1'(z_\alpha) + \tfrac{1}{2}\Big(P_1^{cf}(z_\alpha)\Big)^2P_1''(z_\alpha) + P_1^{cf}(z_\alpha)P_2'(z_\alpha) + P_3(z_\alpha)\Big) + O(n^{-2}).
\end{aligned} \tag{3.13}$$

Similarly,

$$\begin{aligned}
&\phi\Big(z_\alpha + \sum_{j\ge1} n^{-j/2}P_j^{cf}(z_\alpha)\Big)\\
&= \phi\Big(z_\alpha n^{-1/2}P_1^{cf}(z_\alpha) + n^{-1}P_2^{cf}(z_\alpha) + O(n^{-3/2})\Big)
\end{aligned}$$

$$
\begin{aligned}
&= \phi(z_\alpha) - z_\alpha \phi(z_\alpha)\left(n^{-1/2}P_1^{cf}(z_\alpha) + n^{-1}P_2^{cf}(z_\alpha) + O(n^{-3/2})\right) \\
&\quad + \tfrac{1}{2}(z_\alpha^2 - 1)\phi(z_\alpha)\left(n^{-1/2}P_1^{cf}(z_\alpha) + n^{-1}P_2^{cf}(z_\alpha) + O(n^{-3/2})\right)^2 + O(n^{-3/2}) \\
&= \phi(z_\alpha) - n^{-1/2}z_\alpha\phi(z_\alpha)P_1^{cf}(z_\alpha) + n^{-1}\phi(z_\alpha)\left(\tfrac{1}{2}(z_\alpha^2-1)\left(P_1^{cf}(z_\alpha)\right)^2 - z_\alpha P_2^{cf}(z_\alpha)\right) \\
&\quad + O(n^{-3/2}). \qquad (3.14)
\end{aligned}
$$

Multiplying (3.13) and (3.14) yields

$$
\begin{aligned}
&\sum_{i\geq 1} n^{-i/2}P_i\Big(z_\alpha + \sum_{j\geq 1} n^{-j/2}P_j^{cf}(z_\alpha)\Big)\phi\Big(z_\alpha + \sum_{j\geq 1} n^{-j/2}P_j^{cf}(z_\alpha)\Big) \\
&= n^{-1/2}\phi(z_\alpha)P_1(z_\alpha) + n^{-1}\phi(z_\alpha)\left(-z_\alpha P_1(z_\alpha)P_1^{cf}(z_\alpha) + P_1'(z_\alpha)P_1^{cf}(z_\alpha) + P_2(z_\alpha)\right) \\
&\quad + n^{-3/2}\phi(z_\alpha)\Big(-z_\alpha P_1(z_\alpha)P_2^{cf}(z_\alpha) + \tfrac{1}{2}(z_\alpha^2-1)P_1(z_\alpha)\left(P_1^{cf}(z_\alpha)\right)^2 - z_\alpha\left(P_1^{cf}(z_\alpha)\right)^2 P_1'(z_\alpha) \\
&\qquad - z_\alpha P_2(z_\alpha)P_1^{cf}(z_\alpha) + P_2^{cf}(z_\alpha)P_1'(z_\alpha) + \tfrac{1}{2}\left(P_1^{cf}(z_\alpha)\right)^2 P_1''(z_\alpha)\Big) \\
&\qquad + P_1(z_\alpha)P_2'(z_\alpha) + P_3(z_\alpha)\Big) + O(n^{-2}). \qquad (3.15)
\end{aligned}
$$

We now sum (3.12) and (3.15) to obtain

$$
\begin{aligned}
0 &= n^{-1/2}\phi(z_\alpha)\left(P_1^{cf}(z_\alpha) + P_1(z_\alpha)\right) \\
&\quad + n^{-1}\phi(z_\alpha)\left(P_2^{cf}(z_\alpha) - \tfrac{1}{2}z_\alpha\left(P_1^{cf}(z_\alpha)\right)^2 - z_\alpha P_1(z_\alpha)P_1^{cf}(z_\alpha) + P_1^{cf}(z_\alpha)P_1'(z_\alpha) + P_2(z_\alpha)\right) \\
&\quad + n^{-3/2}\phi(z_\alpha)\Big(P_3^{cf}(z_\alpha) - z_\alpha P_1^{cf}(z_\alpha)P_2^{cf}(z_\alpha) + \tfrac{1}{6}(z_\alpha^2-1)\left(P_1^{cf}(z_\alpha)\right)^3 - z_\alpha P_1(z_\alpha)P_2^{cf}(z_\alpha) \\
&\qquad + \tfrac{1}{2}(z_\alpha^2-1)P_1(z_\alpha)\left(P_1^{cf}(z_\alpha)\right)^2 - z_\alpha\left(P_1^{cf}(z_\alpha)\right)^2 P_1'(z_\alpha) - z_\alpha P_2(z_\alpha)P_1^{cf}(z_\alpha) \\
&\qquad + P_2^{cf}(z_\alpha)P_1'(z_\alpha) + \tfrac{1}{2}\left(P_1^{cf}(z_\alpha)\right)^2 P_1''(z_\alpha) + P_1^{cf}(z_\alpha)P_2'(z_\alpha) + P_3(z_\alpha)\Big) + O(n^{-2}).
\end{aligned}
$$

This implies that $P_1^{cf}(z_\alpha) = -P_1(z_\alpha)$ and that

$$
\begin{aligned}
P_2^{cf}(z_\alpha) &= \tfrac{1}{2}z_\alpha\left(P_1^{cf}(z_\alpha)\right)^2 + z_\alpha P_1(z_\alpha)P_1^{cf}(z_\alpha) - P_1^{cf}(z_\alpha)P_1'(z_\alpha) - P_2(z_\alpha) \\
&= P_1(z_\alpha)P_1'(z_\alpha) - \tfrac{1}{2}z_\alpha P_1^2(z_\alpha) - P_2(z_\alpha),
\end{aligned}
$$

confirming (3.9) and (3.10). Furthermore,

$$
\begin{aligned}
P_3^{cf}(z_\alpha) &= z_\alpha P_1^{cf}(z_\alpha)P_2^{cf}(z_\alpha) - \tfrac{1}{6}(z_\alpha^2-1)\left(P_1^{cf}(z_\alpha)\right)^3 + z_\alpha P_1(z_\alpha)P_2^{cf}(z_\alpha) \\
&\quad - \tfrac{1}{2}(z_\alpha^2-1)P_1(z_\alpha)\left(P_1^{cf}(z_\alpha)\right)^2 + z_\alpha\left(P_1^{cf}(z_\alpha)\right)^2 P_1'(z_\alpha) + z_\alpha P_2(z_\alpha)P_1^{cf}(z_\alpha) \\
&\quad - P_2^{cf}(z_\alpha)P_1'(z_\alpha) - \tfrac{1}{2}\left(P_1^{cf}(z_\alpha)\right)^2 P_1''(z_\alpha) - P_1^{cf}(z_\alpha)P_2'(z_\alpha) - P_3(z_\alpha) \\
&= -\tfrac{1}{3}(z_\alpha^2-1)P_1^3(z_\alpha) + z_\alpha P_1^2(z_\alpha)P_1'(z_\alpha) - \tfrac{1}{2}P_1^2(z_\alpha)P_1''(z_\alpha) - z_\alpha P_1(z_\alpha)P_2(z_\alpha) \\
&\quad + P_1(z_\alpha)P_2'(z_\alpha) - P_1'(z_\alpha)P_2^{cf}(z_\alpha) - P_3(z_\alpha).
\end{aligned}
$$

Noting that this last equation holds for all $\alpha \in (0,1)$, and therefore for all $z_\alpha \in \mathbb{R}$, and substituting $P_2^{cf}(z_\alpha)$ by the expression in (3.10), we obtain the required result. □

3.4 Some examples

In this section we provide three examples to illustrate how a parameter (and its estimator) can be formulated in terms of the smooth function model of Bhattacharya and Ghosh (1978). For the first example, where the parameter of interest is the population mean, we also provide three-term Edgeworth expansions of the distributions of $n^{1/2}A_S(\bar{\mathbf{W}}_n)$ and $n^{1/2}A_T(\bar{\mathbf{W}}_n)$.

3.4.1 The mean

Let $X_1, X_2, \ldots, X_n$ denote a sample drawn randomly from a univariate distribution with mean μ and variance $0 < \sigma^2 < \infty$. Define $\bar{X}_n = n^{-1}\sum_{i=1}^n X_i$ and $\widehat{\sigma}_n^2 = n^{-1}\sum_{i=1}^n \left(X_i - \bar{X}_n\right)^2$. Suppose we are interested in estimating $\theta = \mu$. In the smooth function model setting, choose $d = 2$ and set

$$\mathbf{W}_i = \begin{bmatrix} X_i \\ X_i^2 \end{bmatrix}, \qquad i = 1, \ldots, n.$$

Then, assuming that the first two moments of X_1 exist, we have

$$\boldsymbol{\nu} = \mathrm{E}(\mathbf{W}_1) = \begin{bmatrix} \mathrm{E}(X_1) \\ \mathrm{E}(X_1^2) \end{bmatrix} = \begin{bmatrix} \mu \\ \mu^2 + \sigma^2 \end{bmatrix}$$

and

$$\bar{\mathbf{W}}_n = \frac{1}{n}\sum_{i=1}^n \mathbf{W}_i = \begin{bmatrix} n^{-1}\sum_{i=1}^n X_i \\ n^{-1}\sum_{i=1}^n X_i^2 \end{bmatrix} = \begin{bmatrix} \bar{X}_n \\ \bar{X}_n^2 + \widehat{\sigma}_n^2 \end{bmatrix}.$$

For any $\mathbf{w} = (w_1, w_2)' \in \mathbb{R}^2$, choosing

$$g(\mathbf{w}) = w_1 \quad \text{and} \quad h^2(\mathbf{w}) = w_2 - w_1^2$$

ensures that

$$\theta = g(\boldsymbol{\nu}) = \mu \quad \text{and} \quad \beta^2 = h^2(\boldsymbol{\nu}) = \sigma^2.$$

The corresponding estimators for the above quantities are given by

$$\widehat{\theta}_n = g(\bar{\mathbf{W}}_n) = \bar{X}_n,$$
$$\widehat{\beta}_n^2 = h^2(\bar{\mathbf{W}}_n) = \bar{X}_n^2 + \widehat{\sigma}_n^2 - \bar{X}_n^2 = \widehat{\sigma}_n^2.$$

Therefore the two smooth functions of interest are

$$A_S(\bar{\mathbf{W}}_n) = \frac{g(\bar{\mathbf{W}}_n) - g(\boldsymbol{\nu})}{h(\boldsymbol{\nu})} = \frac{\bar{X}_n - \mu}{\sigma}$$

and

$$A_T(\bar{\mathbf{W}}_n) = \frac{g(\bar{\mathbf{W}}_n) - g(\boldsymbol{\nu})}{h(\bar{\mathbf{W}}_n)} = \frac{\bar{X}_n - \mu}{\widehat{\sigma}_n}.$$

Note that $\mathrm{Var}(n^{1/2}A_S(\bar{\mathbf{W}}_n)) = 1$. Also, by the theorem of Section 27.7 of Cramér (1946), we have that $\mathrm{Var}(n^{1/2}A_T(\bar{\mathbf{W}}_n)) = 1 + O(n^{-1/2})$. Hence, both $n^{1/2}A_S(\bar{\mathbf{W}}_n)$ and $n^{1/2}A_T(\bar{\mathbf{W}}_n)$ have asymptotic variance 1.

Assume that Cramér's condition is satisfied and that $\mathrm{E}\left(X_1^{12}\right) < \infty$. The latter condition will ensure that $\mathrm{E}(\|\mathbf{W}_1\|^6) < \infty$, since we chose $\mathbf{W}_i = [X_i, X_i^2]'$. We then have by Theorem 3.1 that

$$\mathrm{P}\left(\frac{n^{1/2}\left(\bar{X}_n - \mu\right)}{\sigma} \le x\right) = \Phi(x) + n^{-1/2}P_1(x)\phi(x) + n^{-1}P_2(x)\phi(x) + n^{-3/2}P_3(x)\phi(x) + O\left(n^{-2}\right), \tag{3.16}$$

with

$$\begin{aligned} P_1(x) &= -\tfrac{1}{6}\kappa_3'(x^2-1), \\ P_2(x) &= -x\left\{\tfrac{1}{24}\kappa_4'(x^2-3) + \tfrac{1}{72}(\kappa_3')^2(x^4-10x^2+15)\right\}, \end{aligned} \tag{3.17}$$

where κ_3' and κ_4' denote the third and fourth *standardised* cumulants of X, respectively (cf. Section A.1.2). The derivations of the polynomials P_1 and P_2 can be found in Hall (1992, p. 45). The third polynomial P_3 is stated and derived in the following lemma. To the best of our knowledge this result seems to be new.

Lemma 3.3. *Under the assumptions of Theorem 3.1 the polynomial P_3 appearing in the expansion in (3.16) is given by*

$$\begin{aligned} P_3(x) = &-\tfrac{1}{120}\kappa_5'\left(x^4-6x^2+3\right) - \tfrac{1}{144}\kappa_4'\kappa_3'\left(x^6-15x^4+45x^2-15\right) \\ &- \tfrac{1}{1296}\left(\kappa_3'\right)^3\left(x^8-28x^6+210x^4-420x^2+105\right). \end{aligned}$$

Proof. Define

$$S_n = n^{1/2}A_S(\bar{\mathbf{W}}_n) = \frac{n^{1/2}\left(\bar{X}_n-\mu\right)}{\sigma} = n^{1/2}\bar{X}_n,$$

where we have assumed, without loss of generality, that $\mu = 0$ and $\sigma^2 = 1$. Hence,

$$\mathrm{E}\left(S_n\right) = n^{1/2}\,\mathrm{E}\left(\bar{X}_n\right) = 0 \quad \text{and} \quad \mathrm{E}\left(S_n^2\right) = n\,\mathrm{E}\left(\bar{X}_n^2\right) = 1.$$

Let $\mu_k = \mathrm{E}(X_1^k)$. From (27.2.2) of Cramér (1946) it follows that

$$\mathrm{E}\left(S_n^3\right) = n^{3/2}\,\mathrm{E}\left(\bar{X}_n^3\right) = \frac{\mu_3}{n^{1/2}} \quad \text{and} \quad \mathrm{E}\left(S_n^4\right) = n^2\,\mathrm{E}\left(\bar{X}_n^4\right) = 3 + \frac{\mu_4-3}{n},$$

and (B.17) in the appendix implies that

$$\mathrm{E}\left(S_n^5\right) = n^{5/2}\,\mathrm{E}\left(\bar{X}_n^5\right) = \frac{10\mu_3}{n^{1/2}} + \frac{\mu_5-10\mu_3}{n^{3/2}}.$$

Denoting the jth cumulant of S_n by $\kappa_{j,n}$, we have from (3.43) in Stuart and Ord (1994) that

$$\begin{aligned} \kappa_{1,n} &= \mathrm{E}\left(S_n\right) = 0, \\ \kappa_{2,n} &= \mathrm{E}\left(S_n^2\right) = 1, \\ \kappa_{3,n} &= \mathrm{E}\left(S_n^3\right) = n^{-1/2}\mu_3, \\ \kappa_{4,n} &= \mathrm{E}\left(S_n^4\right) - 3\left\{\mathrm{E}\left(S_n^2\right)\right\}^2 = n^{-1}\left(\mu_4-3\right), \\ \kappa_{5,n} &= \mathrm{E}\left(S_n^5\right) - 10\,\mathrm{E}\left(S_n^3\right)\mathrm{E}\left(S_n^2\right) = n^{-3/2}\left(\mu_5-10\mu_3\right). \end{aligned}$$

In the notation of the expansion in (3.5) we therefore have the following constants:

$$\begin{aligned}
&k_{1,1}=0, && k_{1,2}=0, && k_{1,3}=0,\\
&k_{2,1}=1, && k_{2,2}=0, && k_{3,1}=\mu_3=\kappa_3',\\
&k_{3,2}=0, && k_{4,1}=\mu_4-3=\kappa_4', && k_{5,1}=\mu_5-10\mu_3=\kappa_5'.
\end{aligned}$$

Substituting these constants in the expression for r_3 in Lemma 3.1 we obtain

$$r_3(u)=\tfrac{1}{120}\kappa_5' u^5+\tfrac{1}{144}\kappa_4'\kappa_3' u^7+\tfrac{1}{1296}(\kappa_3')^3 u^9. \tag{3.18}$$

It is well known that

$$\frac{d^k}{dx^k}\Phi(x)=(-1)^{k-1}He_{k-1}(x)\phi(x), \qquad k\geq 1, \tag{3.19}$$

where He_k denotes the kth Hermite polynomial. The ones we require are

$$\begin{aligned}
He_2(x)&=x^2-1,\\
He_4(x)&=x^4-6x^2+3,\\
He_6(x)&=x^6-15x^4+45x^2-15,\\
He_8(x)&=x^8-28x^6+210x^4-420x^2+105.
\end{aligned}$$

Now, applying relation (3.7) to the expression in (3.18) we obtain

$$\begin{aligned}
P_3(x)\phi(x)&=r_3\left(-\frac{d}{dx}\right)\Phi(x)\\
&=-\tfrac{1}{120}\kappa_5'\frac{d^5}{dx^5}\Phi(x)-\tfrac{1}{144}\kappa_4'\kappa_3'\frac{d^7}{dx^7}\Phi(x)-\tfrac{1}{1296}(\kappa_3')^3\frac{d^9}{dx^9}\Phi(x)\\
&=-\tfrac{1}{120}\kappa_5'\left(x^4-6x^2+3\right)\phi(x)-\tfrac{1}{144}\kappa_4'\kappa_3'\left(x^6-15x^4+45x^2-15\right)\phi(x)\\
&\quad-\tfrac{1}{1296}(\kappa_3')^3\left(x^8-28x^6+210x^4-420x^2+105\right)\phi(x),
\end{aligned}$$

which yields the result. □

For the Studentised smooth function A_T we may write, also by Theorem 3.1 under the assumption $\mathrm{E}\left(X_1^{12}\right)<\infty$,

$$\mathrm{P}\left(\frac{n^{1/2}\left(\bar{X}_n-\mu\right)}{\widehat{\sigma}_n}\leq x\right)=\Phi(x)+n^{-1/2}Q_1(x)\phi(x)+n^{-1}Q_2(x)\phi(x)+n^{-3/2}Q_3(x)\phi(x)+O\left(n^{-2}\right), \tag{3.20}$$

with

$$\begin{aligned}
Q_1(x)&=\tfrac{1}{6}\kappa_3'(2x^2+1),\\
Q_2(x)&=x\left(\tfrac{1}{12}\kappa_4'(x^2-3)-\tfrac{1}{18}\left(\kappa_3'\right)^2(x^4+2x^2-3)-\tfrac{1}{4}(x^2+3)\right).
\end{aligned} \tag{3.21}$$

The polynomials Q_1 and Q_2 above are derived and given in Hall (1992, p. 72ff). In our work we will also require the third polynomial Q_3, a result which has only recently been derived by Finner and Dickhaus (2010). We reproduce their result in Lemma 3.4 in a form more suited for our purposes, along with an alternative proof.

Remark 3.1. Due to the generality of Theorem 3.1, the moment requirements are rather stringent. For the expansion in (3.20) to be valid according to this theorem, we require that $\mathrm{E}\left(X_1^{12}\right) < \infty$. However, for the specific case of Student's t-statistic Hall (1987) demonstrated that the expansion in (3.20) is valid even under the less restrictive condition $\mathrm{E}\left(X_1^6\right) < \infty$, as long as the distribution of X_1 is nonsingular (which is implied by Cramér's condition).

Lemma 3.4. *Suppose the distribution of X_1 is nonsingular. If $\mathrm{E}\left(X_1^6\right) < \infty$, the polynomial Q_3 in the expansion in* (3.20) *is given by*

$$Q_3(x) = -\tfrac{1}{40}\kappa_5'\left(2x^4+8x^2+1\right) - \tfrac{1}{144}\kappa_4'\kappa_3'\left(4x^6-30x^4-90x^2-15\right)$$
$$+\tfrac{1}{1296}\left(\kappa_3'\right)^3\left(8x^8+28x^6-210x^4-525x^2-105\right) + \tfrac{1}{24}\kappa_3'\left(2x^6-3x^4-6x^2\right).$$

Proof. Define
$$T_n = n^{1/2}A_T(\bar{\mathbf{W}}_n) = \frac{n^{1/2}\left(\bar{X}_n-\mu\right)}{\widehat{\sigma}_n}.$$
Throughout the proof we will assume without loss of generality that $\mu = \mathrm{E}(X_1) = 0$ and $\sigma^2 = \mathrm{E}(X_1^2) = 1$. It is easy to see that T_n can then be expressed as

$$T_n = \frac{n^{1/2}\bar{X}_n}{\widehat{\sigma}_n} = n^{1/2}\bar{X}_n\left\{1+\frac{1}{n}\sum_{i=1}^{n}\left(X_i^2-1\right)-\bar{X}_n^2\right\}^{-1/2}.$$

By the central limit theorem, under the assumption that the X_i are i.i.d. random variables with mean 0 and variance 1, it follows that

$$\sqrt{n}\bar{X}_n \xrightarrow{d} N(0,1) \quad \text{and} \quad \sqrt{n}\left\{\frac{1}{n}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\} \xrightarrow{d} N(0,1),$$

implying that both quantities are of order $O_p(1)$. Hence,

$$\bar{X}_n = O_p(n^{-1/2}) \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^{n}\left(X_i^2-1\right) = O_p(n^{-1/2}).$$

The reader should note at this point that we will not attempt to find asymptotic expressions for the cumulants of T_n, but rather for the cumulants of a Taylor approximation of T_n. To be concise, suppose we may write $T_n = T_{nr}+O_p(n^{-r/2})$. By the arguments in Section 2.7 of Hall (1992), Edgeworth expansions of the distributions of T_n and T_{nr} will typically disagree only in terms of order $n^{-r/2}$ or smaller.

In the notation of Theorem 2.1 of Hall (1992), define T_{nr} as the leading terms of a Taylor series expansion of T_n such that $T_n = T_{nr}+O_p(n^{-r/2})$. In our case, by Taylor series expansion of the function $(1+x)^{-1/2}$ about zero, we may then write $T_n = T_{n,4}+O_p(n^{-2})$, where

$$T_{n,4} := n^{1/2}\bar{X}_n\Bigg(1-\tfrac{1}{2}n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)+\tfrac{1}{2}\bar{X}_n^2+\tfrac{3}{8}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^2$$
$$-\tfrac{3}{4}\bar{X}_n^2\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}-\tfrac{5}{16}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^3\Bigg) =: \sum_{k=1}^{6}T_{n,4}^{\{k\}}.$$

Since $\mu = 0$ it follows that $\mathrm{E}(T_{n,4}^{\{1\}}) = n^{1/2}\mathrm{E}(\bar{X}_n) = 0$. Furthermore, by (27.2.2) of Cramér (1946) we have that

$$\mathrm{E}\left(T_{n,4}^{\{3\}}\right) = \tfrac{1}{2}n^{1/2}\mathrm{E}\left(\bar{X}_n^3\right) = \tfrac{1}{2}n^{-3/2}\mu_3 = \tfrac{1}{2}n^{-3/2}\kappa_3',$$

since $\kappa_3' = \mu_3/\sigma^3 = \mu_3$ (cf. Section A.1.2). We also have by (B.19), (B.20), (B.21) and (B.24) that

$$\mathrm{E}\left(T_{n,4}^{\{2\}}\right) = -\tfrac{1}{2}n^{1/2}\mathrm{E}\left(\bar{X}_n\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}\right) = -\tfrac{1}{2}n^{-1/2}\mu_3 = -\tfrac{1}{2}n^{-1/2}\kappa_3',$$

$$\mathrm{E}\left(T_{n,4}^{\{4\}}\right) = \tfrac{3}{8}n^{1/2}\mathrm{E}\left(\bar{X}_n\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^2\right) = \tfrac{3}{8}n^{-3/2}\left(\mu_5-2\mu_3\right) = \tfrac{3}{8}n^{-3/2}\left(\kappa_5'+8\kappa_3'\right),$$

$$\begin{aligned}\mathrm{E}\left(T_{n,4}^{\{5\}}\right) = -\tfrac{3}{4}n^{1/2}\mathrm{E}\left(\bar{X}_n^3\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}\right) &= -\tfrac{9}{4}n^{-3/2}\mu_3 + O(n^{-5/2})\\ &= -\tfrac{9}{4}n^{-3/2}\kappa_3' + O(n^{-5/2}),\end{aligned}$$

$$\begin{aligned}\mathrm{E}\left(T_{n,4}^{\{6\}}\right) = -\tfrac{5}{16}n^{1/2}\mathrm{E}\left(\bar{X}_n\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^3\right) &= -\tfrac{15}{16}n^{-3/2}\mu_3\left(\mu_4-1\right) + O(n^{-5/2})\\ &= -\tfrac{15}{16}n^{-3/2}\kappa_3'\left(\kappa_4'+2\right) + O(n^{-5/2}),\end{aligned}$$

where we have made use of the fact that $\kappa_4' = \mu_4 - 3$ and $\kappa_5' = \mu_5 - 10\mu_3$. Combining these results we see that

$$\mathrm{E}\left(T_{n,4}\right) = -\tfrac{1}{2}n^{-1/2}\kappa_3' + \tfrac{1}{16}n^{-3/2}\left(6\kappa_5' - 15\kappa_4'\kappa_3' - 10\kappa_3'\right) + O(n^{-5/2}).$$

From Hall (1992, p. 73) we have that the second moment of a Taylor approximation of T_n is given by

$$\mathrm{E}\left(T_n^2\right) = 1 + n^{-1}\left(2\left(\kappa_3'\right)^2 + 3\right) + O(n^{-2}).$$

Write $T_n^3 =: T_{n,4}^3 + O_p(n^{-2})$. Then $T_{n,4}^3$ is given by

$$\begin{aligned}T_{n,4}^3 = n^{3/2}\bar{X}_n^3\Bigg(&1 - \tfrac{3}{2}n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right) + \tfrac{3}{4}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^2 - \tfrac{1}{8}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^3\\ &+ \tfrac{3}{2}\bar{X}_n^2 - \tfrac{3}{2}\bar{X}_n^2\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\} + \tfrac{9}{8}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^2 - \tfrac{9}{8}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^3\\ &- \tfrac{9}{4}\bar{X}_n^2\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\} - \tfrac{15}{16}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^3\Bigg)\\ = n^{3/2}\bar{X}_n^3\Bigg(&1 - \tfrac{3}{2}n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right) + \tfrac{15}{8}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^2 - \tfrac{35}{16}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^3\\ &+ \tfrac{3}{2}\bar{X}_n^2 - \tfrac{15}{4}\bar{X}_n^2\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}\Bigg)\\ =: \sum_{k=7}^{12} T_{n,4}^{\{k\}}&.\end{aligned}$$

From (27.2.2) and (27.2.3) of Cramér (1946) it follows that

$$\mathrm{E}\left(T_{n,4}^{\{7\}}\right) = \tfrac{1}{2}n^{3/2}\,\mathrm{E}\left(\bar{X}_n^3\right) = \tfrac{1}{2}n^{-1/2}\mu_3 = \tfrac{1}{2}n^{-1/2}\kappa_3',$$

and

$$\mathrm{E}\left(T_{n,4}^{\{11\}}\right) = \tfrac{3}{2}n^{3/2}\,\mathrm{E}\left(\bar{X}_n^5\right) = 15n^{-3/2}\mu_3 = 15n^{-3/2}\kappa_3'.$$

By (B.24), (B.25), (B.26) and (B.27),

$$\begin{aligned}\mathrm{E}\left(T_{n,4}^{\{8\}}\right) = -\tfrac{3}{2}n^{3/2}\,\mathrm{E}\left(\bar{X}_n^3\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}\right) &= -\tfrac{9}{2}n^{-1/2}\mu_3 - \tfrac{3}{2}n^{-3/2}\left(\mu_5-4\mu_3\right)\\ &= -\tfrac{9}{2}n^{-1/2}\kappa_3' - \tfrac{3}{2}n^{-3/2}\left(\kappa_5'+6\kappa_3'\right),\end{aligned}$$

$$\begin{aligned}\mathrm{E}\left(T_{n,4}^{\{9\}}\right) = \tfrac{15}{8}n^{3/2}\,\mathrm{E}\left(\bar{X}_n^3\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^2\right) &= \tfrac{15}{8}n^{-3/2}\left(3\mu_5+7\mu_4\mu_3-13\mu_3\right)+O(n^{-3/2})\\ &= \tfrac{15}{8}n^{-3/2}\left(3\kappa_5'+7\kappa_4'\kappa_3'+38\kappa_3'\right)+O(n^{-3/2}),\end{aligned}$$

$$\begin{aligned}\mathrm{E}\left(T_{n,4}^{\{10\}}\right) = -\tfrac{35}{16}n^{3/2}\,\mathrm{E}\left(\bar{X}_n^3\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^3\right) &= -\tfrac{35}{16}n^{-3/2}\left(9\mu_4\mu_3+6\mu_3^3-9\mu_3\right)\\ &= -\tfrac{35}{16}n^{-3/2}\left(9\kappa_4'\kappa_3'+6\left(\kappa_3'\right)^3+18\kappa_3'\right),\end{aligned}$$

$$\mathrm{E}\left(T_{n,4}^{\{12\}}\right) = -\tfrac{15}{4}n^{3/2}\,\mathrm{E}\left(\bar{X}_n^5\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}\right) = -\tfrac{225}{4}n^{-3/2}\mu_3 = -\tfrac{225}{4}n^{-3/2}\kappa_3'.$$

Collecting terms we have that

$$\mathrm{E}\left(T_{n,4}^3\right) = -\tfrac{7}{2}n^{-1/2}\kappa_3' + \tfrac{1}{16}n^{-3/2}\left(66\kappa_5'-105\kappa_4'\kappa_3'-210(\kappa_3')^3-294\kappa_3'\right)+O(n^{-5/2}).$$

Again, Hall (1992, p. 73) shows that the fourth moment of a Taylor approximation of T_n is given by

$$\mathrm{E}\left(T_n^4\right) = 3+n^{-1}\left(28\left(\kappa_3'\right)^2-2\kappa_4'+24\right)+O(n^{-2}).$$

Lastly, write $T_n^5 =: T_{n,4}^5 + O_p(n^{-2})$, where

$$\begin{aligned}T_{n,4}^5 = n^{5/2}\bar{X}_n^5\Bigg(&1-\tfrac{5}{2}n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)+\tfrac{10}{4}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^2-\tfrac{10}{8}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^3\\ &+\tfrac{5}{2}\bar{X}_n^2-5\bar{X}_n^2\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}+\tfrac{15}{8}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^2-\tfrac{15}{4}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^3\\ &-\tfrac{15}{4}\bar{X}_n^2\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}-\tfrac{25}{16}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^3\Bigg)\\ = n^{5/2}\bar{X}_n^5\Bigg(&1-\tfrac{5}{2}n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)+\tfrac{35}{8}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^2-\tfrac{105}{16}\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^3\\ &+\tfrac{5}{2}\bar{X}_n^2-\tfrac{35}{4}\bar{X}_n^2\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}\Bigg)\\ =: \sum_{k=13}^{18}T_{n,4}^{\{k\}}.&\end{aligned}$$

From (B.17), (B.18), (B.27), (B.28), (B.29) and (B.30) it follows that

$$\mathrm{E}\left(T_{n,4}^{\{13\}}\right) = n^{5/2}\,\mathrm{E}\left(\bar{X}_n^5\right) = 10n^{-1/2}\mu_3 + n^{-3/2}\left(\mu_5 - 10\mu_3\right)$$
$$= 10n^{-1/2}\kappa_3' + n^{-3/2}\kappa_5',$$

$$\mathrm{E}\left(T_{n,4}^{\{14\}}\right) = -\tfrac{5}{2}n^{5/2}\,\mathrm{E}\left(\bar{X}_n^5\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}\right)$$
$$= -\tfrac{75}{2}n^{-1/2}\mu_3 - \tfrac{5}{2}n^{-3/2}\left\{10\mu_5 + 15\mu_4\mu_3 - 65\mu_3\right\} + O(n^{-5/2})$$
$$= -\tfrac{75}{2}n^{-1/2}\kappa_3' - \tfrac{5}{2}n^{-3/2}\left\{10\kappa_5' + 15\kappa_4'\kappa_3' + 80\kappa_3'\right\} + O(n^{-5/2}),$$

$$\mathrm{E}\left(T_{n,4}^{\{15\}}\right) = \tfrac{35}{8}n^{5/2}\,\mathrm{E}\left(\bar{X}_n^5\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^2\right) = \tfrac{35}{8}n^{-3/2}\left\{15\mu_5 + 70\mu_4\mu_3 + 20\mu_3^3 - 100\mu_3\right\}$$
$$= \tfrac{35}{8}n^{-3/2}\left\{15\kappa_5' + 70\kappa_4'\kappa_3' + 20(\kappa_3')^3 + 260\kappa_3'\right\},$$

$$\mathrm{E}\left(T_{n,4}^{\{16\}}\right) = -\tfrac{105}{16}n^{5/2}\,\mathrm{E}\left(\bar{X}_n^5\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}^3\right) = -\tfrac{105}{16}n^{-3/2}\left\{45\mu_4\mu_3 + 60\mu_3^3 - 45\mu_3\right\}$$
$$= -\tfrac{105}{16}n^{-3/2}\left\{45\kappa_4'\kappa_3' + 60(\kappa_3')^3 + 90\kappa_3'\right\},$$

$$\mathrm{E}\left(T_{n,4}^{\{17\}}\right) = \tfrac{5}{2}n^{5/2}\,\mathrm{E}\left(\bar{X}_n^7\right) = \tfrac{525}{2}n^{-1/2}\mu_3 + O(n^{-5/2})$$
$$= \tfrac{525}{2}n^{-1/2}\kappa_3' + O(n^{-5/2}),$$

$$\mathrm{E}\left(T_{n,4}^{\{18\}}\right) = -\tfrac{35}{4}n^{5/2}\,\mathrm{E}\left(\bar{X}_n^7\left\{n^{-1}\sum_{i=1}^{n}\left(X_i^2-1\right)\right\}\right) = -\tfrac{3675}{4}n^{-3/2}\mu_3 = -\tfrac{3675}{4}n^{-3/2}\kappa_3'.$$

Therefore,

$$\mathrm{E}\left(T_{n,4}^5\right) = -\tfrac{55}{2}n^{-1/2}\kappa_3' + \tfrac{1}{16}n^{-3/2}\left(666\kappa_5' - 425\kappa_4'\kappa_3' - 4900(\kappa_3')^3 - 4950\kappa_3'\right) + O(n^{-5/2}).$$

Now, denoting the jth cumulant of $T_{n,4}$ by $\kappa_{j,n,4}$, we have from (3.42) in Stuart and Ord (1994) and the results above that

$$\kappa_{1,n,4} = \mathrm{E}\left(T_{n,4}\right)$$
$$= -\tfrac{1}{2}n^{-1/2}\kappa_3' + \tfrac{1}{16}n^{-3/2}\left(6\kappa_5' - 15\kappa_4'\kappa_3' - 10\kappa_3'\right) + O(n^{-5/2}),$$
$$\kappa_{2,n,4} = \mathrm{E}\left(T_{n,4}^2\right) - \left\{\mathrm{E}\left(T_{n,4}\right)\right\}^2$$
$$= 1 + \tfrac{1}{4}n^{-1}\left(7(\kappa_3')^2 + 12\right) + O(n^{-2}),$$
$$\kappa_{3,n,4} = \mathrm{E}\left(T_{n,4}^3\right) - 3\,\mathrm{E}\left(T_{n,4}^2\right)\mathrm{E}\left(T_{n,4}\right) + 2\left\{\mathrm{E}\left(T_{n,4}\right)\right\}^3$$
$$= -2n^{-1/2}\kappa_3' + \tfrac{1}{8}n^{-3/2}\left(24\kappa_5' - 30\kappa_4'\kappa_3' - 83(\kappa_3')^3 - 96\kappa_3'\right) + O(n^{-5/2}),$$
$$\kappa_{4,n,4} = \mathrm{E}\left(T_{n,4}^4\right) - 4\,\mathrm{E}\left(T_{n,4}^3\right)\mathrm{E}\left(T_{n,4}\right) - 3\left\{\mathrm{E}\left(T_{n,4}^2\right)\right\}^2 + 12\,\mathrm{E}\left(T_{n,4}^2\right)\left\{\mathrm{E}\left(T_{n,4}\right)\right\}^2 - 6\left\{\mathrm{E}\left(T_{n,4}\right)\right\}^4$$
$$= 2n\left(6(\kappa_3')^2 - \kappa_4' + 3\right) + O(n^{-2}),$$
$$\kappa_{5,n,4} = \mathrm{E}\left(T_{n,4}^5\right) - 5\,\mathrm{E}\left(T_{n,4}^4\right)\mathrm{E}\left(T_{n,4}\right) - 10\,\mathrm{E}\left(T_{n,4}^3\right)\mathrm{E}\left(T_{n,4}^2\right) + 20\,\mathrm{E}\left(T_{n,4}^3\right)\left\{\mathrm{E}\left(T_{n,4}\right)\right\}^2$$
$$+ 30\left\{\mathrm{E}\left(T_{n,4}^2\right)\right\}^2\mathrm{E}\left(T_{n,4}\right) - 60\,\mathrm{E}\left(T_{n,4}^2\right)\left\{\mathrm{E}\left(T_{n,4}\right)\right\}^3 + 24\left\{\mathrm{E}\left(T_{n,4}\right)\right\}^5$$
$$= n^{-3/2}\left(6\kappa_5' + 20\kappa_4'\kappa_3' - 105(\kappa_3')^3 - 60\kappa_3'\right) + O(n^{-5/2}).$$

which are expansions (of the type seen in (3.5)) of the cumulants of $T_{n,4}$. In the notation of (3.5) we therefore have the following constants:

$$\begin{aligned}
k_{1,1} &= 0,\\
k_{1,2} &= -\tfrac{1}{2}\kappa'_3,\\
k_{1,3} &= \tfrac{1}{16}\left(6\kappa'_5 - 15\kappa'_4\kappa'_3 - 10\kappa'_3\right),\\
k_{2,1} &= 1,\\
k_{2,2} &= \tfrac{1}{4}\left(7(\kappa'_3)^2 + 12\right),\\
k_{3,1} &= -2\kappa'_3,\\
k_{3,2} &= \tfrac{1}{8}\left(24\kappa'_5 - 30\kappa'_4\kappa'_3 - 83(\kappa'_3)^3 - 96\kappa'_3\right),\\
k_{4,1} &= 12(\kappa'_3)^2 - 2\kappa'_4 + 6,\\
k_{5,1} &= 6\kappa'_5 + 20\kappa'_4\kappa'_3 - 105(\kappa'_3)^3 - 60\kappa'_3.
\end{aligned} \tag{3.22}$$

The values of $k_{1,2}$, $k_{2,2}$, $k_{3,1}$ and $k_{4,1}$ are confirmed by those given by Hall (1992) on p. 73.

Notice in Theorem 2.1 of Hall that, whenever $j + l \le r + 2$, the $k_{j,l}$ do not depend on the choice of r (the order of accuracy of the Taylor approximation T_{nr}). As we chose $r = 4$ and since it holds for all $k_{j,l}$ above that $j + l \le 6$, the $k_{j,l}$ above do not depend on terms in T_n of order $O_p(n^{-2})$ or smaller. We may thus argue that these constants agree with those in similar expansions of cumulants of T_n.

Now, applying relation (3.7) to the expression in Lemma 3.1 and making use of (3.19) we obtain

$$\begin{aligned}
Q_3(x)\phi(x) &= r_3\left(-\frac{d}{dx}\right)\Phi(x)\\
&= k_{1,3}\left(-\frac{d}{dx}\Phi(x)\right) + \tfrac{1}{6}\left(k_{3,2} + 3k_{2,2}k_{1,2} + k_{1,2}^3\right)\left(-\frac{d^3}{dx^3}\Phi(x)\right)\\
&\quad + \tfrac{1}{120}\left(k_{5,1} + 5k_{4,1}k_{1,2} + 10k_{3,1}k_{2,2} + 10k_{3,1}k_{1,2}^2\right)\left(-\frac{d^5}{dx^5}\Phi(x)\right)\\
&\quad + \tfrac{1}{144}k_{3,1}\left(k_{4,1} + 2k_{3,1}k_{1,2}\right)\left(-\frac{d^7}{dx^7}\Phi(x)\right) + \tfrac{1}{1296}k_{3,1}^3\left(-\frac{d^9}{dx^9}\Phi(x)\right)\\
&= -k_{1,3}\phi(x) - \tfrac{1}{6}\left(k_{3,2} + 3k_{2,2}k_{1,2} + k_{1,2}^3\right)(x^2 - 1)\phi(x)\\
&\quad - \tfrac{1}{120}\left(k_{5,1} + 5k_{4,1}k_{1,2} + 10k_{3,1}k_{2,2} + 10k_{3,1}k_{1,2}^2\right)\left(x^4 - 6x^2 + 3\right)\phi(x)\\
&\quad - \tfrac{1}{144}k_{3,1}\left(k_{4,1} + 2k_{3,1}k_{1,2}\right)\left(x^6 - 15x^4 + 45x^2 - 15\right)\phi(x)\\
&\quad - \tfrac{1}{1296}k_{3,1}^3\left(x^8 - 28x^6 + 210x^4 - 420x^2 + 105\right)\phi(x).
\end{aligned}$$

Finally, by substituting the $k_{j,l}$ in the last expression by their respective expressions given in (3.22), we obtain

$$\begin{aligned}
Q_3(x)\phi(x) = &-\tfrac{1}{40}\kappa'_5\left(2x^4 + 8x^2 + 1\right)\phi(x) - \tfrac{1}{144}\kappa'_4\kappa'_3\left(4x^6 - 30x^4 - 90x^2 - 15\right)\phi(x)\\
&+ \tfrac{1}{1296}\left(\kappa'_3\right)^3\left(8x^8 + 28x^6 - 210x^4 - 525x^2 - 105\right)\phi(x) + \tfrac{1}{24}\kappa'_3\left(2x^6 - 3x^4 - 6x^2\right)\phi(x),
\end{aligned}$$

which is the required result. □

3.4.2 The variance

Let $X_1, X_2, \ldots, X_n$ be independent, identically distributed random variables from a univariate distribution with mean μ and variance σ^2. Suppose we are interested in estimating $\theta = \sigma^2$, the population variance. In the above setting, choose $d = 4$ and set $\mathbf{W}_i = [X_i, X_i^2, X_i^3, X_i^4]'$. Then

$$\boldsymbol{\nu} = \mathrm{E}(\mathbf{W}_1) = \begin{bmatrix} \mathrm{E}(X_1) \\ \mathrm{E}(X_1^2) \\ \mathrm{E}(X_1^3) \\ \mathrm{E}(X_1^4) \end{bmatrix} = \begin{bmatrix} \mu \\ \mu^2 + \sigma^2 \\ \mathrm{E}(X_1^3) \\ \mathrm{E}(X_1^4) \end{bmatrix}$$

and

$$\bar{\mathbf{W}}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbf{W}_i = \begin{bmatrix} n^{-1}\sum_{i=1}^n X_i \\ n^{-1}\sum_{i=1}^n X_i^2 \\ n^{-1}\sum_{i=1}^n X_i^3 \\ n^{-1}\sum_{i=1}^n X_i^4 \end{bmatrix} = \begin{bmatrix} \bar{X}_n \\ \bar{X}_n^2 + \hat{\sigma}^2 \\ n^{-1}\sum_{i=1}^n X_i^3 \\ n^{-1}\sum_{i=1}^n X_i^4 \end{bmatrix}.$$

Set

$$g(\mathbf{w}) = w_2 - w_1^2,$$
$$h^2(\mathbf{w}) = w_4 - 4w_1w_3 + 6w_1^2w_2 - 3w_1^4 - \left(w_2 - w_1^2\right)^2,$$

for any $\mathbf{w} = (w_1, w_2, w_3, w_4)' \in \mathbb{R}^4$. This ensures that

$$\theta = g(\boldsymbol{\nu}) = \sigma^2,$$
$$\beta^2 = h^2(\boldsymbol{\nu}) = \mathrm{E}\left(X_1^4\right) - 4\mu\mathrm{E}\left(X_1^3\right) + 6\mu^2\mathrm{E}\left(X_1^2\right) - 3\mu^4 - \left(\mathrm{E}\left(X_1^2\right) - \mu^2\right)^2 = \mathrm{E}\left(X_1 - \mu\right)^4 - \sigma^4.$$

Estimators for the above quantities are

$$\hat{\theta}_n = g(\bar{\mathbf{W}}_n) = \frac{1}{n}\sum_{i=1}^n X_i^2 - \bar{X}_n^2 = \hat{\sigma}_n^2,$$
$$\hat{\beta}_n^2 = h^2(\bar{\mathbf{W}}_n) = \frac{1}{n}\sum_{i=1}^n \left(X_i - \bar{X}_n\right)^4 - \hat{\sigma}_n^4.$$

This means that, when our target is the population variance, the two smooth functions of interest are

$$A_S(\bar{\mathbf{W}}_n) = \frac{g(\bar{\mathbf{W}}_n) - g(\boldsymbol{\nu})}{h(\boldsymbol{\nu})} = \frac{\hat{\sigma}_n^2 - \sigma^2}{\sqrt{\mathrm{E}\left(X_1 - \mu\right)^4 - \sigma^4}}$$

and

$$A_T(\bar{\mathbf{W}}_n) = \frac{g(\bar{\mathbf{W}}_n) - g(\boldsymbol{\nu})}{h(\bar{\mathbf{W}}_n)} = \frac{\hat{\sigma}_n^2 - \sigma^2}{\sqrt{n^{-1}\sum_{i=1}^n \left(X_i - \bar{X}_n\right)^4 - \hat{\sigma}_n^4}}.$$

Note that we have made use of the fact that the asymptotic variance of $\hat{\sigma}_n^2$ is given by

$$\frac{\mathrm{E}(X_1 - \mu)^4 - \sigma^4}{n},$$

which is correct to the order n^{-2} (cf. Cramér, 1946, p. 348). Consequently, the asymptotic variance of $n^{1/2}A_S(\bar{\mathbf{W}}_n)$ is 1. Also, by the results of Chapter 28 of Cramér (1946), we know that $n^{1/2}A_T(\bar{\mathbf{W}}_n)$ has a limiting standard normal distribution.

One-term Edgeworth expansions for the distributions of $n^{1/2}A_S(\bar{\mathbf{W}}_n)$ and $n^{1/2}A_T(\bar{\mathbf{W}}_n)$ can be found in Hall (1992, p. 76).

3.4.3 Pearson's correlation coefficient

Let $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ be independent, identically distributed random variables from any bivariate distribution with mean vector $\left(\mu_X,\mu_Y\right)$ and variance vector $\left(\sigma_X^2,\sigma_Y^2\right)$. define the joint *central (population) moments* of X_1 and Y_1 as

$$\mu_{ik} = \mathrm{E}\left(\left(X_1-\mu_X\right)^i\left(Y_1-\mu_Y\right)^k\right).$$

Note that $\mu_{10}=0$, $\mu_{01}=0$, $\mu_{20}=\mathrm{Var}(X_1)$, $\mu_{02}=\mathrm{Var}(Y_1)$ and $\mu_{11}=\mathrm{Cov}(X_1,Y_1)$. Also, define the joint *central sample moments* of X_1 and Y_1 as

$$\widehat{\mu}_{ik} = \frac{1}{n}\sum_{j=1}^{n}\left(X_j-\bar{X}\right)^i\left(Y_j-\bar{Y}\right)^k,$$

where $\bar{X}=n^{-1}\sum_{j=1}^{n}X_j$ and $\bar{Y}=n^{-1}\sum_{j=1}^{n}Y_j$.

Suppose we are interested in estimating Pearson's correlation coefficient, i.e., the parameter of interest is

$$\rho = \frac{\sigma_{XY}}{\sigma_X\sigma_Y} = \frac{\mu_{11}}{\sqrt{\mu_{20}\mu_{02}}},$$

where $\sigma_{XY}=\mathrm{Cov}(X_1,Y_1)$. The usual estimator for for ρ is the sample correlation coefficient

$$\widehat{\rho}_n = \frac{\widehat{\mu}_{11}}{\sqrt{\widehat{\mu}_{20}\widehat{\mu}_{02}}}.$$

Now, according to Cramér (1946, p. 359), *if no particular distribution is imposed on* (X_1,Y_1), an approximation of the variance of $\widehat{\rho}$ is given by

$$\mathrm{Var}\left(\widehat{\rho}_n\right) \approx \frac{\rho^2}{4n}\left(\frac{\mu_{40}}{\mu_{20}^2}+\frac{\mu_{04}}{\mu_{02}^2}+\frac{2\mu_{22}}{\mu_{20}\mu_{02}}+\frac{4\mu_{22}}{\mu_{11}^2}-\frac{4\mu_{31}}{\mu_{11}\mu_{20}}-\frac{4\mu_{13}}{\mu_{11}\mu_{02}}\right).$$

This approximation is correct to the order $n^{-3/2}$, by the theorem of Section 27.7 of Cramér (1946).

In the smooth function model setting, choose $d=8$ and set

$$\mathbf{W}_i = \begin{bmatrix} (X_i-\mu_X)^2 \\ (Y_i-\mu_Y)^2 \\ (X_i-\mu_X)^4 \\ (Y_i-\mu_Y)^4 \\ (X_i-\mu_X)(Y_i-\mu_Y) \\ (X_i-\mu_X)^2(Y_i-\mu_Y)^2 \\ (X_i-\mu_X)^3(Y_i-\mu_Y) \\ (X_i-\mu_X)(Y_i-\mu_Y)^3 \end{bmatrix}.$$

Assuming all the concerned moments exist, we have

$$\boldsymbol{\nu} = \mathrm{E}(\mathbf{W}_1) = \begin{bmatrix}\mu_{20}, & \mu_{02}, & \mu_{40}, & \mu_{04}, & \mu_{11}, & \mu_{22}, & \mu_{31}, & \mu_{13}\end{bmatrix}'$$

and

$$\bar{\mathbf{W}}_n = \frac{1}{n}\sum_{i=1}^{n}\mathbf{W}_i = \begin{bmatrix}\widehat{\mu}_{20}, & \widehat{\mu}_{02}, & \widehat{\mu}_{40}, & \widehat{\mu}_{04}, & \widehat{\mu}_{11}, & \widehat{\mu}_{22}, & \widehat{\mu}_{31}, & \widehat{\mu}_{13}\end{bmatrix}'.$$

Choosing

$$\begin{aligned}
g(\mathbf{w}) &= \frac{w_5}{\sqrt{w_1 w_2}},\\
h^2(\mathbf{w}) &= \frac{g(\mathbf{w})^2}{4}\left(\frac{w_3}{w_1^2} + \frac{w_4}{w_2^2} + \frac{2w_6}{w_1 w_2} + \frac{4w_6}{w_5^2} - \frac{4w_7}{w_5 w_1} - \frac{4w_8}{w_5 w_2}\right)
\end{aligned}$$

for $\mathbf{w} = (w_1, \ldots, w_8)' \in \mathbb{R}^8$, will ensure that

$$\begin{aligned}
\theta = g(\boldsymbol{\nu}) &= \frac{\mu_{11}}{\sqrt{\mu_{20}\mu_{02}}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \rho,\\
\beta^2 = h^2(\boldsymbol{\nu}) &= \frac{\rho^2}{4}\left(\frac{\mu_{40}}{\mu_{20}^2} + \frac{\mu_{04}}{\mu_{02}^2} + \frac{2\mu_{22}}{\mu_{20}\mu_{02}} + \frac{4\mu_{22}}{\mu_{11}^2} - \frac{4\mu_{31}}{\mu_{11}\mu_{20}} - \frac{4\mu_{13}}{\mu_{11}\mu_{02}}\right).
\end{aligned} \tag{3.23}$$

Estimators for the above parameters are given by

$$\begin{aligned}
\widehat{\theta}_n = g(\bar{\mathbf{W}}_n) &= \frac{\widehat{\mu}_{11}}{\sqrt{\widehat{\mu}_{20}\widehat{\mu}_{02}}} = \widehat{\rho}_n,\\
\widehat{\beta}^2 = h^2(\bar{\mathbf{W}}_n) &= \frac{\widehat{\rho}_n^2}{4}\left(\frac{\widehat{\mu}_{40}}{\widehat{\mu}_{20}^2} + \frac{\widehat{\mu}_{04}}{\widehat{\mu}_{02}^2} + \frac{2\widehat{\mu}_{22}}{\widehat{\mu}_{20}\widehat{\mu}_{02}} + \frac{4\widehat{\mu}_{22}}{\widehat{\mu}_{11}^2} - \frac{4\widehat{\mu}_{31}}{\widehat{\mu}_{11}\widehat{\mu}_{20}} - \frac{4\widehat{\mu}_{13}}{\widehat{\mu}_{11}\widehat{\mu}_{02}}\right).
\end{aligned}$$

Note again that the asymptotic variance of $\widehat{\theta} = \widehat{\rho}$ is given by $n^{-1}\beta^2$, with β^2 as in (3.23). The two smooth functions of interest are

$$A_S(\bar{\mathbf{W}}_n) = \frac{g(\bar{\mathbf{W}}_n) - g(\boldsymbol{\nu})}{h(\boldsymbol{\nu})} = \frac{2(\widehat{\rho}_n - \rho)}{\rho}\left(\frac{\mu_{40}}{\mu_{20}^2} + \frac{\mu_{04}}{\mu_{02}^2} + \frac{2\mu_{22}}{\mu_{20}\mu_{02}} + \frac{4\mu_{22}}{\mu_{11}^2} - \frac{4\mu_{31}}{\mu_{11}\mu_{20}} - \frac{4\mu_{13}}{\mu_{11}\mu_{02}}\right)^{-1/2}$$

and

$$A_T(\bar{\mathbf{W}}_n) = \frac{g(\bar{\mathbf{W}}_n) - g(\boldsymbol{\nu})}{h(\bar{\mathbf{W}}_n)} = \frac{2(\widehat{\rho}_n - \rho)}{\widehat{\rho}_n}\left(\frac{\widehat{\mu}_{40}}{\widehat{\mu}_{20}^2} + \frac{\widehat{\mu}_{04}}{\widehat{\mu}_{02}^2} + \frac{2\widehat{\mu}_{22}}{\widehat{\mu}_{20}\widehat{\mu}_{02}} + \frac{4\widehat{\mu}_{22}}{\widehat{\mu}_{11}^2} - \frac{4\widehat{\mu}_{31}}{\widehat{\mu}_{11}\widehat{\mu}_{20}} - \frac{4\widehat{\mu}_{13}}{\widehat{\mu}_{11}\widehat{\mu}_{02}}\right)^{-1/2}.$$

This illustrates that the correlation coefficient may be expressed in terms of the smooth function model of Bhattacharya and Ghosh (1978). Of course, we shall not attempt the daunting task of deriving Edgeworth expansions for $n^{1/2}A_S(\bar{\mathbf{W}}_n)$ and $n^{1/2}A_T(\bar{\mathbf{W}}_n)$. An Edgeworth expansion for the distribution $n^{1/2}A_S(\bar{\mathbf{W}}_n)$ with terms up to order n^{-1} has been derived in the last decade by Ogasawara (2006). As the expression is rather involved and not of direct interest to us, we do not reproduce it here.

3.5 Edgeworth expansion in the linear regression setup

Suppose we observe pairs $\mathscr{X}_n = \{(x_1, Y_1), \ldots, (x_n, Y_n)\}$ generated by the simple linear regression model

$$Y_i = c' + x_i d + \varepsilon_i,$$

where c' and d are unknown, nonrandom constants and $\{\varepsilon_1, \ldots, \varepsilon_n\}$ is a sequence of iid random variables from an unknown distribution with zero mean and constant variance $0 < \sigma^2 < \infty$. Throughout we assume that the x_i are fixed. Setting $c = c' + d\bar{x}_n$, with $\bar{x}_n = n^{-1}\sum_{i=1}^n x_i$, we may rewrite the above model as

$$Y_i = c + (x_i - \bar{x}_n)d + \varepsilon_i.$$

The least-squares estimators for d and c are given by

$$\widehat{d}_n = \frac{1}{n\sigma_{x,n}^2}\sum_{i=1}^n (x_i - \bar{x}_n)Y_i \quad \text{and} \quad \widehat{c}_n = \bar{Y}_n = \frac{1}{n}\sum_{i=1}^n Y_i,$$

where $\sigma_{x,n}^2 = n^{-1}\sum_{i=1}^n (x_i - \bar{x}_n)^2 > 0$. For convenience $\widehat{d}_n$ may be rewritten in terms of the error terms ε_i as

$$\widehat{d}_n = d + \frac{1}{n\sigma_{x,n}^2}\sum_{i=1}^n (x_i - \bar{x}_n)\varepsilon_i.$$

The residuals resulting from this model fit are given by

$$e_i := Y_i - \bar{Y}_n - (x_i - \bar{x}_n)\widehat{d}_n = (\varepsilon_i - \bar{\varepsilon}) - (x_i - \bar{x}_n)(\widehat{d}_n - d),$$

where $\bar{\varepsilon} = \sum_{i=1}^n \varepsilon_i$. A natural estimator for σ^2 is then the mean squared residuals, viz.

$$\begin{aligned}\widehat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^n e_i^2 &= \frac{1}{n}\sum_{i=1}^n \left((\varepsilon_i - \bar{\varepsilon}) - (x_i - \bar{x}_n)(\widehat{d}_n - d)\right)^2 \\ &= \frac{1}{n}\sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^2 - \sigma_{x,n}^2\left(\widehat{d}_n - d\right)^2.\end{aligned}$$

Also, define

$$\begin{aligned}\gamma_{x,n} &= \frac{1}{n\sigma_{x,n}^3}\sum_{i=1}^n (x_i - \bar{x}_n)^3, \\ \kappa_{x,n} &= \frac{1}{n\sigma_{x,n}^4}\sum_{i=1}^n (x_i - \bar{x}_n)^4 - 3, \\ \tau_{x,n} &= \frac{1}{n\sigma_{x,n}^5}\sum_{i=1}^n (x_i - \bar{x}_n)^5 - 10\gamma_{x,n}.\end{aligned}$$

It is clear that $\widehat{d}_n$ has expectation d and variance $\sigma^2/(n\sigma_{x,n}^2)$. Hence, the objective is to obtain a three-term Edgeworth expansion of the distribution of the Studentised statistic

$$T_n = \frac{n^{1/2}(\widehat{d}_n - d)\sigma_{x,n}}{\widehat{\sigma}_n} = \frac{n^{1/2}(\widehat{d}_n - d)\sigma_{x,n}}{\sqrt{\frac{1}{n}\sum_{i=1}^n \varepsilon_i^2 - \bar{\varepsilon}^2 - \left(\widehat{d}_n - d\right)^2\sigma_{x,n}^2}},$$

which is asymptotically pivotal. In the spirit of Bhattacharya and Ghosh (1978), we may express T_n as a smooth function of means. Define

$$\mathbf{W}_i = \begin{bmatrix} \varepsilon_i \\ (x_i - \bar{x}_n)\varepsilon_i / \sigma_{x,n} \\ \varepsilon_i^2 \end{bmatrix}, \qquad i = 1, \ldots, n,$$

so that $\boldsymbol{\nu} = \mathrm{E}(\mathbf{W}_i) = (0, 0, \sigma^2)'$ and

$$\bar{\mathbf{W}}_n = \frac{1}{n} \sum_{i=1}^{n} \mathbf{W}_i = \begin{bmatrix} \bar{\varepsilon} \\ \left(\widehat{d}_n - d\right) \sigma_{x,n} \\ \frac{1}{n} \sum_{i=1}^{n} \varepsilon_i^2 \end{bmatrix}.$$

For any $\mathbf{w} = (w_1, w_2, w_3)' \in \mathbb{R}^3$, choosing

$$g(\mathbf{w}) = w_2 \qquad \text{and} \qquad h^2(\mathbf{w}) = w_3 - w_1^2 - w_2^2$$

yields

$$\theta = g(\boldsymbol{\nu}) = 0 \qquad \text{and} \qquad \beta^2 = h^2(\boldsymbol{\nu}) = \sigma^2,$$

with corresponding estimators

$$\begin{aligned} \widehat{\theta}_n &= g(\bar{\mathbf{W}}_n) = \left(\widehat{d}_n - d\right) \sigma_{x,n}, \\ \widehat{\beta}_n^2 &= h^2(\bar{\mathbf{W}}_n) = \widehat{\sigma}_n^2. \end{aligned}$$

As before, we may now write

$$T_n = n^{1/2} A_T(\bar{\mathbf{W}}_n) = \frac{n^{1/2} \left(g(\bar{\mathbf{W}}_n) - g(\boldsymbol{\nu})\right)}{h(\bar{\mathbf{W}}_n)},$$

so that T_n is expressed as a smooth function of means of independent variables.

It is now tempting to use Theorem 3.1 to obtain an Edgeworth expansion for T_n. However, the results of Theorem 3.1 is based on the assumption that the $\mathbf{W}_i$ are independent and *identically* distributed. Clearly, although they are mutually independent, the $\mathbf{W}_i$ depend on the design points $\{x_i\}$ and are therefore not identically distributed.

To address this issue, Hall (1992, Section 5.4) demonstrates that an Edgeworth expansion of the distribution of T_n may indeed be obtained under some additional assumptions. Assume that $\mathrm{E}(|\varepsilon_1|^{10}) < \infty$ and that ε_1 satisfies Cramér's condition. Suppose $\{x_i\}$ is a sequence of independent realisations of a nondegenerate random variable X with $\mathrm{E}(|X|^5) < \infty$. Then, for a class of sequences $\{x_i\}$ arising with probability one,

$$\begin{aligned} &\mathrm{P}\left(\frac{n^{1/2}(\widehat{d}_n - d)\sigma_{x,n}}{\widehat{\sigma}_n} \leq u \right) \\ &= \Phi(u) + n^{-1/2} Q_{1,n}(u)\phi(u) + n^{-1} Q_{2,n}(u)\phi(u) + n^{-3/2} Q_{3,n}(u)\phi(u) + O(n^{-2}), \end{aligned} \tag{3.24}$$

uniformly in $u \in \mathbb{R}$. The first two polynomials appearing in (3.24) are given by

$$\begin{aligned} Q_{1,n}(u) &= -\tfrac{1}{6} \kappa_3' \gamma_{x,n} He_2(u), \\ Q_{2,n}(u) &= -\tfrac{1}{24} \kappa_4' \kappa_{x,n} He_3(u) - \tfrac{1}{72} (\kappa_3')^2 \gamma_{x,n}^2 He_5(u) - \tfrac{1}{4}(u^2 + 5)u, \end{aligned} \tag{3.25}$$

with κ'_j denoting the jth cumulant of ε_1/σ and $He_j(u)$ the jth Hermite polynomial. The polynomial $Q_{3,n}$, which apparently does not appear in the existing literature, is stated and derived in the following lemma.

Lemma 3.5. *The polynomial $Q_{3,n}$ in (3.24) is given by*

$$Q_{3,n}(u) = -\tfrac{1}{120}\kappa'_5\left\{\tau_{x,n}He_4(u) - 30\gamma_{x,n}He_2(u)\right\} - \tfrac{1}{144}\kappa'_4\kappa'_3\left\{\kappa_{x,n}\gamma_{x,n}He_6(u) + 45\gamma_{x,n}He_2(u)\right\}$$
$$- \tfrac{1}{1296}(\kappa'_3)^3\gamma^3_{x,n}He_8(u) - \tfrac{1}{24}\kappa'_3\gamma_{x,n}(u^2-1)u^4,$$

where κ'_j denotes the jth cumulant of ε_1/σ and $He_j(u)$ the jth Hermite polynomial.

Proof. We follow the method described in Hall (1992, Section 2.4) and assume throughout, without loss of generality, that $\sigma^2 = 1$. Note that one may write

$$T_n = \frac{n^{1/2}(\widehat{d}_n - d)\sigma_{x,n}}{\widehat{\sigma}_n} = \frac{n^{1/2}(\widehat{d}_n - d)\sigma_{x,n}}{\sqrt{1 + \frac{1}{n}\sum_{i=1}^n(\varepsilon_i^2 - 1) - \bar{\varepsilon}^2 - \left(\widehat{d}_n - d\right)^2\sigma^2_{x,n}}}.$$

Note that by Taylor expansion about $(0,0)$ one has

$$\frac{x}{\sqrt{1+y-x^2}} = x - \tfrac{1}{2}xy + \tfrac{1}{2}x^3 + \tfrac{3}{8}xy^2 - \tfrac{3}{4}x^3y - \tfrac{5}{16}xy^3 + \cdots.$$

Therefore, since $\widehat{d}_n - d = O_p(n^{-1/2})$ and $n^{-1}\sum_{i=1}^n(\varepsilon_i^2-1) - \bar{\varepsilon}^2 = O_p(n^{-1/2})$,

$$\begin{aligned}
T_n &= n^{1/2}\sigma_{x,n}\Bigg\{(\widehat{d}_n - d) - \tfrac{1}{2}(\widehat{d}_n - d)\Big\{\frac{1}{n}\sum_{i=1}^n(\varepsilon_i^2-1) - \bar{\varepsilon}^2\Big\} + \tfrac{3}{8}(\widehat{d}_n - d)\Big\{\frac{1}{n}\sum_{i=1}^n(\varepsilon_i^2-1) - \bar{\varepsilon}^2\Big\}^2 \\
&\quad + \tfrac{1}{2}\sigma^2_{x,n}(\widehat{d}_n - d)^3 - \tfrac{3}{4}\sigma^2_{x,n}(\widehat{d}_n - d)^3\Big\{\frac{1}{n}\sum_{i=1}^n(\varepsilon_i^2-1) - \bar{\varepsilon}^2\Big\} - \tfrac{5}{16}(\widehat{d}_n - d)\Big\{\frac{1}{n}\sum_{i=1}^n(\varepsilon_i^2-1) - \bar{\varepsilon}^2\Big\}^3\Bigg\} \\
&\quad + O_p(n^{-2}) \\
&= n^{1/2}\sigma_{x,n}\Bigg\{(\widehat{d}_n - d) - \tfrac{1}{2}(\widehat{d}_n - d)\Big\{\frac{1}{n}\sum_{i=1}^n(\varepsilon_i^2-1)\Big\} + \tfrac{1}{2}\bar{\varepsilon}^2(\widehat{d}_n - d) + \tfrac{3}{8}(\widehat{d}_n - d)\Big\{\frac{1}{n}\sum_{i=1}^n(\varepsilon_i^2-1)\Big\}^2 \\
&\quad - \tfrac{3}{4}\bar{\varepsilon}^2(\widehat{d}_n - d)\Big\{\frac{1}{n}\sum_{i=1}^n(\varepsilon_i^2-1)\Big\} + \tfrac{1}{2}\sigma^2_{x,n}(\widehat{d}_n - d)^3 - \tfrac{3}{4}\sigma^2_{x,n}(\widehat{d}_n - d)^3\Big\{\frac{1}{n}\sum_{i=1}^n(\varepsilon_i^2-1)\Big\} \\
&\quad - \tfrac{5}{16}(\widehat{d}_n - d)\Big\{\frac{1}{n}\sum_{i=1}^n(\varepsilon_i^2-1)\Big\}^3\Bigg\} + O_p(n^{-2}) \\
&=: \widetilde{T}_n + O_p(n^{-2}).
\end{aligned}$$

By the arguments in Section 2.7 of Hall (1992), Edgeworth expansions of the distributions of $\widetilde{T}_n$ and T_n will disagree only in the terms of order n^{-2} or smaller. Hence, to obtain the expansion in (3.24) we may use $\widetilde{T}_n$ as a proxy for T_n. Using the results given in Appendices C and B, it may be shown by arguments similar to those used in the proof of Lemma 3.4 that

$$\begin{aligned}
&\mathrm{E}\left(\widetilde{T}_n\right) = \tfrac{1}{2}n^{-3/2}\kappa'_3\gamma_{x,n}, \\
&\mathrm{E}\left(\widetilde{T}_n^2\right) = 1 + 4n^{-1} + O(n^{-2}), \\
&\mathrm{E}\left(\widetilde{T}_n^3\right) = n^{-1/2}\kappa'_3\gamma_{x,n} - \tfrac{3}{8}n^{-3/2}(4\kappa'_5 - 5\kappa'_4\kappa'_3 - 30\kappa'_3)\gamma_{x,n} + O(n^{-2}), \\
&\mathrm{E}\left(\widetilde{T}_n^4\right) = 3 + n^{-1}(30 + \kappa'_4\kappa_{x,n}) + O(n^{-2}), \\
&\mathrm{E}\left(\widetilde{T}_n^5\right) = 10n^{-1/2}\kappa'_3\gamma_{x,n} + \tfrac{1}{4}n^{-3/2}\left(4\kappa'_5\tau_{x,n} - (60\kappa'_5 - 75\kappa'_4\kappa'_3 - 700\kappa'_3)\gamma_{x,n}\right) + O(n^{-2}).
\end{aligned}$$

Therefore the first five cumulants of $\widetilde{T}_n$ are

$$\begin{aligned}
\kappa_{1,n} &= \tfrac{1}{2} n^{-3/2} \kappa_3' \gamma_{x,n}, \\
\kappa_{2,n} &= 1 + 4n^{-1} + O(n^{-2}), \\
\kappa_{3,n} &= n^{-1/2} \kappa_3' \gamma_{x,n} - \tfrac{1}{8} n^{-3/2} \left(12\kappa_5' - 15\kappa_4'\kappa_3' - 78\kappa_3'\right) \gamma_{x,n} + O(n^{-2}), \\
\kappa_{4,n} &= n^{-1} \left(6 + \kappa_4' \kappa_{x,n}\right) + O(n^{-2}), \\
\kappa_{5,n} &= n^{-3/2} \left(\kappa_5' \tau_{x,n} + 30\kappa_3' \gamma_{x,n}\right) + O(n^{-2}),
\end{aligned}$$

whence we obtain the following constants (in the notation of (3.5)):

$$\begin{aligned}
k_{1,2} &= 0, \\
k_{1,3} &= \tfrac{1}{2} \kappa_3' \gamma_{x,n}, \\
k_{2,2} &= 4, \\
k_{3,1} &= \kappa_3' \gamma_{x,n}, \\
k_{3,2} &= -\tfrac{1}{8} \left(12\kappa_5' - 15\kappa_4'\kappa_3' - 78\kappa_3'\right) \gamma_{x,n}, \\
k_{4,1} &= 6 + \kappa_4' \kappa_{x,n}, \\
k_{5,1} &= \kappa_5' \tau_{x,n} + 30\kappa_3' \gamma_{x,n}.
\end{aligned} \tag{3.26}$$

Substituting these constants in (3.8), the polynomials in (3.25) may be verified easily using relation (3.7). Now, applying relation (3.7) to the expression in Lemma 3.1 we obtain

$$\begin{aligned}
Q_{3,n}(x)\phi(x) &= r_3\left(-\frac{d}{dx}\right)\Phi(x) \\
&= k_{1,3}\left(-\frac{d}{dx}\Phi(x)\right) + \tfrac{1}{6}\left(k_{3,2} + 3k_{2,2}k_{1,2} + k_{1,2}^3\right)\left(-\frac{d^3}{dx^3}\Phi(x)\right) \\
&\quad + \tfrac{1}{120}\left(k_{5,1} + 5k_{4,1}k_{1,2} + 10k_{3,1}k_{2,2} + 10k_{3,1}k_{1,2}^2\right)\left(-\frac{d^5}{dx^5}\Phi(x)\right) \\
&\quad + \tfrac{1}{144}k_{3,1}\left(k_{4,1} + 2k_{3,1}k_{1,2}\right)\left(-\frac{d^7}{dx^7}\Phi(x)\right) + \tfrac{1}{1296}k_{3,1}^3\left(-\frac{d^9}{dx^9}\Phi(x)\right) \\
&= -k_{1,3}\phi(x) - \tfrac{1}{6}\left(k_{3,2} + 3k_{2,2}k_{1,2} + k_{1,2}^3\right)(x^2-1)\phi(x) \\
&\quad - \tfrac{1}{120}\left(k_{5,1} + 5k_{4,1}k_{1,2} + 10k_{3,1}k_{2,2} + 10k_{3,1}k_{1,2}^2\right)\left(x^4 - 6x^2 + 3\right)\phi(x) \\
&\quad - \tfrac{1}{144}k_{3,1}\left(k_{4,1} + 2k_{3,1}k_{1,2}\right)\left(x^6 - 15x^4 + 45x^2 - 15\right)\phi(x) \\
&\quad - \tfrac{1}{1296}k_{3,1}^3\left(x^8 - 28x^6 + 210x^4 - 420x^2 + 105\right)\phi(x).
\end{aligned}$$

Finally, by substituting the $k_{j,l}$ in the last expression by their respective expressions given in (3.26), we obtain the required result. □

3.6 The smooth function model in the bootstrap world

Hall (1992) showed that Edgeworth and Cornish-Fisher expansions are also valid for bootstrap quantities that can be expressed in the smooth function model framework. Although his result is stated for the traditional n/n bootstrap, Chung and Lee (2001) showed that the results can be generalised to include valid Edgeworth and Cornish-Fisher expansions under the m/n bootstrap. In this section we briefly present the findings of Chung and Lee (2001).

Denote by $\mathscr{W}_n = \{\mathbf{W}_1, \ldots, \mathbf{W}_n\}$ the random sample consisting of d-dimensional random vectors defined in Section 3.1. Let $\mathscr{W}_m^* = \{\mathbf{W}_1^*, \mathbf{W}_2^*, \ldots, \mathbf{W}_m^*\}$ be a sample of size m drawn randomly with replacement from $\mathscr{W}_n$. Note that $\mathrm{E}^*(\mathbf{W}_1^*) = \bar{\mathbf{W}}_n = n^{-1}\sum_{i=1}^n \mathbf{W}_i$ and let $\bar{\mathbf{W}}_{m,n}^* = m^{-1}\sum_{i=1}^m \mathbf{W}_i^*$. Here, E^* denotes expectation over the conditional law of $\mathscr{W}_m^*$ given $\mathscr{W}_n$. Now, define the bootstrap quantities

$$\widehat{\theta}_m^* = g\left(\bar{\mathbf{W}}_{m,n}^*\right) \quad \text{and} \quad \widehat{\beta}_m^* = h\left(\bar{\mathbf{W}}_{m,n}^*\right),$$

where g and h are the same smooth functions defined in Section 3.1. The bootstrap versions of our two functions A_S and A_T are then respectively defined as

$$\widehat{A}_S(\mathbf{w}) = \frac{g(\mathbf{w}) - g(\bar{\mathbf{W}}_n)}{h(\bar{\mathbf{W}}_n)} \quad \text{and} \quad \widehat{A}_T(\mathbf{w}) = \frac{g(\mathbf{w}) - g(\bar{\mathbf{W}}_n)}{h(\mathbf{w})}.$$

Note that for both of these functions it holds that $\widehat{A}_S(\bar{\mathbf{W}}_n) = \widehat{A}_T(\bar{\mathbf{W}}_n) = 0$.

We state the results of Chung and Lee (2001) in the following two lemmas, in which the function $\widehat{A}$ stands for either $\widehat{A}_S$ or $\widehat{A}_T$. Like before, the polynomials $\widehat{P}_j$ and $\widehat{P}_j^{cf}$ resulting from the two functions $\widehat{A}_S$ and $\widehat{A}_T$ will generally not be the same.

Theorem 3.3 (Lemma 1 of Chung and Lee, 2001)**.** *Suppose that $m = O(n)$ and that $m \to \infty$ as $n \to \infty$. Assume that the function $\widehat{A}$ has sufficiently many continuous derivatives in an open neighbourhood of $\boldsymbol{\nu} = \mathrm{E}(\mathbf{W}_1)$, that $\mathbf{W}_1$ has sufficiently many finite moments, and that $\mathbf{W}_1$ satisfies Cramér's condition in* (3.2)*. Then, for $j \geq 1$, it holds that*

$$\mathrm{P}^*\left(m^{1/2}\widehat{A}(\bar{\mathbf{W}}_{m,n}^*) \leq x\right) = \Phi(x) + m^{-1/2}\widehat{P}_{1,n}(x)\phi(x) + \cdots + m^{-j/2}\widehat{P}_{j,n}(x)\phi(x) + O_p(m^{-(j+1)/2}),$$

uniformly in x, where the $\widehat{P}_{j,n}$ are obtained by substituting population moments appearing in the P_j defined in Theorem 3.1 for sample moments based on $\mathscr{W}_n$.

Proof. See Chung and Lee (2001, p. 236–238).

As a short illustrative example, consider again the case of the mean discussed in Section 3.4.1. Then

$$\mathbf{W}_i^* = \begin{bmatrix} X_i^* \\ (X_i^*)^2 \end{bmatrix}, \qquad i = 1, \ldots, n,$$

so that we have the standardised bootstrap quantity

$$\widehat{A}_S(\bar{\mathbf{W}}_{m,n}^*) = \frac{g(\bar{\mathbf{W}}_{m,n}^*) - g(\bar{\mathbf{W}}_n)}{h(\bar{\mathbf{W}}_n)} = \frac{\bar{X}_{m,n}^* - \bar{X}_n}{\widehat{\sigma}_n},$$

where $\bar{X}_{m,n}^* = m^{-1}\sum_{i=1}^m X_i^*$. Under the conditions of Theorem 3.3 we may now write, e.g.,

$$\begin{aligned}\mathrm{P}^*\left(m^{1/2}\widehat{A}_S(\bar{\mathbf{W}}_{m,n}^*) \leq x\right) &= \mathrm{P}^*\left(\frac{m^{1/2}\left(\bar{X}_{m,n}^* - \bar{X}_n\right)}{\widehat{\sigma}_n} \leq x\right) \\ &= \Phi(x) + m^{-1/2}\widehat{P}_{1,n}(x)\phi(x) + m^{-1}\widehat{P}_{2,n}(x)\phi(x) + O_p(m^{-3/2}),\end{aligned}$$

uniformly in x, where we have from (3.17) that

$$\widehat{P}_{1,n}(x) = -\tfrac{1}{6}\widehat{\kappa}'_{3,n}(x^2-1),$$
$$\widehat{P}_{2,n}(x) = -x\left\{\tfrac{1}{24}\widehat{\kappa}'_{4,n}(x^2-3) + \tfrac{1}{72}(\widehat{\kappa}'_{3,n})^2(x^4-10x^2+15)\right\},$$

with $\widehat{\kappa}'_{3,n} = n^{-1}\sum_{i=1}^{n}(X_i-\bar{X}_n)^3/\widehat{\sigma}_n^3$ and $\widehat{\kappa}'_{4,n} = n^{-1}\sum_{i=1}^{n}(X_i-\bar{X}_n)^4/\widehat{\sigma}_n^4 - 3$. Similarly,

$$\mathrm{P}^*\left(m^{1/2}\widehat{A}_T(\bar{\mathbf{W}}^*_{m,n}) \le x\right) = \Phi(x) + m^{-1/2}\widehat{Q}_{1,n}(x)\phi(x) + m^{-1}\widehat{Q}_{2,n}(x)\phi(x) + O_p(m^{-3/2}),$$

uniformly in x, where we have from (3.21) that

$$\widehat{Q}_1(x) = \tfrac{1}{6}\widehat{\kappa}'_3(2x^2+1),$$
$$\widehat{Q}_2(x) = x\left(\tfrac{1}{12}\widehat{\kappa}'_4(x^2-3) - \tfrac{1}{18}\left(\widehat{\kappa}'_3\right)^2(x^4+2x^2-3) - \tfrac{1}{4}(x^2+3)\right).$$

As shown in the following lemma, valid Cornish-Fisher expansions of bootstrap quantiles may also be obtained.

Theorem 3.4 (Lemma 2 of Chung and Lee, 2001)**.** *Define*

$$\widehat{w}_{n,m,\alpha} = \inf\left\{x : \mathrm{P}^*\left(m^{1/2}\widehat{A}(\bar{\mathbf{W}}^*_{m,n}) \le x\right) \ge \alpha\right\}.$$

Under the assumptions of Theorem 3.3 there exist polynomials $\widehat{P}^{cf}_{1,n},\ldots,\widehat{P}^{cf}_{j,n}$ such that

$$\widehat{w}_{n,m,\alpha} = z_\alpha + m^{-1/2}\widehat{P}^{cf}_{1,n}(z_\alpha) + \cdots + m^{-j/2}\widehat{P}^{cf}_{j,n}(z_\alpha) + O_p(m^{-(j+1)/2}),$$

uniformly in $\alpha \in (\varepsilon, 1-\varepsilon)$ for any $\varepsilon \in (0, 1/2)$, where the $\widehat{P}^{cf}_{j,n}$ are obtained by substituting population moments appearing in the P^{cf}_j defined in Theorem 3.2 for sample moments based on $\mathscr{W}_n$.

Proof. See Chung and Lee (2001, p. 238).

Note that the relations between the sample quantities $\widehat{P}^{cf}_{j,n}$ and $\widehat{P}_{j,n}$ are the same as the relations between population quantities P^{cf}_j and P_j given in Section 3.3. Hence, we have from (3.9) and (3.10), for example, that

$$\begin{aligned}\widehat{P}^{cf}_{1,n}(x) &= -\widehat{P}_{1,n}(x),\\ \widehat{P}^{cf}_{2,n}(x) &= \widehat{P}_{1,n}(x)\widehat{P}'_{1,n}(x) - \tfrac{1}{2}x\widehat{P}^2_{1,n}(x) - \widehat{P}_{2,n}(x),\end{aligned} \tag{3.27}$$

for all $x \in \mathbb{R}$.

Chapter 4

New percentile confidence bounds

In this chapter we present a new method based on sample splitting that can be used to construct bootstrap confidence bounds for an unknown population parameter θ. This method is then used to derive two new *percentile* confidence bounds—a hybrid version and a backwards version. We derive, under some regularity assumptions, the asymptotic coverage probabilities of the new hybrid and backwards percentile bounds. Among others, it is shown that the new *hybrid* precentile bound has coverage error of $O(n^{-1})$, compared to the coverage error of $O(n^{-1/2})$ of the standard hybrid percentile bound $\widehat{\mathscr{I}}_H$. As far as the new *backwards* percentile bound is concerned, we show that it has coverage error of $O(n^{-1/2})$, but in some cases $O(n^{-1})$.

In the arguments of Hall (1988), the order of coverage error of confidence bounds is primarily determined by a *random* distance, e.g.,

$$\widehat{\theta}_n - \theta = O_p(n^{-1/2}),$$

where $\widehat{\theta}_n$ is some estimator for the parameter θ. The rationale behind our idea rests upon the construction of a confidence bound in such a way that the order of coverage error is essentially determined by a *constant* distance, which is typically of the form

$$\mathrm{E}(\widehat{\theta}_n - \theta) = O(n^{-1}).$$

This may be accomplished by *splitting* the sample into two independent sets. The method of construction relies partly on the fact that, if Y and Z are two independent random variables in $\mathbb{R}$ and we let $\Psi(z) := \mathrm{P}(Y \geq z)$, $z \in \mathbb{R}$, we may write

$$\mathrm{P}(Y \geq Z) = \mathrm{E}(\Psi(Z)). \tag{4.1}$$

That is, if we determine an expression for $\mathrm{P}(Y \geq z)$ in terms of a deterministic z, replace z by a random variable Z, and then take the expected value, we will end up calculating $\mathrm{P}(Y \geq Z)$.

The proof of (4.1) is straightforward:

$$\mathrm{E}\left(\mathrm{I}(Y \geq Z) \,|\, Z = z\right) = \mathrm{E}\left(\mathrm{I}(Y \geq z) \,|\, Z = z\right) = \mathrm{E}\left(\mathrm{I}(Y \geq z)\right) = \mathrm{P}(Y \geq z) = \Psi(z).$$

Since this holds for all $z \in \mathbb{R}$, we have

$$\mathrm{E}(\mathrm{I}(Y \ge Z) \mid Z) = \Psi(Z) \text{ a.s.},$$

which implies that the expected value of the random quantity on the left-hand side will be equal to the expected value of the random quantity on the right-hand side. Noting that the expectation on the left-hand side simplifies to

$$\mathrm{E}[\mathrm{E}(\mathrm{I}(Y \ge Z) \mid Z)] = \mathrm{E}(\mathrm{I}(Y \ge Z)) = \mathrm{P}(Y \ge Z),$$

we have the result (4.1).

4.1 Notation

Let $\mathscr{X}_n = \{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ be a random sample consisting of p-dimensional random vectors drawn from an unknown p-dimensional distribution function F, depending on some unknown scalar parameter θ. In the context of the smooth function model described in Chapter 3, set $\mathbf{W}_k = (f_1(\mathbf{X}_k), \ldots, f_d(\mathbf{X}_k))$, $k = 1, \ldots, n$, where $f_1, \ldots, f_d$ are real-valued Borel-measurable functions on $\mathbb{R}^p$. Define $\boldsymbol{v} = \mathrm{E}(\mathbf{W}_1)$. Assume that the parameter of interest can be written in the form $\theta = g(\boldsymbol{v})$, where $g : \mathbb{R}^d \to \mathbb{R}$ is a known smooth, Borel-measurable function.

Our new method involves splitting the sample into two disjoint sets, say

$$\mathscr{W}_\ell := \{\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_\ell\},$$

for some integer $2 \le \ell \le n-2$, and

$$\mathscr{W}_r := \{\mathbf{W}_{\ell+1}, \mathbf{W}_{\ell+2}, \ldots, \mathbf{W}_n\},$$

which has $r := n-\ell$ elements. Note that $\mathscr{W}_\ell$ and $\mathscr{W}_r$ can be viewed as two *independent* random samples of sizes ℓ and r drawn from the distribution F. Define the following two estimators for θ:

$$\widehat{\theta}_\ell = g\left(\bar{\mathbf{W}}_\ell\right) \quad \text{and} \quad \widehat{\theta}_r = g\left(\bar{\mathbf{W}}_r\right),$$

where $\bar{\mathbf{W}}_\ell = \ell^{-1}\sum_{k=1}^{\ell}\mathbf{W}_k$ and $\bar{\mathbf{W}}_r = r^{-1}\sum_{k=\ell+1}^{r}\mathbf{W}_k$. Throughout we will assume that the asymptotic variance of $\ell^{1/2}\widehat{\theta}_\ell$ is given by $\beta^2 = h^2(\boldsymbol{v})$ for some known smooth, Borel-measurable function $h : \mathbb{R}^d \to \mathbb{R}$. Two possible estimators for β are

$$\widehat{\beta}_\ell = h\left(\bar{\mathbf{W}}_\ell\right) \quad \text{and} \quad \widehat{\beta}_r = h\left(\bar{\mathbf{W}}_r\right).$$

Define

$$A_S(\mathbf{w}) := \frac{g(\mathbf{w}) - g(\boldsymbol{v})}{h(\boldsymbol{v})},$$

so that

$$A_S\left(\bar{\mathbf{W}}_\ell\right) = \frac{g(\bar{\mathbf{W}}_\ell) - g(\boldsymbol{v})}{h(\boldsymbol{v})} = \frac{\widehat{\theta}_\ell - \theta}{\beta}.$$

Note that $\mathrm{Var}(\ell^{1/2}A_S(\bar{\mathbf{W}}_\ell)) \to 1$ as $\ell \to \infty$. Assume that $\mathbf{W}_1$ satisfies Cramér's continuity condition, i.e.,

$$\limsup_{\|\mathbf{t}\|\to\infty} |\chi(\mathbf{t})| < 1, \tag{4.2}$$

where χ denotes the characteristic function of $\mathbf{W}_1$. Then, if g and h are sufficiently smooth in a neighbourhood of $\boldsymbol{\nu}$ and $\mathbf{W}_1$ has sufficiently many finite moments, we have by Theorem 3.1 that

$$\begin{aligned} \mathrm{P}\left(\ell^{1/2}A_S\left(\bar{\mathbf{W}}_\ell\right) \le x\right) &= \mathrm{P}\left(\ell^{1/2}\left(\widehat{\theta}_\ell - \theta\right)/\beta \le x\right) \\ &= \Phi(x) + \ell^{-1/2}P_1(x)\phi(x) + \ell^{-1}P_2(x)\phi(x) + O(\ell^{-3/2}), \end{aligned} \tag{4.3}$$

uniformly in all x, with P_1 and P_2 determining the above Edgeworth expansion as described in Chapter 3. Note that the remainder term $O(\ell^{-3/2})$ is non-random.

Let $\mathscr{W}^*_{m,r} = \{\mathbf{W}^*_1, \mathbf{W}^*_2, \ldots, \mathbf{W}^*_m\}$ denote a *resample* of size m drawn randomly *with replacement* from $\mathscr{W}_r$, i.e., $\mathscr{W}^*_{m,r}$ is an m/r bootstrap sample drawn from $\mathscr{W}_r$. Now define an m/r bootstrap replication of θ as

$$\widehat{\theta}^*_{m,r} = g\left(\bar{\mathbf{W}}^*_{m,r}\right),$$

where $\bar{\mathbf{W}}^*_{m,r} = m^{-1}\sum_{k=1}^m \mathbf{W}^*_k$. Now define the smooth function

$$\widehat{A}_S(\mathbf{w}) := \frac{g(\mathbf{w}) - g(\bar{\mathbf{W}}_r)}{h(\bar{\mathbf{W}}_r)},$$

so that

$$S^*_{m,r} := m^{1/2}\widehat{A}_S\left(\bar{\mathbf{W}}^*_{m,r}\right) = \frac{m^{1/2}\left(g\left(\bar{\mathbf{W}}^*_{m,r}\right) - g\left(\bar{\mathbf{W}}_r\right)\right)}{h\left(\bar{\mathbf{W}}_r\right)} = \frac{m^{1/2}\left(\widehat{\theta}^*_{m,r} - \widehat{\theta}_r\right)}{\widehat{\beta}_r}.$$

Note that $\mathrm{Var}_*(S^*_{m,r}) \to 1$ a.s. as $m \to \infty$. Assuming that the function $\widehat{A}_S$ has sufficiently many bounded derivatives in a neighbourhood of $\boldsymbol{\nu}$, that $\mathbf{W}_1$ has sufficiently many finite moments and satisfies (4.2). Then, if we denote the α-level quantile of $S^*_{m,r}$ by $\widehat{\xi}_{m,r,\alpha}$, i.e.,

$$\widehat{\xi}_{m,r,\alpha} := \inf\left\{x : \mathrm{P}^*\left(S^*_{m,r} \le x\right) \ge \alpha\right\},$$

we have by Theorem 3.4 that

$$\widehat{\xi}_{m,r,\alpha} = z_\alpha + m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha) + O_p(m^{-3/2}),$$

uniformly in $\alpha \in (\varepsilon, 1-\varepsilon)$ for any $\varepsilon \in (0, \frac{1}{2})$, with $\widehat{P}^{cf}_{1,r}$ and $\widehat{P}^{cf}_{2,r}$ the appropriate Cornish-Fisher polynomials described in Chapter 3.

The two procedures introduced below are based on this Cornish-Fisher expansion of $\widehat{\xi}_{m,r,\alpha}$.

4.2 Hybrid percentile bound $\widehat{\mathscr{I}}^N_H$

In Section 2.4.2 we stated that the *standard hybrid percentile* $100(1-\alpha)\%$ upper confidence bound for θ is given by

$$\widehat{\mathscr{I}}_H(\alpha) = \left(-\infty, \widehat{\theta}_n - n^{-1/2}\widehat{\beta}_n\widehat{\xi}_\alpha\right],$$

with $\widehat{\theta}_n$, $\widehat{\beta}_n$ and $\widehat{\xi}_\alpha$ as defined in Section 2.1. Recall that this standard confidence bound attains the nominal coverage probability $1-\alpha$ with a coverage error of order $O(n^{-1/2})$.

We now propose the following modified hybrid percentile $100(1-\alpha)\%$ upper confidence bound for θ, which, under some regularity conditions, attains the nominal coverage probability $1-\alpha$ *with a coverage error of order $O(n^{-1})$, compared to the coverage error of $O(n^{-1/2})$ of $\widehat{\mathscr{I}}_H(\alpha)$*. This result is stated and proved in Theorem 4.1 below.

New procedure

We suggest the following hybrid percentile $100(1-\alpha)\%$ upper confidence bound for θ:

$$\widehat{\mathscr{I}}_H^N(m,\alpha) := \left(-\infty,\, \widehat{\theta}_\ell - \ell^{-1/2}\widehat{\beta}_r\widetilde{\xi}_{m,r,\alpha}\right],$$

where

$$\widetilde{\xi}_{m,r,\alpha} = z_\alpha + m^{-1/2}\widehat{P}_{1,r}^{cf}(z_\alpha) + m^{-1}\widehat{P}_{2,r}^{cf}(z_\alpha).$$

To investigate the asymptotic behaviour of this new confidence bound, the following assumptions are needed:

(A1) $\mathrm{E}\left(\widehat{\beta}_r - \beta\right) = O(r^{-1})$,

(A2) $\mathrm{E}\left\{\left(\widehat{\beta}_r - \beta\right)^4\right\} = O(r^{-2})$,

(A3) $\mathrm{E}\left(\widehat{P}_{1,r}(x) - P_1(x)\right) = O(r^{-1})$,

(A4) $\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x) - P_1(x)\right)^4\right\} = O(r^{-2})$,

(A5) $\mathrm{E}\left(\widehat{P}'_{1,r}(x) - P'_1(x)\right) = O(r^{-1})$,

(A6) $\mathrm{E}\left\{\left(\widehat{P}'_{1,r}(x) - P'_1(x)\right)^4\right\} = O(r^{-2})$,

(A7) $\mathrm{E}\left\{\left(\widehat{P}_{2,r}(x) - P_2(x)\right)^i\right\} = O(r^{-1}), \qquad i = 1,2.$

Theorem 4.1. *Suppose that $\mathbf{W}_1$ satisfies (4.2) and has sufficiently many finite moments. Also, assume that g and h are continuously differentiable up to a sufficiently high order in an open neighborhood of $\mathbf{v}$. Then, if $m = \ell = O(r)$ and $\ell \to \infty$ as $n \to \infty$, we have under assumptions (A1)–(A7) that*

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\right) = 1 - \alpha + \frac{C_\theta(z_\alpha)}{r} + O(\ell^{-3/2}), \tag{4.4}$$

where $C_\theta(z_\alpha)$ is the coefficient of r^{-1} in a power series expansion of

$$-z_\alpha\phi(z_\alpha)\beta^{-1}\mathrm{E}(\widehat{\beta}_r - \beta) + \tfrac{1}{2}z_\alpha^3\phi(z_\alpha)\beta^{-2}\mathrm{E}\{(\widehat{\beta}_r - \beta)^2\}.$$

Moreover, if we choose $\ell = \lfloor \gamma n^\psi \rfloor$ for some $\gamma > 0$ and $\frac{2}{3} < \psi < 1$, then

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\right) = \begin{cases} 1 - \alpha + \dfrac{C_\theta(z_\alpha)}{n} + O(n^{-(2-\psi)} + n^{-3\psi/2}) & \text{if } C_\theta(z_\alpha) \neq 0, \\ 1 - \alpha + O(n^{-3\psi/2}) & \text{if } C_\theta(z_\alpha) = 0. \end{cases} \tag{4.5}$$

In the case where $\psi = 1$ and $0 < \gamma < 1$,

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\right) = 1-\alpha + \frac{C_\theta(z_\alpha)}{(1-\gamma)n} + O(n^{-3/2}).$$

Proof. Noting that $\mathscr{X}_\ell$ and $\mathscr{X}_r$ are independent, and hence also $\widehat{\theta}_\ell$ and $\widehat{\beta}_r \widetilde{\xi}_{m,r,\alpha}$, we may apply the idea stated in (4.1) to rewrite the coverage probability of the confidence bound $\widehat{\mathscr{I}}^N(m,\alpha)$ as

$$\begin{aligned}
\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(m,\alpha)\right) &= \mathrm{P}\left(\theta \le \widehat{\theta}_\ell - \ell^{-1/2}\widehat{\beta}_r\widetilde{\xi}_{m,r,\alpha}\right)\\
&= \mathrm{P}\left(\ell^{1/2}\left(\widehat{\theta}_\ell - \theta\right) \ge \widehat{\beta}_r\widetilde{\xi}_{m,r,\alpha}\right)\\
&= \mathrm{E}\left(\Psi\left(\widehat{\beta}_r\widetilde{\xi}_{m,r,\alpha}\right)\right),
\end{aligned}$$

with

$$\begin{aligned}
\Psi(x) &:= \mathrm{P}\left(\ell^{1/2}\left(\widehat{\theta}_\ell - \theta\right) \ge x\right)\\
&= 1 - \mathrm{P}\left(\frac{\ell^{1/2}\left(\widehat{\theta}_\ell - \theta\right)}{\beta} \le \frac{x}{\beta}\right)\\
&= 1 - \left[\Phi\left(\frac{x}{\beta}\right) + \ell^{-1/2}R_1\left(\frac{x}{\beta}\right) + \ell^{-1}R_2\left(\frac{x}{\beta}\right)\right] + O\left(\ell^{-3/2}\right), \qquad (4.6)
\end{aligned}$$

where the last step follows from (4.3) with $R_j(x) = P_j(x)\phi(x)$. Since the error term does not depend on x, we may write

$$\Psi\left(\widehat{\beta}_r\widetilde{\xi}_{m,r,\alpha}\right) = 1 - \left[\Phi\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha}\right) + \ell^{-1/2}R_1\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha}\right) + \ell^{-1}R_2\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha}\right)\right] + O\left(\ell^{-3/2}\right). \quad (4.7)$$

Note that the remainder term $O\left(\ell^{-3/2}\right)$ is still non-random.

We will now inspect each term in the above expression separately. First, Taylor expansion of Φ about z_α yields

$$\begin{aligned}
\Phi\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha}\right) &= \Phi(z_\alpha) + \phi(z_\alpha)\left[\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha} - z_\alpha\right] - \tfrac{1}{2}z_\alpha\phi(z_\alpha)\left[\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha} - z_\alpha\right]^2 + \delta_{1,m,r}\\
&= \alpha + \phi(z_\alpha)\left[\frac{(\widehat{\beta}_r - \beta)}{\beta}\widetilde{\xi}_{m,r,\alpha} + \widetilde{\xi}_{m,r,\alpha} - z_\alpha\right]\\
&\quad - \tfrac{1}{2}z_\alpha\phi(z_\alpha)\left[\frac{(\widehat{\beta}_r - \beta)}{\beta}\widetilde{\xi}_{m,r,\alpha} + \widetilde{\xi}_{m,r,\alpha} - z_\alpha\right]^2 + \delta_{1,m,r}\,,
\end{aligned}$$

where $\delta_{1,m,r} = \frac{1}{3!}\Phi^{(3)}\left(\Theta_{m,r}\right)\left[\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,\alpha} - z_\alpha\right]^3$ for some $\Theta_{m,r}$ between $\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,\alpha}$ and z_α. Hence,

$$\begin{aligned}
\Phi\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha}\right) &= \alpha + \phi(z_\alpha)\left[\frac{(\widehat{\beta}_r - \beta)}{\beta}\left\{z_\alpha + m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right\}\right.\\
&\quad \left. + \left\{m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right\}\right]\\
&\quad - \tfrac{1}{2}z_\alpha\phi(z_\alpha)\left[\frac{(\widehat{\beta}_r - \beta)}{\beta}\left\{z_\alpha + m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right\}\right]^2 + \delta_{1,m,r}
\end{aligned}$$

$$
\begin{aligned}
&= \alpha + z_\alpha \phi(z_\alpha)\frac{(\widehat{\beta}_r-\beta)}{\beta} - \tfrac{1}{2} z_\alpha^3 \phi(z_\alpha)\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2} \\
&\quad + \phi(z_\alpha)\left\{m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right\} - \tfrac{1}{2} m^{-1} z_\alpha \phi(z_\alpha)\left(\widehat{P}^{cf}_{1,r}(z_\alpha)\right)^2 \\
&\quad + \sum_{i=1}^{7} A_{i,m,r}(\alpha) + \delta_{1,m,r},
\end{aligned} \tag{4.8}
$$

where

$$
A_{1,m,r}(\alpha) = \phi(z_\alpha)\frac{(\widehat{\beta}_r-\beta)}{\beta}\left\{m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right\}, \tag{4.9}
$$

$$
A_{2,m,r}(\alpha) = -\tfrac{1}{2} z_\alpha \phi(z_\alpha)\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2}\left\{m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right\}^2, \tag{4.10}
$$

$$
A_{3,m,r}(\alpha) = -\tfrac{1}{2} z_\alpha \phi(z_\alpha) m^{-2}\left\{\widehat{P}^{cf}_{2,r}(z_\alpha)\right\}^2, \tag{4.11}
$$

$$
A_{4,m,r}(\alpha) = -z_\alpha^2 \phi(z_\alpha)\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2}\left\{m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right\}, \tag{4.12}
$$

$$
A_{5,m,r}(\alpha) = -z_\alpha^2 \phi(z_\alpha)\frac{(\widehat{\beta}_r-\beta)}{\beta}\left\{m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right\}, \tag{4.13}
$$

$$
A_{6,m,r}(\alpha) = -z_\alpha \phi(z_\alpha)\frac{(\widehat{\beta}_r-\beta)}{\beta}\left\{m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right\}^2, \tag{4.14}
$$

$$
A_{7,m,r}(\alpha) = -z_\alpha \phi(z_\alpha) m^{-3/2}\widehat{P}^{cf}_{1,r}(z_\alpha)\widehat{P}^{cf}_{2,r}(z_\alpha). \tag{4.15}
$$

Substituting the Cornish-Fisher polynomials $\widehat{P}^{cf}_{1,r}$ and $\widehat{P}^{cf}_{2,r}$ by their respective expressions in terms of Edgeworth polynomials (given explicitly in (3.27)), (4.8) becomes

$$
\begin{aligned}
\Phi\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha}\right) &= \alpha + z_\alpha \phi(z_\alpha)\frac{(\widehat{\beta}_r-\beta)}{\beta} - \tfrac{1}{2} z_\alpha^3 \phi(z_\alpha)\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2} - m^{-1/2}\phi(z_\alpha)\widehat{P}_{1,r}(z_\alpha) \\
&\quad + m^{-1}\phi(z_\alpha)\widehat{P}_{1,r}(z_\alpha)\left\{\widehat{P}'_{1,r}(z_\alpha) - z_\alpha \widehat{P}_{1,r}(z_\alpha)\right\} - m^{-1}\phi(z_\alpha)\widehat{P}_{2,r}(z_\alpha) \\
&\quad + \sum_{i=1}^{7} A_{i,m,r}(\alpha) + \delta_{1,m,r}.
\end{aligned} \tag{4.16}
$$

By Taylor expansion of R_1 about z_α, the term $\ell^{-1/2} R_1\left(\widehat{\beta}_r \beta^{-1}\widetilde{\xi}_{m,r,\alpha}\right)$ in (4.7) becomes

$$
\begin{aligned}
&\ell^{-1/2} R_1\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha}\right) \\
&= \ell^{-1/2} R_1(z_\alpha) + \ell^{-1/2} R_1'(z_\alpha)\left[\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha} - z_\alpha\right] + \ell^{-1/2}\delta_{2,m,r} \\
&= \ell^{-1/2}\phi(z_\alpha) P_1(z_\alpha) + \ell^{-1/2}\phi(z_\alpha)\left(P_1'(z_\alpha) - z_\alpha P_1(z_\alpha)\right)\left[\frac{(\widehat{\beta}_r-\beta)}{\beta}\widetilde{\xi}_{m,r,\alpha} + \widetilde{\xi}_{m,r,\alpha} - z_\alpha\right] \\
&\quad + \ell^{-1/2}\delta_{2,m,r} \\
&= \ell^{-1/2}\phi(z_\alpha) P_1(z_\alpha) + \ell^{-1/2}\phi(z_\alpha)\left(P_1'(z_\alpha) - z_\alpha P_1(z_\alpha)\right) \\
&\quad \times \left[\frac{(\widehat{\beta}_r-\beta)}{\beta}\left\{z_\alpha + m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right\} + m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right] \\
&\quad + \ell^{-1/2}\delta_{2,m,r} \\
&= \ell^{-1/2}\phi(z_\alpha) P_1(z_\alpha) + \ell^{-1/2} m^{-1/2}\phi(z_\alpha)\widehat{P}^{cf}_{1,r}(z_\alpha)\left(P_1'(z_\alpha) - z_\alpha P_1(z_\alpha)\right) + A_{8,m,r}(\alpha) + \ell^{-1/2}\delta_{2,m,r}
\end{aligned}
$$

$$= \ell^{-1/2}\phi(z_\alpha)P_1(z_\alpha) - \ell^{-1/2}m^{-1/2}\phi(z_\alpha)\widehat{P}_{1,r}(z_\alpha)\left(P_1'(z_\alpha) - z_\alpha P_1(z_\alpha)\right) + A_{8,m,r}(\alpha) + \ell^{-1/2}\delta_{2,m,r}\,, \tag{4.17}$$

where

$$A_{8,m,r}(\alpha) = \ell^{-1/2}\phi(z_\alpha)\left(P_1'(z_\alpha) - z_\alpha P_1(z_\alpha)\right)\Bigg[\frac{(\widehat{\beta}_r - \beta)}{\beta}\left\{z_\alpha + m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right\}$$
$$+ m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\Bigg] \tag{4.18}$$

and $\delta_{2,m,r} = \frac{1}{2}R_1''\left(\Theta_{m,r}\right)\left[\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,\alpha} - z_\alpha\right]^2$ for some $\Theta_{m,r}$ between $\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,\alpha}$ and z_α.

Lastly, also by Taylor expansion about z_α, we have that

$$\ell^{-1}R_2\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha}\right) = \ell^{-1}R_2(z_\alpha) + A_{9,m,r}(\alpha) + \ell^{-1}\delta_{3,m,r}$$
$$= \ell^{-1}\phi(z_\alpha)P_2(z_\alpha) + A_{9,m,r}(\alpha) + \ell^{-1}\delta_{3,m,r}\,, \tag{4.19}$$

where

$$A_{9,m,r}(\alpha) = \ell^{-1}R_2'(z_\alpha)\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha} - z_\alpha\right)$$
$$= \ell^{-1}\phi(z_\alpha)\left(P_2'(z_\alpha) - z_\alpha P_2(z_\alpha)\right)\left\{\frac{(\widehat{\beta}_r - \beta)}{\beta}\widetilde{\xi}_{m,r,\alpha} + \widetilde{\xi}_{m,r,\alpha} - z_\alpha\right\}$$
$$= \ell^{-1}\phi(z_\alpha)\left(P_2'(z_\alpha) - z_\alpha P_2(z_\alpha)\right)\Bigg\{\frac{(\widehat{\beta}_r - \beta)}{\beta}\left(z_\alpha + m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right)$$
$$+ m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\Bigg\} \tag{4.20}$$

and $\delta_{3,m,r} = \frac{1}{2}R_2''\left(\Theta_{m,r}\right)\left[\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,\alpha} - z_\alpha\right]^2$ for some $\Theta_{m,r}$ between $\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,\alpha}$ and z_α.

Substituting (4.16), (4.17) and (4.19) in (4.7) yields

$$\Psi(\widehat{\beta}_r\widetilde{\xi}_{m,r,\alpha}) = 1 - \alpha - z_\alpha\phi(z_\alpha)\frac{(\widehat{\beta}_r - \beta)}{\beta} + \frac{1}{2}z_\alpha^3\phi(z_\alpha)\frac{(\widehat{\beta}_r - \beta)^2}{\beta^2} + m^{-1/2}\phi(z_\alpha)\widehat{P}_{1,r}(z_\alpha)$$
$$- m^{-1}\phi(z_\alpha)\widehat{P}_{1,r}(z_\alpha)\left\{\widehat{P}'_{1,r}(z_\alpha) + z_\alpha\widehat{P}_{1,r}(z_\alpha)\right\} + m^{-1}\phi(z_\alpha)\widehat{P}_{2,r}(z_\alpha)$$
$$- \ell^{-1/2}\phi(z_\alpha)P_1(z_\alpha) + \ell^{-1/2}m^{-1/2}\phi(z_\alpha)\widehat{P}_{1,r}(z_\alpha)\left(P_1'(z_\alpha) - z_\alpha P_1(z_\alpha)\right)$$
$$- \ell^{-1}\phi(z_\alpha)P_2(z_\alpha) + O\left(\ell^{-3/2}\right)$$
$$- \sum_{i=1}^{9}A_{i,\ell,r}(\alpha) - \delta_{1,m,r} - \ell^{-1/2}\delta_{2,m,r} - \ell^{-1}\delta_{3,m,r}.$$

Setting $m = \ell$ and rearranging terms we rewrite the above expression in the following convenient form:

$$\Psi(\widehat{\beta}_r\widetilde{\xi}_{\ell,r,\alpha}) = 1 - \alpha - z_\alpha\phi(z_\alpha)\frac{(\widehat{\beta}_r - \beta)}{\beta} + \frac{1}{2}z_\alpha^3\phi(z_\alpha)\frac{(\widehat{\beta}_r - \beta)^2}{\beta^2}$$
$$+ \ell^{-1/2}\phi(z_\alpha)\left(\widehat{P}_{1,r}(z_\alpha) - P_1(z_\alpha)\right) + \ell^{-1}z_\alpha\phi(z_\alpha)\widehat{P}_{1,r}(z_\alpha)\left(\widehat{P}_{1,r}(z_\alpha) - P_1(z_\alpha)\right)$$
$$- \ell^{-1}\phi(z_\alpha)\widehat{P}_{1,r}(z_\alpha)\left(\widehat{P}'_{1,r}(z_\alpha) - P_1'(z_\alpha)\right) + \ell^{-1}\phi(z_\alpha)\left(\widehat{P}_{2,r}(z_\alpha) - P_2(z_\alpha)\right)$$
$$+ O(\ell^{-3/2}) - \sum_{i=1}^{9}A_{i,\ell,r}(\alpha) - \delta_{1,\ell,r} - \ell^{-1/2}\delta_{2,\ell,r} - \ell^{-1}\delta_{3,\ell,r}.$$

The coverage probability of $\widehat{\mathscr{I}}^N(\ell,\alpha)$ can now be obtained by taking the expected value of $\Psi(\widehat{\beta}_r\widetilde{\xi}_{\ell,r,\alpha})$. From the assumptions of Theorem 4.1 and results (i) and (ii) of Lemma D.1 in the Appendix it follows that

$$\begin{aligned}
\mathrm{P}\Big(\theta\in\widehat{\mathscr{I}}_H^N(\ell,\alpha)\Big) &= \mathrm{E}\left(\Psi(\widehat{\beta}_r\widetilde{\xi}_{\ell,r,\alpha})\right)\\
&= 1-\alpha-z_\alpha\phi(z_\alpha)\frac{\mathrm{E}(\widehat{\beta}_r-\beta)}{\beta}+\tfrac{1}{2}z_\alpha^3\phi(z_\alpha)\frac{\mathrm{E}(\widehat{\beta}_r-\beta)^2}{\beta^2}\\
&\quad-\sum_{i=1}^{9}\mathrm{E}\left(A_{i,\ell,r}(\alpha)\right)-\mathrm{E}\left(\delta_{1,\ell,r}\right)-\ell^{-1/2}\,\mathrm{E}\left(\delta_{2,\ell,r}\right)-\ell^{-1}\,\mathrm{E}\left(\delta_{3,\ell,r}\right)\\
&\quad+O(\ell^{-1/2}r^{-1})+O(\ell^{-3/2}).
\end{aligned}\tag{4.21}$$

Now, by results (v) and (ix) of Lemma D.1 we have that

$$\begin{aligned}
\mathrm{E}\left(A_{1,\ell,r}(\alpha)\right) &= -\phi(z_\alpha)\frac{1}{\beta}\ell^{-1/2}\,\mathrm{E}\left((\widehat{\beta}_r-\beta)\widehat{P}_{1,r}(z_\alpha)\right)+\phi(z_\alpha)\frac{1}{\beta}\ell^{-1}\,\mathrm{E}\left((\widehat{\beta}_r-\beta)\widehat{P}^{cf}_{2,r}(z_\alpha)\right)\\
&= O(\ell^{-1/2}r^{-1}).
\end{aligned}\tag{4.22}$$

Noting that, by the c_r-inequality (see, e.g., Loève, 1977, p. 157),

$$\begin{aligned}
\mathrm{E}\left(\left|A_{2,\ell,r}(\alpha)\right|\right) &= \tfrac{1}{2}\,|z_\alpha|\,\phi(z_\alpha)\,\mathrm{E}\left\{\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2}\left(\ell^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha)+\ell^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right)^2\right\}\\
&\le 2\ell^{-1}\,|z_\alpha|\,\phi(z_\alpha)\,\mathrm{E}\left\{\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2}\left(\widehat{P}^{cf}_{1,r}(z_\alpha)\right)^2\right\}\\
&\quad+2\ell^{-2}\,|z_\alpha|\,\phi(z_\alpha)\,\mathrm{E}\left\{\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2}\left(\widehat{P}^{cf}_{2,r}(z_\alpha)\right)^2\right\},\\
&= 2\ell^{-1}\,|z_\alpha|\,\phi(z_\alpha)\frac{1}{\beta}\,\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\left(\widehat{P}_{1,r}(z_\alpha)\right)^2\right\}+O(\ell^{-2}),
\end{aligned}$$

it follows from (viii) of Lemma D.1 that

$$\mathrm{E}\left(A_{2,\ell,r}(\alpha)\right)\le\left|\mathrm{E}\left(A_{2,\ell,r}(\alpha)\right)\right|\le\mathrm{E}\left(\left|A_{2,\ell,r}(\alpha)\right|\right)=O(\ell^{-1}r^{-1}).\tag{4.23}$$

Also,

$$\mathrm{E}\left(A_{3,\ell,r}(\alpha)\right)=-\tfrac{1}{2}z_\alpha\phi(z_\alpha)\ell^{-2}\,\mathrm{E}\left\{\left(\widehat{P}^{cf}_{2,r}(z_\alpha)\right)^2\right\}=O(\ell^{-2}).\tag{4.24}$$

From result (vii) of Lemma D.1 we have that

$$\begin{aligned}
\mathrm{E}\left(A_{4,\ell,r}(\alpha)\right) &= z_\alpha^2\phi(z_\alpha)\frac{1}{\beta^2}\ell^{-1/2}\,\mathrm{E}\left((\widehat{\beta}_r-\beta)^2\widehat{P}_{1,r}(z_\alpha)\right)-z_\alpha^2\phi(z_\alpha)\frac{1}{\beta^2}\ell^{-1}\,\mathrm{E}\left((\widehat{\beta}_r-\beta)^2\widehat{P}^{cf}_{2,r}(z_\alpha)\right)\\
&= O(\ell^{-1/2}r^{-1}),
\end{aligned}\tag{4.25}$$

and from result (v) and (ix) of the same lemma,

$$\begin{aligned}
\mathrm{E}\left(A_{5,\ell,r}(\alpha)\right) &= z_\alpha^2\phi(z_\alpha)\frac{1}{\beta}\ell^{-1/2}\,\mathrm{E}\left((\widehat{\beta}_r-\beta)\widehat{P}_{1,r}(z_\alpha)\right)-z_\alpha^2\phi(z_\alpha)\frac{1}{\beta}\ell^{-1}\,\mathrm{E}\left((\widehat{\beta}_r-\beta)\widehat{P}^{cf}_{2,r}(z_\alpha)\right)\\
&= O(\ell^{-1/2}r^{-1}).
\end{aligned}\tag{4.26}$$

Since $\mathbf{W}_1$ has sufficiently many finite moments, we may write

$$\begin{aligned}
\mathrm{E}\left(A_{6,\ell,r}(\alpha)\right) &= -z_\alpha\phi(z_\alpha)\frac{1}{\beta}\ell^{-1}\,\mathrm{E}\left\{(\widehat{\beta}_r-\beta)\widehat{P}^2_{1,r}(z_\alpha)\right\} - z_\alpha\phi(z_\alpha)\frac{1}{\beta}\ell^{-2}\,\mathrm{E}\left\{(\widehat{\beta}_r-\beta)\left(\widehat{P}^{cf}_{2,r}(z_\alpha)\right)^2\right\} \\
&\quad + 2z_\alpha\phi(z_\alpha)\frac{1}{\beta}\ell^{-3/2}\,\mathrm{E}\left\{(\widehat{\beta}_r-\beta)\widehat{P}_{1,r}(z_\alpha)\widehat{P}^{cf}_{2,r}(z_\alpha)\right\} \\
&= -z_\alpha\phi(z_\alpha)\frac{1}{\beta}\ell^{-1}\,\mathrm{E}\left\{(\widehat{\beta}_r-\beta)\widehat{P}^2_{1,r}(z_\alpha)\right\} + O(\ell^{-3/2}).
\end{aligned}$$

Therefore, by (vi) of Lemma D.1,

$$\mathrm{E}\left(A_{6,\ell,r}(\alpha)\right) = O(\ell^{-1}r^{-1}) + O(\ell^{-3/2}). \tag{4.27}$$

Also,

$$\mathrm{E}\left(A_{7,\ell,r}(\alpha)\right) = z_\alpha\phi(z_\alpha)\ell^{-3/2}\,\mathrm{E}\left(\widehat{P}_{1,r}(z_\alpha)\widehat{P}^{cf}_{2,r}(z_\alpha)\right) = O(\ell^{-3/2}). \tag{4.28}$$

Finally, from (v) and (ix) of Lemma D.1 we have that

$$\begin{aligned}
\mathrm{E}\left(A_{8,\ell,r}(\alpha)\right) &= \ell^{-1/2}\phi(z_\alpha)\left(P_1'(z_\alpha)-P_1(z_\alpha)\right)\ \left\{z_\alpha\frac{1}{\beta}\mathrm{E}\left(\widehat{\beta}_r-\beta\right) - \ell^{-1/2}\frac{1}{\beta}\mathrm{E}\left((\widehat{\beta}_r-\beta)\,\widehat{P}_{1,r}(z_\alpha)\right)\right. \\
&\quad \left. +\ell^{-1}\frac{1}{\beta}\mathrm{E}\left((\widehat{\beta}_r-\beta)\,\widehat{P}^{cf}_{2,r}(z_\alpha)\right) + \ell^{-1}\,\mathrm{E}\left(\widehat{P}^{cf}_{2,r}(z_\alpha)\right)\right\}, \\
&= \ell^{-1/2}\phi(z_\alpha)\left(P_1'(z_\alpha)-P_1(z_\alpha)\right) \\
&\quad \times\left\{O(r^{-1}) - O(\ell^{-1/2}r^{-1}) + O(\ell^{-1}r^{-1}) + \ell^{-1}\left(P^{cf}_2(z_\alpha)+O(r^{-1})\right)\right\} \\
&= O(\ell^{-3/2}) + O(\ell^{-1/2}r^{-1})
\end{aligned} \tag{4.29}$$

and

$$\begin{aligned}
\mathrm{E}\left(A_{9,\ell,r}(\alpha)\right) &= \ell^{-1}\phi(z_\alpha)\left(P_2'(z_\alpha)-P_2(z_\alpha)\right)\ \left\{z_\alpha\frac{1}{\beta}\mathrm{E}\left(\widehat{\beta}_r-\beta\right) - \ell^{-1/2}\frac{1}{\beta}\mathrm{E}\left((\widehat{\beta}_r-\beta)\,\widehat{P}_{1,r}(z_\alpha)\right)\right. \\
&\quad \left. +\ell^{-1}\frac{1}{\beta}\mathrm{E}\left((\widehat{\beta}_r-\beta)\,\widehat{P}^{cf}_{2,r}(z_\alpha)\right) - \ell^{-1/2}\,\mathrm{E}\left(\widehat{P}_{1,r}(z_\alpha)\right) + \ell^{-1}\,\mathrm{E}\left(\widehat{P}^{cf}_{2,r}(z_\alpha)\right)\right\}, \\
&= \ell^{-1}\phi(z_\alpha)\left(P_2'(z_\alpha)-P_2(z_\alpha)\right)\left\{O(r^{-1}) - O(\ell^{-1/2}r^{-1}) + O(\ell^{-1}r^{-1})\right. \\
&\quad \left. -\ell^{-1/2}\left(P^{cf}_1(z_\alpha)+O(r^{-1})\right) + \ell^{-1}\left(P^{cf}_2(z_\alpha)+O(r^{-1})\right)\right\} \\
&= O(\ell^{-3/2}) + O(\ell^{-1/2}r^{-1}).
\end{aligned} \tag{4.30}$$

Concerning the Lagrange remainder terms $\delta_{j,\ell,r}$, note that it holds for any $\Theta_{\ell,r}$ between $\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{\ell,r,\alpha}$ and z_α that

$$\begin{aligned}
\mathrm{E}\left(\left|\delta_{1,\ell,r}\right|\right) &= \tfrac{1}{3!}\,\mathrm{E}\left\{\left|\Phi^{(3)}\left(\Theta_{\ell,r}\right)\right|\left|\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{\ell,r,\alpha}-z_\alpha\right|^3\right\} \\
&\le \tfrac{1}{3!}\sup_{-\infty<x<\infty}\left|\Phi^{(3)}(x)\right|\mathrm{E}\left\{\left|\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{\ell,r,\alpha}-z_\alpha\right|^3\right\}. \\
&= \tfrac{1}{3!}\sup_{-\infty<x<\infty}\left|(x^2-1)\phi(x)\right|\mathrm{E}\left\{\left|\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{\ell,r,\alpha}-z_\alpha\right|^3\right\}.
\end{aligned}$$

Now because x^2-1 is a polynomial of degree 2, we have that

$$\sup_{-\infty<x<\infty}\left|(x^2-1)\phi(x)\right| < \infty.$$

It therefore follows from result (ii) of Lemma D.2 that

$$\left|\mathrm{E}\left(\delta_{1,\ell,r}\right)\right| \leq \mathrm{E}\left(\left|\delta_{1,\ell,r}\right|\right) = O(r^{-3/2} + \ell^{-3/2}). \tag{4.31}$$

Also,

$$\begin{aligned}
\mathrm{E}\left(\left|\delta_{2,\ell,r}\right|\right) &= \tfrac{1}{2}\,\mathrm{E}\left\{\left|R_1''\left(\Theta_{\ell,r}\right)\right|\left(\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{\ell,r,\alpha} - z_\alpha\right)^2\right\} \\
&\leq \tfrac{1}{2}\sup_{-\infty<x<\infty}\left|R_1''(x)\right|\mathrm{E}\left\{\left(\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{\ell,r,\alpha} - z_\alpha\right)^2\right\} \\
&= \tfrac{1}{2}\sup_{-\infty<x<\infty}\left|\left\{(x^2-1)P_1(x) - 2xP_1'(x) + P_1''(x)\right\}\phi(x)\right|\mathrm{E}\left\{\left(\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{\ell,r,\alpha} - z_\alpha\right)^2\right\}
\end{aligned}$$

Now because $(x^2-1)P_1(x) - 2xP_1'(x) + P_1''(x)$ is a polynomial of degree 4 (P_1 is a polynomial of degree 2 by Theorem 3.1), we have that

$$\sup_{-\infty<x<\infty}\left|\left\{(x^2-1)P_1(x) - 2xP_1'(x) + P_1''(x)\right\}\phi(x)\right| < \infty.$$

Hence, by result (i) of Lemma D.2, we have

$$\ell^{-1/2}\left|\mathrm{E}\left(\delta_{2,\ell,r}\right)\right| \leq \ell^{-1/2}\,\mathrm{E}\left(\left|\delta_{2,\ell,r}\right|\right) = O(\ell^{-1/2}r^{-1} + \ell^{-3/2}). \tag{4.32}$$

Similarly,

$$\begin{aligned}
\mathrm{E}\left(\left|\delta_{3,\ell,r}\right|\right) &= \tfrac{1}{2}\,\mathrm{E}\left\{\left|R_2''\left(\Theta_{\ell,r}\right)\right|\left(\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{\ell,r,\alpha} - z_\alpha\right)^2\right\} \\
&\leq \tfrac{1}{2}\sup_{-\infty<x<\infty}\left|R_2''(x)\right|\mathrm{E}\left\{\left(\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{\ell,r,\alpha} - z_\alpha\right)^2\right\} \\
&= \tfrac{1}{2}\sup_{-\infty<x<\infty}\left|\left\{(x^2-1)P_2(x) - 2xP_2'(x) + P_2''(x)\right\}\phi(x)\right|\mathrm{E}\left\{\left(\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{\ell,r,\alpha} - z_\alpha\right)^2\right\}
\end{aligned}$$

Now $(x^2-1)P_2(x) - 2xP_2'(x) + P_2''(x)$ is a polynomial of degree 7 (cf. Theorem 3.1), which implies that

$$\sup_{-\infty<x<\infty}\left|\left\{(x^2-1)P_2(x) - 2xP_2'(x) + P_2''(x)\right\}\phi(x)\right| < \infty.$$

and therefore, by result (i) of Lemma D.2,

$$\ell^{-1}\left|\mathrm{E}\left(\delta_{3,\ell,r}\right)\right| \leq \ell^{-1}\,\mathrm{E}\left(\left|\delta_{3,\ell,r}\right|\right) = O(\ell^{-1}r^{-1} + \ell^{-3/2}). \tag{4.33}$$

From these results (4.21) becomes

$$\begin{aligned}
\mathrm{P}\left(\theta \in \mathscr{I}_H^N(\ell,\alpha)\right) = \mathrm{E}\left(\Psi(\widehat{\beta}_r\widetilde{\xi}_{\ell,r,\alpha})\right) &= 1 - \alpha - z_\alpha\phi(z_\alpha)\frac{\mathrm{E}(\widehat{\beta}_r - \beta)}{\beta} + \tfrac{1}{2}z_\alpha^3\phi(z_\alpha)\frac{\mathrm{E}(\widehat{\beta}_r - \beta)^2}{\beta^2} \\
&\quad + O(\ell^{-1/2}r^{-1} + \ell^{-3/2}) \\
&= 1 - \alpha + \frac{C_\theta(z_\alpha)}{r} + O(r^{-3/2} + \ell^{-1/2}r^{-1} + \ell^{-3/2}) \\
&= 1 - \alpha + \frac{C_\theta(z_\alpha)}{r} + O(\ell^{-3/2}),
\end{aligned}$$

where we have made use of the assumption that $r^{-1} = O(\ell^{-1})$. This proves (4.4).

If $C_\theta(z_\alpha) \neq 0$, note that for $\frac{2}{3} < \psi < 1$ one may write

$$\frac{C_\theta(z_\alpha)}{r} = \frac{C_\theta(z_\alpha)}{n} + O(n^{-(2-\psi)}).$$

Using this fact the coverage probability in (4.4) becomes

$$1 - \alpha + \frac{C_\theta(z_\alpha)}{n} + O(n^{-(2-\psi)} + n^{-3\psi/2}).$$

Inspection reveals that the error term is $O(n^{-3\psi/2})$ if $\frac{2}{3} < \psi \le \frac{4}{5}$ and $O(n^{-(2-\psi)})$ if $\frac{4}{5} \le \psi < 1$. The proof for $\psi = 1$ follows trivially from (4.4) by noting that $r^{-1} = (1-\gamma)^{-1}n^{-1} + O(n^{-2})$.

If $C_\theta(z_\alpha) = 0$, (4.5) follows directly from (4.4). □

4.3 Backwards percentile bound $\widehat{\mathscr{I}}_B^N$

As seen in Section 2.4.1, the standard *backwards percentile* $100(1-\alpha)\%$ upper confidence bound for θ is defined by

$$\widehat{\mathscr{I}}_B(\alpha) = \left(-\infty, \widehat{\theta}_n + n^{-1/2}\widehat{\beta}_n\widehat{\xi}_{1-\alpha}\right],$$

with $\widehat{\theta}_n$, $\widehat{\beta}_n$ and $\widehat{\xi}_{1-\alpha}$ as defined in Section 2.1. Recall that this standard confidence bound attains the nominal coverage probability $1-\alpha$ with a coverage error of order $O(n^{-1/2})$.

We now propose the following modified backwards percentile $100(1-\alpha)\%$ upper confidence bound for θ, which, under some regularity conditions, attains the nominal coverage probability $1-\alpha$ *with a coverage error of order* $O(n^{-1/2})$. This result is stated and proved in Theorem 4.2 below. However, we will show (in Section 7.2.2) that in some situations the coverage error of this bound reduces to $O(n^{-1})$.

New procedure
We suggest the following backwards percentile $100(1-\alpha)\%$ upper confidence bound for θ:

$$\widehat{\mathscr{I}}_B^N(m,\alpha) := \left(-\infty, \widehat{\theta}_\ell + \ell^{-1/2}\widehat{\beta}_r\widetilde{\xi}_{m,r,1-\alpha}\right],$$

where

$$\widetilde{\xi}_{m,r,1-\alpha} = z_{1-\alpha} + m^{-1/2}\widehat{P}^{cf}_{1,r}(z_{1-\alpha}) + m^{-1}\widehat{P}^{cf}_{2,r}(z_{1-\alpha}).$$

Theorem 4.2. *Under the assumptions of Theorem 4.1 it follows that*

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_B^N(\ell,\alpha)\right) = 1 - \alpha + \frac{K_1(z_\alpha)}{\ell^{1/2}} + \frac{K_2(z_\alpha)}{\ell} + \frac{C_\theta(z_\alpha)}{r} + O(\ell^{-3/2}),$$

where

$$K_1(z_\alpha) = -2P_1(z_\alpha)\phi(z_\alpha),$$

$$K_2(z_\alpha) = P_1(z_\alpha)K_1'(z_\alpha).$$

Further, if we choose $\ell = \lfloor \gamma n \rfloor$ *for some* $0 < \gamma < 1$*, then*

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_B^N(\ell,\alpha)\right) = 1 - \alpha + \frac{K_1(z_\alpha)}{(\gamma n)^{1/2}} + O(n^{-1}).$$

In the case where $K_1(z_\alpha) = K_2(z_\alpha) = 0$*, the results of Theorem 4.1 hold also for* $\widehat{\mathscr{I}}_B^N$.

Remark 4.1. In Section 7.2.2 we apply the results of this theorem to the case where the parameter of interest is the mean of a univariate population. We also derive exact expressions for the constants $C_\theta(z_\alpha)$, $K_1(z_\alpha)$ and $K_2(z_\alpha)$ appearing in this theorem. A case where $K_1(z_\alpha) = K_2(z_\alpha) = 0$ is given in that section.

Proof of Theorem 4.2. Noting that $\mathscr{X}_\ell$ and $\mathscr{X}_r$ are independent, and hence also $\widehat{\theta}_\ell$ and $\widehat{\beta}_r\widetilde{\xi}_{m,r,\alpha}$, we may apply the idea stated in (4.1) to rewrite the coverage probability of the confidence bound $\widehat{\mathscr{I}}_B^N(m,\alpha)$ as

$$\begin{aligned}\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_B^N(m,\alpha)\right) &= \mathrm{P}\left(\theta \le \widehat{\theta}_\ell + \ell^{-1/2}\widehat{\beta}_r\widetilde{\xi}_{m,r,1-\alpha}\right)\\ &= \mathrm{P}\left(\ell^{1/2}\left(\widehat{\theta}_\ell - \theta\right) \ge -\widehat{\beta}_r\widetilde{\xi}_{m,r,1-\alpha}\right)\\ &= \mathrm{E}\left(\Psi\left(-\widehat{\beta}_r\widetilde{\xi}_{m,r,1-\alpha}\right)\right),\end{aligned}$$

with Ψ defined as in (4.6). Analogous to (4.7), an Edgeworth expansion of $\Psi\left(-\widehat{\beta}_r\widetilde{\xi}_{m,r,1-\alpha}\right)$ is given by

$$\begin{aligned}\Psi\left(-\widehat{\beta}_r\widetilde{\xi}_{m,r,1-\alpha}\right) = 1 - &\left[\Phi\left(\frac{-\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,1-\alpha}\right) + \ell^{-1/2}R_1\left(\frac{-\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,1-\alpha}\right) + \ell^{-1}R_2\left(\frac{-\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,1-\alpha}\right)\right]\\ &+ O\left(\ell^{-3/2}\right),\end{aligned}$$

where $R_j(x) = P_j(x)\phi(x)$. Recall that the $P_j(x)$ are odd for even j and even for odd j, i.e., $R_1(-x) = R_1(x)$ and $R_2(-x) = -R_2(x)$ for all $x \in \mathbb{R}$. Using this, together with the fact that $\Phi(-x) = 1 - \Phi(x)$, we may simplify the expansion to

$$\Psi\left(-\widehat{\beta}_r\widetilde{\xi}_{m,r,1-\alpha}\right) = \Phi\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,1-\alpha}\right) - \ell^{-1/2}R_1\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,1-\alpha}\right) + \ell^{-1}R_2\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,1-\alpha}\right) + O\left(\ell^{-3/2}\right). \tag{4.34}$$

Now, as in the proof of Theorem 4.1, we will analyse each term in the above expression by means of Taylor expansion. From (4.16) we have that

$$\begin{aligned}&\Phi\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,1-\alpha}\right)\\ &= 1 - \alpha + z_{1-\alpha}\phi(z_{1-\alpha})\frac{(\widehat{\beta}_r - \beta)}{\beta} - \tfrac{1}{2}z_{1-\alpha}^3\phi(z_{1-\alpha})\frac{(\widehat{\beta}_r - \beta)^2}{\beta^2} - m^{-1/2}\phi(z_{1-\alpha})\widehat{P}_{1,r}(z_{1-\alpha})\\ &\quad + m^{-1}\phi(z_{1-\alpha})\widehat{P}_{1,r}(z_{1-\alpha})\left\{\widehat{P}'_{1,r}(z_{1-\alpha}) - z_{1-\alpha}\widehat{P}_{1,r}(z_{1-\alpha})\right\} - m^{-1}\phi(z_{1-\alpha})\widehat{P}_{2,r}(z_{1-\alpha})\\ &\quad + \sum_{i=1}^{7}A_{i,m,r}(1-\alpha) + \delta_{1,m,r},\end{aligned} \tag{4.35}$$

where $\delta_{1,m,r} = \frac{1}{3!}\Phi^{(3)}\left(\Theta_{m,r}\right)\left[\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,1-\alpha} - z_{1-\alpha}\right]^3$ for some $\Theta_{m,r}$ between $\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,1-\alpha}$ and $z_{1-\alpha}$ and the $A_{i,m,r}$ are defined in (4.9) through (4.15). Also, from (4.17),

$$\begin{aligned}\ell^{-1/2}R_1\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,1-\alpha}\right) &= \ell^{-1/2}\phi(z_{1-\alpha})P_1(z_{1-\alpha})\\ &\quad - \ell^{-1/2}m^{-1/2}\phi(z_{1-\alpha})\widehat{P}_{1,r}(z_{1-\alpha})\left(P'_1(z_{1-\alpha}) - z_{1-\alpha}P_1(z_{1-\alpha})\right)\\ &\quad + A_{8,m,r}(1-\alpha) + \ell^{-1/2}\delta_{2,m,r}\,,\end{aligned} \tag{4.36}$$

with $A_{8,m,r}$ defined in (4.18) and $\delta_{2,m,r} = \frac{1}{2}R_1''(\Theta_{m,r})\left[\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,1-\alpha} - z_{1-\alpha}\right]^2$ for some $\Theta_{m,r}$ between $\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,1-\alpha}$ and $z_{1-\alpha}$. Lastly, from (4.19),

$$\ell^{-1}R_2\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,1-\alpha}\right) = \ell^{-1}\phi(z_{1-\alpha})P_2(z_{1-\alpha}) + A_{9,m,r}(1-\alpha) + \ell^{-1}\delta_{3,m,r}\,, \tag{4.37}$$

where $\delta_{3,m,r} = \frac{1}{2}R_2''(\Theta_{m,r})\left[\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,1-\alpha} - z_{1-\alpha}\right]^2$ for some $\Theta_{m,r}$ between $\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,1-\alpha}$ and $z_{1-\alpha}$. $A_{9,m,r}$ is defined in (4.20). Now, using (4.35), (4.36) and (4.37), then (4.34) becomes

$$\begin{aligned}
\Psi\left(-\widehat{\beta}_r\widetilde{\xi}_{m,r,1-\alpha}\right) = {} & 1-\alpha + z_{1-\alpha}\phi(z_{1-\alpha})\frac{(\widehat{\beta}_r-\beta)}{\beta} - \tfrac{1}{2}z_{1-\alpha}^3\phi(z_{1-\alpha})\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2} \\
& - m^{-1/2}\phi(z_{1-\alpha})\widehat{P}_{1,r}(z_{1-\alpha}) \\
& + m^{-1}\phi(z_{1-\alpha})\widehat{P}_{1,r}(z_{1-\alpha})\left\{\widehat{P}'_{1,r}(z_{1-\alpha}) - z_{1-\alpha}\widehat{P}_{1,r}(z_{1-\alpha})\right\} \\
& - m^{-1}\phi(z_{1-\alpha})\widehat{P}_{2,r}(z_{1-\alpha}) - \ell^{-1/2}\phi(z_{1-\alpha})P_1(z_{1-\alpha}) \\
& + \ell^{-1/2}m^{-1/2}\phi(z_{1-\alpha})\widehat{P}_{1,r}(z_{1-\alpha})\left(P_1'(z_{1-\alpha}) - z_{1-\alpha}P_1(z_{1-\alpha})\right) \\
& + \ell^{-1}\phi(z_{1-\alpha})P_2(z_{1-\alpha}) + O(\ell^{-3/2}) \\
& + \sum_{i=1}^{7}A_{i,m,r}(1-\alpha) - A_{8,m,r}(1-\alpha) + A_{9,m,r}(1-\alpha) \\
& + \delta_{1,m,r} - \ell^{-1/2}\delta_{2,m,r} + \ell^{-1}\delta_{3,m,r}.
\end{aligned}$$

Setting $m = \ell$ yields

$$\begin{aligned}
\Psi\left(-\widehat{\beta}_r\widehat{\xi}_{\ell,r,1-\alpha}\right) = {} & 1-\alpha + z_{1-\alpha}\phi(z_{1-\alpha})\frac{(\widehat{\beta}_r-\beta)}{\beta} - \tfrac{1}{2}z_{1-\alpha}^3\phi(z_{1-\alpha})\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2} \\
& - \ell^{-1/2}\phi(z_{1-\alpha})\left\{\widehat{P}_{1,r}(z_{1-\alpha}) + P_1(z_{1-\alpha})\right\} \\
& - \ell^{-1}\phi(z_{1-\alpha})z_{1-\alpha}\widehat{P}_{1,r}(z_{1-\alpha})\left\{\widehat{P}_{1,r}(z_{1-\alpha}) + P_1(z_{1-\alpha})\right\} \\
& + \ell^{-1}\phi(z_{1-\alpha})\widehat{P}_{1,r}(z_{1-\alpha})\left\{\widehat{P}'_{1,r}(z_{1-\alpha}) + P_1'(z_{1-\alpha})\right\} \\
& - \ell^{-1}\phi(z_{1-\alpha})\left\{\widehat{P}_{2,r}(z_{1-\alpha}) - P_2(z_{1-\alpha})\right\} + O(\ell^{-3/2}) \\
& + \sum_{i=1}^{7}A_{i,\ell,r}(1-\alpha) - A_{8,\ell,r}(1-\alpha) + A_{9,\ell,r}(1-\alpha) \\
& + \delta_{1,\ell,r} - \ell^{-1/2}\delta_{2,\ell,r} + \ell^{-1}\delta_{3,\ell,r}.
\end{aligned}$$

Now, taking the expected value of the above expression, we have from (4.22)–(4.33) and from assumption (A7) of this theorem that

$$\begin{aligned}
\mathrm{E}\left\{\Psi\left(-\widehat{\beta}_r\widehat{\xi}_{\ell,r,1-\alpha}\right)\right\} = {} & 1-\alpha + z_{1-\alpha}\phi(z_{1-\alpha})\mathrm{E}\left(\frac{\widehat{\beta}_r-\beta}{\beta}\right) - \tfrac{1}{2}z_{1-\alpha}^3\phi(z_{1-\alpha})\mathrm{E}\left(\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2}\right) \\
& - \ell^{-1/2}\phi(z_{1-\alpha})\mathrm{E}\left\{\widehat{P}_{1,r}(z_{1-\alpha}) + P_1(z_{1-\alpha})\right\} \\
& - \ell^{-1}\phi(z_{1-\alpha})z_{1-\alpha}\,\mathrm{E}\left(\widehat{P}_{1,r}(z_{1-\alpha})\left\{\widehat{P}_{1,r}(z_{1-\alpha}) + P_1(z_{1-\alpha})\right\}\right) \\
& + \ell^{-1}\phi(z_{1-\alpha})\mathrm{E}\left(\widehat{P}_{1,r}(z_{1-\alpha})\left\{\widehat{P}'_{1,r}(z_{1-\alpha}) + P_1'(z_{1-\alpha})\right\}\right) \\
& + O(\ell^{-1/2}r^{-1} + \ell^{-3/2}).
\end{aligned}$$

Noting that ϕ, P_1 and $\widehat{P}_1$ are even functions (and consequently that P_1' and $\widehat{P}_1'$ are uneven functions), we have, since $z_{1-\alpha} = -z_\alpha$, that

$$\begin{aligned}
&\mathrm{E}\left\{\Psi\left(-\widehat{\beta}_r\widehat{\xi}_{\ell,r,1-\alpha}\right)\right\}\\
&= 1-\alpha - z_\alpha\phi(z_\alpha)\mathrm{E}\left(\frac{\widehat{\beta}_r-\beta}{\beta}\right) + \tfrac{1}{2}z_\alpha^3\phi(z_\alpha)\mathrm{E}\left(\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2}\right) - \ell^{-1/2}\phi(z_\alpha)\mathrm{E}\left\{\widehat{P}_{1,r}(z_\alpha)+P_1(z_\alpha)\right\}\\
&\quad + \ell^{-1}z_\alpha\phi(z_\alpha)\mathrm{E}\left(\widehat{P}_{1,r}(z_\alpha)\left\{\widehat{P}_{1,r}(z_\alpha)+P_1(z_\alpha)\right\}\right) - \ell^{-1}\phi(z_\alpha)\mathrm{E}\left(\widehat{P}_{1,r}(z_\alpha)\left\{\widehat{P}'_{1,r}(z_\alpha)+P'_1(z_\alpha)\right\}\right)\\
&\quad + O(\ell^{-3/2})\\
&= 1-\alpha - z_\alpha\phi(z_\alpha)\mathrm{E}\left(\frac{\widehat{\beta}_r-\beta}{\beta}\right) + \tfrac{1}{2}z_\alpha^3\phi(z_\alpha)\mathrm{E}\left(\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2}\right) - \ell^{-1/2}\phi(z_\alpha)\mathrm{E}\left\{\widehat{P}_{1,r}(z_\alpha)-P_1(z_\alpha)\right\}\\
&\quad - 2\ell^{-1/2}\phi(z_\alpha)P_1(z_\alpha) + \ell^{-1}z_\alpha\phi(z_\alpha)\mathrm{E}\left(\widehat{P}_{1,r}(z_\alpha)\left\{\widehat{P}_{1,r}(z_\alpha)-P_1(z_\alpha)\right\}\right)\\
&\quad + 2\ell^{-1}z_\alpha\phi(z_\alpha)P_1(z_\alpha)\mathrm{E}\left(\widehat{P}_{1,r}(z_\alpha)\right) - \ell^{-1}\phi(z_\alpha)\mathrm{E}\left(\widehat{P}_{1,r}(z_\alpha)\left\{\widehat{P}'_{1,r}(z_\alpha)-P'_1(z_\alpha)\right\}\right)\\
&\quad - 2\ell^{-1}\phi(z_\alpha)P'_1(z_\alpha)\mathrm{E}\left(\widehat{P}_{1,r}(z_\alpha)\right) + O(\ell^{-3/2}).
\end{aligned}$$

By assumption (A3) and results (i) and (ii) of Lemma D.1,

$$\begin{aligned}
\mathrm{E}\left\{\Psi\left(-\widehat{\beta}_r\widehat{\xi}_{\ell,r,1-\alpha}\right)\right\} &= 1-\alpha - z_\alpha\phi(z_\alpha)\mathrm{E}\left(\frac{\widehat{\beta}_r-\beta}{\beta}\right) + \tfrac{1}{2}z_\alpha^3\phi(z_\alpha)\mathrm{E}\left(\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2}\right)\\
&\quad - 2\ell^{-1/2}\phi(z_\alpha)P_1(z_\alpha) + 2\ell^{-1}z_\alpha\phi(z_\alpha)P_1^2(z_\alpha)\\
&\quad - 2\ell^{-1}\phi(z_\alpha)P'_1(z_\alpha)P_1(z_\alpha) + O(\ell^{-1/2}r^{-1} + \ell^{-1}r^{-1} + \ell^{-3/2})\\
&= 1-\alpha - z_\alpha\phi(z_\alpha)\mathrm{E}\left(\frac{\widehat{\beta}_r-\beta}{\beta}\right) + \tfrac{1}{2}z_\alpha^3\phi(z_\alpha)\mathrm{E}\left(\frac{(\widehat{\beta}_r-\beta)^2}{\beta^2}\right)\\
&\quad - 2\ell^{-1/2}\phi(z_\alpha)P_1(z_\alpha) - 2\ell^{-1}\phi(z_\alpha)P_1(z_\alpha)\left\{P'_1(z_\alpha - z_\alpha P_1(z_\alpha)\right\} + O(\ell^{-3/2}),
\end{aligned}$$

which is the required result. □

4.4 Equal-tailed percentile confidence intervals

One-sided upper and lower confidence bounds may be used to construct two-sided, equal-tailed confidence intervals. For example, in the notation of Chapter 2, the standard *hybrid* percentile $(1-2\alpha)$-level confidence interval for θ is given by

$$\widehat{\mathscr{I}}_H(\alpha)\backslash\widehat{\mathscr{I}}_H(1-\alpha) = \left(\widehat{\theta}_n - n^{-1/2}\widehat{\sigma}_n\widehat{\xi}_{1-\alpha},\, \widehat{\theta}_n - n^{-1/2}\widehat{\sigma}_n\widehat{\xi}_\alpha\right].$$

Likewise, the standard *backwards* percentile $(1-2\alpha)$-level confidence interval for θ is given by

$$\widehat{\mathscr{I}}_B(\alpha)\backslash\widehat{\mathscr{I}}_B(1-\alpha) = \left(\widehat{\theta}_n + n^{-1/2}\widehat{\sigma}_n\widehat{\xi}_\alpha,\, \widehat{\theta}_n + n^{-1/2}\widehat{\sigma}_n\widehat{\xi}_{1-\alpha}\right].$$

The order of coverage error of both these intervals is typically $O(n^{-1})$ (see Hall, 1988, p. 949). Moreover, Hall (1988) shows that intervals constructed from the bias-corrected and accelerated bias-corrected bounds also incur coverage errors of order $O(n^{-1})$.

An equal-tailed $(1-2\alpha)$-level confidence interval for θ based on the newly proposed *hybrid* percentile bound $\widehat{\mathscr{I}}_H^N$ is

$$\widehat{\mathscr{I}}_H^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_H^N(\ell,1-\alpha) = \left(\widehat{\theta}_\ell - \ell^{-1/2}\widehat{\beta}_r\widetilde{\xi}_{m,r,1-\alpha},\ \widehat{\theta}_\ell - \ell^{-1/2}\widehat{\beta}_r\widetilde{\xi}_{m,r,\alpha}\right].$$

Under the assumptions of Theorem 4.1, the coverage probability of this interval is typically

$$\begin{aligned}\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_H^N(\ell,1-\alpha)\right) &= \mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\right) - \mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,1-\alpha)\right)\\ &= 1-2\alpha + \frac{C_\theta(z_\alpha)-C_\theta(z_{1-\alpha})}{r} + O(\ell^{-3/2}).\end{aligned}$$

Noting that $C_\theta(z_\alpha) = -C_\theta(z_{1-\alpha})$, the above coverage probability becomes

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_H^N(\ell,1-\alpha)\right) = 1-2\alpha + \frac{2C_\theta(z_\alpha)}{r} + O(\ell^{-3/2}).$$

Hence, if we choose $\ell = \lfloor \gamma n^\psi \rfloor$ for some $\gamma > 0$ and $\frac{2}{3} < \psi \le 1$, this new hybrid percentile bound has coverage error $O(n^{-1})$, which is comparable to the coverage error of the standard hybrid percentile interval.

Similarly, an equal-tailed $(1-2\alpha)$-level confidence interval for θ based on the newly proposed *backwards* percentile bound $\widehat{\mathscr{I}}_B^N$ is given by

$$\widehat{\mathscr{I}}_B^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_B^N(\ell,1-\alpha) = \left(\widehat{\theta}_\ell + \ell^{-1/2}\widehat{\beta}_r\widetilde{\xi}_{m,r,\alpha},\ \widehat{\theta}_\ell + \ell^{-1/2}\widehat{\beta}_r\widetilde{\xi}_{m,r,1-\alpha}\right],$$

which by Theorem 4.2 typically has coverage probability

$$\begin{aligned}&\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_B^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_B^N(\ell,1-\alpha)\right)\\ &= 1-2\alpha + \frac{K_1(z_\alpha)-K_1(z_{1-\alpha})}{\ell^{1/2}} + \frac{K_2(z_\alpha)-K_2(z_{1-\alpha})}{\ell} + \frac{C_\theta(z_\alpha)-C_\theta(z_{1-\alpha})}{r} + O(\ell^{-3/2}).\end{aligned}$$

As $K_1(z_\alpha)$ is determined by $-\phi(z_\alpha)\mathrm{E}\{\widehat{P}_{1,r}(z_\alpha) + P_1(z_\alpha)\}$, we have that $K_1(z_\alpha) = K_1(z_{1-\alpha})$. Therefore,

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_B^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_B^N(\ell,1-\alpha)\right) = 1-2\alpha + \frac{2K_2(z_\alpha)}{\ell} + \frac{2C_\theta(z_\alpha)}{r} + O(\ell^{-3/2}),$$

since $K_2(z_\alpha) = -K_2(z_{1-\alpha})$. If we now choose $\ell = \lfloor \gamma n \rfloor$ for some $0 < \gamma < 1$, then

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_B^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_B^N(\ell,1-\alpha)\right) = 1-2\alpha + \frac{1}{n}\left(\frac{2K_2(z_\alpha)}{\gamma} + \frac{2C_\theta(z_\alpha)}{1-\gamma}\right) + O(n^{-3/2}).$$

Hence, for proper choices of ℓ the new equal-tailed backwards percentile interval has the same order of coverage error as that of the standard backwards percentile interval, namely $O(n^{-1})$. Note that, in contrast to one-sided confidence *bounds* constructed by means of the backwards method, additional assumptions (such as symmetry; see Section 7.2.2 for an example) are not needed to achieve this order of coverage error.

Chapter 5

New percentile-t confidence bounds

Using the same method of splitting the sample presented in Chapter 4, we derive two new *percentile-t* confidence bounds. We demonstrate that, typically, the new *hybrid* precentile-t bound leads to a coverage error of $O(n^{-3/2})$ and in some cases even to $O(n^{-2})$. This is an improvement over the standard percentile-t bootstrap bound $\widehat{\mathscr{J}}$ discussed in Chapter 2, which has coverage error of $O(n^{-1})$. We also show that the new *backwards* percentile-t bound has coverage error of $O(n^{-1/2})$, but in some cases $O(n^{-2})$.

5.1 Notation

Let $\mathscr{X}_n = \{\mathbf{X}_1,\mathbf{X}_2,\ldots,\mathbf{X}_n\}$ be a random sample consisting of p-dimensional random vectors drawn from an unknown p-dimensional distribution function F, depending on some unknown scalar parameter θ. In the context of the smooth function model described in Chapter 3, set $\mathbf{W}_k = (f_1(\mathbf{X}_k),\ldots,f_d(\mathbf{X}_k))$, $k = 1,\ldots,n$, where $f_1,\ldots,f_d$ are real-valued Borel-measurable functions on $\mathbb{R}^p$. Define $\boldsymbol{\nu} = \mathrm{E}(\mathbf{W}_1)$. Assume that the parameter of interest can be written in the form $\theta = g(\boldsymbol{\nu})$, where $g : \mathbb{R}^d \to \mathbb{R}$ is a known smooth, Borel-measurable function.

Our new method involves splitting the sample into two disjoint sets, say

$$\mathscr{W}_\ell := \{\mathbf{W}_1,\mathbf{W}_2,\ldots,\mathbf{W}_\ell\},$$

for some integer $2 \le \ell \le n-2$, and

$$\mathscr{W}_r := \{\mathbf{W}_{\ell+1},\mathbf{W}_{\ell+2},\ldots,\mathbf{W}_n\},$$

which has $r := n-\ell$ elements. Note that $\mathscr{W}_\ell$ and $\mathscr{W}_r$ can be viewed as two *independent* random samples of sizes ℓ and r drawn from the distribution F. Define the following two estimators for θ:

$$\widehat{\theta}_\ell = g\left(\bar{\mathbf{W}}_\ell\right) \quad \text{and} \quad \widehat{\theta}_r = g\left(\bar{\mathbf{W}}_r\right),$$

where $\bar{\mathbf{W}}_\ell = \ell^{-1}\sum_{k=1}^{\ell}\mathbf{W}_k$ and $\bar{\mathbf{W}}_r = r^{-1}\sum_{k=\ell+1}^{r}\mathbf{W}_k$. Throughout we will assume that the asymptotic variance of $\ell^{1/2}\widehat{\theta}_\ell$ is given by $\beta^2 = h^2(\boldsymbol{\nu})$ for some known smooth, Borel-measurable function $h:\mathbb{R}^d \to \mathbb{R}$. Two possible estimators for β are

$$\widehat{\beta}_\ell = h\left(\bar{\mathbf{W}}_\ell\right) \quad \text{and} \quad \widehat{\beta}_r = h\left(\bar{\mathbf{W}}_r\right).$$

Define

$$A_T(\mathbf{w}) := \frac{g(\mathbf{w}) - g(\boldsymbol{\nu})}{h(\mathbf{w})},$$

so that

$$A_T\left(\bar{\mathbf{W}}_\ell\right) = \frac{g(\bar{\mathbf{W}}_\ell) - g(\boldsymbol{\nu})}{h(\bar{\mathbf{W}}_\ell)} = \frac{\widehat{\theta}_\ell - \theta}{\widehat{\beta}_\ell}.$$

Note that $\mathrm{Var}(\ell^{1/2}A_T(\bar{\mathbf{W}}_\ell)) \to 1$ as $\ell \to \infty$. Assume that $\mathbf{W}_1$ satisfies Cramér's continuity condition, i.e.,

$$\limsup_{\|\mathbf{t}\|\to\infty} |\chi(\mathbf{t})| < 1, \tag{5.1}$$

where χ denotes the characteristic function of $\mathbf{W}_1$. Then, if g and h are sufficiently smooth in a neighbourhood of $\boldsymbol{\nu}$ and $\mathbf{W}_1$ has sufficiently many bounded moments, we have by Theorem 3.1 that

$$\begin{aligned} \mathrm{P}\left(\ell^{1/2}A_T\left(\bar{\mathbf{W}}_\ell\right) \le x\right) &= \mathrm{P}\left(\ell^{1/2}\left(\widehat{\theta}_\ell - \theta\right)/\widehat{\beta}_\ell \le x\right) \\ &= \Phi(x) + \ell^{-1/2}Q_1(x)\phi(x) + \ell^{-1}Q_2(x)\phi(x) + \ell^{-3/2}Q_3(x)\phi(x) + O(\ell^{-2}), \end{aligned} \tag{5.2}$$

uniformly in x, with Q_1, Q_2 and Q_3 determining the above Edgeworth expansion as described in Chapter 3. Note that the remainder term $O(\ell^{-2})$ is non-random.

Let $\mathscr{W}^*_{m,r} = \{\mathbf{W}^*_1, \mathbf{W}^*_2, \ldots, \mathbf{W}^*_m\}$ denote a *resample* of size m drawn randomly *with replacement* from $\mathscr{W}_r$, i.e., $\mathscr{W}^*_{m,r}$ is an m/r bootstrap sample drawn from $\mathscr{W}_r$. Now define the m/r bootstrap replications of θ and β as

$$\widehat{\theta}^*_{m,r} = g\left(\bar{\mathbf{W}}^*_{m,r}\right) \quad \text{and} \quad \widehat{\beta}^*_{m,r} = h\left(\bar{\mathbf{W}}^*_{m,r}\right)$$

respectively, where $\bar{\mathbf{W}}^*_{m,r} = m^{-1}\sum_{i=1}^{m}\mathbf{W}^*_i$. Now define the smooth function

$$\widehat{A}_T(\mathbf{w}) := \frac{g(\mathbf{w}) - g(\bar{\mathbf{W}}_r)}{h(\mathbf{w})},$$

so that

$$T^*_{m,r} := m^{1/2}\widehat{A}_T\left(\bar{\mathbf{W}}^*_{m,r}\right) = \frac{m^{1/2}\left(g\left(\bar{\mathbf{W}}^*_{m,r}\right) - g\left(\bar{\mathbf{W}}_r\right)\right)}{h\left(\bar{\mathbf{W}}^*_{m,r}\right)} = \frac{m^{1/2}\left(\widehat{\theta}^*_{m,r} - \widehat{\theta}_r\right)}{\widehat{\beta}^*_{m,r}}.$$

Note that $\mathrm{Var}_*(T^*_{m,r}) \to 1$ a.s. as $m \to \infty$. Assume that the function $\widehat{A}_T$ has sufficiently many bounded derivatives in a neighbourhood of $\boldsymbol{\nu}$, that $\mathbf{W}_1$ has sufficiently many finite moments and satisfies (5.1). Then, if we denote the α-level quantile of $T^*_{m,r}$ by $\widehat{\eta}_{m,r,\alpha}$, i.e.,

$$\widehat{\eta}_{m,r,\alpha} := \inf\left\{x : \mathrm{P}^*\left(T^*_{m,r} \le x\right) \ge \alpha\right\},$$

we have by Theorem 3.4 that

$$\widehat{\eta}_{m,r,\alpha} = z_\alpha + m^{-1/2}\widehat{Q}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{Q}^{cf}_{2,r}(z_\alpha) + m^{-3/2}\widehat{Q}^{cf}_{3,r}(z_\alpha) + O_p(m^{-2}),$$

uniformly in $\alpha \in (\varepsilon, 1-\varepsilon)$ for any $\varepsilon \in (0, \frac{1}{2})$, with $\widehat{Q}_1^{cf}$, $\widehat{Q}_2^{cf}$ and $\widehat{Q}_3^{cf}$ the appropriate Cornish-Fisher polynomials described in Chapter 3.

The two procedures introduced below are based on this Cornish-Fisher expansion of $\widehat{\eta}_{m,r,\alpha}$.

5.2 Hybrid percentile-t bound $\widehat{\mathscr{I}}_H^N$

Recall from Section 2.5 that the standard *hybrid percentile-t* $100(1-\alpha)\%$ upper confidence bound for θ is given by

$$\widehat{\mathscr{I}}(\alpha) = \left(-\infty, \widehat{\theta}_n - n^{-1/2}\widehat{\beta}_n\widehat{\eta}_\alpha\right],$$

with $\widehat{\theta}_n$, $\widehat{\beta}_n$ and $\widehat{\eta}_\alpha$ as defined in Section 2.1. Typically, this standard confidence bound has a coverage error of order $O(n^{-1})$.

We now propose the following modified hybrid percentile $100(1-\alpha)\%$ upper confidence bound for θ, which, under some regularity conditions, attains the nominal coverage probability $1-\alpha$ *with a coverage error of order* $O(n^{-3/2})$. This result is stated and proved in Theorem 5.1 below.

New procedure

We suggest the following hybrid percentile-t $100(1-\alpha)\%$ upper confidence bound for θ:

$$\widehat{\mathscr{I}}_H^N(m,\alpha) := \left(-\infty, \widehat{\theta}_\ell - \ell^{-1/2}\widehat{\beta}_\ell\widetilde{\eta}_{m,r,\alpha}\right],$$

where

$$\widetilde{\eta}_{m,r,\alpha} = z_\alpha + m^{-1/2}\widehat{Q}_{1,r}^{cf}(z_\alpha) + m^{-1}\widehat{Q}_{2,r}^{cf}(z_\alpha) + m^{-3/2}\widehat{Q}_{3,r}^{cf}(z_\alpha).$$

To investigate the asymptotic behaviour of this new confidence bound, the following assumptions are needed:

(B1) $\quad \mathrm{E}\left(\widehat{Q}_{1,r}(x) - Q_1(x)\right) = O(r^{-1}),$

(B2) $\quad \mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x) - Q_1(x)\right)^4\right\} = O(r^{-2}),$

(B3) $\quad \mathrm{E}\left\{\left(\widehat{Q}'_{1,r}(x) - Q'_1(x)\right)^i\right\} = O(r^{-1}), \qquad i = 1,2,$

(B4) $\quad \mathrm{E}\left\{\left(\widehat{Q}''_{1,r}(x) - Q''_1(x)\right)^i\right\} = O(r^{-1}), \qquad i = 1,2,$

(B5) $\quad \mathrm{E}\left\{\left(\widehat{Q}_{2,r}(x) - Q_2(x)\right)^i\right\} = O(r^{-1}), \qquad i = 1,2,$

(B6) $\quad \mathrm{E}\left\{\left(\widehat{Q}'_{2,r}(x) - Q'_2(x)\right)^i\right\} = O(r^{-1}), \qquad i = 1,2,$

(B7) $\quad \mathrm{E}\left\{\widehat{Q}_{3,r}(x) - Q_3(x)\right\} = O(r^{-1}).$

Theorem 5.1. *Suppose that* $\mathbf{W}_1$ *satisfies* (5.1) *and has sufficiently many finite moments and that* g *and* h *have sufficiently many continuous derivatives in an open neighborhood of* $\boldsymbol{\nu}$. *Then, if* $m = \ell = O(r)$ *and* $\ell \to \infty$ *as* $n \to \infty$, *we have under assumptions (B1)–(B7) that*

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\right) = 1-\alpha+\frac{D_\theta(z_\alpha)}{\ell^{1/2} r}+O(\ell^{-2}), \tag{5.3}$$

where $D_\theta(z_\alpha)$ *is the coefficient of* r^{-1} *in a power series expansion of*

$$\phi(z_\alpha)\mathrm{E}\left\{\widehat{Q}_{1,r}(z_\alpha)-Q_1(z_\alpha)\right\}.$$

Moreover, if we choose $\ell = \lfloor \gamma n^\psi \rfloor$ *for some* $\gamma > 0$ *and* $\frac{2}{3} < \psi < 1$, *then*

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\right) = \begin{cases} 1-\alpha+\dfrac{D_\theta(z_\alpha)}{\gamma^{1/2} n^{(2+\psi)/2}}+O(n^{-(4-\psi)/2}+n^{-2\psi}) & \text{if } D_\theta(z_\alpha)\neq 0, \\ 1-\alpha+O(n^{-2\psi}) & \text{if } D_\theta(z_\alpha)=0. \end{cases} \tag{5.4}$$

In the case where $\psi = 1$ *and* $0 < \gamma < 1$,

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\right) = 1-\alpha+\frac{D_\theta(z_\alpha)}{\gamma^{1/2}(1-\gamma) n^{3/2}}+O(n^{-2}).$$

Remark 5.1. As will be shown in Section 7.2.3, it might occur naturally that $D_\theta(z_\alpha) = 0$. In such cases the order of coverage error is reduced to $O(n^{-2\psi})$, for $\frac{2}{3} < \psi \le 1$.

Proof of Theorem 5.1. Noting that $\mathscr{W}_\ell$ and $\mathscr{W}_r$ are independent, and hence also $(\widehat{\theta}_\ell - \theta)/\widehat{\beta}_\ell$ and $\widetilde{\eta}_{m,r,\alpha}$, we may apply the idea stated in (4.1) to rewrite the coverage probability of the confidence bound $\widehat{\mathscr{I}}_H^N(m,\alpha)$ as

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(m,\alpha)\right) = \mathrm{P}\left(\theta \le \widehat{\theta}_\ell - \ell^{-1/2}\widehat{\beta}_\ell \widetilde{\eta}_{m,r,\alpha}\right) = \mathrm{P}\left(\frac{\ell^{1/2}\left(\widehat{\theta}_\ell-\theta\right)}{\widehat{\beta}_\ell} \ge \widetilde{\eta}_{m,r,\alpha}\right) = \mathrm{E}\left(\Psi_t\left(\widetilde{\eta}_{m,r,\alpha}\right)\right),$$

with

$$\begin{aligned} \Psi_t(x) &:= \mathrm{P}\left(\frac{\ell^{1/2}\left(\widehat{\theta}_\ell-\theta\right)}{\widehat{\beta}_\ell} \ge x\right) \\ &= 1-\mathrm{P}\left(\frac{\ell^{1/2}\left(\widehat{\theta}_\ell-\theta\right)}{\widehat{\beta}_\ell} \le x\right) \\ &= 1-\left[\Phi(x)+\ell^{-1/2}Q_1(x)\phi(x)+\ell^{-1}Q_2(x)\phi(x)+\ell^{-3/2}Q_3(x)\phi(x)\right]+O(\ell^{-2}), \end{aligned} \tag{5.5}$$

where the last step follows from (5.2). Since the error term does not depend on x, we may write

$$\begin{aligned} \Psi_t\left(\widetilde{\eta}_{m,r,\alpha}\right) = 1 &- \Phi\left(\widetilde{\eta}_{m,r,\alpha}\right)-\ell^{-1/2}Q_1\left(\widetilde{\eta}_{m,r,\alpha}\right)\phi\left(\widetilde{\eta}_{m,r,\alpha}\right) \\ &- \ell^{-1}Q_2\left(\widetilde{\eta}_{m,r,\alpha}\right)\phi\left(\widetilde{\eta}_{m,r,\alpha}\right)-\ell^{-3/2}Q_3\left(\widetilde{\eta}_{m,r,\alpha}\right)\phi\left(\widetilde{\eta}_{m,r,\alpha}\right)+O\left(\ell^{-2}\right). \end{aligned} \tag{5.6}$$

Note that the remainder term $O\left(\ell^{-2}\right)$ is non-random.

We will now inspect each term in (5.6) separately. Taylor expansion of Φ about z_α yields

$$\begin{aligned} \Phi\left(\widetilde{\eta}_{m,r,\alpha}\right) = \Phi(z_\alpha) &+ \phi(z_\alpha)\left\{\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right\}-\tfrac{1}{2}z_\alpha\phi(z_\alpha)\left\{\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right\}^2 \\ &+ \tfrac{1}{6}\left(z_\alpha^2-1\right)\phi(z_\alpha)\left\{\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right\}^3+\delta_{4,m,r}, \end{aligned}$$

where $\delta_{4,m,r} = \frac{1}{4!}\Phi^{(4)}\left(\Theta_{m,r}\right)\left(\widetilde{\eta}_{m,r,\alpha} - z_\alpha\right)^4$ is a Lagrange remainder term for some $\Theta_{m,r}$ between $\widetilde{\eta}_{m,r,\alpha}$ and z_α. Hence,

$$\begin{aligned}
\Phi\left(\widetilde{\eta}_{m,r,\alpha}\right) &= \alpha + \phi(z_\alpha)\left\{m^{-1/2}\widehat{Q}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{Q}^{cf}_{2,r}(z_\alpha) + m^{-3/2}\widehat{Q}^{cf}_{3,r}(z_\alpha)\right\} \\
&\quad - \tfrac{1}{2}z_\alpha\phi(z_\alpha)\left\{m^{-1/2}\widehat{Q}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{Q}^{cf}_{2,r}(z_\alpha) + m^{-3/2}\widehat{Q}^{cf}_{3,r}(z_\alpha)\right\}^2 \\
&\quad + \tfrac{1}{6}\left(z_\alpha^2 - 1\right)\phi(z_\alpha)\left\{m^{-1/2}\widehat{Q}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{Q}^{cf}_{2,r}(z_\alpha) + m^{-3/2}\widehat{Q}^{cf}_{3,r}(z_\alpha)\right\}^3 + \delta_{4,m,r} \\
&= \alpha + m^{-1/2}\phi(z_\alpha)\widehat{Q}^{cf}_{1,r}(z_\alpha) + m^{-1}\phi(z_\alpha)\left\{\widehat{Q}^{cf}_{2,r}(z_\alpha) - \tfrac{1}{2}z_\alpha\left(\widehat{Q}^{cf}_{1,r}(z_\alpha)\right)^2\right\} \\
&\quad + m^{-3/2}\phi(z_\alpha)\left\{\widehat{Q}^{cf}_{3,r}(z_\alpha) - z_\alpha\widehat{Q}^{cf}_{1,r}(z_\alpha)\widehat{Q}^{cf}_{2,r}(z_\alpha) + \tfrac{1}{6}\left(z_\alpha^2 - 1\right)\left(\widehat{Q}^{cf}_{1,r}(z_\alpha)\right)^3\right\} \\
&\quad + \sum_{k=1}^{6} B_{k,m,r}(\alpha) + \delta_{4,m,r}\,, \qquad (5.7)
\end{aligned}$$

where

$$\begin{aligned}
B_{1,m,r}(\alpha) = m^{-2}\phi(z_\alpha)\Big\{ &- z_\alpha\widehat{Q}^{cf}_{1,r}(z_\alpha)\widehat{Q}^{cf}_{3,r}(z_\alpha) - \tfrac{1}{2}z_\alpha\left(\widehat{Q}^{cf}_{2,r}(z_\alpha)\right)^2 \\
&+ \tfrac{1}{2}(z_\alpha^2 - 1)\left(\widehat{Q}^{cf}_{1,r}(z_\alpha)\right)^2\widehat{Q}^{cf}_{2,r}(z_\alpha)\Big\}, \qquad (5.8)
\end{aligned}$$

$$\begin{aligned}
B_{2,m,r}(\alpha) = m^{-5/2}\phi(z_\alpha)\Big\{ &- z_\alpha\widehat{Q}^{cf}_{2,r}(z_\alpha)Q^{cf}_{3,r}(z_\alpha) + \tfrac{1}{2}(z_\alpha^2 - 1)\left(\widehat{Q}^{cf}_{1,r}(z_\alpha)\right)^2\widehat{Q}^{cf}_{3,r}(z_\alpha) \\
&+ \tfrac{1}{6}(z_\alpha^2 - 1)\widehat{Q}^{cf}_{1,r}(z_\alpha)\left(\widehat{Q}^{cf}_{2,r}(z_\alpha)\right)^2\Big\}, \qquad (5.9)
\end{aligned}$$

$$\begin{aligned}
B_{3,m,r}(\alpha) = m^{-3}\phi(z_\alpha)\Big\{ &- \tfrac{1}{2}z_\alpha\left(\widehat{Q}^{cf}_{3,r}(z_\alpha)\right)^2 + \tfrac{1}{6}(z_\alpha^2 - 1)\left(\widehat{Q}^{cf}_{2,r}(z_\alpha)\right)^3 \\
&+ (z_\alpha^2 - 1)\widehat{Q}^{cf}_{1,r}(z_\alpha)\widehat{Q}^{cf}_{2,r}(z_\alpha)\widehat{Q}^{cf}_{3,r}(z_\alpha)\Big\}, \qquad (5.10)
\end{aligned}$$

$$B_{4,m,r}(\alpha) = m^{-7/2}\phi(z_\alpha)\left\{\tfrac{1}{2}(z_\alpha^2 - 1)\left(\widehat{Q}^{cf}_{2,r}(z_\alpha)\right)^2\widehat{Q}^{cf}_{3,r}(z_\alpha) + \tfrac{1}{2}(z_\alpha^2 - 1)\widehat{Q}^{cf}_{1,r}(z_\alpha)\left(\widehat{Q}^{cf}_{3,r}(z_\alpha)\right)^2\right\}, \qquad (5.11)$$

$$B_{5,m,r}(\alpha) = \tfrac{1}{2}m^{-4}(z_\alpha^2 - 1)\phi(z_\alpha)\widehat{Q}^{cf}_{2,r}(z_\alpha)\left(\widehat{Q}^{cf}_{3,r}(z_\alpha)\right)^2, \qquad (5.12)$$

$$B_{6,m,r}(\alpha) = \tfrac{1}{6}m^{-9/2}(z_\alpha^2 - 1)\phi(z_\alpha)\left(\widehat{Q}^{cf}_{3,r}(z_\alpha)\right)^3. \qquad (5.13)$$

Substituting the Cornish-Fisher polynomials $\widehat{Q}^{cf}_{1,r}$, $\widehat{Q}^{cf}_{2,r}$ and $\widehat{Q}^{cf}_{3,r}$ by their respective expressions in terms of Edgeworth polynomials (given explicitly in (3.9), (3.10) and (3.11)), then (5.7) becomes

$$\begin{aligned}
\Phi\left(\widetilde{\eta}_{m,r,\alpha}\right) &= \alpha - m^{-1/2}\phi(z_\alpha)\widehat{Q}_{1,r}(z_\alpha) \\
&+ m^{-1}\phi(z_\alpha)\left\{\widehat{Q}_{1,r}(z_\alpha)\widehat{Q}'_{1,r}(z_\alpha) - \tfrac{1}{2}z_\alpha\widehat{Q}^2_{1,r}(z_\alpha) - \widehat{Q}_{2,r}(z_\alpha) - \tfrac{1}{2}z_\alpha\widehat{Q}^2_{1,r}(z_\alpha)\right\} \\
&+ m^{-3/2}\phi(z_\alpha)\Big\{-\tfrac{1}{3}(z_\alpha^2 - 1)\widehat{Q}^3_{1,r}(z_\alpha) + \tfrac{3}{2}z_\alpha\widehat{Q}^2_{1,r}(z_\alpha)\widehat{Q}'_{1,r}(z_\alpha) - \tfrac{1}{2}\widehat{Q}^2_{1,r}(z_\alpha)\widehat{Q}''_{1,r}(z_\alpha) \\
&- z_\alpha\widehat{Q}_{1,r}(z_\alpha)\widehat{Q}_{2,r}(z_\alpha) + \widehat{Q}_{1,r}(z_\alpha)\widehat{Q}'_{2,r}(z_\alpha) - \widehat{Q}_{1,r}(z_\alpha)\left(\widehat{Q}'_{1,r}(z_\alpha)\right)^2 + \widehat{Q}'_{1,r}(z_\alpha)\widehat{Q}_{2,r}(z_\alpha) \\
&- \widehat{Q}_{3,r}(z_\alpha) + z_\alpha\widehat{Q}_{1,r}(z_\alpha)\left(\widehat{Q}_{1,r}(z_\alpha)\widehat{Q}'_{1,r}(z_\alpha) - \tfrac{1}{2}z_\alpha\widehat{Q}^2_{1,r}(z_\alpha) - \widehat{Q}_{2,r}(z_\alpha)\right) \\
&- \tfrac{1}{6}\left(z_\alpha^2 - 1\right)\widehat{Q}^3_{1,r}(z_\alpha)\Big\} + \sum_{k=1}^{6} B_{k,m,r}(\alpha) + \delta_{4,m,r}. \qquad (5.14)
\end{aligned}$$

By Taylor expansion of $Q_1(x)\phi(x)$ about z_α, the term $Q_1\left(\widetilde{\eta}_{m,r,\alpha}\right)\phi\left(\widetilde{\eta}_{m,r,\alpha}\right)$ in (5.6) becomes

$$
\begin{aligned}
&Q_1\left(\widetilde{\eta}_{m,r,\alpha}\right)\phi\left(\widetilde{\eta}_{m,r,\alpha}\right)\\
&= Q_1(z_\alpha)\phi(z_\alpha) + \phi(z_\alpha)\left\{Q_1'(z_\alpha) - z_\alpha Q_1(z_\alpha)\right\}\left(\widetilde{\eta}_{m,r,\alpha} - z_\alpha\right)\\
&\quad + \tfrac{1}{2}\phi(z_\alpha)\left\{(z_\alpha^2-1)Q_1(z_\alpha) - 2z_\alpha Q_1'(z_\alpha) + Q_1''(z_\alpha)\right\}\left(\widetilde{\eta}_{m,r,\alpha} - z_\alpha\right)^2 + \delta_{5,m,r}\\
&= Q_1(z_\alpha)\phi(z_\alpha) + \phi(z_\alpha)\left\{Q_1'(z_\alpha) - z_\alpha Q_1(z_\alpha)\right\}\left(m^{-1/2}\widehat{Q}_{1,r}^{cf}(z_\alpha) + m^{-1}\widehat{Q}_{2,r}^{cf}(z_\alpha) + m^{-3/2}\widehat{Q}_{3,r}^{cf}(z_\alpha)\right)\\
&\quad + \tfrac{1}{2}\phi(z_\alpha)\left\{(z_\alpha^2-1)Q_1(z_\alpha) - 2z_\alpha Q_1'(z_\alpha) + Q_1''(z_\alpha)\right\}\\
&\qquad\qquad \times\left(m^{-1/2}\widehat{Q}_{1,r}^{cf}(z_\alpha) + m^{-1}\widehat{Q}_{2,r}^{cf}(z_\alpha) + m^{-3/2}\widehat{Q}_{3,r}^{cf}(z_\alpha)\right)^2 + \delta_{5,m,r}\\
&= Q_1(z_\alpha)\phi(z_\alpha) - m^{-1/2}\phi(z_\alpha)\left\{Q_1'(z_\alpha) - z_\alpha Q_1(z_\alpha)\right\}\widehat{Q}_{1,r}(z_\alpha)\\
&\quad + m^{-1}\phi(z_\alpha)\left\{Q_1'(z_\alpha) - z_\alpha Q_1(z_\alpha)\right\}\left\{\widehat{Q}_{1,r}(z_\alpha)\widehat{Q}_{1,r}'(z_\alpha) - \tfrac{1}{2}z_\alpha\widehat{Q}_{1,r}^2(z_\alpha) - \widehat{Q}_{2,r}(z_\alpha)\right\}\\
&\quad + \tfrac{1}{2}m^{-1}\phi(z_\alpha)\left\{(z_\alpha^2-1)Q_1(z_\alpha) - 2z_\alpha Q_1'(z_\alpha) + Q_1''(z_\alpha)\right\}\widehat{Q}_{1,r}^2(z_\alpha) + \sum_{k=7}^{11} B_{k,m,r}(\alpha) + \delta_{5,m,r}\,,
\end{aligned}
\tag{5.15}
$$

where

$$B_{7,m,r}(\alpha) = m^{-3/2}\phi(z_\alpha)\left\{Q_1'(z_\alpha) - z_\alpha Q_1(z_\alpha)\right\}\widehat{Q}_{3,r}^{cf}(z_\alpha), \tag{5.16}$$

$$B_{8,m,r}(\alpha) = m^{-3/2}\phi(z_\alpha)\left\{(z_\alpha^2-1)Q_1(z_\alpha) - 2z_\alpha Q_1'(z_\alpha) + Q_1''(z_\alpha)\right\}\widehat{Q}_{1,r}^{cf}(z_\alpha)\widehat{Q}_{2,r}^{cf}(z_\alpha), \tag{5.17}$$

$$B_{9,m,r}(\alpha) = m^{-2}\phi(z_\alpha)\left\{(z_\alpha^2-1)Q_1(z_\alpha) - 2z_\alpha Q_1'(z_\alpha) + Q_1''(z_\alpha)\right\} \tag{5.18}$$

$$\times\left\{\tfrac{1}{2}\left(\widehat{Q}_{2,r}^{cf}(z_\alpha)\right)^2 + \widehat{Q}_{1,r}^{cf}(z_\alpha)\widehat{Q}_{3,r}^{cf}(z_\alpha)\right\}, \tag{5.19}$$

$$B_{10,m,r}(\alpha) = m^{-5/2}\phi(z_\alpha)\left\{(z_\alpha^2-1)Q_1(z_\alpha) - 2z_\alpha Q_1'(z_\alpha) + Q_1''(z_\alpha)\right\}\widehat{Q}_{2,r}^{cf}(z_\alpha)\widehat{Q}_{3,r}^{cf}(z_\alpha), \tag{5.20}$$

$$B_{11,m,r}(\alpha) = m^{-3}\phi(z_\alpha)\left\{(z_\alpha^2-1)Q_1(z_\alpha) - 2z_\alpha Q_1'(z_\alpha) + Q_1''(z_\alpha)\right\}\left(\widehat{Q}_{3,r}^{cf}(z_\alpha)\right)^2, \tag{5.21}$$

and $\delta_{5,m,r} = \frac{1}{3!}\frac{d^3}{dx^3}\left[Q_1(x)\phi(x)\right]_{x=\Theta_{m,r}}\left(\widetilde{\eta}_{m,r,\alpha} - z_\alpha\right)^3$ for some $\Theta_{m,r}$ between $\widetilde{\eta}_{m,r,\alpha}$ and z_α.

Also by Taylor expansion of $Q_2(x)\phi(x)$ about z_α, the term $Q_2\left(\widetilde{\eta}_{m,r,\alpha}\right)\phi\left(\widetilde{\eta}_{m,r,\alpha}\right)$ appearing in (5.6) becomes

$$
\begin{aligned}
&Q_2\left(\widetilde{\eta}_{m,r,\alpha}\right)\phi\left(\widetilde{\eta}_{m,r,\alpha}\right)\\
&= Q_2(z_\alpha)\phi(z_\alpha) + \phi(z_\alpha)\left\{Q_2'(z_\alpha) - z_\alpha Q_2(z_\alpha)\right\}\left(\widetilde{\eta}_{m,r,\alpha} - z_\alpha\right) + \delta_{6,m,r}\\
&= Q_2(z_\alpha)\phi(z_\alpha) + \phi(z_\alpha)\left\{Q_2'(z_\alpha) - z_\alpha Q_2(z_\alpha)\right\}\left(m^{-1/2}\widehat{Q}_{1,r}^{cf}(z_\alpha) + m^{-1}\widehat{Q}_{2,r}^{cf}(z_\alpha) + m^{-3/2}\widehat{Q}_{3,r}^{cf}(z_\alpha)\right)\\
&\quad + \delta_{6,m,r}\\
&= Q_2(z_\alpha)\phi(z_\alpha) - m^{-1/2}\phi(z_\alpha)\left\{Q_2'(z_\alpha) - z_\alpha Q_2(z_\alpha)\right\}\widehat{Q}_{1,r}(z_\alpha) + B_{12,m,r}(\alpha) + \delta_{6,m,r}\,,
\end{aligned}
\tag{5.22}
$$

where

$$B_{12,m,r}(\alpha) = \phi(z_\alpha)\left\{Q_2'(z_\alpha) - z_\alpha Q_2(z_\alpha)\right\}\left(m^{-1}\widehat{Q}_{2,r}^{cf}(z_\alpha) + m^{-3/2}\widehat{Q}_{3,r}^{cf}(z_\alpha)\right) \tag{5.23}$$

and $\delta_{6,m,r} = \frac{1}{2}\frac{d^2}{dx^2}\left[Q_2(x)\phi(x)\right]_{x=\Theta_{m,r}}\left(\widetilde{\eta}_{m,r,\alpha} - z_\alpha\right)^2$ for some $\Theta_{m,r}$ between $\widetilde{\eta}_{m,r,\alpha}$ and z_α.

Furthermore, by Taylor expansion about z_α, the term $Q_3\left(\widetilde{\eta}_{m,r,\alpha}\right)\phi\left(\widetilde{\eta}_{m,r,\alpha}\right)$ in (5.6) becomes

$$\begin{aligned}
&Q_3\left(\widetilde{\eta}_{m,r,\alpha}\right)\phi\left(\widetilde{\eta}_{m,r,\alpha}\right)\\
&= Q_3(z_\alpha)\phi(z_\alpha)+\phi(z_\alpha)\left\{Q_3'(z_\alpha)-z_\alpha Q_3(z_\alpha)\right\}\left(\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right)+\delta_{7,m,r}\\
&= Q_3(z_\alpha)\phi(z_\alpha)+\phi(z_\alpha)\left\{Q_3'(z_\alpha)-z_\alpha Q_3(z_\alpha)\right\}\left(m^{-1/2}\widehat{Q}^{cf}_{1,r}(z_\alpha)+m^{-1}\widehat{Q}^{cf}_{2,r}(z_\alpha)+m^{-3/2}\widehat{Q}^{cf}_{3,r}(z_\alpha)\right)\\
&\quad +\delta_{7,m,r}\\
&= Q_3(z_\alpha)\phi(z_\alpha)+B_{13,m,r}(\alpha)+\delta_{7,m,r}\,, \qquad (5.24)
\end{aligned}$$

where

$$B_{13,m,r}(\alpha)=\phi(z_\alpha)\left\{Q_3'(z_\alpha)-z_\alpha Q_3(z_\alpha)\right\}\left(m^{-1/2}\widehat{Q}^{cf}_{1,r}(z_\alpha)+m^{-1}\widehat{Q}^{cf}_{2,r}(z_\alpha)+m^{-3/2}\widehat{Q}^{cf}_{3,r}(z_\alpha)\right) \quad (5.25)$$

and $\delta_{7,m,r}=\frac{1}{2}\frac{d^2}{dx^2}\left[Q_3(x)\phi(x)\right]_{x=\Theta_{m,r}}\left(\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right)^2$ for some $\Theta_{m,r}$ between $\widetilde{\eta}_{m,r,\alpha}$ and z_α.

Now, substituting (5.14), (5.15), (5.22) and (5.24) in (5.6) we obtain

$$\begin{aligned}
&\Psi_t\left(\widetilde{\eta}_{m,r,\alpha}\right)\\
&=1-\alpha+m^{-1/2}\phi(z_\alpha)\widehat{Q}_{1,r}(z_\alpha)\\
&\quad -m^{-1}\phi(z_\alpha)\left\{\widehat{Q}_{1,r}(z_\alpha)\widehat{Q}'_{1,r}(z_\alpha)-\tfrac{1}{2}z_\alpha\widehat{Q}^2_{1,r}(z_\alpha)-\widehat{Q}_{2,r}(z_\alpha)-\tfrac{1}{2}z_\alpha\widehat{Q}^2_{1,r}(z_\alpha)\right\}\\
&\quad -m^{-3/2}\phi(z_\alpha)\left\{-\tfrac{1}{3}(z_\alpha^2-1)\widehat{Q}^3_{1,r}(z_\alpha)+\tfrac{3}{2}z_\alpha\widehat{Q}^2_{1,r}(z_\alpha)\widehat{Q}'_{1,r}(z_\alpha)-\tfrac{1}{2}\widehat{Q}^2_{1,r}(z_\alpha)\widehat{Q}''_{1,r}(z_\alpha)\right.\\
&\quad -z_\alpha\widehat{Q}_{1,r}(z_\alpha)\widehat{Q}_{2,r}(z_\alpha)+\widehat{Q}_{1,r}(z_\alpha)\widehat{Q}'_{2,r}(z_\alpha)-\widehat{Q}_{1,r}(z_\alpha)\left(\widehat{Q}'_{1,r}(z_\alpha)\right)^2+\widehat{Q}'_{1,r}(z_\alpha)\widehat{Q}_{2,r}(z_\alpha)\\
&\quad \left.-\widehat{Q}_{3,r}(z_\alpha)+z_\alpha\widehat{Q}_{1,r}(z_\alpha)\left(\widehat{Q}_{1,r}(z_\alpha)\widehat{Q}'_{1,r}(z_\alpha)-\tfrac{1}{2}z_\alpha\widehat{Q}^2_{1,r}(z_\alpha)-\widehat{Q}_{2,r}(z_\alpha)\right)-\tfrac{1}{6}\left(z_\alpha^2-1\right)\widehat{Q}^3_{1,r}(z_\alpha)\right\}\\
&\quad -\ell^{-1/2}Q_1(z_\alpha)\phi(z_\alpha)+\ell^{-1/2}m^{-1/2}\phi(z_\alpha)\left\{Q_1'(z_\alpha)-z_\alpha Q_1(z_\alpha)\right\}\widehat{Q}_{1,r}(z_\alpha)\\
&\quad -\ell^{-1/2}m^{-1}\phi(z_\alpha)\left\{Q_1'(z_\alpha)-z_\alpha Q_1(z_\alpha)\right\}\left\{\widehat{Q}_{1,r}(z_\alpha)\widehat{Q}'_{1,r}(z_\alpha)-\tfrac{1}{2}z_\alpha\widehat{Q}^2_{1,r}(z_\alpha)-\widehat{Q}_{2,r}(z_\alpha)\right\}\\
&\quad -\tfrac{1}{2}\ell^{-1/2}m^{-1}\phi(z_\alpha)\left\{(z_\alpha^2-1)Q_1(z_\alpha)-2z_\alpha Q_1'(z_\alpha)+Q_1''(z_\alpha)\right\}\widehat{Q}^2_{1,r}(z_\alpha)\\
&\quad -\ell^{-1}Q_2(z_\alpha)\phi(z_\alpha)+\ell^{-1}m^{-1/2}\phi(z_\alpha)\left\{Q_2'(z_\alpha)-z_\alpha Q_2(z_\alpha)\right\}\widehat{Q}_{1,r}(z_\alpha)-\ell^{-3/2}Q_3(z_\alpha)\phi(z_\alpha)\\
&\quad -\sum_{k=1}^{6}B_{k,m,r}(\alpha)-\ell^{-1/2}\sum_{k=7}^{11}B_{k,m,r}(\alpha)-\ell^{-1}B_{12,m,r}(\alpha)-\ell^{-3/2}B_{13,m,r}(\alpha)\\
&\quad -\delta_{4,m,r}-\ell^{-1/2}\delta_{5,m,r}-\ell^{-1}\delta_{6,m,r}-\ell^{-3/2}\delta_{7,m,r}+O(\ell^{-2}).
\end{aligned}$$

Setting $m=\ell$ and and grouping like terms we may rewrite the expression above in the following convenient form:

$$\begin{aligned}
&\Psi_t\left(\widetilde{\eta}_{\ell,r,\alpha}\right)\\
&=1-\alpha+\ell^{-1/2}\phi(z_\alpha)\left\{\widehat{Q}_{1,r}(z_\alpha)-Q_1(z_\alpha)\right\}\\
&\quad +\ell^{-1}\phi(z_\alpha)\left\{\widehat{Q}_{2,r}(z_\alpha)-Q_2(z_\alpha)\right\}-\ell^{-1}\phi(z_\alpha)\left\{\widehat{Q}'_{1,r}(z_\alpha)-Q_1'(z_\alpha)\right\}\widehat{Q}_{1,r}(z_\alpha)\\
&\quad +\ell^{-1}z_\alpha\phi(z_\alpha)\left\{\widehat{Q}_{1,r}(z_\alpha)-Q_1(z_\alpha)\right\}\widehat{Q}_{1,r}(z_\alpha)+\tfrac{1}{2}\ell^{-3/2}\phi(z_\alpha)(z_\alpha^2-1)\left\{\widehat{Q}_{1,r}(z_\alpha)-Q_1(z_\alpha)\right\}\widehat{Q}^2_{1,r}(z_\alpha)\\
&\quad -\ell^{-3/2}z_\alpha\phi(z_\alpha)\left\{\widehat{Q}_{1,r}(z_\alpha)-Q_1(z_\alpha)\right\}\left\{\widehat{Q}_{1,r}(z_\alpha)\widehat{Q}'_{1,r}(z_\alpha)-\tfrac{1}{2}z_\alpha\widehat{Q}^2_{1,r}(z_\alpha)-\widehat{Q}_{2,r}(z_\alpha)\right\}
\end{aligned}$$

$$
\begin{aligned}
&+\ell^{-3/2}\phi(z_\alpha)\{\widehat{Q}_{3,r}(z_\alpha)-Q_3(z_\alpha)\}+\ell^{-3/2}z_\alpha\phi(z_\alpha)\{\widehat{Q}_{2,r}(z_\alpha)-Q_2(z_\alpha)\}\widehat{Q}_{1,r}(z_\alpha)\\
&-\ell^{-3/2}\phi(z_\alpha)\left\{\widehat{Q}'_{2,r}(z_\alpha)-Q'_2(z_\alpha)\right\}\widehat{Q}_{1,r}(z_\alpha)-\ell^{-3/2}\phi(z_\alpha)\left\{\widehat{Q}'_{1,r}(z_\alpha)-Q'_1(z_\alpha)\right\}\widehat{Q}^2_{1,r}(z_\alpha)\\
&+\ell^{-3/2}\phi(z_\alpha)\left\{\widehat{Q}'_{1,r}(z_\alpha)-Q'_1(z_\alpha)\right\}\left\{\widehat{Q}_{1,r}(z_\alpha)\widehat{Q}'_{1,r}(z_\alpha)-\tfrac{1}{2}z_\alpha\widehat{Q}^2_{1,r}(z_\alpha)-\widehat{Q}_{2,r}(z_\alpha)\right\}\\
&+\tfrac{1}{2}\ell^{-3/2}\phi(z_\alpha)\left\{\widehat{Q}''_{1,r}(z_\alpha)-Q''_1(z_\alpha)\right\}\widehat{Q}^2_{1,r}(z_\alpha)\\
&-\sum_{k=1}^{6}B_{k,\ell,r}(\alpha)-\ell^{-1/2}\sum_{k=7}^{11}B_{k,\ell,r}(\alpha)-\ell^{-1}B_{12,\ell,r}(\alpha)-\ell^{-3/2}B_{13,\ell,r}(\alpha)\\
&-\delta_{4,\ell,r}-\ell^{-1/2}\delta_{5,\ell,r}-\ell^{-1}\delta_{6,\ell,r}-\ell^{-3/2}\delta_{7,\ell,r}+O(\ell^{-2}).
\end{aligned}
$$

The coverage probability of $\mathscr{I}^N_H(\ell,\alpha)$ may now be obtained by taking the expected value of $\Psi_t(\widetilde{\eta}_{\ell,r,\alpha})$. Applying the results of Lemma D.3 we may write this coverage probability as

$$
\begin{aligned}
&\mathrm{P}\left(\theta\in\mathscr{I}^N_H(\ell,\alpha)\right)=\mathrm{E}\left(\Psi_t(\widetilde{\eta}_{\ell,r,\alpha})\right)\\
&=1-\alpha+\ell^{-1/2}\phi(z_\alpha)\mathrm{E}\{\widehat{Q}_{1,r}(z_\alpha)-Q_1(z_\alpha)\}+O(\ell^{-1}r^{-1})+O(\ell^{-2})\\
&\quad-\sum_{k=1}^{6}\mathrm{E}\left(B_{k,\ell,r}(\alpha)\right)-\ell^{-1/2}\sum_{k=7}^{11}\mathrm{E}\left(B_{k,\ell,r}(\alpha)\right)-\ell^{-1}\mathrm{E}\left(B_{12,\ell,r}(\alpha)\right)-\ell^{-3/2}\mathrm{E}\left(B_{13,\ell,r}(\alpha)\right)\\
&\quad-\mathrm{E}\left(\delta_{4,\ell,r}\right)-\ell^{-1/2}\mathrm{E}\left(\delta_{5,\ell,r}\right)-\ell^{-1}\mathrm{E}\left(\delta_{6,\ell,r}\right)-\ell^{-3/2}\mathrm{E}\left(\delta_{7,\ell,r}\right). \qquad (5.26)
\end{aligned}
$$

We will now evaluate the expected values of the terms containing $B_{k,m,r}$, of which full expressions are given in (5.8)–(5.13), (5.16)–(5.21), (5.23) and (5.25). Note that the $B_{k,m,r}$ depend on the Cornish-Fisher polynomials $\widehat{Q}^{cf}_{1,r}$, $\widehat{Q}^{cf}_{2,r}$ and $\widehat{Q}^{cf}_{3,r}$, which in turn depend on moments of $\mathbf{W}_1$ up to the fifth order (see Theorem 3.1). Under the assumption of Theorem 5.1 that $\mathbf{W}_1$ has a sufficiently large number of finite moments, we may assume that all the expected values $\mathrm{E}\left(B_{k,\ell,r}(\alpha)\right)$, $k=1,2,\ldots,13$, will be finite. Hence we have that

$$
\begin{aligned}
&\mathrm{E}\left(B_{1,\ell,r}(\alpha)\right)=O(\ell^{-2}), && \mathrm{E}\left(B_{2,\ell,r}(\alpha)\right)=O(\ell^{-5/2}), && \mathrm{E}\left(B_{3,\ell,r}(\alpha)\right)=O(\ell^{-3}),\\
&\mathrm{E}\left(B_{4,\ell,r}(\alpha)\right)=O(\ell^{-7/2}), && \mathrm{E}\left(B_{5,\ell,r}(\alpha)\right)=O(\ell^{-4}), && \mathrm{E}\left(B_{6,\ell,r}(\alpha)\right)=O(\ell^{-9/2}),\\
&\mathrm{E}\left(B_{7,\ell,r}(\alpha)\right)=O(\ell^{-3/2}), && \mathrm{E}\left(B_{8,\ell,r}(\alpha)\right)=O(\ell^{-3/2}), && \mathrm{E}\left(B_{9,\ell,r}(\alpha)\right)=O(\ell^{-2}),\\
&\mathrm{E}\left(B_{10,\ell,r}(\alpha)\right)=O(\ell^{-5/2}), && \mathrm{E}\left(B_{11,\ell,r}(\alpha)\right)=O(\ell^{-3}), && \mathrm{E}\left(B_{12,\ell,r}(\alpha)\right)=O(\ell^{-1}),\\
&\mathrm{E}\left(B_{13,\ell,r}(\alpha)\right)=O(\ell^{-1/2}). && && \qquad (5.27)
\end{aligned}
$$

Substituting these results in (5.26) yields

$$
\begin{aligned}
\mathrm{P}\left(\theta\in\mathscr{I}^N_H(\ell,\alpha)\right)&=1-\alpha+\ell^{-1/2}\phi(z_\alpha)\mathrm{E}\{\widehat{Q}_{1,r}(z_\alpha)-Q_1(z_\alpha)\}+O(\ell^{-1}r^{-1})+O(\ell^{-2})\\
&\quad-\mathrm{E}\left(\delta_{4,\ell,r}\right)-\ell^{-1/2}\mathrm{E}\left(\delta_{5,\ell,r}\right)-\ell^{-1}\mathrm{E}\left(\delta_{6,\ell,r}\right)-\ell^{-3/2}\mathrm{E}\left(\delta_{7,\ell,r}\right). \qquad (5.28)
\end{aligned}
$$

We now treat the Lagrange remainder terms $\delta_{j,\ell,r}$. It holds for any $\Theta_{\ell,r}$ between $\widetilde{\eta}_{\ell,r,\alpha}$ and z_α that

$$
\begin{aligned}
\left|\mathrm{E}\left(\delta_{4,\ell,r}\right)\right|\le\mathrm{E}\left|\delta_{4,\ell,r}\right|&=\tfrac{1}{4!}\mathrm{E}\left\{\left|\Phi^{(4)}\left(\Theta_{\ell,r}\right)\right|\left(\widetilde{\eta}_{\ell,r,\alpha}-z_\alpha\right)^4\right\}\\
&\le\tfrac{1}{4!}\sup_{-\infty<x<\infty}\left|\Phi^{(4)}(x)\right|\mathrm{E}\left\{\left(\widetilde{\eta}_{\ell,r,\alpha}-z_\alpha\right)^4\right\}.\\
&\le\tfrac{1}{4!}\sup_{-\infty<x<\infty}\left|(x^3-3x)\phi(x)\right|\mathrm{E}\left\{\left(\widetilde{\eta}_{\ell,r,\alpha}-z_\alpha\right)^4\right\}.
\end{aligned}
$$

Note that

$$\sup_{-\infty<x<\infty} \left|(x^3-3x)\phi(x)\right| < \infty.$$

It therefore follows from result (iii) of Lemma D.4 that

$$\mathrm{E}\left(\delta_{4,\ell,r}\right) \leq \left|\mathrm{E}\left(\delta_{4,\ell,r}\right)\right| = O(\ell^{-2}). \tag{5.29}$$

Also,

$$\begin{aligned}
&\left|\mathrm{E}\left(\delta_{5,\ell,r}\right)\right| \leq \mathrm{E}\left|\delta_{5,\ell,r}\right| \\
&= \tfrac{1}{3!}\,\mathrm{E}\left\{\left|\frac{d^3}{dx^3}\left[Q_1(x)\phi(x)\right]_{x=\Theta_{m,r}}\right|\left|\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right|^3\right\} \\
&\leq \tfrac{1}{3!}\sup_{-\infty<s<\infty}\left|\frac{d^3}{dx^3}\left[Q_1(x)\phi(x)\right]_{x=s}\right|\mathrm{E}\left\{\left|\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right|^3\right\} \\
&= \tfrac{1}{3!}\sup_{-\infty<x<\infty}\left|\left\{-(x^3-3x)Q_1(x)+3(x^2-1)Q_1'(x)-3xQ_1''(x)+Q_1^{(3)}(x)\right\}\phi(x)\right|\mathrm{E}\left\{\left|\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right|^3\right\}
\end{aligned}$$

Since

$$\sup_{-\infty<x<\infty}\left|\left\{-(x^3-3x)Q_1(x)+3(x^2-1)Q_1'(x)-3xQ_1''(x)+Q_1^{(3)}(x)\right\}\phi(x)\right| < \infty,$$

we have by result (ii) of Lemma D.4 that

$$\ell^{-1/2}\,\mathrm{E}\left(\delta_{5,\ell,r}\right) \leq \ell^{-1/2}\left|\mathrm{E}\left(\delta_{5,\ell,r}\right)\right| = O(\ell^{-2}). \tag{5.30}$$

Similarly,

$$\begin{aligned}
&\left|\mathrm{E}\left(\delta_{6,\ell,r}\right)\right| \leq \mathrm{E}\left|\delta_{6,\ell,r}\right| \\
&= \tfrac{1}{2}\,\mathrm{E}\left\{\left|\frac{d^2}{dx^2}\left[Q_2(x)\phi(x)\right]_{x=\Theta_{m,r}}\right|\left(\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right)^2\right\} \\
&= \tfrac{1}{2}\sup_{-\infty<s<\infty}\left|\frac{d^2}{dx^2}\left[Q_2(x)\phi(x)\right]_{x=s}\right|\mathrm{E}\left\{\left(\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right)^2\right\} \\
&= \tfrac{1}{2}\sup_{-\infty<x<\infty}\left|\left\{(x^2-1)Q_2(x)-2xQ_2'(x)+Q_2''(x)\right\}\phi(x)\right|\mathrm{E}\left\{\left(\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right)^2\right\}.
\end{aligned}$$

Noting that

$$\sup_{-\infty<x<\infty}\left|\left\{(x^2-1)Q_2(x)-2xQ_2'(x)+Q_2''(x)\right\}\phi(x)\right| < \infty,$$

result (i) of Lemma D.4 yields

$$\ell^{-1}\,\mathrm{E}\left(\delta_{6,\ell,r}\right) \leq \ell^{-1}\left|\mathrm{E}\left(\delta_{6,\ell,r}\right)\right| = O(\ell^{-2}). \tag{5.31}$$

In exactly the same way it can be proved that

$$\ell^{-3/2}\,\mathrm{E}\left(\delta_{7,\ell,r}\right) = O(\ell^{-5/2}). \tag{5.32}$$

Combining results (5.29)–(5.32) with (5.28), we have that

$$\begin{aligned}
\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\right) &= 1-\alpha+\ell^{-1/2}\phi(z_\alpha)\mathrm{E}\left\{\widehat{Q}_{1,r}(z_\alpha)-Q_1(z_\alpha)\right\}+O(\ell^{-1}r^{-1}+\ell^{-2})\\
&= 1-\alpha+\frac{D_\theta(z_\alpha)}{\ell^{1/2}r}+O(\ell^{-1/2}r^{-3/2}+\ell^{-1}r^{-1}+\ell^{-2})\\
&= 1-\alpha+\frac{D_\theta(z_\alpha)}{\ell^{1/2}r}+O(\ell^{-2}),
\end{aligned}$$

where we use $r^{-1}=O(\ell^{-1})$, which proves the first part of the theorem.

To obtain the remaining results, note that, if $D_\theta(z_\alpha)\neq 0$, one may write for $\frac{2}{3}<\psi<1$ that

$$\frac{D_\theta(z_\alpha)}{\ell^{1/2}r}=\frac{D_\theta(z_\alpha)}{\gamma^{1/2}n^{(2+\psi)/2}}+O(n^{-(4-\psi)/2}).$$

Hence, the coverage probability in (5.3) becomes

$$1-\alpha+\frac{D_\theta(z_\alpha)}{\gamma^{1/2}n^{(2+\psi)/2}}+O(n^{-(4-\psi)/2}+n^{-2\psi}).$$

Inspection reveals that the error term reduces to $O(n^{-2\psi})$ if $\frac{2}{3}<\psi\leq\frac{4}{5}$ and to $O(n^{-(4-\psi)/2})$ if $\frac{4}{5}\leq\psi<1$. The proof for $\psi=1$ follows readily from (5.3) by noting that $\ell^{-1/2}r^{-1}=\gamma^{-1/2}(1-\gamma)^{-1}n^{-3/2}+O(n^{-5/2})$. If $D_\theta(z_\alpha)=0$, (5.4) follows immediately from (5.3). □

5.3 Backwards percentile-t bound $\widehat{\mathscr{I}}_B^N$

We propose the following backwards percentile $100(1-\alpha)\%$ upper confidence bound for θ, which under standard conditions attains the nominal coverage probability $1-\alpha$ *with a coverage error of order* $O(n^{-1/2})$. This result is stated and proved in the Theorem 5.2 below.

New procedure
We suggest the following hybrid percentile-t $100(1-\alpha)\%$ upper confidence bound for θ:

$$\widehat{\mathscr{I}}_B^N(m,\alpha):=\left(-\infty,\widehat{\theta}_\ell+\ell^{-1/2}\widehat{\beta}_\ell\widetilde{\eta}_{m,r,1-\alpha}\right],$$

where

$$\widetilde{\eta}_{m,r,1-\alpha}=z_{1-\alpha}+m^{-1/2}\widehat{Q}^{cf}_{1,r}(z_{1-\alpha})+m^{-1}\widehat{Q}^{cf}_{2,r}(z_{1-\alpha})+m^{-3/2}\widehat{Q}^{cf}_{3,r}(z_{1-\alpha}).$$

Theorem 5.2. *Under the assumptions of Theorem 5.1 it follows that*

$$\mathrm{P}\left(\theta\in\widehat{\mathscr{I}}_B^N(\ell,\alpha)\right)=1-\alpha+\frac{K_3(z_\alpha)}{\ell^{1/2}}+\frac{K_4(z_\alpha)}{\ell}+\frac{K_5(z_\alpha)}{\ell^{3/2}}-\frac{D_\theta(z_\alpha)}{\ell^{1/2}r}+O(\ell^{-2}),$$

where

$$\begin{aligned}
K_3(z_\alpha)&=-2Q_1(z_\alpha)\phi(z_\alpha),\\
K_4(z_\alpha)&=Q_1(z_\alpha)K_3'(z_\alpha),\\
K_5(z_\alpha)&=\tfrac{1}{2}Q_1^2(z_\alpha)K_3''(z_\alpha)+Q_2^{cf}(z_\alpha)K_3'(z_\alpha)-2Q_3(z_\alpha)\phi(z_\alpha).
\end{aligned}$$

Furthermore, if we choose $\ell = \lfloor \gamma n \rfloor$ *for some* $0 < \gamma < 1$*, then*

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{J}}_B^N(\ell,\alpha)\right) = 1-\alpha + \frac{K_3(z_\alpha)}{(\gamma n)^{1/2}} + O(n^{-1}).$$

In the case where $K_3(z_\alpha) = K_4(z_\alpha) = K_5(z_\alpha) = 0$*, all the results of Theorem 5.1 hold for* $\widehat{\mathscr{J}}_B^N$*.*

Proof. Noting that $\mathscr{X}_\ell$ and $\mathscr{X}_r$ are independent, and hence also $(\widehat{\theta}_\ell - \theta)/\widehat{\beta}_\ell$ and $\widetilde{\eta}_{m,r,1-\alpha}$, we may apply the idea stated in (4.1) to rewrite the coverage probability of the confidence bound $\widehat{\mathscr{J}}_B^N(m,\alpha)$ as

$$\begin{aligned}
\mathrm{P}\left(\theta \in \widehat{\mathscr{J}}_B^N(m,\alpha)\right) &= \mathrm{P}\left(\theta \le \widehat{\theta}_\ell + \ell^{-1/2}\widehat{\beta}_\ell \widetilde{\eta}_{m,r,1-\alpha}\right) \\
&= \mathrm{P}\left(\frac{\ell^{1/2}\left(\widehat{\theta}_\ell - \theta\right)}{\widehat{\beta}_\ell} \ge -\widetilde{\eta}_{m,r,1-\alpha}\right) \\
&= \mathrm{E}\left(\Psi_t\left(-\widetilde{\eta}_{m,r,1-\alpha}\right)\right),
\end{aligned}$$

with Ψ_t as defined in (5.5). Since the error term in (5.5) does not depend on x, we may write

$$\begin{aligned}
\Psi_t\left(-\widetilde{\eta}_{m,r,1-\alpha}\right) = 1 &- \Phi\left(-\widetilde{\eta}_{m,r,1-\alpha}\right) - \ell^{-1/2}Q_1\left(-\widetilde{\eta}_{m,r,1-\alpha}\right)\phi\left(-\widetilde{\eta}_{m,r,1-\alpha}\right) \\
&- \ell^{-1}Q_2\left(-\widetilde{\eta}_{m,r,1-\alpha}\right)\phi\left(-\widetilde{\eta}_{m,r,1-\alpha}\right) - \ell^{-3/2}Q_3\left(-\widetilde{\eta}_{m,r,1-\alpha}\right)\phi\left(-\widetilde{\eta}_{m,r,1-\alpha}\right) \\
&+ O\left(\ell^{-2}\right).
\end{aligned}$$

Recall from Theorem 3.1 that the $Q_j(x)$ are odd for even j and even for odd j. Now since $\phi(x)$ is even, it follows that $Q_j(x)\phi(x)$ is odd if j is even and even if j is odd. Using this, together with the fact that $\Phi(-x) = 1-\Phi(x)$, we may simplify the above expansion to

$$\begin{aligned}
\Psi_t\left(-\widetilde{\eta}_{m,r,1-\alpha}\right) = \Phi\left(\widetilde{\eta}_{m,r,1-\alpha}\right) &- \ell^{-1/2}Q_1\left(\widetilde{\eta}_{m,r,1-\alpha}\right)\phi\left(\widetilde{\eta}_{m,r,1-\alpha}\right) \\
&+ \ell^{-1}Q_2\left(\widetilde{\eta}_{m,r,1-\alpha}\right)\phi\left(\widetilde{\eta}_{m,r,1-\alpha}\right) - \ell^{-3/2}Q_3\left(\widetilde{\eta}_{m,r,1-\alpha}\right)\phi\left(\widetilde{\eta}_{m,r,1-\alpha}\right) \\
&+ O\left(\ell^{-2}\right).
\end{aligned} \tag{5.33}$$

Substituting (5.14), (5.15), (5.22) and (5.24) in (5.33) we obtain

$$\begin{aligned}
&\Psi_t\left(-\widetilde{\eta}_{m,r,1-\alpha}\right) \\
&= 1-\alpha - m^{-1/2}\phi(z_{1-\alpha})\widehat{Q}_{1,r}(z_{1-\alpha}) \\
&\quad + m^{-1}\phi(z_{1-\alpha})\left\{\widehat{Q}_{1,r}(z_{1-\alpha})\widehat{Q}'_{1,r}(z_{1-\alpha}) - \tfrac{1}{2}z_{1-\alpha}\widehat{Q}^2_{1,r}(z_{1-\alpha}) - \widehat{Q}_{2,r}(z_{1-\alpha}) - \tfrac{1}{2}z_{1-\alpha}\widehat{Q}^2_{1,r}(z_{1-\alpha})\right\} \\
&\quad + m^{-3/2}\phi(z_{1-\alpha})\Big\{-\tfrac{1}{3}(z^2_{1-\alpha}-1)\widehat{Q}^3_{1,r}(z_{1-\alpha}) + \tfrac{3}{2}z_{1-\alpha}\widehat{Q}^2_{1,r}(z_{1-\alpha})\widehat{Q}'_{1,r}(z_{1-\alpha}) \\
&\qquad - \tfrac{1}{2}\widehat{Q}^2_{1,r}(z_{1-\alpha})\widehat{Q}''_{1,r}(z_{1-\alpha}) - z_{1-\alpha}\widehat{Q}_{1,r}(z_{1-\alpha})\widehat{Q}_{2,r}(z_{1-\alpha}) + \widehat{Q}_{1,r}(z_{1-\alpha})\widehat{Q}'_{2,r}(z_{1-\alpha}) \\
&\qquad - \widehat{Q}_{1,r}(z_{1-\alpha})\left(\widehat{Q}'_{1,r}(z_{1-\alpha})\right)^2 + \widehat{Q}'_{1,r}(z_{1-\alpha})\widehat{Q}_{2,r}(z_{1-\alpha}) - \widehat{Q}_{3,r}(z_{1-\alpha}) \\
&\qquad + z_{1-\alpha}\widehat{Q}_{1,r}(z_{1-\alpha})\left(\widehat{Q}_{1,r}(z_{1-\alpha})\widehat{Q}'_{1,r}(z_{1-\alpha}) - \tfrac{1}{2}z_{1-\alpha}\widehat{Q}^2_{1,r}(z_{1-\alpha}) - \widehat{Q}_{2,r}(z_{1-\alpha})\right) \\
&\qquad - \tfrac{1}{6}\left(z^2_{1-\alpha}-1\right)\widehat{Q}^3_{1,r}(z_{1-\alpha})\Big\} \\
&\quad - \ell^{-1/2}Q_1(z_{1-\alpha})\phi(z_{1-\alpha}) + \ell^{-1/2}m^{-1/2}\phi(z_{1-\alpha})\left\{Q'_1(z_{1-\alpha}) - z_{1-\alpha}Q_1(z_{1-\alpha})\right\}\widehat{Q}_{1,r}(z_{1-\alpha}) \\
&\quad - \ell^{-1/2}m^{-1}\phi(z_{1-\alpha})\left\{Q'_1(z_{1-\alpha}) - z_{1-\alpha}Q_1(z_{1-\alpha})\right\}
\end{aligned}$$

$$\times \left\{\widehat{Q}_{1,r}(z_{1-\alpha})\widehat{Q}'_{1,r}(z_{1-\alpha}) - \tfrac{1}{2} z_{1-\alpha}\widehat{Q}^2_{1,r}(z_{1-\alpha}) - \widehat{Q}_{2,r}(z_{1-\alpha})\right\}$$
$$- \tfrac{1}{2}\ell^{-1/2} m^{-1}\phi(z_{1-\alpha})\left\{(z^2_{1-\alpha}-1)Q_1(z_{1-\alpha}) - 2z_{1-\alpha}Q'_1(z_{1-\alpha}) + Q''_1(z_{1-\alpha})\right\}\widehat{Q}^2_{1,r}(z_{1-\alpha})$$
$$+ \ell^{-1}Q_2(z_{1-\alpha})\phi(z_{1-\alpha}) - \ell^{-1}m^{-1/2}\phi(z_{1-\alpha})\left\{Q'_2(z_{1-\alpha}) - z_{1-\alpha}Q_2(z_{1-\alpha})\right\}\widehat{Q}_{1,r}(z_{1-\alpha})$$
$$- \ell^{-3/2}Q_3(z_{1-\alpha})\phi(z_{1-\alpha}) + \sum_{k=1}^{6} B_{k,m,r}(1-\alpha) - \ell^{-1/2}\sum_{k=7}^{11} B_{k,m,r}(1-\alpha) + \ell^{-1}B_{12,m,r}(1-\alpha)$$
$$- \ell^{-3/2}B_{13,m,r}(1-\alpha) + \delta_{4,m,r} - \ell^{-1/2}\delta_{5,m,r} + \ell^{-1}\delta_{6,m,r} - \ell^{-3/2}\delta_{7,m,r} + O(\ell^{-2}).$$

Setting $m = \ell$ and grouping like terms, we may rewrite the above expression as

$$\begin{aligned}
\Psi_t\left(-\widetilde{\eta}_{\ell,r,\alpha}\right) = 1 - \alpha &- \ell^{-1/2}\phi(z_{1-\alpha})\left\{\widehat{Q}_{1,r}(z_{1-\alpha}) + Q_1(z_{1-\alpha})\right\} \\
&- \ell^{-1}\phi(z_{1-\alpha})\left\{\widehat{Q}_{2,r}(z_{1-\alpha}) - Q_2(z_{1-\alpha})\right\} \\
&+ \ell^{-1}\phi(z_{1-\alpha})\left\{\widehat{Q}'_{1,r}(z_{1-\alpha}) + Q'_1(z_{1-\alpha})\right\}\widehat{Q}_{1,r}(z_{1-\alpha}) \\
&- \ell^{-1}z_{1-\alpha}\phi(z_{1-\alpha})\left\{\widehat{Q}_{1,r}(z_{1-\alpha}) + Q_1(z_{1-\alpha})\right\}\widehat{Q}_{1,r}(z_{1-\alpha}) \\
&- \tfrac{1}{2}\ell^{-3/2}\phi(z_{1-\alpha})(z^2_{1-\alpha}-1)\left\{\widehat{Q}_{1,r}(z_{1-\alpha}) + Q_1(z_{1-\alpha})\right\}\widehat{Q}^2_{1,r}(z_{1-\alpha}) \\
&+ \ell^{-3/2}z_{1-\alpha}\phi(z_{1-\alpha})\left\{\widehat{Q}_{1,r}(z_{1-\alpha}) + Q_1(z_{1-\alpha})\right\} \\
&\qquad \times \left\{\widehat{Q}_{1,r}(z_{1-\alpha})\widehat{Q}'_{1,r}(z_{1-\alpha}) - \tfrac{1}{2}z_{1-\alpha}\widehat{Q}^2_{1,r}(z_{1-\alpha}) - \widehat{Q}_{2,r}(z_{1-\alpha})\right\} \\
&- \ell^{-3/2}\phi(z_{1-\alpha})\left\{\widehat{Q}_{3,r}(z_{1-\alpha}) + Q_3(z_{1-\alpha})\right\} \\
&- \ell^{-3/2}z_{1-\alpha}\phi(z_{1-\alpha})\left\{\widehat{Q}_{2,r}(z_{1-\alpha}) - Q_2(z_{1-\alpha})\right\}\widehat{Q}_{1,r}(z_{1-\alpha}) \\
&+ \ell^{-3/2}\phi(z_{1-\alpha})\left\{\widehat{Q}'_{2,r}(z_{1-\alpha}) - Q'_2(z_{1-\alpha})\right\}\widehat{Q}_{1,r}(z_{1-\alpha}) \\
&+ \ell^{-3/2}z_{1-\alpha}\phi(z_{1-\alpha})\left\{\widehat{Q}'_{1,r}(z_{1-\alpha}) + Q'_1(z_{1-\alpha})\right\}\widehat{Q}^2_{1,r}(z_{1-\alpha}) \\
&- \ell^{-3/2}\phi(z_{1-\alpha})\left\{\widehat{Q}'_{1,r}(z_{1-\alpha}) + Q'_1(z_{1-\alpha})\right\} \\
&\qquad \times \left\{\widehat{Q}_{1,r}(z_{1-\alpha})\widehat{Q}'_{1,r}(z_{1-\alpha}) - \tfrac{1}{2}z_{1-\alpha}\widehat{Q}^2_{1,r}(z_{1-\alpha}) - \widehat{Q}_{2,r}(z_{1-\alpha})\right\} \\
&- \tfrac{1}{2}\ell^{-3/2}\phi(z_{1-\alpha})\left\{\widehat{Q}''_{1,r}(z_{1-\alpha}) + Q''_1(z_{1-\alpha})\right\}\widehat{Q}^2_{1,r}(z_{1-\alpha}) \\
&+ \sum_{k=1}^{6} B_{k,\ell,r}(1-\alpha) - \ell^{-1/2}\sum_{k=7}^{11} B_{k,\ell,r}(1-\alpha) + \ell^{-1}B_{12,\ell,r}(1-\alpha) \\
&- \ell^{-3/2}B_{13,\ell,r}(1-\alpha) + \delta_{4,\ell,r} - \ell^{-1/2}\delta_{5,\ell,r} + \ell^{-1}\delta_{6,\ell,r} - \ell^{-3/2}\delta_{7,\ell,r} + O(\ell^{-2}).
\end{aligned}$$

The coverage probability of $\widehat{\mathscr{I}}^N_B(\ell,\alpha)$ may now be obtained by taking the expected value of $\Psi_t(-\widetilde{\eta}_{\ell,r,\alpha})$. That is,

$$\begin{aligned}
&\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}^N_B(\ell,\alpha)\right) = \mathrm{E}\left(\Psi_t(-\widetilde{\eta}_{\ell,r,\alpha})\right) \\
&= 1 - \alpha - \ell^{-1/2}\phi(z_{1-\alpha})\mathrm{E}\left(\left\{\widehat{Q}_{1,r}(z_{1-\alpha}) + Q_1(z_{1-\alpha})\right\}\right) \\
&\quad - \ell^{-1}\phi(z_{1-\alpha})\mathrm{E}\left(\left\{\widehat{Q}_{2,r}(z_{1-\alpha}) - Q_2(z_{1-\alpha})\right\}\right) \\
&\quad + \ell^{-1}\phi(z_{1-\alpha})\mathrm{E}\left(\left\{\widehat{Q}'_{1,r}(z_{1-\alpha}) + Q'_1(z_{1-\alpha})\right\}\widehat{Q}_{1,r}(z_{1-\alpha})\right) \\
&\quad - \ell^{-1}z_{1-\alpha}\phi(z_{1-\alpha})\mathrm{E}\left(\left\{\widehat{Q}_{1,r}(z_{1-\alpha}) + Q_1(z_{1-\alpha})\right\}\widehat{Q}_{1,r}(z_{1-\alpha})\right) \\
&\quad - \tfrac{1}{2}\ell^{-3/2}\phi(z_{1-\alpha})(z^2_{1-\alpha}-1)\mathrm{E}\left(\left\{\widehat{Q}_{1,r}(z_{1-\alpha}) + Q_1(z_{1-\alpha})\right\}\widehat{Q}^2_{1,r}(z_{1-\alpha})\right)
\end{aligned}$$

$$
\begin{aligned}
&+\ell^{-3/2}z_{1-\alpha}\phi(z_{1-\alpha})\mathrm{E}\Big(\{\widehat{Q}_{1,r}(z_{1-\alpha})+Q_1(z_{1-\alpha})\}\,\widehat{Q}^{cf}_{2,r}(z_{1-\alpha})\\
&-\ell^{-3/2}\phi(z_{1-\alpha})\mathrm{E}\left(\{\widehat{Q}_{3,r}(z_{1-\alpha})+Q_3(z_{1-\alpha})\}\right)\\
&-\ell^{-3/2}z_{1-\alpha}\phi(z_{1-\alpha})\mathrm{E}\left(\{\widehat{Q}_{2,r}(z_{1-\alpha})-Q_2(z_{1-\alpha})\}\,\widehat{Q}_{1,r}(z_{1-\alpha})\right)\\
&+\ell^{-3/2}\phi(z_{1-\alpha})\mathrm{E}\left(\left\{\widehat{Q}'_{2,r}(z_{1-\alpha})-Q'_2(z_{1-\alpha})\right\}\widehat{Q}_{1,r}(z_{1-\alpha})\right)\\
&+\ell^{-3/2}z_{1-\alpha}\phi(z_{1-\alpha})\mathrm{E}\left(\left\{\widehat{Q}'_{1,r}(z_{1-\alpha})+Q'_1(z_{1-\alpha})\right\}\widehat{Q}^2_{1,r}(z_{1-\alpha})\right)\\
&-\ell^{-3/2}\phi(z_{1-\alpha})\mathrm{E}\Big(\left\{\widehat{Q}'_{1,r}(z_{1-\alpha})+Q'_1(z_{1-\alpha})\right\}\widehat{Q}^{cf}_{2,r}(z_{1-\alpha})\\
&-\tfrac{1}{2}\ell^{-3/2}\phi(z_{1-\alpha})\mathrm{E}\left(\left\{\widehat{Q}''_{1,r}(z_{1-\alpha})+Q''_1(z_{1-\alpha})\right\}\widehat{Q}^2_{1,r}(z_{1-\alpha})\right)\\
&+\sum_{k=1}^{6}B_{k,\ell,r}(1-\alpha)-\ell^{-1/2}\sum_{k=7}^{11}B_{k,\ell,r}(1-\alpha)+\ell^{-1}B_{12,\ell,r}(1-\alpha)\\
&-\ell^{-3/2}B_{13,\ell,r}(1-\alpha)+\delta_{4,\ell,r}-\ell^{-1/2}\delta_{5,\ell,r}+\ell^{-1}\delta_{6,\ell,r}-\ell^{-3/2}\delta_{7,\ell,r}+O(\ell^{-2}).
\end{aligned}
$$

Under the assumptions of the theorem, applying the results of Lemma D.3 and recalling the results in (5.27), (5.29), (5.30), (5.31) and (5.32), the coverage probability reduces to

$$
\begin{aligned}
\mathrm{P}\left(\theta\in\mathscr{J}^N_B(\ell,\alpha)\right)=1-\alpha&-\ell^{-1/2}\phi(z_{1-\alpha})\mathrm{E}\left(\widehat{Q}_{1,r}(z_{1-\alpha})+Q_1(z_{1-\alpha})\right)\\
&+\ell^{-1}\phi(z_{1-\alpha})\mathrm{E}\left(\left\{\widehat{Q}'_{1,r}(z_{1-\alpha})+Q'_1(z_{1-\alpha})\right\}\widehat{Q}_{1,r}(z_{1-\alpha})\right)\\
&-\ell^{-1}z_{1-\alpha}\phi(z_{1-\alpha})\mathrm{E}\left(\{\widehat{Q}_{1,r}(z_{1-\alpha})+Q_1(z_{1-\alpha})\}\,\widehat{Q}_{1,r}(z_{1-\alpha})\right)\\
&-\tfrac{1}{2}\ell^{-3/2}\phi(z_{1-\alpha})(z^2_{1-\alpha}-1)\mathrm{E}\left(\{\widehat{Q}_{1,r}(z_{1-\alpha})+Q_1(z_{1-\alpha})\}\,\widehat{Q}^2_{1,r}(z_{1-\alpha})\right)\\
&+\ell^{-3/2}z_{1-\alpha}\phi(z_{1-\alpha})\mathrm{E}\Big(\{\widehat{Q}_{1,r}(z_{1-\alpha})+Q_1(z_{1-\alpha})\}\,\widehat{Q}^{cf}_{2,r}(z_{1-\alpha})\\
&-\ell^{-3/2}\phi(z_{1-\alpha})\mathrm{E}\left(\widehat{Q}_{3,r}(z_{1-\alpha})+Q_3(z_{1-\alpha})\right)\\
&+\ell^{-3/2}z_{1-\alpha}\phi(z_{1-\alpha})\mathrm{E}\left(\left\{\widehat{Q}'_{1,r}(z_{1-\alpha})+Q'_1(z_{1-\alpha})\right\}\widehat{Q}^2_{1,r}(z_{1-\alpha})\right)\\
&-\ell^{-3/2}\phi(z_{1-\alpha})\mathrm{E}\Big(\left\{\widehat{Q}'_{1,r}(z_{1-\alpha})+Q'_1(z_{1-\alpha})\right\}\widehat{Q}^{cf}_{2,r}(z_{1-\alpha})\\
&-\tfrac{1}{2}\ell^{-3/2}\phi(z_{1-\alpha})\mathrm{E}\left(\left\{\widehat{Q}''_{1,r}(z_{1-\alpha})+Q''_1(z_{1-\alpha})\right\}\widehat{Q}^2_{1,r}(z_{1-\alpha})\right)\\
&+O(\ell^{-1}r^{-1}+\ell^{-2}).
\end{aligned}
$$

Recall that $z_{1-\alpha}=-z_\alpha$. Since ϕ is a symmetric, even function, it holds that $\phi(z_{1-\alpha})=\phi(-z_\alpha)=\phi(z_\alpha)$. Also, since Q_j are polynomials which are odd for even j and even for odd j, it follows that

$$
Q'_j(z_{1-\alpha})=Q'_j(-z_\alpha)=\begin{cases}Q'_j(z_\alpha) & \text{if } j \text{ is even,}\\ -Q'_j(z_\alpha) & \text{if } j \text{ is odd,}\end{cases}
$$

and

$$
Q''_j(z_{1-\alpha})=Q''_j(-z_\alpha)=\begin{cases}-Q''_j(z_\alpha) & \text{if } j \text{ is even,}\\ Q''_j(z_\alpha) & \text{if } j \text{ is odd.}\end{cases}
$$

Also note that

$$
\begin{aligned}
\widehat{Q}^{cf}_{2,r}(z_{1-\alpha}) &= \widehat{Q}_{1,r}(z_{1-\alpha})\widehat{Q}'_{1,r}(z_{1-\alpha}) - \tfrac{1}{2}z_{1-\alpha}\widehat{Q}^2_{1,r}(z_{1-\alpha}) - \widehat{Q}_{2,r}(z_{1-\alpha}) \\
&= -\widehat{Q}_{1,r}(z_\alpha)\widehat{Q}'_{1,r}(z_\alpha) + \tfrac{1}{2}z_\alpha\widehat{Q}^2_{1,r}(z_\alpha) + \widehat{Q}_{2,r}(z_\alpha) \\
&= -\widehat{Q}^{cf}_{2,r}(z_\alpha).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}^N_B(\ell,\alpha)\right) &= 1-\alpha-\ell^{-1/2}\phi(z_\alpha)\mathrm{E}\left(\widehat{Q}_{1,r}(z_\alpha)+Q_1(z_\alpha)\right) \\
&\quad -\ell^{-1}\phi(z_\alpha)\mathrm{E}\left(\left\{\widehat{Q}'_{1,r}(z_\alpha)+Q'_1(z_\alpha)\right\}\widehat{Q}_{1,r}(z_\alpha)\right) \\
&\quad +\ell^{-1}z_\alpha\phi(z_\alpha)\mathrm{E}\left(\{\widehat{Q}_{1,r}(z_\alpha)+Q_1(z_\alpha)\}\widehat{Q}_{1,r}(z_\alpha)\right) \\
&\quad -\tfrac{1}{2}\ell^{-3/2}\phi(z_\alpha)(z_\alpha^2-1)\mathrm{E}\left(\{\widehat{Q}_{1,r}(z_\alpha)+Q_1(z_\alpha)\}\widehat{Q}^2_{1,r}(z_\alpha)\right) \\
&\quad +\ell^{-3/2}z_\alpha\phi(z_\alpha)\mathrm{E}\left(\{\widehat{Q}_{1,r}(z_\alpha)+Q_1(z_\alpha)\}\widehat{Q}^{cf}_{2,r}(z_\alpha)\right) \\
&\quad -\ell^{-3/2}\phi(z_\alpha)\mathrm{E}\left(\widehat{Q}_{3,r}(z_\alpha)+Q_3(z_\alpha)\right) \\
&\quad +\ell^{-3/2}z_\alpha\phi(z_\alpha)\mathrm{E}\left(\left\{\widehat{Q}'_{1,r}(z_\alpha)+Q'_1(z_\alpha)\right\}\widehat{Q}^2_{1,r}(z_\alpha)\right) \\
&\quad -\ell^{-3/2}\phi(z_\alpha)\mathrm{E}\left(\left\{\widehat{Q}'_{1,r}(z_\alpha)+Q'_1(z_\alpha)\right\}\widehat{Q}^{cf}_{2,r}(z_\alpha)\right) \\
&\quad -\tfrac{1}{2}\ell^{-3/2}\phi(z_\alpha)\mathrm{E}\left(\left\{\widehat{Q}''_{1,r}(z_\alpha)+Q''_1(z_\alpha)\right\}\widehat{Q}^2_{1,r}(z_\alpha)\right) \\
&\quad +O(\ell^{-1}r^{-1}+\ell^{-2}), \\
&= 1-\alpha-2\ell^{-1/2}\phi(z_\alpha)Q_1(z_\alpha)-\ell^{-1/2}\phi(z_\alpha)\mathrm{E}\left(\widehat{Q}_{1,r}(z_\alpha)-Q_1(z_\alpha)\right) \\
&\quad -2\ell^{-1}\phi(z_\alpha)Q_1(z_\alpha)\left(Q'_1(z_\alpha)-z_\alpha Q_1(z_\alpha)\right) \\
&\quad -\ell^{-3/2}\phi(z_\alpha)Q^2_1(z_\alpha)\left\{(z_\alpha^2-1)Q_1(z_\alpha)-2z_\alpha Q'_1(z_\alpha)+Q''_1(z_\alpha)\right\} \\
&\quad -2\ell^{-3/2}\phi(z_\alpha)Q^{cf}_2\left\{Q'_1(z_\alpha)-z_\alpha Q_1(z_\alpha)\right\}-2\ell^{-3/2}\phi(z_\alpha)Q_3(z_\alpha) \\
&\quad +O(\ell^{-1/2}r^{-1}+\ell^{-1}r^{-1}+\ell^{-3/2}r^{-1}+\ell^{-2}) \\
&= 1-\alpha+\frac{K_3(z_\alpha)}{\ell^{1/2}}-\frac{D_\theta(z_\alpha)}{\ell^{1/2}r}+\frac{K_4(z_\alpha)}{\ell}+\frac{K_5(z_\alpha)}{\ell^{3/2}}+O(\ell^{-2})
\end{aligned}
$$

where we made use of the assumption $\ell = O(r)$, i.e., $r^{-1} = O(\ell^{-1})$. □

5.4 Equal-tailed percentile-t confidence intervals

As noted earlier, one-sided upper and lower confidence bounds may be used to construct equal-tailed confidence intervals. For example, in the notation of Chapter 2 the standard bootstrap percentile-t $(1-2\alpha)$-level confidence interval for θ is given by

$$
\widehat{\mathscr{I}}(\alpha)\backslash\widehat{\mathscr{I}}(1-\alpha) = \left(\widehat{\theta}_n - n^{-1/2}\widehat{\sigma}_n\widehat{\eta}_{1-\alpha},\, \widehat{\theta}_n - n^{-1/2}\widehat{\sigma}_n\widehat{\eta}_\alpha\right].
$$

The order of coverage error of this interval is typically $O(n^{-1})$. Hall (1988, p. 949) shows that in the case where the parameter of interest is the mean of a univariate population, this interval has a reduced coverage error of $O(n^{-2})$ if $\kappa'_3 = \kappa'_4 = 0$, where κ'_j denotes the jth standardised cumulant of the underlying population.

We now show that equal-tailed confidence intervals with a reduced coverage error of $O(n^{-2})$ may be obtained using the newly proposed hybrid percentile-t bound $\widehat{\mathscr{I}}_H^N$ *without assumptions such as* $\kappa_3' = \kappa_4' = 0$. We have from Theorem 5.1 that the confidence interval

$$\widehat{\mathscr{I}}_H^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_H^N(\ell,1-\alpha) = \left(\widehat{\theta}_\ell - \ell^{-1/2}\widehat{\beta}_\ell\widetilde{\eta}_{m,r,1-\alpha},\, \widehat{\theta}_\ell - \ell^{-1/2}\widehat{\beta}_\ell\widetilde{\eta}_{m,r,\alpha}\right]$$

typically has coverage probability

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_H^N(\ell,1-\alpha)\right) = 1-2\alpha + \frac{D_\theta(z_\alpha) - D_\theta(z_{1-\alpha})}{\ell^{1/2}r} + O(\ell^{-2}),$$

where $D_\theta(z_\alpha)$ is the coefficient of r^{-1} in a power series expansion of

$$\phi(z_\alpha)\mathrm{E}\left\{\widehat{Q}_{1,r}(z_\alpha) - Q_1(z_\alpha)\right\}.$$

Recalling that ϕ, Q_1 and $\widehat{Q}_{1,r}$ are even functions, it follows immediately that $D_\theta(z_\alpha) = D_\theta(z_{1-\alpha})$, so that

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_H^N(\ell,1-\alpha)\right) = 1-2\alpha + O(\ell^{-2}).$$

If we now choose $\ell = \lfloor \gamma n^\psi \rfloor$ for some $\gamma > 0$ and $\frac{2}{3} < \psi \le 1$, then

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_H^N(\ell,1-\alpha)\right) = 1-2\alpha + O(n^{-2\psi}),$$

which for $\psi = 1$ becomes

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_H^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_H^N(\ell,1-\alpha)\right) = 1-2\alpha + O(n^{-2}).$$

Unlike the standard percentile-t, this result holds even when $\kappa_3' \neq 0$ or $\kappa_4' \neq 0$. This constitutes a significant improvement over the standard percentile-t interval $\widehat{\mathscr{I}}(\alpha)\backslash\widehat{\mathscr{I}}(1-\alpha)$ in terms of coverage probability.

In a similar fashion it may be shown that the new backwards percentile-t confidence interval

$$\widehat{\mathscr{I}}_B^N(\alpha)\backslash\widehat{\mathscr{I}}_B^N(1-\alpha) = \left(\widehat{\theta}_\ell + \ell^{-1/2}\widehat{\beta}_\ell\widetilde{\eta}_{m,r,\alpha},\, \widehat{\theta}_\ell + \ell^{-1/2}\widehat{\beta}_\ell\widetilde{\eta}_{m,r,1-\alpha}\right]$$

has coverage probability

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_B^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_B^N(\ell,1-\alpha)\right) = 1-2\alpha + \frac{2K_4(z_\alpha)}{\ell} + O(\ell^{-1/2}r^{-1}).$$

Choosing $\ell = \lfloor \gamma n \rfloor$ for some $0 < \gamma < 1$ then yields

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{I}}_B^N(\ell,\alpha)\backslash\widehat{\mathscr{I}}_B^N(\ell,1-\alpha)\right) = 1-2\alpha + \frac{2K_4(z_\alpha)}{\gamma n} + O(n^{-3/2}),$$

so that this interval is second-order accurate.

Chapter 6

Summary of main results

To summarise the results discussed thus far, we provide in the following tables the typical coverage errors of the standard and newly proposed confidence bounds and intervals. The coverage errors displayed here are those obtained when choosing $\ell = \lfloor \gamma n^{\psi} \rfloor$ for best choices of $\gamma > 0$ and $\frac{2}{3} < \psi \leq 1$.

Table 6.1: Typical coverage errors of the one-sided confidence bounds.

Bound	Standard	New	See Remark
Hybrid percentile	$O(n^{-1/2})$	$O(n^{-1})$	
Backwards percentile	$O(n^{-1/2})$	$O(n^{-1/2})$	**6.1**
Hybrid percentile-t	$O(n^{-1})$	$O(n^{-3/2})$	**6.2**
Backwards percentile-t	n/a	$O(n^{-1/2})$	**6.3**

Table 6.2: Typical coverage errors of the two-sided equal-tailed confidence intervals.

Interval	Standard	New
Hybrid percentile	$O(n^{-1})$	$O(n^{-1})$
Backwards percentile	$O(n^{-1})$	$O(n^{-1})$
Hybrid percentile-t	$O(n^{-1})$	$O(n^{-2})$
Backwards percentile-t	n/a	$O(n^{-1})$

Remark 6.1. In some natural situations the coverage error of the new backwards percentile bound $\widehat{\mathscr{I}}_B^N$ reduces to $O(n^{-1})$. One such example will be discussed in Section 7.2.2, where the parameter of interest is the mean of a univariate population with a distribution *symmetric* about zero.

Remark 6.2. If, in the notation of Chapter 5, it holds that $\mathrm{E}\{\widehat{Q}_{1,r}(z_\alpha) - Q_1(z_\alpha)\} = O(n^{-3/2})$, Theorem 5.1 states that the coverage error of the new hybrid percentile-t bound $\widehat{\mathscr{J}}_H^N$ typically reduces to $O(n^{-2})$. This is the case, for example, if the parameter of interest is the mean of a univariate population with a distribution *symmetric* about zero. This will be shown in Section 7.2.3.

Remark 6.3. In some situations the coverage error of the new backwards percentile-t bound $\widehat{\mathscr{I}}_B^N$ reduces to $O(n^{-2})$, corresponding to that of the hybrid percentile-t bound $\widehat{\mathscr{I}}_H^N$. An example of such a situation is given in Section 7.2.4.

Chapter 7

An illustrative application

The results derived in Chapters 4 and 5 hold in general for any parameter θ which can be expressed as a smooth function of means (i.e., in the regular smooth function model of Bhattacharya and Ghosh, 1978). This includes parameters such as the mean or variance of a population, the correlation coefficient between two random variables, and the ratio of two population means. In this chapter we provide a detailed discussion for the specific case where the parameter of interest is the *mean* of a univariate population. For this case we derive the exact coefficient of the leading term in an asymptotic expansion of the coverage error of each newly proposed bound.

7.1 The expectation of an asymptotic approximation

To be able to derive rigorously exact asymptotic expressions for the expectations in Theorems 4.1, 4.2, 5.1 and 5.2, and the corresponding assumptions (A1)–(A7) and (B1)–(B7), calls for a special form of the so-called "delta method". One convenient result (Hurt, 1976) states formal conditions under which the expectation of a Taylor approximation of a bounded function g of statistics accurately approximates the expectation of the function itself up to an arbitrary order. Theorem 7.1 below extends the result derived by Hurt (1976) in that it allows the restriction of boundedness of g to be relaxed. Furthermore, the theorem is also a generalisation of a result by Cramér (1946).

Before proving the theorem, we first state a well-known fact that we will use repeatedly throughout this chapter. For completeness we also provide a short proof.

Lemma 7.1. *Suppose X is a one-dimensional random variable with finite moments up to order s, for some integer $s \geq 1$. If*

$$\mathrm{E}(|X|^s) = O(n^{-s/2}),$$

then

$$\mathrm{E}\left(|X|^q\right) = O(n^{-q/2}), \quad \textit{for all } q = 1, 2, \ldots, s.$$

Proof. It is well known (cf. Loève, 1977, p. 156) that, for all $q = 1, 2, \ldots, s$,

$$\left\{\mathrm{E}\left(|X|^q\right)\right\}^{1/q} \leq \left\{\mathrm{E}\left(|X|^s\right)\right\}^{1/s},$$

or, equivalently,

$$\mathrm{E}\left(|X|^q\right) \le \left\{\mathrm{E}\left(|X|^s\right)\right\}^{q/s}.$$

Hence, for any $q = 1, 2, \ldots, s$, we may write

$$\mathrm{E}\left(|X|^q\right) \le \left\{O(n^{-s/2})\right\}^{q/s} = O(n^{-q/2}). \qquad \square$$

We now move on to the main result of this section.

Theorem 7.1. *For any positive integer s, let $g : \mathbb{R}^q \to \mathbb{R}$ be a function having bounded $(s+1)$-order partial derivatives in an open neighbourhood of some point $\boldsymbol{v} \in \mathbb{R}^q$. Suppose $\mathbf{V}$ is a q-vector of real-valued statistics (determined by a sample of size n) such that $|g(\mathbf{V})| \le Cn^{\delta/2}$ a.s. for $n \ge n_0$, with $n_0 \ge 1$, $C > 0$ and $\delta \ge 0$ some finite constants. If $\mathbf{V}$ has finite moments up to order $k = (2s) \vee (\delta + s + 1)$ and $\mathrm{E}\,|V_i - v_i|^k = O(n^{-k/2})$, $i = 1, \ldots, q$, then*

$$\mathrm{E}\{g(\mathbf{V})\} = g(\boldsymbol{v}) + \sum_{1 \le |\alpha| \le s} \frac{1}{\alpha!} \mathrm{E}\left\{(\mathbf{V} - \boldsymbol{v})^\alpha\right\} \partial^\alpha g(\boldsymbol{v}) + O(n^{-(s+1)/2}),$$

where $|\alpha| = \alpha_1 + \cdots + \alpha_q$, $\alpha! = \alpha_1! \cdots \alpha_q!$, $(\mathbf{V} - \boldsymbol{v})^\alpha = \prod_{i=1}^q (V_i - v_i)^{\alpha_i}$, and

$$\partial^\alpha g(\boldsymbol{v}) = \frac{\partial^{|\alpha|}}{\partial v_1^{\alpha_1} \cdots \partial v_q^{\alpha_q}} g(v_1, \ldots, v_q),$$

for any multi-index $\alpha = (\alpha_1, \ldots, \alpha_q) \in \mathbb{N}_0^q$.

Proof. Define the open neighbourhood on which g has bounded $(s+1)$-order partial derivatives by

$$Q_{\boldsymbol{\varepsilon}} := \left\{\mathbf{x} \in \mathbb{R}^q : |x_i - v_i| < \varepsilon_i \text{ for all } i = 1, \ldots, q\right\},$$

for some positive $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_q) \in \mathbb{R}^q$. By the Boole and Chebyshev inequalities it follows that

$$\mathrm{P}(\mathbf{V} \notin Q_{\boldsymbol{\varepsilon}}) \le \sum_{i=1}^q \mathrm{P}(|V_i - v_i| \ge \varepsilon_i) \le \sum_{i=1}^q \frac{\mathrm{E}\left\{|V_i - v_i|^k\right\}}{\varepsilon_i^k} = O(n^{-k/2}),$$

a fact that we will be using repeatedly below.

Now, observe that, since $|g(\mathbf{V})| \le Cn^{\delta/2}$ a.s., it follows that

$$\begin{aligned}
\left|\int_{\{\mathbf{V} \in Q_{\boldsymbol{\varepsilon}}^c\}} (g(\mathbf{V}) - g(\boldsymbol{v}))\, d\mathrm{P}\right| &\le \int_{\{\mathbf{V} \in Q_{\boldsymbol{\varepsilon}}^c\}} |g(\mathbf{V})|\, d\mathrm{P} + |g(\boldsymbol{v})| \int_{\{\mathbf{V} \in Q_{\boldsymbol{\varepsilon}}^c\}} d\mathrm{P} \\
&= \int_{\{\mathbf{V} \in Q_{\boldsymbol{\varepsilon}}^c\}} |g(\mathbf{V})|\, d\mathrm{P} + |g(\boldsymbol{v})| \mathrm{P}(\mathbf{V} \notin Q_{\boldsymbol{\varepsilon}}) \\
&\le Cn^{\delta/2} \mathrm{P}(\mathbf{V} \notin Q_{\boldsymbol{\varepsilon}}) + O(n^{-k/2}) \\
&= O(n^{-(k-\delta)/2}),
\end{aligned}$$

so that

$$\mathrm{E}\{g(\mathbf{V}) - g(\boldsymbol{v})\} = \int_{\{\mathbf{V} \in Q_{\boldsymbol{\varepsilon}}\}} (g(\mathbf{V}) - g(\boldsymbol{v}))\, d\mathrm{P} + O(n^{-(k-\delta)/2}). \tag{7.1}$$

For the error term in this expression to be $O(n^{-(s+1)/2})$, we require $k \ge \delta + s + 1$.

Now, by Taylor expansion of g about $\boldsymbol{v}$ we may write

$$\int_{\{\mathbf{V}\in Q_\varepsilon\}} (g(\mathbf{V}) - g(\boldsymbol{v}))\,d\,\mathrm{P} = \sum_{1\le|\alpha|\le s} \frac{1}{\alpha!}(\partial^\alpha g)(\boldsymbol{v}) \int_{\{\mathbf{V}\in Q_\varepsilon\}} (\mathbf{V}-\boldsymbol{v})^\alpha d\,\mathrm{P} + R_n, \tag{7.2}$$

where

$$R_n = \sum_{|\alpha|=s+1} \frac{1}{\alpha!} \int_{\{\mathbf{V}\in Q_\varepsilon\}} (\mathbf{V}-\boldsymbol{v})^\alpha\,(\partial^\alpha g)(\boldsymbol{v} + \xi(\mathbf{V}-\boldsymbol{v}))\,d\,\mathrm{P},$$

for some intermediate point $\xi = \xi(\mathbf{V}, \boldsymbol{v}) \in (0,1)$. Now, for any α such that $1 \le |\alpha| \le s$, we have by a generalisation of the Hölder inequality (see, e.g., Finner, 1992) that

$$\begin{aligned}
\int_{\{\mathbf{V}\in Q_\varepsilon^c\}} |\mathbf{V}-\boldsymbol{v}|^\alpha d\,\mathrm{P} &\le \left(\int_{\{\mathbf{V}\in Q_\varepsilon^c\}} \left(|\mathbf{V}-\boldsymbol{v}|^\alpha\right)^2 d\,\mathrm{P}\right)^{1/2} \left(\int_{\{\mathbf{V}\in Q_\varepsilon^c\}} d\,\mathrm{P}\right)^{1/2} \\
&\le \left(\prod_{\alpha_i\neq 0} \left(\int_{\{\mathbf{V}\in Q_\varepsilon^c\}} |V_i - v_i|^{2|\alpha|} d\,\mathrm{P}\right)^{\alpha_i/|\alpha|}\right)^{1/2} O(n^{-k/4}) \\
&\le \left(\prod_{\alpha_i\neq 0} \left(\mathrm{E}\left\{|V_i - v_i|^{2|\alpha|}\right\}\right)^{\alpha_i/|\alpha|}\right)^{1/2} O(n^{-k/4}) \\
&= \left(\prod_{\alpha_i\neq 0} \left(O(n^{-|\alpha|})\right)^{\alpha_i/|\alpha|}\right)^{1/2} O(n^{-k/4}) \\
&= \left(O(n^{-|\alpha|})\right)^{1/2} O(n^{-k/4}) = O(n^{-(k+2)/4}),
\end{aligned}$$

so that

$$\begin{aligned}
\int_{\{\mathbf{V}\in Q_\varepsilon\}} (\mathbf{V}-\boldsymbol{v})^\alpha d\,\mathrm{P} &= \mathrm{E}\left\{(\mathbf{V}-\boldsymbol{v})^\alpha\right\} - \int_{\{\mathbf{V}\in Q_\varepsilon^c\}} (\mathbf{V}-\boldsymbol{v})^\alpha d\,\mathrm{P} \\
&= \mathrm{E}\left\{(\mathbf{V}-\boldsymbol{v})^\alpha\right\} + O(n^{-(k+2)/4}).
\end{aligned} \tag{7.3}$$

Note that we have assumed here that each element in $\mathbf{V}$ has at least $2\max|\alpha| = 2s$ finite moments. Hence we require $k \ge 2s$.

We now treat the remainder term R_n, again by applying the generalised Hölder inequality. Since by assumption g has bounded $(s+1)$-order partial derivatives on $\boldsymbol{Q_\varepsilon}$, there exists, for any α such that $|\alpha| = s+1$, a finite nonnegative constant K such that

$$\begin{aligned}
&\left|\int_{\{\mathbf{V}\in Q_\varepsilon\}} (\mathbf{V}-\boldsymbol{v})^\alpha (\partial^\alpha g)(\boldsymbol{v} + \xi(\mathbf{V}-\boldsymbol{v}))\,d\,\mathrm{P}\right| \\
&\le K \int_{\{\mathbf{V}\in Q_\varepsilon\}} |\mathbf{V}-\boldsymbol{v}|^\alpha d\,\mathrm{P} \\
&\le K \prod_{\alpha_i\neq 0} \left(\int_{\{\mathbf{V}\in Q_\varepsilon\}} |V_i - v_i|^{|\alpha|} d\,\mathrm{P}\right)^{\alpha_i/|\alpha|} \\
&\le K \prod_{\alpha_i\neq 0} \left(\mathrm{E}\left\{|V_i - v_i|^{|\alpha|}\right\}\right)^{\alpha_i/|\alpha|} = O(n^{-(s+1)/2}).
\end{aligned}$$

Hence, $R_n = O(n^{-(s+1)/2})$, which is *nonrandom*. Combining this result with (7.1), (7.2) and (7.3) proves the theorem. □

The following corollary is a special case of Theorem 7.1 which we will require repeatedly below. It is stated here for convenience.

Corollary 7.1. *Let $g : \mathbb{R}^2 \to \mathbb{R}$ be a function having bounded fourth order partial derivatives in an open neighbourhood of some point $(v_1, v_2) \in \mathbb{R}^2$. Suppose (V_1, V_2) is a vector of real-valued statistics (determined by a sample of size n) such that $|g(V_1, V_2)| \leq Cn^{\delta/2}$ a.s. for $n \geq n_0$, with $n_0 \geq 1$, $C > 0$ and $\delta \geq 0$ some finite constants. Also suppose that V_1 and V_2 have finite moments up to order $k = 6 \vee (\delta + 4)$ and $\mathrm{E}\,|V_i - v_i|^k = O(n^{-k/2})$, $i = 1, 2$.*

If we assume that

(a) $\mathrm{E}(V_i - v_i) = \frac{C_i}{n} + O(n^{-2}), \qquad i = 1, 2,$

(b) $\mathrm{E}\left\{(V_i - v_i)^2\right\} = \frac{D_i}{n} + O(n^{-2}), \qquad i = 1, 2,$

(c) $\mathrm{E}\{(V_1 - v_1)(V_2 - v_2)\} = \frac{E_{12}}{n} + O(n^{-2}),$

(d) $\mathrm{E}\left\{(V_1 - v_1)^i (V_2 - v_2)^j\right\} = O(n^{-2}), \qquad i + j = 3,\ i = 0, 1, 2, 3,$

for finite constants C_1, C_2, D_1, D_2 and E_{12}, then

$$\mathrm{E}\left\{g\left(V_{1,n}, V_{2,n}\right) - g(v_1, v_2)\right\} = \frac{1}{n}\Big\{g_1(v_1, v_2)C_1 + g_2(v_1, v_2)C_2 + \tfrac{1}{2}g_{11}(v_1, v_2)D_1 + \tfrac{1}{2}g_{22}(v_1, v_2)D_2 + g_{12}(v_1, v_2)E_{12}\Big\} + O(n^{-2}),$$

where

$$g_1(x, y) = \frac{\partial}{\partial x}g(x, y), \qquad g_2(x, y) = \frac{\partial}{\partial y}g(x, y),$$

$$g_{11}(x, y) = \frac{\partial^2}{\partial x^2}g(x, y), \qquad g_{12}(x, y) = \frac{\partial^2}{\partial x \partial y}g(x, y), \qquad g_{22}(x, y) = \frac{\partial^2}{\partial y^2}g(x, y).$$

Proof. Take $q = 2$ and $s = 3$ in Theorem 7.1. □

7.1.1 Moments of standardised sample cumulants

We now illustrate how Theorem 7.1 may be applied to obtain useful asymptotic expressions for the expectation of three particular standardised sample cumulants, which are required in Section 7.2. For a more detailed account of cumulants and other quantities of interest, see Appendix A.

Throughout the rest of this section, $\{X_1, X_2, \ldots, X_n\}$ denotes a random sample from a univariate population with mean μ. Define the kth *central moment* of X_1 by

$$\mu_k = \mathrm{E}\left\{\left(X_1 - \mu\right)^k\right\}$$

with correspsonding *central sample moment*

$$m_k = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}_n\right)^k,$$

where $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$. Also define

$$\sigma^2 = \mu_2 \quad \text{and} \quad \hat{\sigma}_n^2 = m_2.$$

Skewness (κ_3')

The *skewness* of X_1 is defined as the third standardised cumulant of X_1, i.e.,

$$\kappa_3' = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{\mu_2^{3/2}},$$

with sample version

$$\widehat{\kappa}_{3,n}' = \frac{m_3}{\widehat{\sigma}_n^3} = \frac{m_3}{m_2^{3/2}}.$$

Lemma 7.2. *Suppose that* $\mathrm{E}(|X_1|^k) < \infty$ *for some sufficiently large* k *and that* $\mu_2 > 0$. *Then*

$$\begin{aligned}\mathrm{E}\left(\widehat{\kappa}_{3,n}' - \kappa_3'\right) &= -\frac{1}{8n}\left\{12\kappa_5' - 15\kappa_4'\kappa_3' + 54\kappa_3'\right\} + O(n^{-2}) \\ &= \frac{\mu_3\left(21\sigma^4 + 15\mu_4\right) - 12\sigma^2\mu_5}{8\sigma^7 n} + O(n^{-2}),\end{aligned}$$

and $\mathrm{E}\{(\widehat{\kappa}_{3,n}' - \kappa_3')^4\} = O(n^{-2})$.

Proof. In the context of Corollary 7.1, set

$$V_1 = m_3, \qquad \nu_1 = \mu_3, \qquad V_2 = m_2 = \widehat{\sigma}_n^2, \qquad \nu_2 = \mu_2 = \sigma^2,$$

and let

$$h(x,y) = \frac{x}{y^{3/2}}.$$

We would now like to apply Corollary 7.1 to the functions $g = h, h^2, h^3, h^4$.

By Lemma A.1 in the Appendix we have for this choice of g that

$$|h(V_1,V_2)| = |h(m_3,m_2)| = \left|\frac{m_3}{m_2^{3/2}}\right| \le \sqrt{n}.$$

This implies that

$$|h(V_1,V_2)|^j \le n^{j/2}, \qquad j = 1,2,\ldots.$$

Since we need to apply Corollary 7.1 to h^4, we require $\delta = 4$. Noting that $k = 6 \vee (\delta + 4) = 8$, we must have that

$$\mathrm{E}\left\{(V_1 - \nu_1)^8\right\} = \mathrm{E}\left\{\left(m_3 - \mu_3\right)^8\right\} = O(n^{-4})$$

and

$$\mathrm{E}\left\{(V_2 - \nu_2)^8\right\} = \mathrm{E}\left\{\left(m_2 - \mu_2\right)^8\right\} = O(n^{-4}),$$

which are both confirmed by (27.5.5) of Cramér (1946) under the assumption of sufficiently many finite moments. Finally, given that $\mu_2 > 0$, the functions h, h^2, h^3 and h^4 all have bounded derivatives up to the fourth order in an open neighbourhood of $(\nu_1,\nu_2) = (\mu_3,\mu_2)$.

In the context of Corollary 7.1, we have the following constants which we obtain from the results of Lemma A.6, Lemma A.7 and Lemma A.9:

$$C_1 = -3\mu_3, \qquad D_1 = \mu_6 - 6\mu_4\mu_2 - \mu_3^2 + 9\mu_2^3, \qquad E_{12} = \mu_5 - 4\mu_3\mu_2,$$

$$C_2 = -\mu_2, \qquad D_2 = \mu_4 - \mu_2^2.$$

Also note that (by the same lemmas) item *(d)* of Corollary 7.1 is satisfied.

Therefore, since

$$h_1(x,y) = \frac{1}{y^{3/2}}, \qquad h_2(x,y) = -\frac{3x}{2y^{5/2}},$$
$$h_{11}(x,y) = 0, \qquad h_{12}(x,y) = -\frac{3}{2y^{5/2}}, \qquad h_{22}(x,y) = \frac{15x}{4y^{7/2}},$$

we may by Corollary 7.1 write

$$\begin{aligned}
\mathrm{E}\left(\widehat{\kappa}'_{3,n}\right) &= \mathrm{E}\left(\frac{m_3}{m_2^{3/2}}\right) \\
&= \frac{\mu_3}{\mu_2^{3/2}} + \frac{1}{n}\left\{\frac{1}{\mu_2^{3/2}}C_1 - \frac{3\mu_3}{2\mu_2^{5/2}}C_2 + 0\cdot D_1 + \frac{15\mu_3}{8\mu_2^{7/2}}D_2 - \frac{3}{2\mu_2^{5/2}}E_{12}\right\} + O(n^{-2}) \qquad (7.4)\\
&= \kappa'_3 + \frac{1}{n}\left\{-\frac{3\mu_3}{\sigma^3} + \frac{3\sigma^2\mu_3}{2\sigma^5} + \frac{15\mu_3\left(\mu_4-\sigma^4\right)}{8\sigma^7} - \frac{3\left(\mu_5 - 4\sigma^2\mu_3\right)}{2\sigma^5}\right\} + O(n^{-2}) \\
&= \kappa'_3 + \frac{\mu_3\left(21\sigma^4 + 15\mu_4\right) - 12\sigma^2\mu_5}{8\sigma^7 n} + O(n^{-2}),
\end{aligned}$$

which proves the first part of the theorem.

Since

$$\frac{\partial}{\partial x}h^2(x,y) = \frac{2x}{y^3}, \qquad \frac{\partial}{\partial y}h^2(x,y) = -\frac{3x^2}{y^4},$$
$$\frac{\partial^2}{\partial x^2}h^2(x,y) = \frac{2}{y^3}, \qquad \frac{\partial^2}{\partial x \partial y}h^2(x,y) = -\frac{6x}{y^4}, \qquad \frac{\partial^2}{\partial y^2}h^2(x,y) = \frac{12x^2}{y^5},$$

we have from Corollary 7.1 that

$$\mathrm{E}\left\{\left(\widehat{\kappa}'_{3,n}\right)^2\right\} = \mathrm{E}\left(\frac{m_3^2}{m_2^3}\right) = \left(\kappa'_3\right)^2 + \frac{1}{n}\left\{\frac{2\mu_3}{\mu_2^3}C_1 - \frac{3\mu_3^2}{\mu_2^4}C_2 + \frac{1}{\mu_2^3}D_1 + \frac{6\mu_3^2}{\mu_2^5}D_2 - \frac{6\mu_3}{\mu_2^4}E_{12}\right\} + O(n^{-2}). \qquad (7.5)$$

For $(\kappa'_3)^3$ we need the derivatives of h^3, which are

$$\frac{\partial}{\partial x}h^3(x,y) = \frac{3x^2}{y^{9/2}}, \qquad \frac{\partial}{\partial y}h^3(x,y) = -\frac{9x^3}{2y^{11/2}},$$
$$\frac{\partial^2}{\partial x^2}h^3(x,y) = \frac{6x}{y^{9/2}}, \qquad \frac{\partial^2}{\partial x \partial y}h^3(x,y) = -\frac{27x^2}{2y^{11/2}}, \qquad \frac{\partial^2}{\partial y^2}h^3(x,y) = \frac{99x^3}{4y^{13/2}}.$$

Hence, by Corollary 7.1, we have

$$\begin{aligned}
\mathrm{E}\left\{\left(\widehat{\kappa}'_{3,n}\right)^3\right\} &= \mathrm{E}\left(\frac{m_3^3}{m_2^{9/2}}\right) \\
&= \left(\kappa'_3\right)^3 + \frac{1}{n}\left\{\frac{3\mu_3^2}{\mu_2^{9/2}}C_1 - \frac{9\mu_3^3}{2\mu_2^{11/2}}C_2 + \frac{3\mu_3}{\mu_2^{9/2}}D_1 + \frac{99\mu_3^3}{8\mu_2^{13/2}}D_2 - \frac{27\mu_3^2}{2\mu_2^{11/2}}E_{12}\right\} + O(n^{-2}). \qquad (7.6)
\end{aligned}$$

Finally, it follows from

$$\frac{\partial}{\partial x}h^4(x,y) = \frac{4x^3}{y^6}, \qquad \frac{\partial}{\partial y}h^4(x,y) = -\frac{6x^4}{y^7},$$
$$\frac{\partial^2}{\partial x^2}h^4(x,y) = \frac{12x^2}{y^6}, \qquad \frac{\partial^2}{\partial x \partial y}h^4(x,y) = -\frac{24x^3}{y^7}, \qquad \frac{\partial^2}{\partial y^2}h^4(x,y) = \frac{42x^4}{y^8},$$

and Corollary 7.1 that

$$\mathrm{E}\left\{\left(\widehat{\kappa}'_{3,n}\right)^4\right\} = \mathrm{E}\left(\frac{m_3^4}{m_2^6}\right)$$
$$= \left(\kappa'_3\right)^4 + \frac{1}{n}\left\{\frac{4\mu_3^3}{\mu_2^6}C_1 - \frac{6\mu_3^4}{\mu_2^7}C_2 + \frac{12\mu_3^2}{\mu_2^6}D_1 + \frac{42\mu_3^4}{\mu_2^8}D_2 - \frac{24\mu_3^3}{\mu_2^7}E_{12}\right\} + O(n^{-2}). \quad (7.7)$$

Using (7.4), (7.5), (7.6) and (7.7) we obtain

$$\mathrm{E}\left\{\left(\widehat{\kappa}'_{3,n} - \kappa'_3\right)^4\right\} = \mathrm{E}\left\{\left(\widehat{\kappa}'_{3,n}\right)^4\right\} - 4\kappa'_3\,\mathrm{E}\left\{\left(\widehat{\kappa}'_{3,n}\right)^3\right\} + 6\left(\kappa'_3\right)^2\mathrm{E}\left\{\left(\widehat{\kappa}'_{3,n}\right)^2\right\} - 4\left(\kappa'_3\right)^3\mathrm{E}\left(\widehat{\kappa}'_{3,n}\right) + \left(\kappa'_3\right)^4$$
$$= \left(\kappa'_3\right)^4 + \frac{1}{n}\left\{\frac{4\mu_3^3}{\mu_2^6}C_1 - \frac{6\mu_3^4}{\mu_2^7}C_2 + \frac{12\mu_3^2}{\mu_2^6}D_1 + \frac{42\mu_3^4}{\mu_2^8}D_2 - \frac{24\mu_3^3}{\mu_2^7}E_{12}\right\}$$
$$- 4\left(\kappa'_3\right)^4 - \frac{4\mu_3}{n\mu_2^{3/2}}\left\{\frac{3\mu_3^2}{\mu_2^{9/2}}C_1 - \frac{9\mu_3^3}{2\mu_2^{11/2}}C_2 + \frac{3\mu_3}{\mu_2^{9/2}}D_1 + \frac{99\mu_3^3}{8\mu_2^{13/2}}D_2 - \frac{27\mu_3^2}{2\mu_2^{11/2}}E_{12}\right\}$$
$$+ 6\left(\kappa'_3\right)^4 + \frac{6\mu_3^2}{n\mu_2^3}\left\{\frac{2\mu_3}{\mu_2^3}C_1 - \frac{3\mu_3^2}{\mu_2^4}C_2 + \frac{1}{\mu_2^3}D_1 + \frac{6\mu_3^2}{\mu_2^5}D_2 - \frac{6\mu_3}{\mu_2^4}E_{12}\right\}$$
$$- 4\left(\kappa'_3\right)^4 - \frac{4\mu_3^3}{n\mu_2^{9/2}}\left\{\frac{1}{\mu_2^{3/2}}C_1 - \frac{3\mu_3}{2\mu_2^{5/2}}C_2 + \frac{15\mu_3}{8\mu_2^{7/2}}D_2 - \frac{3}{2\mu_2^{5/2}}E_{12}\right\}$$
$$+ \left(\kappa'_3\right)^4 + O(n^{-2})$$
$$= O(n^{-2}).$$

□

Excess kurtosis (κ'_4)

The *excess kurtosis* of X_1 is defined as the fourth standardised cumulant of X_1, i.e.,

$$\kappa'_4 = \frac{\mu_4}{\sigma^4} - 3 = \frac{\mu_4}{\mu_2^2} - 3,$$

with sample version

$$\widehat{\kappa}'_{4,n} = \frac{m_4}{\widehat{\sigma}_n^4} - 3 = \frac{m_4}{m_2^2} - 3.$$

Lemma 7.3. *Suppose that* $\mathrm{E}(|X_1|^k) < \infty$ *for some sufficiently large* k *and that* $\mu_2 > 0$. *Then*

$$\mathrm{E}\left(\widehat{\kappa}'_{4,n} - \kappa'_4\right) = -\frac{1}{n}\left\{2\kappa'_6 - 3(\kappa'_4)^2 + 15\kappa'_4 + 12(\kappa'_3)^2 + 6\right\} + O(n^{-2})$$
$$= \frac{-2\mu_6\mu_2 - 3\mu_4\mu_2^2 + 3\mu_4^2 + 8\mu_3^2\mu_2 + 6\mu_2^4}{n\mu_2^4} + O(n^{-2}),$$

and $\mathrm{E}\{(\widehat{\kappa}'_{4,n} - \kappa'_4)^2\} = O(n^{-1})$.

Proof. In the context of Corollary 7.1, set

$$V_1 = m_4, \qquad v_1 = \mu_4, \qquad V_2 = m_2 = \widehat{\sigma}_n^2, \qquad v_2 = \mu_2 = \sigma^2,$$

and let

$$h(x, y) = \frac{x}{y^2}.$$

We would now like to apply Corollary 7.1 to the functions $g = h, h^2$.

For this choice of h it follows from Lemma A.1 that

$$|h(V_1, V_2)| = |h(m_4, m_2)| = \left|\frac{m_4}{m_2^2}\right| \le n.$$

This implies that

$$|h(V_1, V_2)|^j \le n^j, \qquad j = 1, 2, \dots.$$

Since we need to apply Corollary 7.1 to h^2, we require $\delta = 4$. Noting that $k = 6 \vee (\delta + 4) = 8$, we must have that

$$\mathrm{E}\left\{(V_1 - \nu_1)^8\right\} = \mathrm{E}\left\{(m_4 - \mu_4)^8\right\} = O(n^{-4})$$

and

$$\mathrm{E}\left\{(V_2 - \nu_2)^8\right\} = \mathrm{E}\left\{(m_2 - \mu_2)^8\right\} = O(n^{-4}),$$

which are both confirmed by (27.5.5) of Cramér (1946) under the assumption of sufficiently many finite moments. Finally, given that $\mu_2 > 0$, the functions h and h^2 both have bounded derivatives up to the fourth order in an open neighbourhood of $(\nu_1, \nu_2) = (\mu_4, \mu_2)$.

In the context of Corollary 7.1, we have the following constants which we obtain from the results of Lemmas A.6, A.8 and A.10:

$$C_1 = -4\mu_4 + 6\mu_2^2, \qquad D_1 = \mu_8 - 8\mu_5\mu_3 - \mu_4^2 + 16\mu_3^2\mu_2, \qquad E_{12} = \mu_6 - \mu_4\mu_2 - 4\mu_3^2,$$
$$C_2 = -\mu_2, \qquad D_2 = \mu_4 - \mu_2^2.$$

Also note that (by the same lemmas) item *(d)* of Corollary 7.1 is satisfied.

Therefore, since

$$h_1(x,y) = \frac{1}{y^2}, \qquad h_2(x,y) = -\frac{2x}{y^3},$$
$$h_{11}(x,y) = 0, \qquad h_{12}(x,y) = -\frac{2}{y^3}, \qquad h_{22}(x,y) = \frac{6x}{y^4},$$

we may by Corollary 7.1 write

$$\begin{aligned}
&\mathrm{E}\left(\frac{m_4}{m_2^2}\right)\\
&= \frac{\mu_4}{\mu_2^2} + \frac{1}{n}\left\{\frac{1}{\mu_2^2}C_1 - \frac{2\mu_4}{\mu_2^3}C_2 + \tfrac{1}{2}\cdot 0 \cdot D_1 + \frac{6\mu_4}{2\mu_2^4}D_2 - \frac{2}{\mu_2^3}E_{12}\right\} + O(n^{-2})\\
&= \frac{\mu_4}{\mu_2^2} + \frac{1}{n\mu_2^4}\left\{-4\mu_4\mu_2^2 + 6\mu_2^4 + 2\mu_4\mu_2^2 + 3\mu_4^2 - 3\mu_4\mu_2^2 - 2\mu_6\mu_2 + 2\mu_4\mu_2^2 + 8\mu_3^2\mu_2\right\} + O(n^{-2})\\
&= \frac{\mu_4}{\mu_2^2} + \frac{-2\mu_6\mu_2 - 3\mu_4\mu_2^2 + 3\mu_4^2 + 8\mu_3^2\mu_2 + 6\mu_2^4}{n\mu_2^4} + O(n^{-2}).
\end{aligned}$$

Noting that $\widehat{\kappa}'_{4,n} = m_4/m_2^2 - 3$ and $\kappa'_4 = \mu_4/\mu_2^2 - 3$ we have the first result.

Since

$$\frac{\partial}{\partial x}h^2(x,y) = \frac{2x}{y^4}, \qquad \frac{\partial}{\partial y}h^2(x,y) = -\frac{4x^2}{y^5},$$
$$\frac{\partial^2}{\partial x^2}h^2(x,y) = \frac{2}{y^4}, \qquad \frac{\partial^2}{\partial x \partial y}h^2(x,y) = -\frac{8x}{y^5}, \qquad \frac{\partial^2}{\partial y^2}h^2(x,y) = \frac{20x^2}{y^6},$$

we have from Corollary 7.1 that

$$\mathrm{E}\left(\frac{m_4^2}{m_2^4}\right) = \frac{\mu_4^2}{\mu_2^4} + \frac{1}{n}\left\{\frac{2\mu_4}{\mu_2^4}C_1 - \frac{4\mu_4^2}{\mu_2^5}C_2 + \frac{1}{\mu_2^4}D_1 + \frac{10\mu_4^2}{\mu_2^6}D_2 - \frac{8\mu_4}{\mu_2^5}E_{12}\right\} + O(n^{-2})$$
$$= \frac{\mu_4^2}{\mu_2^4} + \frac{1}{n\mu_2^6}\Big\{2\mu_4\mu_2^2\left(-4\mu_4 + 6\mu_2^2\right) + 4\mu_4^2\mu_2^2 + \mu_2^2\left(\mu_8 - 8\mu_5\mu_3 - \mu_4^2 + 16\mu_3^2\mu_2\right)$$
$$+ 10\mu_4^2\left(\mu_4 - \mu_2^2\right) - 8\mu_4\mu_2\left(\mu_6 - \mu_4\mu_2 - 4\mu_3^2\right)\Big\} + O(n^{-2})$$
$$= \frac{\mu_4^2}{\mu_2^4} + \frac{1}{n\mu_2^6}\Big\{\mu_8\mu_2^2 - 8\mu_6\mu_4\mu_2 - 8\mu_5\mu_3\mu_2^2 - 7\mu_4^2 + 12\mu_4^2\mu_2^4 + 10\mu_4^3 + 32\mu_4\mu_3^2\mu_2$$
$$+ 16\mu_3^2\mu_2^3\Big\} + O(n^{-2}).$$

Hence,

$$\mathrm{E}\left\{\left(\widehat{\kappa}'_{4,n} - \kappa'_4\right)^2\right\} = \mathrm{E}\left\{\left(\frac{m_4}{m_2^2} - \frac{\mu_4}{\mu_2^2}\right)^2\right\} = \mathrm{E}\left(\frac{m_4^2}{m_2^4}\right) - \frac{2\mu_4}{\mu_2^2}\mathrm{E}\left(\frac{m_4}{m_2}\right) + \frac{\mu_4^2}{\mu_2^4}$$
$$= \frac{\mu_4^2}{\mu_2^4} - \frac{2\mu_4^2}{\mu_2^4} + \frac{\mu_4^2}{\mu_2^4} + O(n^{-1})$$
$$= O(n^{-1}).$$

□

Fifth standardised cumulant (κ'_5)

The fifth standardised cumulant of X_1 is given by

$$\kappa'_5 = \frac{\mu_5}{\sigma^5} - \frac{10\mu_3}{\sigma^3} = \frac{\mu_5}{\mu_2^{5/2}} - \frac{10\mu_3}{\mu_2^{3/2}} = \frac{\mu_5}{\mu_2^{5/2}} - 10\kappa'_3,$$

with sample version

$$\widehat{\kappa}'_{5,n} = \frac{m_5}{m_2^{5/2}} - 10\widehat{\kappa}'_{3,n}.$$

Lemma 7.4. *Suppose that* $\mathrm{E}(|X_1|^k) < \infty$ *for some sufficiently large* k *and that* $\mu_2 > 0$. *Then*

$$\mathrm{E}\left(\widehat{\kappa}'_{5,n} - \kappa'_5\right) = O(n^{-1}).$$

Proof. In the context of Theorem 7.1, set

$$V_1 = m_5, \qquad v_1 = \mu_5, \qquad V_2 = m_2 = \widehat{\sigma}_n^2, \qquad v_2 = \mu_2 = \sigma^2,$$

and let

$$h(x,y) = \frac{x}{y^{5/2}}.$$

We would now like to apply Theorem 7.1 to the function $g = h$. We wish to determine only the constant term $g(\boldsymbol{v})$, so we may set $s = 2$.

For this choice of h it follows from Lemma A.1 that

$$|h(V_1, V_2)| = |h(m_5, m_2)| = \left|\frac{m_5}{m_2^{5/2}}\right| \le n^{3/2}.$$

Hence we require $\delta = 3$ in Theorem 7.1. Noting that $k = (2s) \vee (\delta + s + 1) = 4 \vee 6 = 6$, we must have that

$$\mathrm{E}\{(V_1 - v_1)^6\} = \mathrm{E}\left\{(m_5 - \mu_5)^6\right\} = O(n^{-3})$$

and

$$\mathrm{E}\{(V_2 - v_2)^6\} = \mathrm{E}\left\{(m_2 - \mu_2)^6\right\} = O(n^{-3}),$$

which are both confirmed by (27.5.5) of Cramér (1946) under the assumption of sufficiently many finite moments. Finally, given that $\mu_2 > 0$, the functions h has bounded derivatives up to the fourth order in an open neighbourhood of $(v_1, v_2) = (\mu_5, \mu_2)$.

By the results on p. 350 of Cramér (1946), the quantities $\mathrm{E}(m_j - \mu_j)$, $\mathrm{E}\{(m_j - \mu_j)^2\}$ and $\mathrm{E}\{(m_5 - \mu_5)(m_2 - \mu_2)\}$, for $j = 2, 5$, are all $O(n^{-1})$ under the assumption of sufficiently many finite moments. It follows directly from Theorem 7.1 that

$$\mathrm{E}\left(\frac{m_5}{m_2^{5/2}}\right) = \frac{\mu_5}{\mu_2^{5/2}} + O(n^{-1}).$$

Hence, combining this with the results of Lemma 7.2, we have

$$\mathrm{E}\left(\widehat{\kappa}'_{5,n} - \kappa'_5\right) = \mathrm{E}\left(\frac{m_5}{m_2^{5/2}} - \frac{\mu_5}{\mu_2^{5/2}}\right) - 10\,\mathrm{E}\left(\widehat{\kappa}'_3 - \kappa'_3\right) = O(n^{-1}). \qquad \square$$

7.2 Application: new confidence bounds for the mean

Let $X_1, X_2, \ldots, X_n$ denote a sample drawn randomly from a univariate distribution with mean μ and variance $0 < \sigma^2 < \infty$. We would now like to study the newly proposed confidence bounds when the goal is to construct a $100(1-\alpha)\%$ confidence bound for μ. As in Section 3.4.1, set

$$\mathbf{W}_k = \begin{bmatrix} X_k \\ X_k^2 \end{bmatrix}, \qquad k = 1, \ldots, n.$$

In the notation of Sections 4.1 and 5.1 we then have $\boldsymbol{v} = \mathrm{E}(\mathbf{W}_1) = (\mu, \mu^2 + \sigma^2)'$,

$$\bar{\mathbf{W}}_\ell = \begin{bmatrix} \ell^{-1} \sum_{k=1}^{\ell} X_k \\ \ell^{-1} \sum_{k=1}^{\ell} X_k^2 \end{bmatrix} \quad \text{and} \quad \bar{\mathbf{W}}_r = \begin{bmatrix} r^{-1} \sum_{k=\ell+1}^{n} X_k \\ r^{-1} \sum_{k=\ell+1}^{n} X_k^2 \end{bmatrix}.$$

Also, as in Section 3.4.1, let $g(x_1, x_2) = x_1$ and $h^2(x_1, x_2) = x_2 - x_1^2$ so that $\theta = g(\boldsymbol{v}) = \mu$ and $\beta^2 = h^2(\boldsymbol{v}) = \sigma^2$. The appropriate estimators for θ are then given by

$$\bar{X}_\ell := \widehat{\theta}_\ell = g(\bar{\mathbf{W}}_\ell) = \ell^{-1} \sum_{k=1}^{\ell} X_k \quad \text{and} \quad \bar{X}_r := \widehat{\theta}_r = g(\bar{\mathbf{W}}_r) = r^{-1} \sum_{k=\ell+1}^{n} X_k.$$

The estimators for β^2 are

$$\widehat{\sigma}_\ell^2 := \widehat{\beta}_\ell^2 = h^2(\bar{\mathbf{W}}_\ell) = \ell^{-1} \sum_{k=1}^{\ell} (X_k - \bar{X}_\ell)^2 \quad \text{and} \quad \widehat{\sigma}_r^2 := \widehat{\beta}_r^2 = h^2(\bar{\mathbf{W}}_r) = r^{-1} \sum_{k=\ell+1}^{n} (X_k - \bar{X}_r)^2.$$

As discussed in Chapter 3, for the case of the mean it has been shown in the literature that the polynomials P_1 and P_2 in (4.3) are given by

$$\begin{aligned} P_1(x) &= -\tfrac{1}{6}\kappa_3'(x^2-1), \\ P_2(x) &= -x\left\{\tfrac{1}{24}\kappa_4'(x^2-3)+\tfrac{1}{72}(\kappa_3')^2(x^4-10x^2+15)\right\}, \end{aligned} \tag{7.8}$$

where κ_3' and κ_4' denote the third and fourth cumulants of $(X_1-\mu)/\sigma$, respectively. The corresponding sample versions $\widehat{P}_{1,r}$ and $\widehat{P}_{2,r}$ based on the subsample $\mathscr{W}_r = \{\mathbf{W}_{\ell+1}, \mathbf{W}_{\ell+2}, \ldots, \mathbf{W}_n\}$ are given by

$$\begin{aligned} \widehat{P}_{1,r}(x) &= -\tfrac{1}{6}\widehat{\kappa}_{3,r}'(x^2-1), \\ \widehat{P}_{2,r}(x) &= -x\left\{\tfrac{1}{24}\widehat{\kappa}_{4,r}'(x^2-3)+\tfrac{1}{72}(\widehat{\kappa}_{3,r}')^2(x^4-10x^2+15)\right\}, \end{aligned}$$

where

$$\widehat{\kappa}_{3,r}' = \frac{r^{-1}\sum_{k=\ell+1}^{n}(X_k-\bar{X}_r)^3}{\widehat{\sigma}_r^3} \quad \text{and} \quad \widehat{\kappa}_{4,r}' = \frac{r^{-1}\sum_{k=\ell+1}^{n}(X_k-\bar{X}_r)^4}{\widehat{\sigma}_r^4} - 3.$$

For the percentile-t methods, we require expressions for the first three Edgeworth polynomials given in (5.2). As shown in Chapter 3, they are given by

$$\begin{aligned} Q_1(x) &= \tfrac{1}{6}\kappa_3'(2x^2+1), \\ Q_2(x) &= x\left(\tfrac{1}{12}\kappa_4'(x^2-3)-\tfrac{1}{18}\left(\kappa_3'\right)^2(x^4+2x^2-3)-\tfrac{1}{4}(x^2+3)\right), \\ Q_3(x) &= -\tfrac{1}{40}\kappa_5'\left(2x^4+8x^2+1\right)-\tfrac{1}{144}\kappa_4'\kappa_3'\left(4x^6-30x^4-90x^2-15\right) \\ &\quad +\tfrac{1}{1296}\left(\kappa_3'\right)^3\left(8x^8+28x^6-210x^4-525x^2-105\right)+\tfrac{1}{24}\kappa_3'\left(2x^6-3x^4-6x^2\right), \end{aligned} \tag{7.9}$$

with κ_5' denoting the first cumulant of $(X_1-\mu)/\sigma$. The corresponding sample versions $\widehat{Q}_{1,r}$, $\widehat{Q}_{2,r}$ and $\widehat{Q}_{3,r}$ based on the subsample $\mathscr{W}_r$ are given by

$$\begin{aligned} \widehat{Q}_{1,r}(x) &= \tfrac{1}{6}\widehat{\kappa}_{3,r}'(2x^2+1), \\ \widehat{Q}_{2,r}(x) &= x\left(\tfrac{1}{12}\widehat{\kappa}_{4,r}'(x^2-3)-\tfrac{1}{18}\left(\widehat{\kappa}_{3,r}'\right)^2(x^4+2x^2-3)-\tfrac{1}{4}(x^2+3)\right), \\ \widehat{Q}_{3,r}(x) &= -\tfrac{1}{40}\widehat{\kappa}_{5,r}'\left(2x^4+8x^2+1\right)-\tfrac{1}{144}\widehat{\kappa}_{4,r}'\widehat{\kappa}_{3,r}'\left(4x^6-30x^4-90x^2-15\right) \\ &\quad +\tfrac{1}{1296}\left(\widehat{\kappa}_{3,r}'\right)^3\left(8x^8+28x^6-210x^4-525x^2-105\right)+\tfrac{1}{24}\widehat{\kappa}_{3,r}'\left(2x^6-3x^4-6x^2\right), \end{aligned}$$

where

$$\widehat{\kappa}_{5,r}' = \frac{r^{-1}\sum_{k=\ell+1}^{n}(X_k-\bar{X}_r)^5}{\widehat{\sigma}_r^5} - 10\widehat{\kappa}_{3,r}'.$$

7.2.1 Hybrid percentile bound

We may now apply Theorem 4.1 to obtain a $100(1-\alpha)\%$ upper confidence bound for $\theta = \mu$. An expression for the constant $C_\theta(z_\alpha)$ appearing in the asymptotic coverage error is given in the following theorem.

Theorem 7.2. *Let $X_1, X_2, \ldots, X_n$ denote a sample drawn randomly from a univariate distribution with mean $\theta = \mu$ and variance $0 < \sigma^2 < \infty$. Under the assumptions of Theorem 4.1, we have*

$$C_\mu(z_\alpha) = \frac{\{\mu_4 + 3\sigma^4 + z_\alpha^2(\mu_4 - \sigma^4)\} z_\alpha \phi(z_\alpha)}{8\sigma^4} = \tfrac{1}{8}\{\kappa_4' + 6 + z_\alpha^2(\kappa_4' + 2)\} z_\alpha \phi(z_\alpha),$$

where $\mu_4 = \mathrm{E}\{(X_1 - \mu)^4\}$ and κ_4' denotes the fourth cumulant of $(X_1 - \mu)/\sigma$.

Proof. Assume that the moment conditions on X_1 and Cramér's condition required by Theorem 4.1 are met. It is easily verified that the functions g and h also satisfy the conditions of Theorem 4.1.

We will now investigate whether assumptions (A1)–(A7) of Theorem 4.1 are satisfied. Lemma A.11 confirms that assumptions (A1) and (A2) are met, i.e.,

$$\mathrm{E}(\widehat{\sigma}_r - \sigma) = O(r^{-1}) \quad \text{and} \quad \mathrm{E}\{(\widehat{\sigma}_r - \sigma)^4\} = O(r^{-2}).$$

It follows from Lemma 7.2 that

$$\mathrm{E}\left(\widehat{P}_{1,r}(x) - P_1(x)\right) = -\tfrac{1}{6}(x^2 - 1)\mathrm{E}\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right) = O(r^{-1}),$$

$$\mathrm{E}\left(\widehat{P}_{1,r}'(x) - P_1'(x)\right) = -\tfrac{1}{3}x\,\mathrm{E}\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right) = O(r^{-1}),$$

$$\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x) - P_1(x)\right)^4\right\} = \tfrac{1}{6^4}(x^2 - 1)^4\mathrm{E}\left\{\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right)^4\right\} = O(r^{-2}),$$

$$\mathrm{E}\left\{\left(\widehat{P}_{1,r}'(x) - P_1'(x)\right)^4\right\} = \tfrac{1}{3^4}x^4\,\mathrm{E}\left\{\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right)^4\right\} = O(r^{-2}).$$

Lemma 7.3, together with (7.5), implies that

$$\begin{aligned}
&\mathrm{E}\left(\widehat{P}_{2,r}(x) - P_2(x)\right) \\
&= -x\left\{\tfrac{1}{24}(x^2 - 3)\mathrm{E}\left(\widehat{\kappa}_{4,r}' - \kappa_4'\right) + \tfrac{1}{72}(x^4 - 10x^2 + 15)\mathrm{E}\left(\left(\widehat{\kappa}_{3,r}'\right)^2 - (\kappa_3')^2\right)\right\} = O(r^{-1}).
\end{aligned}$$

Lastly it follows from (7.5) and (7.7) that

$$\mathrm{E}\left\{\left(\left(\widehat{\kappa}_{3,r}'\right)^2 - (\kappa_3')^2\right)^2\right\} = \mathrm{E}\left\{\left(\widehat{\kappa}_{3,r}'\right)^4\right\} - 2(\kappa_3')^2\mathrm{E}\left\{\left(\widehat{\kappa}_{3,r}'\right)^2\right\} + (\kappa_3')^4 = O(r^{-1}),$$

so that, by the c_r-inequality and Lemma 7.3, we have

$$\begin{aligned}
&\mathrm{E}\left\{\left(\widehat{P}_{2,r}(x) - P_2(x)\right)^2\right\} \\
&= x^2\,\mathrm{E}\left\{\left(\tfrac{1}{24}(x^2 - 3)\left(\widehat{\kappa}_{4,r}' - \kappa_4'\right) + \tfrac{1}{72}(x^4 - 10x^2 + 15)\left(\left(\widehat{\kappa}_{3,r}'\right)^2 - (\kappa_3')^2\right)\right)^2\right\} \\
&\leq x^2\tfrac{2}{24^2}(x^2 - 3)^2\mathrm{E}\left\{\left(\widehat{\kappa}_{4,r}' - \kappa_4'\right)^2\right\} + x^2\tfrac{2}{72^2}(x^4 - 10x^2 + 15)^2\mathrm{E}\left\{\left(\left(\widehat{\kappa}_{3,r}'\right)^2 - (\kappa_3')^2\right)^2\right\} \\
&= O(r^{-1}).
\end{aligned}$$

Therefore, assumptions (A3)–(A7) are satisfied and we may apply Theorem 4.1.

To find the constant $C_\theta(z_\alpha) = C_\mu(z_\alpha)$, note that it follows from Lemma A.11 that

$$
\begin{aligned}
& -z_\alpha \phi(z_\alpha)\sigma^{-1}\,\mathrm{E}(\widehat{\sigma}_r - \sigma) + \tfrac{1}{2} z_\alpha^3 \phi(z_\alpha)\sigma^{-2}\,\mathrm{E}\left\{(\widehat{\sigma}_r - \sigma)^2\right\} \\
&= z_\alpha \phi(z_\alpha)\sigma^{-1}\left(\frac{\mu_4 + 3\sigma^4}{8\sigma^3 r}\right) + \tfrac{1}{2} z_\alpha^3 \phi(z_\alpha)\sigma^{-2}\left(\frac{\mu_4 - \sigma^4}{4\sigma^2 r}\right) \\
&= \frac{\left\{\mu_4 + 3\sigma^4 + z_\alpha^2\left(\mu_4 - \sigma^4\right)\right\} z_\alpha \phi(z_\alpha)}{8\sigma^4 r} + O(r^{-2}).
\end{aligned}
$$

Hence,

$$
C_\mu(z_\alpha) = \frac{\left\{\mu_4 + 3\sigma^4 + z_\alpha^2\left(\mu_4 - \sigma^4\right)\right\} z_\alpha \phi(z_\alpha)}{8\sigma^4}. \qquad \square
$$

7.2.2 Backwards percentile bound

For the case of the mean the coverage probability of the new backwards percentile bound $\widehat{\mathscr{I}}_B^N(\alpha)$ is given in Theorem 4.2, with constants $C_\mu(z_\alpha)$, $K_1(z_\alpha)$ and $K_2(z_\alpha)$ given by the following theorem.

Theorem 7.3. *Let $X_1, X_2, \dots, X_n$ denote a sample drawn randomly from a univariate distribution with mean $\theta = \mu$ and variance $0 < \sigma^2 < \infty$. Under the assumptions of Theorem 4.2, we have*

$$
\begin{aligned}
C_\mu(z_\alpha) &= \tfrac{1}{8}\left\{\kappa_4' + 6 + z_\alpha^2(\kappa_4' + 2)\right\} z_\alpha \phi(z_\alpha), \\
K_1(z_\alpha) &= \tfrac{1}{3}\kappa_3'\left(z_\alpha^2 - 1\right)\phi(z_\alpha), \\
K_2(z_\alpha) &= \tfrac{1}{18}(\kappa_3')^2 z_\alpha\left(z_\alpha^2 - 1\right)\left(z_\alpha^2 - 3\right)\phi(z_\alpha),
\end{aligned}
$$

with $\kappa_3' = \mathrm{E}\{(X_1 - \mu)^3/\sigma^3\}$ and $\kappa_4' = \mathrm{E}\{(X_1 - \mu)^4/\sigma^4\}$.

Moreover, if the sample $X_1, X_2, \dots, X_n$ is drawn from a symmetric *distribution, then $\kappa_3' = 0$ and consequently $K_1 = K_2 = 0$. Therefore, in the case of symmetry, choosing $\ell = \lfloor \gamma n^\psi \rfloor$ for some $\gamma > 0$ and $\frac{2}{3} < \psi \le 1$ we may write*

$$
\mathrm{P}\left(\mu \in \widehat{\mathscr{I}}_B^N(\ell, \alpha)\right) = \mathrm{P}\left(\mu \in \widehat{\mathscr{I}}_H^N(\ell, \alpha)\right) = 1 - \alpha + \frac{C_\mu(z_\alpha)}{n} + O(n^{-(2-\psi)} + n^{-3\psi/2}).
$$

In the special case where $\psi = 1$,

$$
\mathrm{P}\left(\mu \in \widehat{\mathscr{I}}_B^N(\ell, \alpha)\right) = \mathrm{P}\left(\mu \in \widehat{\mathscr{I}}_H^N(\ell, \alpha)\right) = 1 - \alpha + \frac{C_\mu(z_\alpha)}{(1-\gamma)n} + O(n^{-3/2}).
$$

Proof. In the proof of Theorem 7.2 we have shown that all assumptions of Theorem 4.1 are satisfied. To obtain the constants K_1 and K_2, first note that it follows directly from (7.8) that

$$
K_1(z_\alpha) = -2P_1(z_\alpha)\phi(z_\alpha) = \tfrac{1}{3}\kappa_3'\left(z_\alpha^2 - 1\right)\phi(z_\alpha).
$$

Hence,

$$
K_1'(z_\alpha) = \tfrac{1}{3}\kappa_3'\left\{2z_\alpha - z_\alpha\left(z_\alpha^2 - 1\right)\right\}\phi(z_\alpha) = -\tfrac{1}{3}\kappa_3' z_\alpha\left(z_\alpha^2 - 3\right)\phi(z_\alpha),
$$

so that

$$
K_2(z_\alpha) = K_1'(z_\alpha)P_1(z_\alpha) = \tfrac{1}{18}\left(\kappa_3'\right)^2 z_\alpha\left(z_\alpha^2 - 1\right)\left(z_\alpha^2 - 3\right)\phi(z_\alpha).
$$

The expression for $C_\mu(z_\alpha)$ was derived in the proof of Theorem 7.2. $\square$

7.2.3 Hybrid percentile-t bound

For the case of the mean, the coverage probability of the new hybrid percentile-t bound $\widehat{\mathscr{I}}_H^N(\alpha)$ is given in Theorem 5.1, with constant $D_\mu(z_\alpha)$ given by the following theorem.

Theorem 7.4. *Let $X_1, X_2, \ldots, X_n$ denote a sample drawn randomly from a univariate distribution with mean μ and variance $0 < \sigma^2 < \infty$. Under the assumptions of Theorem 5.1, we have*

$$\begin{aligned} D_\mu(z_\alpha) &= \frac{\{\mu_3(21\sigma^4 + 15\mu_4) - 12\sigma^2\mu_5\}(2z_\alpha^2 + 1)\phi(z_\alpha)}{48\sigma^7} \\ &= -\tfrac{1}{48}\{12\kappa_5' - 15\kappa_4'\kappa_3' + 54\kappa_3'\}(2z_\alpha^2 + 1)\phi(z_\alpha), \end{aligned}$$

with $\mu_k = \mathrm{E}\{(X_1 - \mu)^k\}$ and κ_j' denoting the jth cumulant of $(X_1 - \mu)/\sigma$.

Proof. Assume that the moment conditions on X_1 and Cramér's condition required by Theorem 5.1 are met. It is easily verified that the functions g and h also satisfy the conditions of Theorem 5.1.

We will now investigate whether assumptions (B1)–(B7) of Theorem 5.1 are satisfied. It follows from Lemma 7.2 that

$$\begin{aligned} \mathrm{E}\left(\widehat{Q}_{1,r}(x) - Q_1(x)\right) &= \tfrac{1}{6}(2x^2+1)\,\mathrm{E}\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right) = O(r^{-1}), \\ \mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x) - Q_1(x)\right)^4\right\} &= \tfrac{1}{6^4}(2x^2+1)^4\,\mathrm{E}\left\{\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right)^4\right\} = O(r^{-2}), \\ \mathrm{E}\left(\widehat{Q}_{1,r}'(x) - Q_1'(x)\right) &= \tfrac{2}{3}x\,\mathrm{E}\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right) = O(r^{-1}), \\ \mathrm{E}\left\{\left(\widehat{Q}_{1,r}'(x) - Q_1'(x)\right)^2\right\} &= \tfrac{2^2}{3^2}x^2\,\mathrm{E}\left\{\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right)^2\right\} = O(r^{-1}). \\ \mathrm{E}\left(\widehat{Q}_{1,r}''(x) - Q_1''(x)\right) &= \tfrac{2}{3}\,\mathrm{E}\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right) = O(r^{-1}), \\ \mathrm{E}\left\{\left(\widehat{Q}_{1,r}''(x) - Q_1''(x)\right)^2\right\} &= \tfrac{2^2}{3^2}\,\mathrm{E}\left\{\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right)^2\right\} = O(r^{-1}). \end{aligned}$$

Lemma 7.3, together with (7.5), implies that

$$\mathrm{E}\left(\widehat{Q}_{2,r}(x) - Q_2(x)\right) = x\left(\tfrac{1}{12}(x^2-3)\,\mathrm{E}\left(\widehat{\kappa}_{4,r}' - \kappa_4'\right) - \tfrac{1}{18}(x^4+2x^2-3)\,\mathrm{E}\left\{\left(\widehat{\kappa}_{3,r}'\right)^2 - (\kappa_3')^2\right\}\right) = O(r^{-1}).$$

From (7.5) and (7.7) we have that

$$\mathrm{E}\left\{\left(\left(\widehat{\kappa}_{3,r}'\right)^2 - (\kappa_3')^2\right)^2\right\} = \mathrm{E}\left(\left(\widehat{\kappa}_{3,r}'\right)^4\right) - 2(\kappa_3')^2\,\mathrm{E}\left\{\left(\widehat{\kappa}_{3,r}'\right)^2\right\} + (\kappa_3')^4 = O(r^{-1}),$$

so that, by the c_r-inequality and Lemma 7.3, we have

$$\begin{aligned} &\mathrm{E}\left\{\left(\widehat{Q}_{2,r}(x) - Q_2(x)\right)^2\right\} \\ &= x^2\,\mathrm{E}\left\{\left(\tfrac{1}{12}(x^2-3)\left(\widehat{\kappa}_{4,r}' - \kappa_4'\right) + \tfrac{1}{18}(x^4+2x^2-3)\left(\left(\widehat{\kappa}_{3,r}'\right)^2 - (\kappa_3')^2\right)\right)^2\right\} \\ &\le \tfrac{2}{12^2}x^2(x^2-3)^2\,\mathrm{E}\left\{\left(\widehat{\kappa}_{4,r}' - \kappa_4'\right)^2\right\} + \tfrac{2}{18^2}x^2(x^4+2x^2-3)^2\,\mathrm{E}\left\{\left(\left(\widehat{\kappa}_{3,r}'\right)^2 - (\kappa_3')^2\right)^2\right\} \\ &= O(r^{-1}). \end{aligned}$$

Similarly, since

$$Q_2'(x) = \tfrac{1}{4}\kappa_4'(x^2-1) - \tfrac{1}{18}\left(\kappa_3'\right)^2(5x^4+6x^2-3) - \tfrac{1}{4}(3x^2+3),$$

we have that

$$\mathrm{E}\left(\widehat{Q}_{2,r}'(x) - Q_2'(x)\right) = \tfrac{1}{4}(x^2-1)\,\mathrm{E}\left(\widehat{\kappa}_{4,r}' - \kappa_4'\right) - \tfrac{1}{18}(5x^4+6x^2-3)\,\mathrm{E}\left(\left(\widehat{\kappa}_{3,r}'\right)^2 - \left(\kappa_3'\right)^2\right) = O(r^{-1}),$$

and

$$\begin{aligned}
&\mathrm{E}\left\{\left(\widehat{Q}_{2,r}'(x) - Q_2'(x)\right)^2\right\} \\
&\le \tfrac{2}{4^2}\left(x^2-1\right)^2 \mathrm{E}\left\{\left(\widehat{\kappa}_{4,r}' - \kappa_4'\right)^2\right\} + \tfrac{2}{18^2}\left(x^4+6x^2-3\right)^2 \mathrm{E}\left\{\left(\left(\widehat{\kappa}_{3,r}'\right)^2 - \left(\kappa_3'\right)^2\right)^2\right\} \\
&= O(r^{-1}).
\end{aligned}$$

Finally, note that by the Cauchy-Schwarz inequality and Lemmas 7.2 and 7.3 it follows that

$$\left|\mathrm{E}\left\{\left(\widehat{\kappa}_{4,r}' - \kappa_4'\right)\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right)\right\}\right| \le \sqrt{\mathrm{E}\left\{\left(\widehat{\kappa}_{4,r}' - \kappa_4'\right)^2\right\}\mathrm{E}\left\{\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right)^2\right\}} = O(r^{-1}).$$

This implies that

$$\mathrm{E}\left\{\widehat{\kappa}_{4,r}'\widehat{\kappa}_{3,r}' - \kappa_4'\kappa_3'\right\} = O(r^{-1}).$$

Hence, by Lemmas 7.2, 7.3 and 7.4 we have that

$$\begin{aligned}
&\mathrm{E}\left(\widehat{Q}_{3,r}(x) - Q_3(x)\right) \\
&= -\tfrac{1}{40}\,\mathrm{E}\left(\widehat{\kappa}_{5,r}' - \kappa_5'\right)\left(2x^4+8x^2+1\right) - \tfrac{1}{144}\,\mathrm{E}\left(\widehat{\kappa}_{4,r}'\widehat{\kappa}_{3,r}' - \kappa_4'\kappa_3'\right)\left(4x^6-30x^4-90x^2-15\right) \\
&\quad + \tfrac{1}{1296}\,\mathrm{E}\left\{\left(\widehat{\kappa}_{3,r}'\right)^3 - \left(\kappa_3'\right)^3\right\}\left(8x^8+28x^6-210x^4-525x^2-105\right) \\
&\quad + \tfrac{1}{24}\,\mathrm{E}\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right)\left(2x^6-3x^4-6x^2\right). \\
&= O(r^{-1}).
\end{aligned}$$

Therefore, assumptions (B1)–(B7) are satisfied.

To determine the constant $D_\theta(z_\alpha) = D_\mu(z_\alpha)$, note that we have from Lemma 7.2 that

$$\begin{aligned}
\phi(z_\alpha)\,\mathrm{E}\left\{\widehat{Q}_{1,r}(z_\alpha) - Q_1(z_\alpha)\right\} &= \tfrac{1}{6}(2z_\alpha^2+1)\phi(z_\alpha)\,\mathrm{E}\left(\widehat{\kappa}_{3,r}' - \kappa_3'\right) \\
&= \frac{\left\{\mu_3\left(21\sigma^4+15\mu_4\right) - 12\sigma^2\mu_5\right\}(2z_\alpha^2+1)\phi(z_\alpha)}{48\sigma^7 r} + O(r^{-2}).
\end{aligned}$$

Hence, in the notation of Theorem 5.1, we have

$$D_\mu(z_\alpha) = \frac{\left\{\mu_3\left(21\sigma^4+15\mu_4\right) - 12\sigma^2\mu_5\right\}(2z_\alpha^2+1)\phi(z_\alpha)}{48\sigma^7}. \qquad \square$$

7.2.4 Backwards percentile-t bound

For the case of the mean, the coverage probability of the new backwards percentile-t bound $\widehat{\mathscr{I}}_B^N(\alpha)$ is given in Theorem 5.2, with constants $D_\mu(z_\alpha)$, $K_3(z_\alpha)$, $K_4(z_\alpha)$ and $K_5(z_\alpha)$ given by the following theorem.

Theorem 7.5. *Let $X_1, X_2, \ldots, X_n$ denote a sample drawn randomly from a univariate distribution with mean $\theta = \mu$ and variance $0 < \sigma^2 < \infty$. Under the assumptions of Theorem 5.2, we have*

$$D_\mu(z_\alpha) = -\tfrac{1}{48}\left\{12\kappa_5' - 15\kappa_4'\kappa_3' + 54\kappa_3'\right\}(2z_\alpha^2+1)\phi(z_\alpha),$$

$$K_3(z_\alpha) = -\tfrac{1}{3}\kappa_3'(2z_\alpha^2+1)\phi(z_\alpha),$$

$$K_4(z_\alpha) = \tfrac{1}{18}(\kappa_3')^2 z_\alpha(2z_\alpha^2+1)(2z_\alpha^2-3)\phi(z_\alpha),$$

and

$$K_5(z_\alpha) = \Big\{\tfrac{1}{20}\kappa_5'\left(2z_\alpha^4+8z_\alpha^2+1\right) - \tfrac{1}{24}\kappa_4'\kappa_3'\left(4z_\alpha^4+36z_\alpha^2+5\right) + \tfrac{1}{4}\kappa_3' z_\alpha^2\left(2z_\alpha^2-1\right) + \tfrac{1}{648}(\kappa_3')^3\left(-32z_\alpha^8+176z_\alpha^6+66z_\alpha^4+561z_\alpha^2+96\right)\Big\}\phi(z_\alpha), \quad (7.10)$$

with $\mu_k = \mathrm{E}\{(X-\mu)^k\}$ and κ_j' denoting the jth cumulant of $(X_1-\mu)/\sigma$.

Proof. In the proof of Theorem 7.4 we have shown that all assumptions of Theorem 5.1 are satisfied. Now, note that it follows directly from (7.9) that

$$K_3(z_\alpha) = -2Q_1(z_\alpha)\phi(z_\alpha) = -\tfrac{1}{3}\kappa_3'(2z_\alpha^2+1)\phi(z_\alpha).$$

Since,

$$K_3'(z_\alpha) = \tfrac{1}{3}\kappa_3' z_\alpha(2z_\alpha^2-3)\phi(z_\alpha),$$

we have

$$K_4(z_\alpha) = K_3'(z_\alpha)Q_1(z_\alpha) = \tfrac{1}{18}\left(\kappa_3'\right)^2 z_\alpha(2z_\alpha^2+1)(2z_\alpha^2-3)\phi(z_\alpha).$$

Note that

$$K_3''(z_\alpha) = -\tfrac{1}{3}\kappa_3\left(2z_\alpha^4-9z_\alpha^2+3\right)\phi(z_\alpha).$$

After substituting (3.10) and (7.9) in

$$K_5(z_\alpha) = \tfrac{1}{2}Q_1^2(z_\alpha)K_3''(z_\alpha) + Q_2^{cf}(z_\alpha)K_3'(z_\alpha) - 2Q_3(z_\alpha)\phi(z_\alpha),$$

the required expression for K_5 may be obtained by straightforward algebra.

The expression for $D_\mu(z_\alpha)$ was derived in the proof of Theorem 7.4. □

Remark 7.1. Consider the case where X_1 has a symmetric distribution. Then $\kappa_3' = \kappa_5' = 0$, whence $Q_1(x) = Q_3(x) = 0$, $\forall x \in \mathbb{R}$. Consequently, $K_3(z_\alpha) = K_4(z_\alpha) = K_5(z_\alpha) = D_\theta(z_\alpha) = 0$ so that the coverage error of $\widehat{\mathscr{I}}_B^N(\ell,\alpha)$ reduces to $O(\ell^{-2})$. See Remark 5.1.

Chapter 8

Extension to the slope parameter in the linear regression model

It has been shown in the literature (see Hall, 1992) that the properties of both the *standard* percentile and *standard* percentile-t bootstrap bounds carry over to regression problems. For example, in a simple linear regression model, confidence bounds for the slope parameter constructed using the standard methods $\widehat{\mathscr{I}}_H$ and $\widehat{\mathscr{J}}$ have reduced coverage errors of $O(n^{-1})$ and $O(n^{-3/2})$, respectively. In this chapter we investigate only the performance of our new hybrid percentile-t bound (the two percentile and the backwards percentile-t bounds can be treated similarly) in the linear regression setup. We show that the coverage error of this bound is typically $O(n^{-2})$. To facilitate exposition, we consider only simple linear regression, but the results may be extended to multiple linear regression.

Suppose we observe pairs $\mathscr{X}_n = \{(x_1, Y_1), \ldots, (x_n, Y_n)\}$ generated by the simple linear regression model

$$Y_i = c' + x_i d + \varepsilon_i,$$

where c' and d are unknown, nonrandom constants and $\{\varepsilon_1, \ldots, \varepsilon_n\}$ is a sequence of iid random variables from an unknown distribution with zero mean and constant variance $0 < \sigma^2 < \infty$. Throughout we assume that the x_i are fixed. Setting $c = c' + d\bar{x}_n$, with $\bar{x}_n = n^{-1}\sum_{i=1}^{n} x_i$, we may rewrite the above model as

$$Y_i = c + (x_i - \bar{x}_n)d + \varepsilon_i.$$

Throughout this chapter we denote the jth central moment of ε_1 by $\mu_j = \mathrm{E}(\varepsilon^j)$. Also, denote the jth cumulant of ε_1/σ by κ'_j.

8.1 Confidence bound for the slope parameter

We may now construct our new *hybrid percentile-t* confidence bound for d. As before, split the original sample in two disjoint parts

$$\mathscr{X}_\ell = \{(x_1, Y_1), \ldots, (x_\ell, Y_\ell)\} \quad \text{and} \quad \mathscr{X}_r = \{(x_{\ell+1}, Y_{\ell+1}), \ldots, (x_n, Y_n)\},$$

for some integer $2 \le \ell \le n-2$. Writing $\sigma^2_{x,\ell} = \ell^{-1}\sum_{k=1}^{\ell}(x_k - \bar{x}_\ell)^2$, with $\bar{x}_\ell = \ell^{-1}\sum_{k=1}^{\ell} x_k$, the least-squares estimators (based solely on $\mathscr{X}_\ell$) for d and c are given by

$$\widehat{d}_\ell = \frac{1}{\ell\sigma^2_{x,\ell}}\sum_{k=1}^{\ell}(x_k - \bar{x}_\ell)Y_k \quad \text{and} \quad \widehat{c}_\ell = \bar{Y}_\ell = \ell^{-1}\sum_{k=1}^{\ell} Y_k.$$

Define the following cumulants based on the x_i:

$$\gamma_{x,\ell} = \frac{1}{\ell\sigma^3_{x,\ell}}\sum_{i=1}^{\ell}(x_i - \bar{x}_\ell)^3,$$
$$\kappa_{x,\ell} = \frac{1}{\ell\sigma^4_{x,\ell}}\sum_{i=1}^{\ell}(x_i - \bar{x}_\ell)^4 - 3,$$
$$\tau_{x,\ell} = \frac{1}{\ell\sigma^5_{x,\ell}}\sum_{i=1}^{\ell}(x_i - \bar{x}_\ell)^5 - 10\gamma_{x,\ell}.$$

Let $\gamma_{x,r}$, $\kappa_{x,r}$ and $\tau_{x,r}$ be the same functions of $\mathscr{X}_r$ as $\gamma_{x,\ell}$, $\kappa_{x,\ell}$ and $\tau_{x,\ell}$ are of $\mathscr{X}_\ell$.

Since the variance of $\widehat{d}_\ell$ is $\sigma^2/(\ell\sigma^2_{x,\ell})$, the new $(1-\alpha)$-level percentile-t confidence bound for d (corresponding to $\widehat{\mathscr{J}}^N_H$) is given by

$$\widehat{\mathscr{K}}^N_H(m,\alpha) := \left(-\infty,\ \widehat{d}_\ell - \ell^{-1/2}\sigma^{-1}_{x,\ell}\widehat{\sigma}_\ell\widetilde{\eta}_{m,r,\alpha}\right],$$

where

$$\widehat{\sigma}^2_\ell = \frac{1}{\ell}\sum_{k=1}^{\ell} e_k^2 := \frac{1}{\ell}\sum_{k=1}^{\ell}\left(Y_k - \bar{Y}_\ell - (x_k - \bar{x}_\ell)\widehat{d}_\ell\right)^2,$$

and

$$\widetilde{\eta}_{m,r,\alpha} := z_\alpha + m^{-1/2}\widehat{Q}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{Q}^{cf}_{2,r}(z_\alpha) + m^{-3/2}\widehat{Q}^{cf}_{3,r}(z_\alpha).$$

The Cornish-Fisher polynomials $\widehat{Q}^{cf}_{j,r}$ appearing in this expression are completely determined by the Edgeworth polynomials $\widehat{Q}_{j,r}$ through the relations (3.9), (3.10) and (3.11), where, as shown in Section 3.5, the $\widehat{Q}_{j,r}$ are given by

$$\widehat{Q}_{1,r}(u) = -\tfrac{1}{6}\widetilde{\kappa}'_{3,r}\gamma_{x,r}He_2(u),$$
$$\widehat{Q}_{2,r}(u) = -\tfrac{1}{24}\widetilde{\kappa}'_{4,r}\kappa_{x,r}He_3(u) - \tfrac{1}{72}(\widetilde{\kappa}'_{3,r})^2\gamma^2_{x,r}He_5(u) - \tfrac{1}{4}(u^2+5)u,$$
$$\widehat{Q}_{3,r}(u) = -\tfrac{1}{120}\widetilde{\kappa}'_{5,r}\left\{\tau_{x,r}He_4(u) - 30\gamma_{x,r}He_2(u)\right\} - \tfrac{1}{144}\widetilde{\kappa}'_{4,r}\widetilde{\kappa}'_{3,r}\left\{\kappa_{x,r}\gamma_{x,r}He_6(u) + 45\gamma_{x,r}He_2(u)\right\}$$
$$- \tfrac{1}{1296}(\widetilde{\kappa}'_{3,r})^3\gamma^3_{x,r}He_8(u) - \tfrac{1}{24}\widetilde{\kappa}'_{3,r}\gamma_{x,r}(u^2-1)u^4,$$

with

$$\widetilde{m}_{j,r} = \frac{1}{r}\sum_{k=\ell+1}^{n} e_k^j, \qquad \widetilde{\kappa}'_{3,r} = \frac{\widetilde{m}_{3,r}}{\widetilde{m}^{3/2}_{2,r}}, \qquad \widetilde{\kappa}'_{4,r} = \frac{\widetilde{m}_{4,r}}{\widetilde{m}^2_{2,r}} - 3, \qquad \widetilde{\kappa}'_{5,r} = \frac{\widetilde{m}_{5,r}}{\widetilde{m}^{5/2}_{2,r}} - 10\widetilde{\kappa}'_{3,r}.$$

Theorem 8.1. *Suppose that ε_1 has sufficiently many finite moments and satisfies Cramér's condition. Assume $\limsup_{n\to\infty}\max_{1\le i\le n}|x_i - \bar{x}_n| < \infty$, $\gamma_{x,r} - \gamma_{x,\ell} = O(n^{-(1+\delta)})$ for some $\delta > 0$, $\kappa_{x,r} - \kappa_{x,\ell} = O(n^{-1})$, and $\tau_{x,r} - \tau_{x,\ell} = O(n^{-1})$. Then, if $m = \ell = O(r)$ and $\ell \to \infty$ as $n \to \infty$, we have that*

$$\mathrm{P}\left(d \in \widehat{\mathscr{K}}^N_H(\ell,\alpha)\right) = 1 - \alpha + \frac{E_d(z_\alpha)}{\ell^{1/2}r} + O(\ell^{-2} + \ell^{-1/2}n^{-(1+\delta)}),$$

with

$$E_d(z_\alpha) = \tfrac{1}{48}\gamma_{x,r}\left(12\kappa_5' - 15\kappa_4'\kappa_3' + 66\kappa_3'\right)(z_\alpha^2 - 1)\phi(z_\alpha),$$

where κ_j' *denotes the jth cumulant of* ε_1/σ.
Moreover, if we choose $\ell = \lfloor \gamma n^\psi \rfloor$ *for some* $\gamma > 0$ *and* $\frac{2}{3} < \psi < 1$, *then*

$$\mathrm{P}\left(d \in \widehat{\mathscr{K}}_H^N(\ell,\alpha)\right) = \begin{cases} 1-\alpha + \dfrac{E_d(z_\alpha)}{\gamma^{1/2}n^{(2+\psi)/2}} + O(n^{-\min\{2-\psi/2,\,2\psi,\,1+\delta+\psi/2\}}) & \textit{if } E_d(z_\alpha) \neq 0, \\ 1-\alpha + O(n^{-2\psi} + n^{-(1+\delta+\psi/2)}) & \textit{if } E_d(z_\alpha) = 0. \end{cases}$$

In the case where $\psi = 1$ *and* $0 < \gamma < 1$,

$$\mathrm{P}\left(d \in \widehat{\mathscr{K}}_H^N(\ell,\alpha)\right) = 1-\alpha + \frac{E_d(z_\alpha)}{\gamma^{1/2}(1-\gamma)n^{3/2}} + O(n^{-2} + n^{-(3/2+\delta)}),$$

which becomes

$$\mathrm{P}\left(d \in \widehat{\mathscr{K}}_H^N(\ell,\alpha)\right) = 1-\alpha + O(n^{-2} + n^{-(3/2+\delta)})$$

if ε_1 *has a symmetric distribution around zero.*

Remark 8.1. If the design points are regularly spaced, say $x_i = u\frac{i}{n} + v$, $i = 1,\ldots,n$, for some constants u and v, then the assumptions on the x_i in Theorem 8.1 can easily be verified. In fact, since in this case $\gamma_{x,r} = \gamma_{x,\ell} = 0$, we can take $\delta = \infty$. Consequently, $E_d(z_\alpha) = 0$ so that the coverage error reduces to $O(n^{-2})$, *even if the errors have an asymmetric distribution*.

Proof of Theorem 8.1. By (3.24) one may write the coverage probability of the confidence bound $\widehat{\mathscr{K}}_H^N(m,\alpha)$ as

$$\mathrm{P}\left(d \le \widehat{d}_\ell - \ell^{-1/2}\sigma_{x,\ell}^{-1}\widehat{\sigma}_\ell\widetilde{\eta}_{m,r,\alpha}\right) = \mathrm{E}\left(\Psi_t\left(\widetilde{\eta}_{m,r,\alpha}\right)\right),$$

with

$$\begin{aligned} \Psi_t(x) &:= \mathrm{P}(\ell^{1/2}(\widehat{d}_\ell - d)\sigma_{x,\ell}/\widehat{\sigma}_\ell \ge x) \\ &= 1 - \Phi(x) - \ell^{-1/2}Q_{1,\ell}(x)\phi(x) - \ell^{-1}Q_{2,\ell}(x)\phi(x) - \ell^{-3/2}Q_{3,\ell}(x)\phi(x) + O(\ell^{-2}), \end{aligned}$$

uniformly in x, where, from (3.25) and Lemma 3.5,

$$\begin{aligned} Q_{1,\ell}(u) &= -\tfrac{1}{6}\kappa_3'\gamma_{x,\ell}He_2(u), \\ Q_{2,\ell}(u) &= -\tfrac{1}{24}\kappa_4'\kappa_{x,\ell}He_3(u) - \tfrac{1}{72}(\kappa_3')^2\gamma_{x,\ell}^2He_5(u) - \tfrac{1}{4}(u^2+5)u, \\ Q_{3,\ell}(u) &= -\tfrac{1}{120}\kappa_5'\left\{\tau_{x,\ell}He_4(u) - 30\gamma_{x,\ell}He_2(u)\right\} - \tfrac{1}{144}\kappa_4'\kappa_3'\left\{\kappa_{x,\ell}\gamma_{x,\ell}He_6(u) + 45\gamma_{x,\ell}He_2(u)\right\} \\ &\quad - \tfrac{1}{1296}(\kappa_3')^3\gamma_{x,\ell}^3He_8(u) - \tfrac{1}{24}\kappa_3'\gamma_{x,\ell}(u^2-1)u^4. \end{aligned}$$

Following exactly the same steps as in the proof of Theorem 5.1 yields

$$\mathrm{P}\left(\theta \in \widehat{\mathscr{K}}_H^N(\ell,\alpha)\right) = 1-\alpha + \frac{\phi(z_\alpha)\mathrm{E}\left\{\widehat{Q}_{1,r}(z_\alpha) - Q_{1,\ell}(z_\alpha)\right\}}{\ell^{1/2}} + O(\ell^{-2}). \tag{8.1}$$

The latter result relies on the following properties:

(i) $\mathrm{E}\left\{\{\widehat{Q}_{1,r}(x)-Q_{1,\ell}(x)\right)^{i}\right\}=O(r^{-1}),\ i=1,\ldots,4,$

(ii) $\mathrm{E}\left\{(\widehat{Q}'_{1,r}(x)-Q'_{1,\ell}(x))^{i}\right\}=O(r^{-1}),\ i=1,2,$

(iii) $\mathrm{E}\left\{(\widehat{Q}''_{1,r}(x)-Q''_{1,\ell}(x))^{i}\right\}=O(r^{-1}),\ i=1,2,$

(iv) $\mathrm{E}\left\{(\widehat{Q}_{2,r}(x)-Q_{2,\ell}(x))^{i}\right\}=O(r^{-1}),\ i=1,2,$

(v) $\mathrm{E}\left\{(\widehat{Q}'_{2,r}(x)-Q'_{2,\ell}(x))^{i}\right\}=O(r^{-1}),\ i=1,2,$

(vi) $\mathrm{E}\left\{\widehat{Q}_{3,r}(x)-Q_{3,\ell}(x)\right\}=O(r^{-1}).$

Note that, since $\gamma_{x,r}-\gamma_{x,\ell}=O(n^{-1})$, we have

$$\begin{aligned}\mathrm{E}\left(\widehat{Q}_{1,r}(u)-Q_{1,\ell}(u)\right)&=-\tfrac{1}{6}He_2(u)\mathrm{E}\left(\widetilde{\kappa}'_{3,r}\gamma_{x,r}-\kappa'_3\gamma_{x,\ell}\right)\\&=-\tfrac{1}{6}He_2(u)\mathrm{E}\left(\widetilde{\kappa}'_{3,r}\gamma_{x,r}-\kappa'_3\gamma_{x,r}\right)+O(n^{-1})\\&=-\tfrac{1}{6}He_2(u)\gamma_{x,r}\,\mathrm{E}\left(\widetilde{\kappa}'_{3,r}-\kappa'_3\right)+O(n^{-1})\\&=O(r^{-1}+n^{-1})=O(r^{-1}),\end{aligned}$$

where we made use of Lemma C.9. Analogously,

$$\begin{aligned}\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(u)-Q_{1,\ell}(u)\right)^4\right\}&=-\tfrac{1}{6^4}He_2^4(u)\mathrm{E}\left\{\left(\widetilde{\kappa}'_{3,r}\gamma_{x,r}-\kappa'_3\gamma_{x,\ell}\right)^4\right\}\\&=-\tfrac{1}{6^4}He_2^4(u)\mathrm{E}\left\{\left(\widetilde{\kappa}'_{3,r}\gamma_{x,r}-\kappa'_3\gamma_{x,r}\right)^4\right\}+O(n^{-1})\\&=-\tfrac{1}{6^4}He_2^4(u)\gamma_{x,r}^4\,\mathrm{E}\left\{\left(\widetilde{\kappa}'_{3,r}-\kappa'_3\right)^4\right\}+O(n^{-1})\\&=O(r^{-2}+n^{-1})=O(r^{-1}).\end{aligned}$$

In a similar fashion, properties (i)–(vi) above may easily be verified using the results of Appendix C.

We now derive the constant $E_d(z_\alpha)$. Note that by the result of Lemma C.9 we have that

$$\phi(z_\alpha)\mathrm{E}\left\{\widehat{Q}_{1,r}(z_\alpha)-Q_1(z_\alpha)\right\}=\frac{1}{48r}\left(12\kappa'_5-15\kappa'_4\kappa'_3+66\kappa'_3\right)\gamma_{x,r}(z_\alpha^2-1)\phi(z_\alpha)+O(n^{-(1+\delta)}+r^{-2}),$$

which is the coefficient of $\ell^{-1/2}$ in (8.1). Therefore,

$$E_d(z_\alpha)=\tfrac{1}{48}\gamma_{x,r}\left(12\kappa'_5-15\kappa'_4\kappa'_3+66\kappa'_3\right)(z_\alpha^2-1)\phi(z_\alpha).$$

The remainder of the proof proceeds exactly as that of Theorem 5.1. □

Chapter 9

Simulation study

A modest simulation study was carried out to compare the standard upper bounds $\widehat{\mathscr{I}}_H$, $\widehat{\mathscr{I}}_B$, $\widehat{\mathscr{J}}$ and the upper bound proposed by Chung and Lee (2001), which we denote by C-L, with the newly developed upper bounds $\widehat{\mathscr{I}}_H^N$, $\widehat{\mathscr{I}}_B^N$, $\widehat{\mathscr{J}}_H^N$ and $\widehat{\mathscr{J}}_B^N$, where the parameter of interest is the population mean. Monte Carlo estimates were calculated for the non-coverage probability (NC) and expected size of the upper bound (EUB) resulting from each method. We considered the performance of the different bounds for samples of sizes $n = 50, 100, 200$ drawn from the uniform(0,1), standard Laplace, χ_3^2 and $F_{6,20}$ distributions. The new bounds were evaluated for $\alpha = 5\%$ and different choices of ℓ such that the assumption $\ell = O(r)$ required by the theorems is satisfied. Each entry in Tables 9.1–9.5 is based on 100 000 independent Monte Carlo trials, each comprising 10 000 bootstrap samples. Standard errors were found to be negligibly small and are not reported. All calculations were done in R.

Recall that for distributions with $\kappa_3' = 0$ the standard *percentile* bounds $\widehat{\mathscr{I}}_H$ and $\widehat{\mathscr{I}}_B$ have coverage errors of order $O(n^{-1})$ (see Hall, 1988), which is of the same order as the coverage errors produced by the newly proposed percentile bounds $\widehat{\mathscr{I}}_H^N$ and $\widehat{\mathscr{I}}_B^N$. Therefore, for the two symmetric distributions we report in Tables 9.1 and 9.2 results only for the *percentile-t* type bounds $\widehat{\mathscr{J}}$ and $\widehat{\mathscr{J}}_H^N$, which have coverage errors of order $O(n^{-1})$ and $O(n^{-2})$, respectively. We omit the results for $\widehat{\mathscr{J}}_B^N$, since its behavior is almost identical to that of $\widehat{\mathscr{J}}_H^N$ (see Remark 7.1). We do not consider distributions with $\kappa_4' = 0$ (e.g., the normal distribution), since in this case the various confidence bounds have almost identical performance in terms of coverage error. For the uniform and Laplace distributions $\kappa_4' = -1.2$ and $\kappa_4' = 3$ respectively.

Comparing Tables 9.1 and 9.2 it is evident that, for both the uniform and Laplace distributions, the new bound $\widehat{\mathscr{J}}_H^N$ significantly outperforms the standard percentile-t bound $\widehat{\mathscr{J}}$ in terms of coverage error for all sample sizes considered. This striking performance is visible even for a relatively small sample. Although the upper bound $\widehat{\mathscr{J}}_H^N$ is slightly larger than $\widehat{\mathscr{J}}$ in each case (as expected), a suitable choice of ℓ greatly diminishes this difference. Note that a larger choice of ℓ corresponds to a smaller upper bound, which agrees with the definition of $\widehat{\mathscr{J}}_H^N$.

The results for the skewed distributions presented in Tables 9.3–9.5 show that for most choices of ℓ the newly proposed *percentile* bounds $\widehat{\mathscr{I}}_H^N$ and $\widehat{\mathscr{I}}_B^N$ significantly outperform

Table 9.1: Results of the existing percentile-t method $\widehat{\mathscr{J}}$ for two symmetric distributions.

	$n = 50$		$n = 100$		$n = 200$	
Distribution	NC	EUB	NC	EUB	NC	EUB
Uniform	0.045	0.568	0.048	0.548	0.049	0.534
Laplace	0.059	0.334	0.056	0.234	0.054	0.165

Table 9.2: Results of the new hybrid percentile-t method $\widehat{\mathscr{J}}_H^N$ for two symmetric distributions.

	$n = 50$			$n = 100$			$n = 200$		
Distribution	ℓ	NC	EUB	ℓ	NC	EUB	ℓ	NC	EUB
Uniform	25	0.050	0.598	50	0.050	0.568	100	0.050	0.548
	30	0.050	0.589	60	0.050	0.562	120	0.050	0.544
	35	0.051	0.582	70	0.050	0.557	140	0.050	0.540
	40	0.050	0.577	80	0.050	0.554	160	0.050	0.538
Laplace	30	0.050	0.436	60	0.050	0.304	120	0.050	0.213
	35	0.051	0.402	70	0.050	0.281	140	0.050	0.197
	40	0.050	0.375	80	0.050	0.263	160	0.050	0.185
	45	0.050	0.352	90	0.050	0.248	180	0.050	0.174

the standard percentile bounds $\widehat{\mathscr{I}}_H$ and $\widehat{\mathscr{I}}_B$ in terms of coverage error. Furthermore, it is clear that the bound C-L, which also has coverage error $O(n^{-1})$, performs slightly better than $\widehat{\mathscr{I}}_H^N$, but slightly worse than $\widehat{\mathscr{I}}_B^N$. The performance of the new *percentile-t* bound $\widehat{\mathscr{J}}_H^N$ is comparable to that of the standard percentile-t bound $\widehat{\mathscr{J}}$. We omit the results for $\widehat{\mathscr{J}}_B^N$, as its coverage error $O(n^{-1/2})$ compares poorly to the error $O(n^{-3/2})$ attained by $\widehat{\mathscr{J}}_H^N$ (see Theorem 5.2). Again, the size of the upper bound can be decreased with an appropriate choice of ℓ. Notice that, in agreement with theory, the coverage errors of all considered bounds converge to the nominal coverage error α as the sample size n is increased.

Overall, it is clear that the improvement in coverage accuracy comes at the cost of a larger upper bound. However, by making a suitable choice of ℓ when splitting the sample one may achieve a significantly improved coverage probability with only a slight increase in the magnitude of the upper bound. Ideally, a data-based choice of ℓ is needed which, however, will require deeper analysis and we leave a detailed study for future research.

Table 9.3: Results of the existing methods for two skewed distributions.

		$n=50$		$n=100$		$n=200$	
Distribution	Type	NC	EUB	NC	EUB	NC	EUB
χ^2_3	$\widehat{\mathscr{I}}_H$	0.092	3.537	0.077	3.388	0.068	3.278
	$\widehat{\mathscr{I}}_B$	0.080	3.576	0.068	3.390	0.062	3.289
	C-L	0.064	3.641	0.057	3.436	0.053	3.304
	$\widehat{\mathscr{I}}$	0.056	3.674	0.052	3.453	0.051	3.309
$F_{5,8}$	$\widehat{\mathscr{I}}_H$	0.135	1.612	0.112	1.539	0.096	1.484
	$\widehat{\mathscr{I}}_B$	0.115	1.650	0.097	1.562	0.084	1.498
	C-L	0.090	1.732	0.079	1.599	0.070	1.517
	$\widehat{\mathscr{I}}$	0.080	1.772	0.070	1.627	0.064	1.531

Table 9.4: Results of the new methods for the χ^2_3 distribution.

	$n=50$			$n=100$			$n=200$		
Type	ℓ	NC	EUB	ℓ	NC	EUB	ℓ	NC	EUB
$\widehat{\mathscr{I}}^N_H$	20	0.065	3.820	40	0.058	3.599	80	0.054	3.433
	25	0.068	3.734	50	0.059	3.536	100	0.055	3.388
	30	0.074	3.667	60	0.062	3.489	120	0.055	3.354
	35	0.081	3.610	70	0.065	3.451	140	0.058	3.327
$\widehat{\mathscr{I}}^N_B$	20	0.050	3.909	40	0.046	3.649	80	0.045	3.459
	25	0.055	3.802	50	0.049	3.575	100	0.047	3.408
	30	0.061	3.720	60	0.052	3.520	120	0.049	3.371
	35	0.070	3.651	70	0.057	3.476	140	0.051	3.341
$\widehat{\mathscr{I}^N_H}$	20	0.059	4.174	40	0.053	3.770	80	0.050	3.514
	25	0.059	4.003	50	0.053	3.669	100	0.051	3.451
	30	0.060	3.883	60	0.054	3.597	120	0.051	3.407
	35	0.062	3.791	70	0.054	3.543	140	0.052	3.372

Table 9.5: Results of the new methods for the $F_{5,8}$ distribution.

	$n=50$			$n=100$			$n=200$		
Type	ℓ	NC	EUB	ℓ	NC	EUB	ℓ	NC	EUB
$\widehat{\mathscr{I}}^N_H$	20	0.088	1.748	40	0.073	1.644	80	0.065	1.563
	25	0.096	1.706	50	0.078	1.612	100	0.068	1.540
	30	0.105	1.671	60	0.085	1.588	120	0.072	1.522
	35	0.118	1.640	70	0.093	1.566	140	0.077	1.507
$\widehat{\mathscr{I}}^N_B$	20	0.063	1.828	40	0.052	1.695	80	0.048	1.594
	25	0.075	1.764	50	0.060	1.650	100	0.053	1.563
	30	0.087	1.715	60	0.069	1.617	120	0.059	1.540
	35	0.103	1.672	70	0.080	1.589	140	0.066	1.521
$\widehat{\mathscr{I}^N_H}$	20	0.074	2.070	40	0.062	1.830	80	0.057	1.666
	25	0.078	1.937	50	0.066	1.748	100	0.060	1.616
	30	0.083	1.848	60	0.069	1.693	120	0.062	1.582
	35	0.088	1.781	70	0.073	1.651	140	0.064	1.556

Appendix A

Moments of sample moments and cumulants (iid case)

This appendix constitutes a brief account of sample moments and cumulants, along with some asymptotic expressions for moments of these quantities. These results are required to derive the theoretical results in the main part of the text.

Throughout this appendix X denotes a univariate random variable with distribution function F with mean μ and variance σ^2. Let $X_1, X_2, \ldots, X_n$ be a sample of size n drawn randomly from F.

A.1 Moments, cumulants and k-statistics

We devote this section to formally define the notions of moments and cumulants. As many of the derivations in this appendix make use of Fisher's so-called k-statistics, we here also provide the definition thereof as well as a brief history.

A.1.1 Central moments

For $k = 1, 2, \ldots$, the kth *central (population) moment* of X is defined by

$$\mu_k = \mathrm{E}\left\{(X-\mu)^k\right\}.$$

The corresponding kth *central sample moment* of X (based on the sample $X_1, X_2, \ldots, X_n$) is defined by

$$m_k = \frac{1}{n}\sum_{j=1}^{n}\left(X_j - \bar{X}_n\right)^k,$$

with $\bar{X}_n = n^{-1}\sum_{j=1}^{n} X_j$.

Remark A.1. Note that, by definition, $\sigma^2 = \mu_2$. Accordingly, let

$$\widehat{\sigma}_n^2 := m_2 = \frac{1}{n}\sum_{j=1}^{n}\left(X_j - \bar{X}_n\right)^2.$$

A.1.2 Cumulants

Define the *cumulant-generating function* $K(t)$ of X as the natural logarithm of the characteristic function of X, i.e.,

$$K(t) = \log \mathrm{E}\left(e^{itX}\right), \qquad i^2 = -1.$$

If the jth moment of X exists, the jth *cumulant* κ_j of X is defined as the coefficient of $(it)^j/j!$ in a power series expansion of $K(t)$:

$$K(t) = \sum_{j=1}^{\infty} \kappa_j \frac{(it)^j}{j!}.$$

Texts such as Cramér (1946, p. 187) and Stuart and Ord (1994, p. 90f) provide expressions for the first few cumulants of X in terms of central moments of X. Below we reproduce such expressions for the first eight cumulants of X:

$$\begin{aligned}
\kappa_1 &= 0,\\
\kappa_2 &= \mu_2,\\
\kappa_3 &= \mu_3,\\
\kappa_4 &= \mu_4 - 3\mu_2^2,\\
\kappa_5 &= \mu_5 - 10\mu_3\mu_2\\
\kappa_6 &= \mu_6 - 15\mu_4\mu_2 - 10\mu_3^2 + 30\mu_2^3\\
\kappa_7 &= \mu_7 - 21\mu_5\mu_2 - 35\mu_4\mu_3 + 210\mu_3\mu_2^2,\\
\kappa_8 &= \mu_8 - 28\mu_6\mu_2 - 56\mu_5\mu_3 - 35\mu_4^2 + 420\mu_4\mu_2^2 + 560\mu_3^2\mu_2 - 630\mu_2^4.
\end{aligned} \tag{A.1}$$

The *sample* cumulant $\widehat{\kappa}_{j,n}$ corresponding to κ_j is obtained by substituting in the expression for κ_j the central population moments for the corresponding central sample moments, i.e.,

$$\begin{aligned}
\widehat{\kappa}_{2,n} &= m_2,\\
\widehat{\kappa}_{3,n} &= m_3,\\
\widehat{\kappa}_{4,n} &= m_4 - 3m_2^2, \text{ etc.}
\end{aligned} \tag{A.2}$$

In the main text we are particularly interested in the *standardised cumulants* of X, by which we mean the cumulants of the standardised quantity $Y = (X - \mu)/\sigma$. Similar to the non-standardised versions, the standardised cumulants can be found by inspecting coefficients in a power series expansion of $\log \mathrm{E}(e^{itY})$. Alternatively the jth *standardised* cumulant of X, which we denote by κ'_j, may be obtained by multiplying the jth *non-standardised* cumulant of X by σ^{-j}, i.e.,

$$\kappa'_j = \frac{\kappa_j}{\sigma^j}.$$

This follows from the behaviour of cumulants under affine transformations and hence they are sometimes referred to as *semi-invariants* (Dressel, 1940; Stuart and Ord, 1994, see).

The first few standardised cumulants of X are

$$\kappa_2' = \frac{\kappa_2}{\sigma^2} = 1,$$
$$\kappa_3' = \frac{\kappa_3}{\sigma^3} = \frac{\mu_3}{\sigma^3},$$
$$\kappa_4' = \frac{\kappa_4}{\sigma^4} = \frac{\mu_4}{\sigma^4} - 3.$$

As before, the standardised *sample* cumulant $\widehat{\kappa}'_{j,n}$ corresponding to κ_j' is obtained by substituting in the expression for κ_j' the central population moments for the corresponding central sample moments.

Before moving on, it is interesting and useful to note that, for a fixed sample size n and any $j = 2,3,\ldots$, the quantity $|m_j/\widehat{\sigma}_n^j|$ appearing in the expressions of the standardised cumulants may be bounded by a power of $n^{1/2}$. This fact is stated and proved in the following lemma.

Lemma A.1. *If X_1 has sufficiently many finite moments, then*

$$\left|\frac{m_j}{\widehat{\sigma}_n^j}\right| = \left|\frac{m_j}{m_2^{j/2}}\right| \le n^{(j-2)/2}, \qquad j = 2,3,\ldots.$$

Proof. Following the method of proof given in Cramér (1946, p. 357) for $j = 3$, observe that

$$\begin{aligned}
\left|\frac{m_j}{m_2^{j/2}}\right| &= \left|\frac{n^{-1}\sum_{i=1}^n (X_i - \bar{X}_n)^j}{n^{-j/2}\left(\sum_{j=1}^n (X_j - \bar{X}_n)^2\right)^{j/2}}\right| \\
&\le n^{(j-2)/2}\sum_{i=1}^n \left|\frac{(X_i - \bar{X}_n)^j}{\left(\sum_{j=1}^n (X_j - \bar{X}_n)^2\right)^{j/2}}\right| \\
&= n^{(j-2)/2}\sum_{i=1}^n \left(\frac{(X_i - \bar{X}_n)^2}{\sum_{j=1}^n (X_j - \bar{X}_n)^2}\right)^{j/2} \\
&\le n^{(j-2)/2}\sum_{i=1}^n \frac{(X_i - \bar{X}_n)^2}{\sum_{j=1}^n (X_j - \bar{X}_n)^2} \\
&= n^{(j-2)/2}.
\end{aligned}$$

□

A.1.3 k-statistics

It is well-known that the algebraic process of deriving moments of m_p can become extremely laborious as soon as one moves beyond the simplest cases. This prompted researchers to devise new ways of deriving moments of sample moments, including the use of tables of symmetric functions. Fisher (1930) discovered that these calculations may be simplified considerably by making use of combinatorial methods. His method employs so-called *k-statistics*, which are symmetric functions k_p, $p \ge 1$, defined such that the mean of each k_p equals the pth *non-standardised* population cumulant of X, i.e.,

$$\mathrm{E}\left(k_p\right) = \kappa_p.$$

Obtaining cumulant expressions for the k_p proved to be much simpler than obtaining similar expressions for the moments m_p. The interested reader is referred to Stuart and Ord (1994) who treat the subject of k-statistics extensively.

For the derivations in the following sections, we will require the following expressions of sample moments in terms of k-statistics (cf. Stuart and Ord, 1994, p. 422):

$$\begin{aligned} m_2 &= \frac{n-1}{n} k_2, \\ m_3 &= \frac{(n-1)(n-2)}{n^2} k_3, \\ m_4 &= \frac{(n-1)(n-2)(n-3)}{n^2(n+1)} k_4 + \frac{3(n-1)^3}{n^2(n+1)} k_2^2. \end{aligned} \tag{A.3}$$

From expressions such as these it is straightforward to calculate any moment $\mathrm{E}(m_p^r)$, $r \geq 2$. The only requirement is to have available expressions for the non-central moments $\mathrm{E}(k_p^r)$.

Let $\kappa(p^r)$ denote the the rth cumulant of k_p. Noting that $\kappa(p) = \kappa_p$, we have from Stuart and Ord (1994, p. 88) that

$$\begin{aligned} \mathrm{E}\left(k_p\right) &= \kappa_p, \\ \mathrm{E}\left(k_p^2\right) &= \kappa(p^2) + \kappa_p^2, \\ \mathrm{E}\left(k_p^3\right) &= \kappa(p^3) + 3\kappa(p^2)\kappa_p + \kappa_p^3, \\ \mathrm{E}\left(k_p^4\right) &= \kappa(p^4) + 4\kappa(p^3)\kappa_p + 3\kappa(p^2)^2 + 6\kappa(p^2)\kappa_p^2 + \kappa_p^4, \\ \mathrm{E}\left(k_p^5\right) &= \kappa(p^5) + 5\kappa(p^4)\kappa_p + 10\kappa(p^3)\kappa(p^2) + 10\kappa(p^3)\kappa_p^2 + 15\kappa(p^2)^2\kappa_p \\ &\quad + 10\kappa(p^2)\kappa_p^3 + \kappa_p^5, \\ \mathrm{E}\left(k_p^6\right) &= \kappa(p^6) + 6\kappa(p^5)\kappa_p + 15\kappa(p^4)\kappa(p^2) + 15\kappa(p^4)\kappa_p^2 + 10\kappa(p^3)^2 + 60\kappa(p^3)\kappa(p^2)\kappa_p \\ &\quad + 20\kappa(p^3)\kappa_p^2 + 15\kappa(p^2)^3 + 45\kappa(p^2)^2\kappa_p^2 + 15\kappa(p^2)\kappa_p^4 + \kappa_p^6. \end{aligned} \tag{A.4}$$

Exact expressions for $\kappa(p^r)$ in terms of the cumulants of X have been derived in literature and are given (up to the tenth order) in Stuart and Ord (1994, p. 432ff).

To be able to calculate joint moments of the central sample moments, such as $\mathrm{E}(m_3 m_2^2)$, we will also require expressions for the joint moments of the k-statistics in terms of joint cumulants of X. Kendall (1940) introduced a method for easily deriving such expressions from univariate expressions such as the ones given in (A.4). Cook (1951) used this method to calculate such formulae for moments up to the sixth order, and we reproduce below the ones we require. In these expressions, $\kappa(p^r q^s)$ is given by the coefficient of $(it_1)^r(it_2)^s/(r!s!)$ in a power series expansion of $\log \mathrm{E}(e^{it_1 k_p + it_2 k_q})$.

$$\begin{aligned} \mathrm{E}\left(k_p k_q\right) &= \kappa(pq) + \kappa_p\kappa_q, \\ \mathrm{E}\left(k_p^2 k_q\right) &= \kappa(p^2 q) + \kappa(p^2)\kappa_q + 2\kappa(pq)\kappa_p + \kappa_p^2\kappa_q, \\ \mathrm{E}\left(k_p^3 k_q\right) &= \kappa(p^3 q) + \kappa(p^3)\kappa_q + 3\kappa(p^2 q)\kappa_p + 3\kappa(p^2)\kappa(pq) + 3\kappa(p^2)\kappa_p\kappa_q \\ &\quad + 3\kappa(pq)\kappa_p^2 + \kappa_p^3\kappa_q, \end{aligned} \tag{A.5}$$

$$
\begin{aligned}
\mathrm{E}\left(k_p^2 k_q^2\right) &= \kappa(p^2q^2) + 2\kappa(p^2q)\kappa_q + \kappa(p^2)\kappa_q^2 + \kappa(p^2)\kappa(q^2) + 2\kappa(pq^2)\kappa_p + 2\kappa(pq)^2 \\
&\quad + 4\kappa(pq)\kappa_p\kappa_q + \kappa_p^2\kappa_q^2 + \kappa_p^2\kappa(q^2), \\
\mathrm{E}\left(k_p^4 k_q\right) &= \kappa(p^4q) + \kappa(p^4)\kappa_q + 4\kappa(p^3q)\kappa_p + 4\kappa(p^3)\kappa(pq) + 4\kappa(p^3)\kappa_p\kappa_q \\
&\quad + 6\kappa(p^2q)\kappa(p^2) + 6\kappa(p^2q)\kappa_p^2 + 3\kappa(p^2)^2\kappa_q + 12\kappa(p^2)\kappa(pq)\kappa_p \\
&\quad + 6\kappa(p^2)\kappa_p^2\kappa_q + 4\kappa(pq)\kappa_p^3 + \kappa_p^4\kappa_q.
\end{aligned} \tag{A.6}
$$

A.2 Moments of non-central sample moments

In this section we derive asymptotic expressions for moments of some of the non-central sample moments.

Lemma A.2. *Assume that X has sufficiently many finite moments. Then*

(i)
$$
\begin{aligned}
\mathrm{E}\left(m_3^2\right) &= \kappa_3^2 + \frac{\kappa_6 + 9\kappa_4\kappa_2 + 3\kappa_3^2 + 6\kappa_2^3}{n} + O(n^{-2}) \\
&= \mu_3^2 + \frac{\mu_6 - 6\mu_4\mu_2 - 7\mu_3^2 + 9\mu_2^3}{n} + O(n^{-2}),
\end{aligned}
$$

(ii)
$$
\begin{aligned}
\mathrm{E}\left(m_3^3\right) &= \kappa_3^3 + \frac{3\kappa_3\left(\kappa_6 + 9\kappa_4\kappa_2 + 6\kappa_3^2 + 6\kappa_2^3\right)}{n} + O(n^{-2}) \\
&= \mu_3^3 + \frac{3\mu_3\left(\mu_6 - 6\mu_4\mu_2 - 4\mu_3^2 + 9\mu_2^3\right)}{n} + O(n^{-2}).
\end{aligned}
$$

Proof.

(i) From (A.3) and (A.4) we have

$$
\mathrm{E}\left(m_3^2\right) = \frac{(n-1)^2(n-2)^2}{n^4}\mathrm{E}\left(k_3^2\right) = \frac{(n-1)^2(n-2)^2}{n^4}\left\{\kappa(3^2) + \kappa_3^2\right\}.
$$

Noting that

$$
\frac{(n-1)^2(n-2)^2}{n^4} = 1 - \frac{6}{n} + O(n^{-2}),
$$

it follows from (12.40) of Stuart and Ord (1994) that

$$
\begin{aligned}
\mathrm{E}\left(m_3^2\right) &= \kappa_3^2 - \frac{6\kappa_3^2}{n} + \frac{\kappa_6 + 9\kappa_4\kappa_2 + 9\kappa_3^2 + 6\kappa_2^3}{n} + O(n^{-2}) \\
&= \kappa_3^2 + \frac{\kappa_6 + 9\kappa_4\kappa_2 + 3\kappa_3^2 + 6\kappa_2^3}{n} + O(n^{-2}).
\end{aligned}
$$

Substituting the cumulants for their corresponding expressions in terms of central moments (see (A.1)) yields the second result.

(ii) From (A.3) and (A.4) we have

$$
\mathrm{E}\left(m_3^3\right) = \frac{(n-1)^3(n-2)^3}{n^6}\mathrm{E}\left(k_3^3\right) = \frac{(n-1)^3(n-2)^3}{n^6}\left\{\kappa(3^3) + 3\kappa(3^2)\kappa_3 + \kappa_3^3\right\}.
$$

Noting that

$$
\frac{(n-1)^3(n-2)^3}{n^6} = 1 - \frac{9}{n} + O(n^{-2}),
$$

it follows from (12.40) of Stuart and Ord (1994) that

$$\begin{aligned}
\mathrm{E}\left(m_3^3\right) &= \kappa_3^3 - \frac{9\kappa_3^3}{n} + \frac{3\kappa_3}{n}\left\{\kappa_6 + 9\kappa_4\kappa_2 + 9\kappa_3^2 + 6\kappa_2^3\right\} + O(n^{-2}) \\
&= \kappa_3^3 + \frac{3\kappa_3\left(\kappa_6 + 9\kappa_4\kappa_2 + 6\kappa_3^2 + 6\kappa_2^3\right)}{n} + O(n^{-2}).
\end{aligned}$$

Substituting the cumulants for their corresponding expressions in terms of central moments (see (A.1)) yields the second result. □

Lemma A.3. *Assume that X has sufficiently many finite moments. Then*

(i) $$\mathrm{E}(m_4) = \mu_4 - \frac{4\mu_4 - 6\mu_2^2}{n} + O(n^{-2}),$$

(ii) $$\mathrm{E}\left(m_4^2\right) = \mu_4^2 + \frac{\mu_8 - 8\mu_5\mu_3 + 12\mu_4\mu_2^2 - 9\mu_4^2 + 16\mu_3^2\mu_2}{n} + O(n^{-2}),$$

(iii) $$\mathrm{E}\left(m_4^3\right) = \mu_4^3 + \frac{3\mu_8\mu_4 - 24\mu_5\mu_4\mu_3 - 15\mu_4^3 + 18\mu_4^2\mu_2^2 + 48\mu_4\mu_3^2\mu_2}{n} + O(n^{-2}).$$

Proof.

(i) Follows directly from (27.5.1) of Cramér (1946).

(ii) By (27.5.4) of Cramér (1946) we know that

$$\mathrm{Var}(m_4) = \frac{\mu_8 - 8\mu_5\mu_3 - \mu_4^2 + 16\mu_3^2\mu_2}{n} + O(n^{-2}).$$

Hence, using the result in *(i)*,

$$\begin{aligned}
\mathrm{E}\left(m_4^2\right) &= \mathrm{Var}(m_4) + \{\mathrm{E}(m_4)\}^2 \\
&= \frac{\mu_8 - 8\mu_5\mu_3 - \mu_4^2 + 16\mu_3^2\mu_2}{n} + \left(\mu_4 - \frac{4\mu_4 - 6\mu_2^2}{n}\right)^2 + O(n^{-2}) \\
&= \mu_4^2 + \frac{\mu_8 - 8\mu_5\mu_3 - \mu_4^2 + 16\mu_3^2\mu_2}{n} - \frac{8\mu_4^2 - 12\mu_4\mu_2^2}{n} + O(n^{-2}) \\
&= \mu_4^2 + \frac{\mu_8 - 8\mu_5\mu_3 + 12\mu_4\mu_2^2 - 9\mu_4^2 + 16\mu_3^2\mu_2}{n} + O(n^{-2}).
\end{aligned}$$

(iii) From (A.3) it follows that

$$\begin{aligned}
m_4^3 = {} & \frac{(n-1)^3(n-2)^3(n-3)^3}{n^6(n+1)^3}k_4^3 + \frac{27(n-1)^9}{n^6(n+1)^3}k_2^6 + \frac{9(n-1)^5(n-2)^2(n-3)^2}{n^6(n+1)^3}k_4^2k_2^2 \\
& + \frac{27(n-1)^7(n-2)(n-3)}{n^6(n+1)^3}k_4k_2^4,
\end{aligned}$$

whence, by making use of Taylor approximation and the expressions in (A.4), (A.5) and (A.6),

$$\begin{aligned}
\mathrm{E}\left(m_4^3\right) &= \left(1-\frac{21}{n}\right)\mathrm{E}\left(k_4^3\right)+27\left(1-\frac{12}{n}\right)\mathrm{E}\left(k_2^6\right)+9\left(1-\frac{18}{n}\right)\mathrm{E}\left(k_4^2k_2^2\right)\\
&\quad +27\left(1-\frac{15}{n}\right)\mathrm{E}\left(k_4k_2^4\right)+O(n^{-2})\\
&= \left(1-\frac{21}{n}\right)\{\kappa(4^3)3\kappa(4^2)\kappa_4+\kappa_4^3\}+27\left(1-\frac{12}{n}\right)\{\kappa(2^6)+6\kappa(2^5)\kappa_2+15\kappa(2^4)\kappa(2^2)\\
&\quad +15\kappa(2^4)\kappa_2^2+10\kappa(2^3)^2+60\kappa(2^3)\kappa(2^2)\kappa_2+20\kappa(2^3)\kappa_2^2+15\kappa(2^2)^3\\
&\quad +45\kappa(2^2)^2\kappa_2^2+15\kappa(2^2)\kappa_2^4+\kappa_2^6\}+9\left(1-\frac{18}{n}\right)\{\kappa(4^22^2)+2\kappa(2^24)\kappa_4+\kappa(2^2)\kappa_4^2\\
&\quad +\kappa(2^2)\kappa(4^2)+2\kappa(2\cdot 4^2)\kappa_2+2\kappa(2\cdot 4)^2+4\kappa(2\cdot 4)\kappa_2\kappa_4+\kappa_2^2\kappa_4^2+\kappa_2^2\kappa(4^2)\}\\
&\quad +27\left(1-\frac{15}{n}\right)\{\kappa(2^4\cdot 4)+\kappa(2^4)\kappa_4+4\kappa(2^3\cdot 4)\kappa_2+4\kappa(2^3)\kappa(2\cdot 4)+4\kappa(2^3)\kappa_4\kappa_2\\
&\quad +6\kappa(2^2\cdot 4)\kappa(2^2)+6\kappa(2^2\cdot 4)\kappa_2^2+3\kappa(2^2)^2\kappa_4+12\kappa(2^2)\kappa(2\cdot 4)\kappa_2+6\kappa(2^2)\kappa_2^2\kappa_4\\
&\quad +4\kappa(2\cdot 4)\kappa_2^3+\kappa_4\kappa_2^4\}+O(n^{-2}).
\end{aligned}$$

Using the results in Paragraph 12.6 of Stuart and Ord (1994), this expression reduces to

$$\begin{aligned}
\mathrm{E}\left(m_4^3\right) &= \kappa_4^3+27\kappa_2^6+9\kappa_4^2\kappa_2^2+27\kappa_4\kappa_2^4+\frac{1}{n}\{3\kappa_8\kappa_4+9\kappa_8\kappa_2^2+84\kappa_6\kappa_4\kappa_2+144\kappa_5\kappa_4\kappa_3\\
&\quad +252\kappa_6\kappa_2^3+90\kappa_4^3+828\kappa_4^2\kappa_2^2+648\kappa_4\kappa_3^2\kappa_2+432\kappa_5\kappa_3\kappa_2^2+1908\kappa_4\kappa_2^4\\
&\quad +1944\kappa_3^2\kappa_2^3+702\kappa_2^6\}+O(n^{-2}).
\end{aligned}$$

Note that by (3.38) of Stuart and Ord (1994) we have

$$\mu_4^3=\left(\kappa_4+3\kappa_2^2\right)=\kappa_4^3+27\kappa_2^6+9\kappa_4^2\kappa_2^2+27\kappa_4\kappa_2^4.$$

Using this fact and substituting each cumulant by its corresponding expression in terms of central population moments (see (3.43) in Stuart and Ord, 1994) yield the result. □

Lemma A.4. *Assume that X has sufficiently many finite moments. Then*

(i)
$$\begin{aligned}
\mathrm{E}(m_3m_2) &= \kappa_3\kappa_2+\frac{\kappa_5+2\kappa_3\kappa_2}{n}+O(n^{-2})\\
&= \mu_3\mu_2+\frac{\mu_5-8\mu_3\mu_2}{n}+O(n^{-2}),
\end{aligned}$$

(ii)
$$\begin{aligned}
\mathrm{E}\left(m_3m_2^2\right) &= \kappa_3\kappa_2^2+\frac{2\kappa_5\kappa_2+\kappa_4\kappa_3+9\kappa_3\kappa_2^2}{n}+O(n^{-2})\\
&= \mu_3\mu_2^2+\frac{2\mu_5\mu_2+\mu_4\mu_3-14\mu_3\mu_2^2}{n}+O(n^{-2}).
\end{aligned}$$

(iii)
$$\begin{aligned}
\mathrm{E}\left(m_3^2m_2\right) &= \kappa_3^2\kappa_2+\frac{\kappa_6\kappa_2+2\kappa_5\kappa_3+9\kappa_4\kappa_2^2+14\kappa_3^2\kappa_2+6\kappa_2^4}{n}+O(n^{-2})\\
&= \mu_3^2\mu_2+\frac{\mu_6\mu_2+2\mu_5\mu_3-6\mu_4\mu_2^2-16\mu_3^2\mu_2+9\mu_2^4}{n}+O(n^{-2}),
\end{aligned}$$

Proof.

(i) Multiplying the expressions for m_2 and m_3 in (A.3) and taking expected values, we have from (A.4) that

$$\begin{aligned}\mathrm{E}(m_3m_2) &= \frac{(n-1)^2(n-2)}{n^3}\mathrm{E}(k_3k_2)\\ &= \left(1-\frac{4}{n}+O(n^{-2})\right)\{\kappa(32)+\kappa_3\kappa_2\}.\end{aligned}$$

From (12.47) of Stuart and Ord (1994) it then follows that

$$\begin{aligned}\mathrm{E}(m_3m_2) &= \left(1-\frac{4}{n}\right)\{\kappa_3\kappa_2+\kappa_5+6\kappa_3\kappa_2\}+O(n^{-2})\\ &= \kappa_3\kappa_2-\frac{4\kappa_3\kappa_2}{n}+\frac{\kappa_5+6\kappa_3\kappa_2}{n}+O(n^{-2})\\ &= \kappa_3\kappa_2+\frac{\kappa_5+2\kappa_3\kappa_2}{n}+O(n^{-2}).\end{aligned}$$

Substituting the cumulants for their corresponding expressions in terms of central moments (see (A.1)) yields the second result.

(ii) Multiplying the expressions for m_2^2 and m_3 in (A.3) and taking expected values, we have from (A.4) that

$$\begin{aligned}\mathrm{E}\left(m_3m_2^2\right) &= \frac{(n-1)^3(n-2)}{n^4}\mathrm{E}\left(k_3k_2^2\right)\\ &= \left(1-\frac{5}{n}+O(n^{-2})\right)\left\{\kappa(32^2)+\kappa(2^2)\kappa_3+2\kappa(32)\kappa_2+\kappa_3\kappa_2^2\right\}.\end{aligned}$$

From (12.35), (12.47) and (12.59) of Stuart and Ord (1994) it then follows that

$$\begin{aligned}\mathrm{E}\left(m_3m_2^2\right) &= \left(1-\frac{5}{n}\right)\left(\frac{\kappa_3}{n}\{\kappa_4+2\kappa_2^2\}+\frac{2\kappa_2}{n}\{\kappa_5+6\kappa_3\kappa_2\}+\kappa_3\kappa_2^2\right)+O(n^{-2})\\ &= \kappa_3\kappa_2^2+\frac{2\kappa_5\kappa_2+\kappa_4\kappa_3+9\kappa_3\kappa_2^2}{n}+O(n^{-2}).\end{aligned}$$

Substituting the cumulants for their corresponding expressions in terms of central moments (see (A.1)) yields the second result.

(iii) Multiplying the expressions for m_2 and m_3^2 in (A.3) and taking expected values, we have from (A.4) that

$$\begin{aligned}\mathrm{E}\left(m_3^2m_2\right) &= \frac{(n-1)^3(n-2)^2}{n^5}\mathrm{E}\left(k_3^2k_2\right)\\ &= \left(1-\frac{7}{n}+O(n^{-2})\right)\left\{\kappa(3^22)+\kappa(3^2)\kappa_2+2\kappa(32)\kappa_3+\kappa_3^2\kappa_2\right\}.\end{aligned}$$

From (12.40), (12.47) and (12.63) of Stuart and Ord (1994) it then follows that

$$\begin{aligned}\mathrm{E}\left(m_3^2m_2\right) &= \left(1-\frac{7}{n}\right)\left(\frac{\kappa_2}{n}\{\kappa_6+9\kappa_4\kappa_2+9\kappa_3^2+6\kappa_2^3\}+\frac{2\kappa_3}{n}\{\kappa_5+6\kappa_3\kappa_2\}+\kappa_3^2\kappa_2\right)+O(n^{-2})\\ &= \kappa_3^2\kappa_2+\frac{\kappa_6\kappa_2+2\kappa_5\kappa_3+9\kappa_4\kappa_2^2+14\kappa_3^2\kappa_2+6\kappa_2^4}{n}+O(n^{-2}).\end{aligned}$$

Substituting the cumulants for their corresponding expressions in terms of central moments (see (A.1)) yields the second result. □

Lemma A.5. *Assume that X has sufficiently many finite moments. Then*

(i) $$\mathrm{E}(m_4 m_2) = \left(\kappa_4 + 3\kappa_2^2\right)\kappa_2 + \frac{\kappa_6 + 9\kappa_4\kappa_2 + 6\kappa_3^2 + 3\kappa_2^3}{n} + O(n^{-2})$$
$$= \mu_4\mu_2 + \frac{\mu_6 - 6\mu_4\mu_2 - 4\mu_3^2 + 6\mu_2^3}{n} + O(n^{-2}),$$

(ii) $$\mathrm{E}\left(m_4 m_2^2\right) = (\kappa_4 + 3\kappa_2)\kappa_2^2 + \frac{2\kappa_6\kappa_2 + 27\kappa_4\kappa_2^2 + \kappa_4^2 + 12\kappa_3^2\kappa_2 + 18\kappa_2^4}{n} + O(n^{-2})$$
$$= \mu_4\mu_2^2 + \frac{2\mu_6\mu_2 - 9\mu_4\mu_2^2 + \mu_4^2 - 8\mu_3^2\mu_2 + 6\mu_2^4}{n} + O(n^{-2}),$$

(iii) $$\mathrm{E}\left(m_4^2 m_2\right) = \left(\kappa_4 + 3\kappa_2^2\right)^2\kappa_2 + \frac{1}{n}\{\kappa_8\kappa_2 + 2\kappa_6\kappa_4 + 34\kappa_6\kappa_2^2 + 48\kappa_5\kappa_3\kappa_2 + 12\kappa_4\kappa_3^2$$
$$+ 53\kappa_4^2\kappa_2 + 270\kappa_4\kappa_2^3 + 252\kappa_3^2\kappa_2^2 + 123\kappa_2^5\} + O(n^{-2})$$
$$= \mu_4^2\mu_2 + \frac{1}{n}\{\mu_8\mu_2 + 2\mu_6\mu_4 - 8\mu_5\mu_3\mu_2 - 12\mu_4^2\mu_2 - 8\mu_4\mu_3^2$$
$$+ 12\mu_4\mu_2^3 + 16\mu_3^2\mu_2^2\} + O(n^{-2}).$$

Proof.

(i) (27.5.6) of Cramér (1946) implies that

$$\mathrm{E}\{(m_4 - \mu_4)(m_2 - \mu_2)\} = \frac{\mu_6 - \mu_4\mu_2 - 4\mu_3^2}{n} + O(n^{-2}).$$

Expanding and rearranging terms we obtain

$$\mathrm{E}(m_4 m_2) = \mu_2\,\mathrm{E}(m_4) + \mu_4\,\mathrm{E}(m_2) - \mu_4\mu_2 + \frac{\mu_6 - \mu_4\mu_2 - 4\mu_3^2}{n} + O(n^{-2}),$$

which, by Lemma A.3 above and (27.4.1) of Cramér (1946), reduces to

$$\mathrm{E}(m_4 m_2) = \mu_4\mu_2 + \frac{\mu_6 - 6\mu_4\mu_2 - 4\mu_3^2 + 6\mu_2^3}{n} + O(n^{-2}).$$

This result has been verified using Fisher's method of k-statistics discussed earlier.

(ii) From (A.3) we have that

$$\mathrm{E}\left(m_4 m_2^2\right) = \frac{(n-1)^3(n-2)(n-3)}{n^4(n+1)}\mathrm{E}\left(k_4 k_2^2\right) + \frac{3(n-1)^5}{n^4(n+1)}\mathrm{E}\left(k_2^4\right)$$
$$= \left(1 - \frac{9}{n}\right)\mathrm{E}\left(k_4 k_2^2\right) + 3\left(1 - \frac{6}{n}\right)\mathrm{E}\left(k_2^4\right) + O(n^{-2}).$$

Using (A.4), (A.5) and (A.6), along with the results in Paragraph 12.16 of Stuart and Ord (1994), we obtain

$$\mathrm{E}\left(m_4 m_2^2\right) = \left(1 - \frac{9}{n}\right)\{\kappa(2^2 4) + \kappa(2^2)\kappa_4 + 2\kappa(2\cdot 4)\kappa_2 + \kappa_4\kappa_2^2\}$$
$$+ 3\left(1 - \frac{6}{n}\right)\{\kappa(2^4) + 4\kappa(2^3)\kappa_2 + 3\kappa(2^2)^2 + 6\kappa(2^2)\kappa_2^2 + \kappa_2^4\} + O(n^{-2})$$
$$= \kappa_4\kappa_2^2 + 3\kappa_2^4 + \frac{1}{n}\{\kappa_4^2 + 2\kappa_4\kappa_2^2 + 2\kappa_6\kappa_2 + 16\kappa_4\kappa_2^2 + 12\kappa_3^2\kappa_2\}$$

$$+\frac{3}{n}\left\{6\kappa_4\kappa_2^2+12\kappa_2^4\right\}-\frac{9\kappa_4\kappa_2^2}{n}-\frac{18\kappa_2^4}{n}+O(n^{-2})$$
$$=(\kappa_4+3\kappa_2)\kappa_2^2+\frac{1}{n}\left\{2\kappa_6\kappa_2+27\kappa_4\kappa_2^2+\kappa_4^2+12\kappa_3^2\kappa_2+18\kappa_2^4\right\}.$$

The result in terms of central moments is obtain by substituting all population cumulants with their corresponding expressions given in (A.1).

(iii) Note that by (A.3)

$$m_4^2=\frac{(n-1)^2(n-2)^2(n-3)^2}{n^4(n+1)^2}k_4^2+\frac{9(n-1)^6}{n^4(n+1)^2}k_2^4+\frac{6(n-1)^4(n-2)(n-3)}{n^4(n+1)^2}k_4k_2^2.$$

Therefore,

$$\begin{aligned}\mathrm{E}\left(m_4^2m_2\right)&=\frac{(n-1)^3(n-2)^2(n-3)^2}{n^5(n+1)^2}\mathrm{E}\left(k_4^2k_2\right)+\frac{9(n-1)^7}{n^5(n+1)^2}\mathrm{E}\left(k_2^5\right)\\&\quad+\frac{6(n-1)^5(n-2)(n-3)}{n^5(n+1)^2}\mathrm{E}\left(k_4k_2^3\right)\\&=\left(1-\frac{15}{n}\right)\mathrm{E}\left(k_4^2k_2\right)+9\left(1-\frac{9}{n}\right)\mathrm{E}\left(k_2^5\right)+6\left(1-\frac{12}{n}\right)\mathrm{E}\left(k_4k_2^3\right)+O(n^{-2}).\end{aligned}$$

By (A.4), (A.5) and (A.6) above and the results in Paragraph 12.16 of Stuart and Ord (1994) we have

$$\begin{aligned}\mathrm{E}\left(m_4^2m_2\right)&=\left(1-\frac{15}{n}\right)\left\{\kappa(4^22)+\kappa(4^2)\kappa_2+2\kappa(42)\kappa_4+\kappa_4^2\kappa_2\right\}\\&\quad+9\left(1-\frac{9}{n}\right)\left\{\kappa(2^5)+5\kappa(2^4)\kappa_2+10\kappa(2^3)\kappa(2^2)+10\kappa(2^3)\kappa_2^2+15\kappa(2^2)^2\kappa_2\right.\\&\quad\left.+10\kappa(2^2)\kappa_2^3+\kappa_2^5\right\}+6\left(1-\frac{12}{n}\right)\left\{\kappa(2^34)+\kappa(2^3)\kappa_4+3\kappa(2^24)\kappa_2\right.\\&\quad\left.+3\kappa(2^2)\kappa(2\cdot4)+3\kappa(2^2)\kappa_2\kappa_4+3\kappa(2\cdot4)\kappa_2^2+\kappa_4\kappa_2^3\right\}+O(n^{-2})\\&=\kappa_4^2\kappa_2+9\kappa_2^5+6\kappa_4\kappa_2^3+\frac{1}{n}\left\{\kappa_2\left(\kappa_8+16\kappa_6\kappa_2+48\kappa_5\kappa_3+34\kappa_4^2+72\kappa_4\kappa_2^2\right.\right.\\&\quad\left.\left.+144\kappa_3^2\kappa_2+24\kappa_2^4\right)+2\kappa_4\left(\kappa_6+8\kappa_4\kappa_2+6\kappa_3^2\right)\right\}+\frac{9}{n}\left\{10\kappa_2^3\left(\kappa_4+2\kappa_2^2\right)\right\}\\&\quad+\frac{6}{n}\left\{3\kappa_4\kappa_2\left(\kappa_4+2\kappa_2^2\right)+3\kappa_2^2\left(\kappa_6+8\kappa_4\kappa_2+6\kappa_3^2\right)\right\}-\frac{15\kappa_4^2\kappa_2}{n}-\frac{81\kappa_2^5}{n}\\&\quad-\frac{72\kappa_4\kappa_2^3}{n}+O(n^{-2})\\&=\left(\kappa_4+3\kappa_2^2\right)^2\kappa_2+\frac{1}{n}\left\{\kappa_8\kappa_2+2\kappa_6\kappa_4+34\kappa_6\kappa_2^2+48\kappa_5\kappa_3\kappa_2+12\kappa_4\kappa_3^2\right.\\&\quad\left.+53\kappa_4^2\kappa_2+270\kappa_4\kappa_2^3+252\kappa_3^2\kappa_2^2+123\kappa_2^5\right\}+O(n^{-2}).\end{aligned}$$

The result in terms of central moments is obtain by substituting all population cumulants with their corresponding expressions given in (A.1). □

A.3 Moments of central sample moments

Lemma A.6. *Suppose that X has sufficiently many finite moments. Then*

(i) $\mathrm{E}\left(m_2-\mu_2\right)=-\dfrac{\mu_2}{n}$,

(ii) $\mathrm{E}\left\{\left(m_2-\mu_2\right)^2\right\}=\dfrac{\mu_4-\mu_2^2}{n}+O(n^{-2})$,

(iii) $\mathrm{E}\left\{\left(m_2-\mu_2\right)^3\right\}=O(n^{-2})$.

Proof.

(i) Follows directly from (27.4.1) of Cramér (1946).

(ii) Follows directly from (27.4.2) of Cramér (1946).

(iii) Note that

$$\begin{aligned}\mathrm{E}\left\{\left(m_2-\frac{n-1}{n}\mu_2\right)^3\right\}&=\mathrm{E}\left\{\left(m_2-\mu_2+\frac{\mu_2}{n}\right)^3\right\}\\&=\mathrm{E}\left\{\left(m_2-\mu_2\right)^3\right\}+\frac{3\mu_2}{n}\mathrm{E}\left\{\left(m_2-\mu_2\right)^2\right\}+\frac{3\mu_2^2}{n^2}\mathrm{E}\left(m_2-\mu_2\right)+\frac{\mu_2^3}{n^3}\\&=\mathrm{E}\left\{\left(m_2-\mu_2\right)^3\right\}+O(n^{-2}),\end{aligned}$$

where we made use of the fact that $\mathrm{E}\{(m_2-\mu_2)^2\}=O(n^{-1})$, as mentioned earlier in this lemma. From (27.4.3) of Cramér (1946) we know that the left-hand side of the above expression is $O(n^{-2})$. Hence we must have that $\mathrm{E}\{(m_2-\mu_2)^3\}=O(n^{-2})$. □

Lemma A.7. *Suppose that X has sufficiently many finite moments. Then*

(i) $\mathrm{E}\left(m_3-\mu_3\right)=-\dfrac{3\mu_3}{n}+O(n^{-2})$,

(ii) $\mathrm{E}\left\{\left(m_3-\mu_3\right)^2\right\}=\dfrac{\mu_6-6\mu_4\mu_2-\mu_3^2+9\mu_2^3}{n}+O(n^{-2})$,

(iii) $\mathrm{E}\left\{\left(m_3-\mu_3\right)^3\right\}=O(n^{-2})$.

Proof.

(i) Follows directly from (27.5.1) of Cramér (1946).

(ii) Making use of (27.5.3) and (27.5.4) of Cramér (1946), we have that

$$\begin{aligned}\mathrm{E}\left\{\left(m_3-\mu_3\right)^2\right\}&=\mathrm{E}\left\{\left(m_3-\mathrm{E}(m_3)+O(n^{-1})\right)^2\right\}\\&=\mathrm{E}\left\{(m_3-\mathrm{E}(m_3))^2\right\}+O(n^{-2})\\&=\mathrm{Var}(m_3)+O(n^{-2})\\&=\frac{\mu_6-6\mu_4\mu_2-\mu_3^2+9\mu_2^3}{n}+O(n^{-2}).\end{aligned}$$

(iii) From (27.5.1) of Cramér (1946) we have that

$$\mathrm{E}(m_3) = \frac{(n-1)(n-2)}{n^2}\mu_3 = \mu_3 - \frac{3\mu_3}{n} + O(n^{-2}).$$

This, together with Lemma A.2, implies that

$$\begin{aligned}
\mathrm{E}\left\{(m_3-\mu_3)^3\right\} &= \mathrm{E}\left(m_3^3\right) - 3\mu_3\mathrm{E}\left(m_3^2\right) + 3\mu_3^2\mathrm{E}(m_3) - \mu_3^3 \\
&= \mu_3^3 + \frac{3\mu_3\left(\mu_6 - 6\mu_4\mu_2 - 4\mu_3^2 + 9\mu_2^3\right)}{n} - 3\mu_3^3 - \frac{3\mu_3\left(\mu_6 - 6\mu_4\mu_2 - 7\mu_3^2 + 9\mu_2^3\right)}{n} \\
&\quad + 3\mu_3^3 - \frac{9\mu_3^3}{n} - \mu_3^3 + O(n^{-2}) \\
&= O(n^{-2}).
\end{aligned}$$

□

Lemma A.8. *Suppose that X has sufficiently many finite moments. Then*

(i) $\mathrm{E}(m_4 - \mu_4) = -\dfrac{4\mu_4 - 6\mu_2^2}{n} + O(n^{-2})$,

(ii) $\mathrm{E}\left\{(m_4-\mu_4)^2\right\} = \dfrac{\mu_8 - 8\mu_5\mu_3 - \mu_4^2 + 16\mu_3^2\mu_2}{n} + O(n^{-2})$,

(iii) $\mathrm{E}\left\{(m_4-\mu_4)^3\right\} = O(n^{-2})$.

Proof.

(i) Follows directly from (27.5.1) of Cramér (1946).

(ii) Lemma A.3 implies that

$$\begin{aligned}
&\mathrm{E}\left\{(m_4-\mu_4)^2\right\} \\
&= \mathrm{E}\left(m_4^2\right) - 2\mu_4\mathrm{E}(m_4) + \mu_4^2 \\
&= \mu_4^2 + \frac{\mu_8 - 8\mu_5\mu_3 + 12\mu_4\mu_2^2 - 9\mu_4^2 + 16\mu_3^2\mu_2}{n} - 2\mu_4^2 + \frac{8\mu_4^2 - 12\mu_4\mu_2^2}{n} + \mu_4^2 + O(n^{-2}) \\
&= \frac{\mu_8 - 8\mu_5\mu_3 - \mu_4^2 + 16\mu_3^2\mu_2}{n} + O(n^{-2}).
\end{aligned}$$

(iii) From Lemma A.3 we have that

$$\begin{aligned}
\mathrm{E}\left\{(m_4-\mu_4)^3\right\} &= \mathrm{E}\left(m_4^3\right) - 3\mu_4\mathrm{E}\left(m_4^2\right) + 3\mu_4^2\mathrm{E}(m_4) - \mu_4^3 \\
&= \mu_4^3 + \frac{3\mu_8\mu_4 - 24\mu_5\mu_4\mu_3 - 15\mu_4^3 + 18\mu_4^2\mu_2^2 + 48\mu_4\mu_3^2\mu_2}{n} - 3\mu_4^3 \\
&\quad - \frac{3\mu_4\left(\mu_8 - 8\mu_5\mu_3 + 12\mu_4\mu_2^2 - 9\mu_4^2 + 16\mu_3^2\mu_2\right)}{n} + 3\mu_4^3 \\
&\quad - \frac{3\mu_4^2\left(4\mu_4 - 6\mu_2^2\right)}{n} - \mu_4^3 + O(n^{-2}) \\
&= O(n^{-2}).
\end{aligned}$$

□

Lemma A.9. *Suppose that X has sufficiently many finite moments. Then*

(i) $\mathrm{E}\{(m_3-\mu_3)(m_2-\mu_2)\} = \dfrac{\mu_5-4\mu_3\mu_2}{n}+O(n^{-2}),$

(ii) $\mathrm{E}\left\{(m_3-\mu_3)(m_2-\mu_2)^2\right\} = O(n^{-2}),$

(iii) $\mathrm{E}\left\{(m_3-\mu_3)^2(m_2-\mu_2)\right\} = O(n^{-2}).$

Proof.

(i) Follows directly from (27.5.6) of Cramér (1946).

(ii) Multiplying out the term inside the expected value, it follows from Lemma A.4 and equations (27.4.1), (27.4.2) and (27.5.1) of Cramér (1946) that

$$\begin{aligned}
&\mathrm{E}\left\{(m_3-\mu_3)(m_2-\mu_2)^2\right\}\\
&=\mathrm{E}\left(m_3m_2^2\right)-2\mu_2\,\mathrm{E}(m_3m_2)+\mu_2^2\,\mathrm{E}(m_3)-\mu_3\,\mathrm{E}\left(m_2^2\right)+2\mu_3\mu_2\,\mathrm{E}(m_2)-\mu_3\mu_2^2\\
&=\mu_3\mu_2^2+\frac{2\mu_5\mu_2+\mu_4\mu_3-14\mu_3\mu_2^2}{n}-2\mu_2\left(\mu_3\mu_2+\frac{\mu_5-8\mu_3\mu_2}{n}\right)+\mu_2^2\left(\mu_3-\frac{3\mu_3}{n}\right)\\
&\quad-\mu_3\left(\mu_2^2+\frac{\mu_4-3\mu_2^2}{n}\right)+2\mu_3\mu_2\left(\mu_2-\frac{\mu_2}{n}\right)-\mu_3\mu_2^2+O(n^{-2})\\
&=O(n^{-2}).
\end{aligned}$$

(iii) Multiplying out the product inside the expected value, it follows from Lemma A.2, Lemma A.4 and equations (27.4.1) and (27.5.1) of Cramér (1946) that

$$\begin{aligned}
&\mathrm{E}\left\{(m_3-\mu_3)^2(m_2-\mu_2)\right\}\\
&=\mathrm{E}\left(m_3^2m_2\right)-\mu_2\,\mathrm{E}\left(m_3^2\right)-2\mu_3\,\mathrm{E}(m_3m_2)+2\mu_3\mu_2\,\mathrm{E}(m_3)+\mu_3^2\,\mathrm{E}(m_2)-\mu_3^2\mu_2\\
&=\mu_3^2+\frac{\mu_6\mu_2+2\mu_5\mu_3-6\mu_4\mu_2^2-16\mu_3^2\mu_2+9\mu_2^4}{n}-\mu_2\left(\mu_3^2+\frac{\mu_6-6\mu_4\mu_2-7\mu_3^2+9\mu_2^3}{n}\right)\\
&\quad-2\mu_3\left(\mu_3\mu_2+\frac{\mu_5-8\mu_3\mu_2}{n}\right)+2\mu_3\mu_2\left(\mu_3-\frac{3\mu_3}{n}\right)+\mu_3^2\left(\mu_2-\frac{\mu_2}{n}\right)-\mu_3^2\mu_2+O(n^{-2})\\
&=O(n^{-2}).
\end{aligned}$$

□

Lemma A.10. *Suppose that X has sufficiently many finite moments. Then*

(i) $\mathrm{E}\{(m_4-\mu_4)(m_2-\mu_2)\} = \dfrac{\mu_6-\mu_4\mu_2-4\mu_3^2}{n}+O(n^{-2}),$

(ii) $\mathrm{E}\left\{(m_4-\mu_4)(m_2-\mu_2)^2\right\} = O(n^{-2}),$

(iii) $\mathrm{E}\left\{(m_4-\mu_4)^2(m_2-\mu_2)\right\} = O(n^{-2}).$

Proof.

(i) Follows directly from (27.5.6) of Cramér (1946).

(ii) Using (27.4.1) and (27.4.2) of Cramér (1946) along with the results of Lemmas A.3 and A.5 we may write

$$
\begin{aligned}
&\mathrm{E}\left\{(m_4-\mu_4)(m_2-\mu_2)^2\right\} \\
&= \mathrm{E}\left(m_4 m_2^2\right) - \mu_4\,\mathrm{E}\left(m_2^2\right) - 2\mu_2\,\mathrm{E}(m_4 m_2) + 2\mu_4\mu_2\,\mathrm{E}(m_2) + \mu_2^2\,\mathrm{E}(m_4) - \mu_4\mu_2^2 \\
&= \mu_4\mu_2^2 + \frac{1}{n}\{2\mu_6\mu_2 - 3\mu_4\mu_2^2 + \mu_4^2 + 8\mu_3^2\mu_2 + 6\mu_2^4\} - \mu_4\mu_2^2 - \frac{1}{n}\{\mu_4^2 - 3\mu_4\mu_2^2\} \\
&\quad - 2\mu_4\mu_2^2 - \frac{2\mu_2}{n}\{\mu_6 - 6\mu_4\mu_2 - 4\mu_3^2 + 6\mu_2^3\} + 2\mu_4\mu_2^2 - \frac{1}{n}\{2\mu_4\mu_2^2\} + \mu_4\mu_2^2 \\
&\quad - \frac{\mu_2^2}{n}\{4\mu_4 - 6\mu_2^2\} - \mu_4\mu_2^2 + O(n^{-2}) \\
&= O(n^{-2}).
\end{aligned}
$$

(iii) Note that

$$
\begin{aligned}
&\mathrm{E}\left\{(m_4-\mu_4)^2(m_2-\mu_2)\right\} \\
&= \mathrm{E}\left(m_4^2 m_2\right) - \mu_2\,\mathrm{E}\left(m_4^2\right) - 2\mu_4\,\mathrm{E}(m_4 m_2) + 2\mu_4\mu_2\,\mathrm{E}(m_4) + \mu_4^2\,\mathrm{E}(m_2) - \mu_4^2\mu_2.
\end{aligned}
$$

Hence, by (27.4.1) of Cramér (1946) and the results of Lemmas A.3 and A.5 above, we have

$$
\begin{aligned}
&\mathrm{E}\left\{(m_4-\mu_4)^2(m_2-\mu_2)\right\} \\
&= \mu_4^2\mu_2 + \frac{1}{n}\{\mu_8\mu_2 + 2\mu_6\mu_4 - 8\mu_5\mu_3\mu_2 - 12\mu_4^2\mu_2 - 8\mu_4\mu_3^2 + 12\mu_4\mu_2^3 + 16\mu_3^2\mu_2^2\} \\
&\quad - \mu_4^2\mu_2 - \frac{\mu_2}{n}\{\mu_8 - 8\mu_5\mu_3 + 12\mu_4\mu_2^2 - 9\mu_4^2 + 16\mu_3^2\mu_2\} - 2\mu_4^2\mu_2 \\
&\quad - \frac{2\mu_4}{n}\{\mu_6 - 6\mu_4\mu_2 - 4\mu_3^2 + 6\mu_2^3\} + 2\mu_4^2\mu_2 - \frac{2\mu_4}{n}\{4\mu_4 - 6\mu_2^2\} + \mu_4^2\mu_2 - \frac{\mu_4^2\mu_2}{n} \\
&\quad - \mu_4^2\mu_2 + O(n^{-2}) \\
&= O(n^{-2}).
\end{aligned}
$$

□

Lemma A.11. *Suppose that* $\mathrm{E}(|X|^k) < \infty$ *for some sufficiently large* k*. Then, as* $n \to \infty$*,*

$$\mathrm{E}(\widehat{\sigma}_n - \sigma) = \mathrm{E}\left(\sqrt{m_2} - \sqrt{\mu_2}\right) = -\frac{\mu_4 + 3\sigma^4}{8\sigma^3 n} + O(n^{-2}), \tag{A.7}$$

$$\mathrm{E}(\widehat{\sigma}_n - \sigma)^2 = \mathrm{E}\left(\sqrt{m_2} - \sqrt{\mu_2}\right)^2 = \frac{\mu_4 - \sigma^4}{4\sigma^2 n} + O(n^{-2}), \tag{A.8}$$

and

$$\mathrm{E}(\widehat{\sigma}_n - \sigma)^4 = \mathrm{E}\left(\sqrt{m_2} - \sqrt{\mu_2}\right)^4 = O(n^{-2}). \tag{A.9}$$

Proof. From (27.7.1) and (27.7.2) of Cramér (1946) we have that

$$\mathrm{E}(\widehat{\sigma}_n) = \sigma + O(n^{-1})$$

and

$$\operatorname{Var}(\widehat{\sigma}_n) = \frac{\mu_4 - \sigma^4}{4\sigma^2 n} + O(n^{-2}).$$

Hence,

$$\begin{aligned}
\mathrm{E}\left\{(\widehat{\sigma}_n - \sigma)^2\right\} &= \mathrm{E}\left\{\left(\widehat{\sigma}_n - \mathrm{E}(\widehat{\sigma}_n) + O(n^{-1})\right)^2\right\} \\
&= \mathrm{E}\left\{(\widehat{\sigma}_n - \mathrm{E}(\widehat{\sigma}_n))^2\right\} + 2O(n^{-1})\mathrm{E}(\widehat{\sigma}_n - \mathrm{E}(\widehat{\sigma}_n)) + O(n^{-2}) \\
&= \operatorname{Var}(\widehat{\sigma}_n) + O(n^{-2}) \\
&= \frac{\mu_4 - \sigma^4}{4\sigma^2 n} + O(n^{-2}), \qquad \text{(A.10)}
\end{aligned}$$

which proves (A.8).

To obtain (A.7), first note that the difference $\widehat{\sigma}_n - \sigma$ may be written as

$$\widehat{\sigma}_n - \sigma = \frac{\widehat{\sigma}_n^2 - \sigma^2}{2\sigma} - \frac{(\widehat{\sigma}_n - \sigma)^2}{2\sigma},$$

so that by (A.10) above and (27.4.1) of Cramér (1946) we have

$$\begin{aligned}
\mathrm{E}(\widehat{\sigma}_n - \sigma) &= \frac{1}{2\sigma}\mathrm{E}\left(\widehat{\sigma}_n^2 - \sigma^2\right) - \frac{1}{2\sigma}\mathrm{E}\left\{(\widehat{\sigma}_n - \sigma)^2\right\} \\
&= \frac{-\sigma^2}{2\sigma n} - \frac{\mu_4 - \sigma^4}{8\sigma^3 n} + O(n^{-2}) \\
&= -\frac{\mu_4 + 3\sigma^4}{8\sigma^3 n} + O(n^{-2}).
\end{aligned}$$

Lastly, observing that

$$(\widehat{\sigma}_n - \sigma)^4 \leq \frac{1}{\sigma^4}\left(\widehat{\sigma}_n^2 - \sigma^2\right)^4,$$

we have by (27.5.5) of Cramér (1946) that $\mathrm{E}\left\{(\widehat{\sigma}_n - \sigma)^4\right\} = O(n^{-2})$, which proves (A.9). □

Appendix B

Some moments of weighted sums of sample observations

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be independent and identically distributed random q-vectors, $q \geq 1$, such that $\mathrm{E}(\mathbf{X}_1) = \mathbf{0}$. For a triangular array of numbers in $\mathbb{R}^q$, say $\{\mathbf{w}_i(n), i = 1, \ldots, n\}$, $n \in \mathbb{N}$, our goal is to obtain closed-form asymptotic expressions for moments of the form

$$E_{r_1,\ldots,r_q} := \mathrm{E}\left\{\prod_{k=1}^{q}\left(\frac{1}{n}\sum_{i=1}^{n} w_{ik}X_{ik}\right)^{r_k}\right\}$$

for given integers $r_1, \ldots, r_q \geq 0$, where X_{ik} and $w_{ik} = w_{ik}(n)$ denote the kth elements of $\mathbf{X}_i$ and $\mathbf{w}_i(n)$, respectively. Our results will be expressed in terms of the repeated sums of weights

$$W\left(1^{s_1^{(1)}} \cdots q^{s_q^{(1)}}, \ldots, 1^{s_1^{(\ell)}} \cdots q^{s_q^{(\ell)}}\right) := \frac{1}{{}^nP_\ell}\sum_{I_{\ell,n}}\left\{\prod_{j=1}^{\ell} w_{i_j,1}^{s_1^{(j)}} \cdots w_{i_j,q}^{s_q^{(j)}}\right\}$$

and joint moments

$$\mu\left(1^{s_1} \cdots q^{s_q}\right) = \mathrm{E}\left(X_{11}^{s_1} \cdots X_{1q}^{s_q}\right),$$

where the s_k and $s_k^{(j)}$, $k = 1, \ldots, q$, $j = 1, \ldots, \ell$, are non-negative integers and $I_{\ell,n} = \{(i_1, \ldots, i_\ell) \in \{1, \ldots, n\}^\ell : i_p \neq i_q \forall p \neq q\}$. As usual, ${}^nP_\ell = n!/(n-\ell)!$. For notational conciseness, also define

$$\Sigma\left(1^{s_1^{(1)}} \cdots q^{s_q^{(1)}}, \ldots, 1^{s_1^{(\ell)}} \cdots q^{s_q^{(\ell)}}\right) := W\left(1^{s_1^{(1)}} \cdots q^{s_q^{(1)}}, \ldots, 1^{s_1^{(\ell)}} \cdots q^{s_q^{(\ell)}}\right)\prod_{j=1}^{\ell}\mu\left(1^{s_1^{(j)}} \cdots q^{s_q^{(j)}}\right).$$

So, for example,

$$\Sigma(1^2 2) = W(1^2 2)\mu(1^2 2) = \left(\frac{1}{n}\sum_{i=1}^{n} w_{i,1}^2 w_{i,2}\right)\mathrm{E}\left(X_{11}^2 X_{12}\right)$$

and

$$\Sigma(1^3, 12^2) = W(1^3, 12^2)\mu(1^3)\mu(12^2) = \left(\frac{1}{n(n-1)}\sum_{(i,j)} w_{i,1}^3 w_{j,1} w_{j,2}^2\right)\mathrm{E}\left(X_{11}^3\right)\mathrm{E}\left(X_{11}X_{12}^2\right).$$

Notice that if $w_{ik} = 1$ for all indices i and k, then $W(\cdot) = 1$ (regardless of the argument) and the above two expressions reduce to

$$\Sigma\left(1^2 2\right) = \mathrm{E}\left(X_{11}^2 X_{12}\right) \quad \text{and} \quad \Sigma\left(1^3, 12^2\right) = \mathrm{E}\left(X_{11}^3\right)\mathrm{E}\left(X_{11}X_{12}^2\right).$$

Throughout we assume that $\mathbf{X}_1$ has sufficiently many finite moments.

B.1 Some general results

Lemma B.1. *Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables with $\mathrm{E}(X_1) = 0$ and let $\{w_i(n), i = 1, \ldots, n\}_{n=1}^{\infty}$ be a triangular array of real numbers such that $\max_{1 \le i \le n} |w_i(n)| = O(1)$ as $n \to \infty$. For any integer $r \ge 0$, if $\mathrm{E}(|X_1|^r) < \infty$, we have that*

$$E_r = \mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^n w_i X_i\right)^r\right\} = O(n^{-\lfloor (r+1)/2 \rfloor}),$$

where $w_i = w_i(n)$ and $\lfloor x \rfloor$ denotes the integer part of x.

Proof. Writing $w_i = w_i(n)$, note that

$$\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^n w_i X_i\right)^r\right\} = \frac{1}{n^r}\mathrm{E}\left\{\left(\sum_{i=1}^n w_i X_i\right)^r\right\} = \frac{1}{n^r}\sum_{i_1=1}^n \cdots \sum_{i_r=1}^n w_{i_1}\cdots w_{i_r}\,\mathrm{E}\left(X_{i_1}\cdots X_{i_r}\right). \tag{B.1}$$

First consider the case when r is even. Since $\mathrm{E}(X_1) = 0$, the order (in terms of n) of the sum in (B.1) is determined by terms with $\frac{r}{2}$ repeated sums, which each has $n(n-1)\cdots(n-\frac{r}{2}+1)$ terms depending on the $w_i(n)$. As $\max_{1 \le i \le n} |w_i(n)| = O(1)$, such sums are of order $O(n^{r/2})$. Therefore,

$$\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^n w_i X_i\right)^r\right\} = \frac{1}{n^r}O(n^{r/2}) = O(n^{-r/2}) = O(n^{-\lfloor (r+1)/2 \rfloor}).$$

Likewise, if r is odd, the largest contribution to the sum in (B.1) is given by terms with $\frac{r-1}{2}$ repeated sums, which are of order $O(n^{(r-1)/2})$. Therefore,

$$\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^n w_i X_i\right)^r\right\} = \frac{1}{n^r}O(n^{(r-1)/2}) = O(n^{-(r+1)/2}) = O(n^{-\lfloor (r+1)/2 \rfloor}). \qquad \square$$

The result may be generalised to an expectation of a product of sums such as $E_{r_1,\ldots,r_q}$ defined above. We provide such a generalisation below, of which the result is slightly crude and therefore not necessarily optimal.

Lemma B.2. *Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be independent and identically distributed random q-vectors, $q \ge 1$, with $\mathrm{E}(\mathbf{X}_1) = \mathbf{0}$ and let $\{\mathbf{w}_i(n), i = 1, \ldots, n\}_{n=1}^{\infty}$ be a triangular array of numbers in $\mathbb{R}^q$. Suppose that for $k = 1, \ldots, q$ it holds that $\max_{1 \le i \le n} |w_{ik}(n)| = O(1)$ as $n \to \infty$, where $w_{ik}(n)$ denotes the kth element of $\mathbf{w}_i(n)$. For any set of nonnegative integers $r_1, \ldots, r_q$, if $\mathbf{X}_1$ has sufficiently many finite moments, it follows that*

$$E_{r_1,\ldots,r_q} = \mathrm{E}\left\{\prod_{k=1}^q \left(\frac{1}{n}\sum_{i=1}^n w_{ik} X_{ik}\right)^{r_k}\right\} = O(n^{-(\sum_k r_k)/2}),$$

for given integers $r_1, \ldots, r_q \ge 0$, where $w_{ik} = w_{ik}(n)$.

Proof. Let $R = \sum_{k=1}^{q} r_k$. By a generalised form of Hölder's inequality (see, e.g., Finner, 1992) it follows that

$$\begin{aligned}
\left|E_{r_1,\ldots,r_q}\right| &\le \mathrm{E}\left\{ \prod_{k=1}^{q} \left| \frac{1}{n} \sum_{i=1}^{n} w_{ik} X_{ik} \right|^{r_k} \right\} \\
&\le \prod_{k=1}^{q} \mathrm{E}\left\{ \left| \frac{1}{n} \sum_{i=1}^{n} w_{ik} X_{ik} \right|^{R} \right\}^{r_k/R} \\
&= \prod_{k=1}^{q} \left\{ O(n^{-R/2}) \right\}^{r_k/R} \\
&= O(n^{-R/2}).
\end{aligned}$$

□

For easy reference we state below the special case where $q = 2$.

Corollary B.1. *Let $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ be iid bivariate random variables with $\mathrm{E}(X_1) = \mathrm{E}(Y_1) = 0$. Let $\{v_i(n), i = 1, \ldots, n\}_{n=1}^{\infty}$ and $\{w_i(n), i = 1, \ldots, n\}_{n=1}^{\infty}$ be triangular arrays of real numbers such that $\max_{1\le i\le n} |v_i(n)| = O(1)$ and $\max_{1\le i\le n} |w_i(n)| = O(1)$ as $n \to \infty$. For given integers $r_1, r_2 \ge 0$, if (X_1, Y_1) has sufficiently many finite moments, it follows that*

$$E_{r_1,r_2} = \mathrm{E}\left\{ \left(\frac{1}{n} \sum_i v_i X_i \right)^{r_1} \left(\frac{1}{n} \sum_j w_j Y_j \right)^{r_2} \right\} = O(n^{-(r_1+r_2)/2}),$$

where $v_i = v_i(n)$ and $w_i = w_i(n)$.

In the following sections we will derive more refined results which cannot be obtained from the above lemmas. These results are required in the main text and in Appendix C. For ease of notation we will use the symbols:

$$\sum_{(i,j)} = \sum_{i=1}^{n} \sum_{j\ne i}^{n}, \qquad \sum_{(i,j,k)} = \sum_{i=1}^{n} \sum_{j\ne i}^{n} \sum_{\substack{k\ne i\\ k\ne j}}^{n}, \qquad \sum_{(i,j,k,\ell)} = \sum_{i=1}^{n} \sum_{j\ne i}^{n} \sum_{\substack{k\ne i\\ k\ne j}}^{n} \sum_{\substack{\ell\ne i\\ \ell\ne j\\ \ell\ne k}}^{n}.$$

B.2 Univariate case ($q = 1$)

Let $X_1, X_2, \ldots, X_n$ be iid univariate random variables such that $\mathrm{E}(X_1) = 0$ and let $\{w_i(n), i = 1, \ldots, n\}_{n=1}^{\infty}$ be a triangular array of real numbers such that $\max_{1\le i\le n} |w_i(n)| = O(1)$ as $n \to \infty$. For a few choices of an integer $r \ge 1$, we state and derive asymptotic expressions for

$$E_r = \mathrm{E}\left\{ \left(\frac{1}{n} \sum_{i=1}^{n} w_i X_i \right)^{r} \right\},$$

where $w_i = w_i(n)$.

$$E_2 = \frac{1}{n^2} \mathrm{E}\left\{ \sum_i w_i^2 X_i^2 \right\} = \frac{\Sigma(1^2)}{n}. \tag{B.2}$$

$$E_3 = \frac{1}{n^3}\,\mathrm{E}\left\{\sum_i w_i^3 X_i^3\right\} = \frac{\Sigma(1^3)}{n^2}. \tag{B.3}$$

$$\begin{aligned} E_4 &= \frac{1}{n^4}\,\mathrm{E}\left\{\sum_i w_i^4 X_i^4 + 3\sum_{(i,j)} w_i^2 w_j^2 X_i^2 X_j^2\right\} \\ &= \frac{1}{n^4}\,\mathrm{E}\left\{n\Sigma(1^4) + 3n(n-1)\Sigma(1^2,1^2)\right\} \\ &= \frac{3\Sigma(1^2,1^2)}{n^2} + \frac{\Sigma(1^4) - 3\Sigma(1^2,1^2)}{n^3}. \end{aligned} \tag{B.4}$$

$$\begin{aligned} E_5 &= \frac{1}{n^5}\,\mathrm{E}\left\{\sum_i w_i^5 X_i^5 + 10\sum_{(i,j)} w_i^3 w_j^2 X_i^3 X_j^2\right\} \\ &= \frac{1}{n^5}\left\{n\Sigma(1^5) + 10n(n-1)\Sigma(1^3,1^2)\right\} \\ &= \frac{10\Sigma(1^3,1^2)}{n^3} + \frac{\Sigma(1^5) - 10\Sigma(1^3,1^2)}{n^4}. \end{aligned} \tag{B.5}$$

$$\begin{aligned} E_7 &= \frac{1}{n^7}\,\mathrm{E}\left\{105\sum_{(i,j,k)} w_i^3 w_j^2 w_k^2 X_i^3 X_j^2 X_k^2\right\} + O(n^{-5}) \\ &= \frac{105\Sigma(1^3,1^2,1^2)}{n^4} + O(n^{-5}). \end{aligned} \tag{B.6}$$

B.3 Bivariate case ($q = 2$)

Let $(X_1,Y_1),\dots,(X_n,Y_n)$ be iid bivariate random vectors such that $\mathrm{E}(X_1) = \mathrm{E}(Y_1) = 0$ and let $\{v_i(n), i = 1,\dots,n\}_{n=1}^{\infty}$ and $\{w_i(n), i = 1,\dots,n\}_{n=1}^{\infty}$ be triangular arrays of real numbers such that $\max_{1\le i\le n}|v_i(n)| = O(1)$ and $\max_{1\le i\le n}|w_i(n)| = O(1)$ as $n \to \infty$. For a few choices of integers $r_1, r_2 \ge 0$, we state and derive asymptotic expressions for

$$E_{r_1 r_2} = \mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n} v_i X_i\right)^{r_1}\left(\frac{1}{n}\sum_{j=1}^{n} w_j Y_j\right)^{r_2}\right\},$$

where $v_i = v_i(n)$ and $w_i = w_i(n)$.

$$E_{1,1} = \frac{1}{n^2}\,\mathrm{E}\left\{\sum_i v_i w_i X_i Y_i\right\} = \frac{\Sigma(12)}{n}. \tag{B.7}$$

$$E_{2,1} = \frac{1}{n^3}\,\mathrm{E}\left\{\sum_i v_i^2 w_i X_i^2 Y_i\right\} = \frac{\Sigma(1^2 2)}{n^2}. \tag{B.8}$$

$$\begin{aligned} E_{2,2} &= \frac{1}{n^4}\,\mathrm{E}\left\{\sum_i v_i^2 w_i^2 X_i^2 Y_i^2 + \sum_{(i,j)} v_i^2 w_j^2 X_i^2 Y_j^2 + 2\sum_{(i,j)} v_i w_i v_j w_j X_i Y_i X_j Y_j\right\} \\ &= \frac{1}{n^4}\left\{n\Sigma(1^2 2^2) + n(n-1)\Sigma(1^2,2^2) + 2n(n-1)\Sigma(12,12)\right\} \\ &= \frac{\Sigma(1^2,2^2) + 2\Sigma(12,12)}{n^2} + \frac{\Sigma(1^2 2^2) - \Sigma(1^2,2^2) - 2\Sigma(12,12)}{n^3}. \end{aligned} \tag{B.9}$$

$$
\begin{aligned}
E_{3,1} &= \frac{1}{n^4}\,\mathrm{E}\left\{\sum_i v_i^3 \Sigma_i X_i^3 Y_i + 3\sum_{(i,j)} v_i^2 v_j \Sigma_j X_i^2 X_j Y_j\right\} \\
&= \frac{1}{n^4}\left\{n\Sigma(1^3 2) + 3n(n-1)\Sigma(1^2, 12)\right\} \\
&= \frac{3\Sigma(1^2,12)}{n^2} + \frac{\Sigma(1^3 2) - 3\Sigma(1^2,12)}{n^3}.
\end{aligned} \tag{B.10}
$$

$$
\begin{aligned}
E_{3,2} = \frac{1}{n^5}\,\mathrm{E}\Bigg\{ & \sum_i v_i^3 w_i^2 X_i^3 Y_i^2 + \sum_{(i,j)} v_i^3 w_j^2 X_i^3 Y_j^2 + 3\sum_{(i,j)} v_i^2 v_j w_j^2 X_i^2 X_j Y_j^2 \\
& + 6\sum_{(i,j)} v_i^2 w_i v_j w_j X_i^2 Y_i X_j Y_j \Bigg\} \\
= \frac{1}{n^5} & \left\{n\Sigma(1^3 2^2) + n(n-1)\Sigma(1^3, 2^2) + 3n(n-1)\Sigma(1^2, 12^2) + 6n(n-1)\Sigma(1^2 2, 12)\right\} \\
= \frac{\Sigma(1^3,2^2) + 3\Sigma(1^2,12^2) + 6\Sigma(1^2 2,12)}{n^3} & + \frac{\Sigma(1^3 2^2) - \Sigma(1^3,2^2) - 3\Sigma(1^2,12^2) - 6\Sigma(1^2 2,12)}{n^4}.
\end{aligned} \tag{B.11}
$$

$$
\begin{aligned}
E_{3,3} = \frac{1}{n^6}\,\mathrm{E}\Bigg\{ & \sum_i v_i^3 w_i^3 X_i^3 Y_i^3 + \sum_{(i,j)} v_i^3 w_j^3 X_i^3 Y_j^3 + 3\sum_{(i,j)} v_i^3 w_i w_j^2 X_i^3 Y_i Y_j^2 \\
& + 9\sum_{(i,j)} v_i^2 w_i v_j w_j^2 X_i^2 Y_i X_j Y_j^2 + 9\sum_{(i,j)} v_i^2 w_i^2 v_j w_j X_i^2 Y_i^2 X_j Y_j \\
& + 3\sum_{(i,j)} v_i^2 v_j w_j^3 X_i^2 X_j Y_j^3 + 9\sum_{(i,j,k)} v_i^2 v_j w_j w_k^2 X_i^2 X_j Y_j Y_k^2 \\
& + 6\sum_{(i,j,k)} v_i w_i v_j w_j v_k w_k X_i Y_i X_j Y_j X_k Y_k \Bigg\} \\
= \frac{9\Sigma(1^2,2^2,12) + 6\Sigma(12,12,12)}{n^3} & + O(n^{-4}).
\end{aligned} \tag{B.12}
$$

$$
\begin{aligned}
E_{5,1} = \frac{1}{n^6}\,\mathrm{E}\Bigg\{ & \sum_i v_i^5 w_i X_i^5 Y_i + 5\sum_{(i,j)} v_i^4 v_j w_j X_i^4 X_j Y_j + 10\sum_{(i,j)} v_i^3 v_j^2 w_j X_i^3 X_j^2 Y_j \\
& + 10\sum_{(i,j)} v_i^2 v_j^3 w_j X_i^2 X_j^3 Y_j + 15\sum_{(i,j,k)} v_i^2 v_j^2 v_k w_j X_i^2 X_j^2 X_k Y_j \Bigg\} \\
= \frac{1}{n^6}\Big\{ & n\Sigma(1^5 2) + 5n(n-1)\Sigma(1^4,12) + 10n(n-1)\Sigma(1^3,1^2 2) + 10n(n-1)\Sigma(1^2,1^3 2) \\
& + 15n(n-1)(n-2)\Sigma(1^2,1^2,12)\Big\} \\
= \frac{15\Sigma(1^2,1^2,12)}{n^3} & + \frac{5\Sigma(1^4,12) + 10\Sigma(1^3,1^2 2) + 10\Sigma(1^2,1^3 2) - 45\Sigma(1^2,1^2,12)}{n^4} + O(n^{-5}).
\end{aligned} \tag{B.13}
$$

$$
\begin{aligned}
E_{5,2} &= \frac{1}{n^7}\,\mathrm{E}\Big\{\sum_i v_i^5 w_i^2 X_i^5 Y_i^2 + \sum_{(i,j)} v_i^5 w_j^2 X_i^5 Y_j^2 + 5\sum_{(i,j)} v_i^4 v_j w_j^2 X_i^4 X_j Y_j^2 \\
&\quad + 10\sum_{(i,j)} v_i^3 v_j^2 w_j^2 X_i^3 X_j^2 Y_j^2 + 10\sum_{(i,j)} v_i^2 v_j^3 w_j^2 X_i^2 X_j^3 Y_j^2 \\
&\quad + 10\sum_{(i,j)} v_i^4 w_i v_j w_j X_i^4 Y_i X_j Y_j + 20\sum_{(i,j)} v_i^3 w_i v_j^2 w_j X_i^3 Y_i X_j^2 Y_j \\
&\quad + 10\sum_{(i,j,k)} v_i^3 v_j^2 w_k^2 X_i^3 X_j^2 Y_k^2 + 20\sum_{(i,j,k)} v_i^3 v_j w_j v_k w_k X_i^3 X_j Y_j X_k Y_k \\
&\quad + 15\sum_{(i,j,k)} v_i w_i^2 v_j^2 v_k^2 X_i Y_i^2 X_j^2 X_k^2 + 60\sum_{(i,j,k)} v_i^2 w_i v_j w_j v_k^2 X_i^2 Y_i X_j Y_j X_k^2\Big\} \\
&= \frac{1}{n^7}\Big\{10n(n-1)(n-2)\Sigma(1^3,1^2,2^2) + 20n(n-1)(n-2)\Sigma(1^3,12,12) \\
&\quad + 15n(n-1)(n-2)\Sigma(12^2,1^2,1^2) + 60n(n-1)(n-2)\Sigma(1^2 2,12,1^2)\Big\} + O(n^{-5}) \\
&= \frac{10\Sigma(1^3,1^2,2^2) + 20\Sigma(1^3,12,12) + 15\Sigma(12^2,1^2,1^2) + 60\Sigma(1^2 2,12,1^2)}{n^4} + O(n^{-5}). \qquad \text{(B.14)}
\end{aligned}
$$

$$
\begin{aligned}
E_{5,3} &= \frac{1}{n^8}\,\mathrm{E}\Big\{45\sum_{(i,j,k,\ell)} v_i^2 v_j^2 v_k w_k w_\ell^2 X_i^2 X_j^2 X_k Y_k Y_\ell^2 \\
&\qquad + 60\sum_{(i,j,k,\ell)} v_i^2 v_j w_j v_k w_k v_\ell w_\ell X_i^2 X_j Y_j X_k Y_k X_\ell Y_\ell\Big\} + O(n^{-5}) \\
&= \frac{45\Sigma(1^2,1^2,12,2^2) + 60\Sigma(1^2,12,12,12)}{n^4} + O(n^{-5}). \qquad \text{(B.15)}
\end{aligned}
$$

$$
\begin{aligned}
E_{7,1} &= \frac{1}{n^8}\,\mathrm{E}\Big\{105\sum_{(i,j,k,\ell)} X_i^2 X_j^2 X_k^2 X_\ell Y_\ell\Big\} + O(n^{-5}) \\
&= \frac{1}{n^8}\left\{105n(n-1)(n-2)(n-3)\Sigma(1^2,1^2,1^2,12)\right\} + O(n^{-5}) \\
&= \frac{105\Sigma(1^2,1^2,1^2,12)}{n^4} + O(n^{-5}). \qquad \text{(B.16)}
\end{aligned}
$$

B.3.1 Special case

For the bivariate case above ($q = 2$), consider the special case where $Y_i = X_i^2 - 1$ and $v_i = w_i = 1$, $i = 1,\ldots,n$. Setting $\mu_k = \mathrm{E}(X_1^k)$, we have the following results.

$$E_{5,0} = \frac{10\mu_3}{n^3} + \frac{\mu_5 - 10\mu_3}{n^4}. \qquad \text{(B.17)}$$

$$E_{7,0} = \frac{105\mu_3}{n^4} + O(n^{-5}). \qquad \text{(B.18)}$$

$$E_{1,1} = \frac{\mu_3}{n}. \qquad \text{(B.19)}$$

$$E_{1,2} = \frac{\mu_5 - 2\mu_3}{n^2}. \qquad \text{(B.20)}$$

$$E_{1,3} = \frac{3\mu_3(\mu_4 - 1)}{n^2} + \frac{\mu_7 - 3\mu_5 - 3\mu_4\mu_3 + 6\mu_3}{n^3}. \qquad \text{(B.21)}$$

$$E_{2,1} = \frac{\mu_4 - 1}{n^2}. \tag{B.22}$$

$$E_{2,2} = \frac{\mu_4 + 2\mu_3^2 - 1}{n^2} + \frac{\mu_6 - 3\mu_4 - 2\mu_3^2 + 2}{n^3}. \tag{B.23}$$

$$E_{3,1} = \frac{3\mu_3}{n^2} + \frac{\mu_5 - 4\mu_3}{n^3}. \tag{B.24}$$

$$E_{3,2} = \frac{3\mu_5 + 7\mu_4\mu_3 - 13\mu_3}{n^3} + \frac{\mu_7 - 5\mu_5 - 7\mu_4\mu_3 + 14\mu_3}{n^4}. \tag{B.25}$$

$$E_{3,3} = \frac{9\mu_3(\mu_4 - 1) + 6\mu_3^3}{n^3} + O(n^{-4}). \tag{B.26}$$

$$E_{5,1} = \frac{15\mu_3}{n^3} + \frac{10\mu_5 + 15\mu_4\mu_3 - 65\mu_3}{n^4} + O(n^{-5}). \tag{B.27}$$

$$E_{5,2} = \frac{15\mu_5 + 70\mu_4\mu_3 + 20\mu_3^3 - 100\mu_3}{n^4} + O(n^{-5}). \tag{B.28}$$

$$E_{5,3} = \frac{45\mu_4\mu_3 + 60\mu_3^3 - 45\mu_3}{n^4} + O(n^{-5}). \tag{B.29}$$

$$E_{7,1} = \frac{105\mu_3}{n^4} + O(n^{-5}). \tag{B.30}$$

B.4 Trivariate case ($q = 3$)

Let $(X_1, Y_1, Z_1), \ldots, (X_n, Y_n, Z_n)$ be iid trivariate random vectors such that $\mathrm{E}(X_1) = \mathrm{E}(Y_1) = \mathrm{E}(Z_1) = 0$ and let $\{u_i(n), i = 1, \ldots, n\}_{n=1}^{\infty}$, $\{v_i(n), i = 1, \ldots, n\}_{n=1}^{\infty}$ and $\{w_i(n), i = 1, \ldots, n\}_{n=1}^{\infty}$ be triangular arrays of real numbers such that $\max_{1 \le i \le n} |u_i(n)| = O(1)$, $\max_{1 \le i \le n} |v_i(n)| = O(1)$, and $\max_{1 \le i \le n} |w_i(n)| = O(1)$ as $n \to \infty$. Then,

$$E_{1,1,1} = \frac{1}{n^3} \mathrm{E}\left\{ \sum_i u_i v_i w_i X_i Y_i Z_i \right\} = \frac{\Sigma(123)}{n^2}, \tag{B.31}$$

where $u_i = u_i(n)$, $v_i = v_i(n)$ and $w_i = w_i(n)$.

Appendix C

Moments of sample moments and cumulants (regression)

We provide here some results relating to the expectation of estimators for the slope parameter in a simple linear regression model, as well as expressions for some moments of sample moments based on residuals resulting from the model fit.

This appendix employs the notation of Section 3.5. Recall that $\varepsilon_1, \ldots, \varepsilon_n$ are iid random errors with expectation 0 and variance $0 < \sigma^2 < \infty$, and that the estimator for the slope parameter d is given by

$$\widehat{d}_n = d + \frac{1}{n\sigma_x^2} \sum_{i=1}^{n} \widetilde{x}_i(\varepsilon_i - \bar{\varepsilon}) = d + \frac{1}{n\sigma_x^2} \sum_{i=1}^{n} \widetilde{x}_i \varepsilon_i, \tag{C.1}$$

with $\sigma_x^2 = \sigma_{x,n}^2 = n^{-1} \sum_{i=1}^{n} \widetilde{x}_i^2 > 0$, where we set $\widetilde{x}_i = x_i - \bar{x}_n$ to facilitate notation. Throughout we assume that $\max_{1 \le i \le n} |\widetilde{x}_i| = O(1)$ as $n \to \infty$. Also recall that

$$\gamma_{x,n} = \frac{1}{n\sigma_{x,n}^3} \sum_{i=1}^{n} \widetilde{x}_i^3, \qquad \kappa_{x,n} = \frac{1}{n\sigma_{x,n}^4} \sum_{i=1}^{n} \widetilde{x}_i^4 - 3, \qquad \tau_{x,n} = \frac{1}{n\sigma_{x,n}^5} \sum_{i=1}^{n} \widetilde{x}_i^5 - 10\gamma_{x,n}.$$

and

$$\widetilde{m}_j = \frac{1}{n} \sum_{i=1}^{n} e_i^j, \qquad \widetilde{\kappa}'_{3,n} = \frac{\widetilde{m}_3}{\widetilde{m}_2^{3/2}}, \qquad \widetilde{\kappa}'_{4,n} = \frac{\widetilde{m}_4}{\widetilde{m}_2^2} - 3, \qquad \widetilde{\kappa}'_{5,n} = \frac{\widetilde{m}_5}{\widetilde{m}_2^{5/2}} - 10\widetilde{\kappa}'_{3,n},$$

where

$$e_i = (\varepsilon_i - \bar{\varepsilon}) - \widetilde{x}_i(\widehat{d}_n - d).$$

C.1 Moments of an estimator for the slope parameter

Lemma C.1. *Assume that* $\max_{1 \le i \le n} |x_i - \bar{x}_n| = O(1)$ *as* $n \to \infty$ *and that* ε_1 *has sufficiently many finite moments. Then*

(i) $\mathrm{E}\left(\widehat{d}_n - d\right) = 0,$

(ii) $\mathrm{E}\left\{\left(\widehat{d}_n - d\right)^2\right\} = \dfrac{\mu_2}{n\sigma_x^2},$

(iii) $\mathrm{E}\left\{\left(\widehat{d}_n - d\right)^3\right\} = \dfrac{\mu_3 \gamma_x}{n^2 \sigma_x^3}$,

(iv) $\mathrm{E}\left\{\left(\widehat{d}_n - d\right)^4\right\} = \dfrac{3\mu_2^2}{n^2 \sigma_x^4} + O(n^{-3})$.

Proof.

(i) Follows directly from (C.1) since $\mathrm{E}(\varepsilon_1) = 0$.

(ii) By (B.2) we have that

$$\mathrm{E}\left\{\left(\widehat{d}_n - d\right)^2\right\} = \frac{1}{\sigma_x^4}\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i \varepsilon_i\right)^2\right\} = \frac{1}{n\sigma_x^4}\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i^2\right)\mathrm{E}(\varepsilon_1^2) = \frac{1}{n\sigma_x^2}\mu_2.$$

(iii) From (B.3) it follows that

$$\mathrm{E}\left\{\left(\widehat{d}_n - d\right)^3\right\} = \frac{1}{\sigma_x^6}\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i \varepsilon_i\right)^3\right\} = \frac{1}{n^2\sigma_x^6}\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i^3\right)\mathrm{E}(\varepsilon_1^3) = \frac{\gamma_x}{n^2\sigma_x^3}\mu_3.$$

(iv) By (B.4),

$$\begin{aligned}\mathrm{E}\left\{\left(\widehat{d}_n - d\right)^4\right\} = \frac{1}{\sigma_x^8}\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i \varepsilon_i\right)^4\right\} &= \frac{3}{n^2\sigma_x^8}\left(\frac{1}{n}\sum_i \widetilde{x}_i^2\right)^2\left(\mathrm{E}(\varepsilon_1^2)\right)^2 + O(n^{-3}) \\ &= \frac{3}{n^2\sigma_x^4}\mu_2^2 + O(n^{-3}).\end{aligned}$$

□

In some cases we will not require the coefficient of the leading term, but only the rate of convergence. To that end we state and prove the following general result.

Lemma C.2. *Assume that* $\max_{1\le i\le n}|x_i - \bar{x}_n| = O(1)$ *as* $n\to\infty$ *and that* ε_1 *has sufficiently many finite moments. Then, for any integer* $p \ge 1$,

$$\mathrm{E}\left\{\left(\widehat{d}_n - d\right)^p\right\} = O(n^{-\lfloor (p+1)/2\rfloor}),$$

where $\lfloor x\rfloor$ *denotes the largest integer smaller than or equal to* x.

Proof. Follows directly from Lemma B.1. □

C.2 Moments of sample moments of residuals

Cramér (1946) states in (27.5.3) the well-known fact that, for any integer $p \ge 1$,

$$\mathrm{E}\left(m_p\right) = \mu_p + O(n^{-1}), \tag{C.2}$$

where $m_p = n^{-1}\sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^p$ and $\mu_p = \mathrm{E}(\varepsilon_1^p)$. As is proved below, this result may be extended to the case where the pth sample moment m_p is based on the residuals instead of an iid sample.

Lemma C.3. *If* $\max_{1\le i\le n}|x_i - \bar{x}_n| = O(1)$ *as* $n\to\infty$ *and* ε_1 *has sufficiently many finite moments, then*

$$\mathrm{E}\left(\widetilde{m}_p\right) = \mu_p + O(n^{-1}),$$

for any integer $p \ge 1$.

Proof. Observe that

$$\begin{aligned}
\widetilde{m}_p &= \frac{1}{n}\sum_{i=1}^{n}\{(\varepsilon_i-\bar{\varepsilon})-\widetilde{x}_i(\widehat{d}_n-d)\}^p \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{k=0}^{p}\binom{p}{k}(\varepsilon_i-\bar{\varepsilon})^{p-k}\left\{-\widetilde{x}_i(\widehat{d}_n-d)\right\}^k \\
&= m_p + \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{p}\binom{p}{k}(\varepsilon_i-\bar{\varepsilon})^{p-k}\left\{-\widetilde{x}_i(\widehat{d}_n-d)\right\}^k \\
&= m_p + \sum_{k=1}^{p}(-1)^k\binom{p}{k}\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i^k(\varepsilon_i-\bar{\varepsilon})^{p-k}\right)\left(\widehat{d}_n-d\right)^k, \qquad \text{(C.3)}
\end{aligned}$$

so that, by (C.2),

$$\mathrm{E}\left(\widetilde{m}_p\right) = \mu_p + \sum_{k=1}^{p}(-1)^k\binom{p}{k}\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i^k(\varepsilon_i-\bar{\varepsilon})^{p-k}\right)\left(\widehat{d}_n-d\right)^k\right\} + O(n^{-1}).$$

By Lemma C.2 and the Cauchy-Schwarz inequality we have that

$$\begin{aligned}
\left|\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i^k(\varepsilon_i-\bar{\varepsilon})^{p-k}\right)\left(\widehat{d}_n-d\right)^k\right\}\right| &\le \sqrt{\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i^k(\varepsilon_i-\bar{\varepsilon})^{p-k}\right)^2\right\}\mathrm{E}\left\{\left(\widehat{d}_n-d\right)^{2k}\right\}} \\
&= \sqrt{O(n^{-\lfloor(2k+1)/2\rfloor})} = O(n^{-k/2}).
\end{aligned}$$

Hence,

$$\mathrm{E}\left(\widetilde{m}_p\right) = \mu_p - p\,\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i(\varepsilon_i-\bar{\varepsilon})^{p-1}\right)\left(\widehat{d}_n-d\right)\right\} + O(n^{-1}),$$

which may be written as

$$\begin{aligned}
\mathrm{E}\left(\widetilde{m}_p\right) &= \mu_p - p\,\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i\left\{\sum_{\ell=0}^{p-1}\binom{p-1}{\ell}\varepsilon_i^\ell(-\bar{\varepsilon})^{p-\ell-1}\right\}\right)\left(\widehat{d}_n-d\right)\right\} + O(n^{-1}) \\
&= \mu_p - p\sum_{\ell=0}^{p-1}(-1)^{p-\ell-1}\binom{p-1}{\ell}\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i\varepsilon_i^\ell\right)\bar{\varepsilon}^{p-\ell-1}\left(\widehat{d}_n-d\right)\right\} + O(n^{-1}).
\end{aligned}$$

Now, notice that by Corollary B.1,

$$\begin{aligned}
\left|\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i\varepsilon_i^\ell\right)\bar{\varepsilon}^{p-\ell-1}\left(\widehat{d}_n-d\right)\right\}\right| &\le \sqrt{\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i\varepsilon_i^\ell\right)^2\right\}\mathrm{E}\left\{\bar{\varepsilon}^{2(p-\ell-1)}\left(\widehat{d}_n-d\right)^2\right\}} \\
&= \sqrt{O(n^{-(p-\ell)})} = O(n^{-(p-\ell)/2}).
\end{aligned}$$

This implies that

$$
\begin{aligned}
\mathrm{E}\left(\widetilde{m}_p\right) &= \mu_p - p\,\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i\varepsilon_i^{p-1}\right)\left(\widehat{d}_n - d\right)\right\} + O(n^{-1}) \\
&= \mu_p - \frac{p}{\sigma_x^2}\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i\left(\varepsilon_i^{p-1} - \mu_{p-1}\right)\right)\left(\frac{1}{n}\sum_{k=1}^{n}\widetilde{x}_k\varepsilon_k\right)\right\} + O(n^{-1}),
\end{aligned}
$$

which by (B.7) becomes

$$
\begin{aligned}
\mathrm{E}\left(\widetilde{m}_p\right) &= \mu_p - \frac{p}{n\sigma_x^2}\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i^2\right)\mathrm{E}\left\{\left(\varepsilon_1^{p-1} - \mu_{p-1}\right)\varepsilon_1\right\} + O(n^{-1}) \\
&= \mu_p - \frac{p\mu_p}{n} + O(n^{-1}) \\
&= \mu_p + O(n^{-1}).
\end{aligned}
$$

□

Another useful result given by Cramér (1946) states that

$$
\mathrm{E}\left\{\left(m_p - \mu_p\right)^{2k}\right\} = O(n^{-k}). \tag{C.4}
$$

This result may also be extended to $\widetilde{m}_p$.

Lemma C.4. *If* $\max_{1\le i\le n}|x_i - \bar{x}_n| = O(1)$ *as* $n\to\infty$ *and* ε_1 *has sufficiently many finite moments, then*

$$
\mathrm{E}\left\{\left(\widetilde{m}_p - \mu_p\right)^{2k}\right\} = O(n^{-k}).
$$

for any integers $p, k \ge 1$.

Proof. From (C.3) we have that

$$
\widetilde{m}_p = m_p + \sum_{j=1}^{p}(-1)^j R_{n,p,j}\left(\widehat{d}_n - d\right)^j,
$$

where

$$
R_{n,p,j} = \binom{p}{j}\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i^j(\varepsilon_i - \bar{\varepsilon})^{p-j}.
$$

It is well-known (see, e.g., von Bahr and Esseen, 1965) that, if $X_1, X_2, \ldots, X_m$ is a sequence of random variables, then, for $r \ge 1$,

$$
\mathrm{E}\left(\left|\sum_{i=1}^{m}X_i\right|^r\right) \le m^{r-1}\sum_{i=1}^{m}\mathrm{E}\left(|X_i|^r\right).
$$

Hence,

$$
\mathrm{E}\left\{\left(\widetilde{m}_p - \mu_p\right)^{2k}\right\} \le (p+1)^{2k-1}\mathrm{E}\left\{\left(m_p - \mu_p\right)^{2k}\right\} + (p+1)^{2k-1}\sum_{j=1}^{p}\mathrm{E}\left\{R_{n,p,j}^{2k}\left(\widehat{d}_n - d\right)^{2kj}\right\}. \tag{C.5}
$$

Since

$$
\begin{aligned}
\left|\mathrm{E}\left\{(\widehat{d}_n - d)^{2kj}R_{n,p,j}\ldots\right\}\right| &\le \sqrt{\mathrm{E}\{(\widehat{d}_n - d)^{4kj}\}\,\mathrm{E}\left\{R_{n,p,j}^2\right\}} \\
&= \sqrt{O(n^{-2kj})\,\mathrm{E}\left\{R_{n,p,j}^2\right\}} \\
&= O(n^{-kj}),
\end{aligned}
$$

(C.5) reduces to

$$\mathrm{E}\left\{\left(\widetilde{m}_p-\mu_p\right)^{2k}\right\}=O(n^{-k})+(p+1)^{2k-1}\sum_{j=1}^{p}O(n^{-kj})=O(n^{-k}),$$

where we have made use of (C.4). □

Alternative proof. This proof follows along the same lines as the derivation of (C.4) given in Cramér (1946). From (C.3) we have that

$$\widetilde{m}_p=m_p+\sum_{k=1}^{p}(-1)^k R_{n,p,k}\left(\widehat{d}_n-d\right)^k,$$

with $R_{n,p,k}$ defined above. Writing

$$\left(\widetilde{m}_p-\mu_p\right)=\left(m_p-\mu_p\right)-R_{n,p,1}\left(\widehat{d}_n-d\right)+R_{n,p,2}\left(\widehat{d}_n-d\right)^2+\cdots, \tag{C.6}$$

we see that any power of $\widetilde{m}_p-\mu_p$ is composed of terms of the form $(m_p-\mu_p)^i(\widehat{d}_n-d)^j R_{n,p,k_1}R_{n,p,k_2}\cdots$. We now show that any such term has a mean of order $O(n^{-(i+j)/2})$.

First note that by the Cauchy-Schwarz inequality one may write

$$\left|\mathrm{E}\left\{(m_p-\mu_p)^{2i}(\widehat{d}_n-d)^{2j}\right\}\right|\leq\sqrt{\mathrm{E}\left\{(m_p-\mu_p)^{4i}\right\}\mathrm{E}\left\{(\widehat{d}_n-d)^{4j}\right\}}=O(n^{-(i+j)}),$$

where we have made use of (C.4) and Lemma C.2. Therefore, by the same inequality,

$$\begin{aligned}&\left|\mathrm{E}\left\{(m_p-\mu_p)^{i}(\widehat{d}_n-d)^{j}R_{n,p,k_1}R_{n,p,k_2}\cdots\right\}\right|\\&\leq\sqrt{\mathrm{E}\left\{(m_p-\mu_p)^{2i}(\widehat{d}_n-d)^{2j}\right\}\mathrm{E}\left\{R^2_{n,p,k_1}R^2_{n,p,k_2}\cdots\right\}}\\&=\sqrt{O(n^{-(i+j)})\mathrm{E}\left\{R^2_{n,p,k_1}R^2_{n,p,k_2}\cdots\right\}}\\&=O(n^{-(i+j)/2}),\end{aligned}$$

assuming that ε_1 has sufficiently many finite moments.

It is now clear that it is sufficient to retain the terms

$$\left(m_p-\mu_p\right)-R_{n,p,1}\left(\widehat{d}_n-d\right)$$

in order to calculate the leading term of $\mathrm{E}\{(\widetilde{m}_p-\mu_p)^{2k}\}$. The remaining terms in (C.6) give a contribution of lower order. Therefore we have by the c_r-inequality that

$$\begin{aligned}&\mathrm{E}\left\{\left(m_p-\mu_p-R_{n,p,1}\left(\widehat{d}_n-d\right)\right)^{2k}\right\}\\&\leq 2^{2k-1}\mathrm{E}\left\{\left(m_p-\mu_p\right)^{2k}\right\}+2^{2k-1}\mathrm{E}\left\{R^{2k}_{n,p,1}\left(\widehat{d}_n-d\right)^{2k}\right\}\\&=O(n^{-k}),\end{aligned}$$

which proves the theorem. □

Lemma C.5. *Assume that* $\max_{1\leq i\leq n}|x_i-\bar{x}_n|=O(1)$ *as* $n\to\infty$ *and that* ε_1 *has sufficiently many finite moments. Then*

(i) $\quad\mathrm{E}\left(\widetilde{m}_2\right)=\mu_2-\dfrac{2\mu_2}{n},$

(ii) $\quad\mathrm{E}\left(\widetilde{m}_2^2\right)=\mu_2^2+\dfrac{\mu_4-5\mu_2^2}{n}+O(n^{-2}).$

Proof.

(i) Since

$$\widetilde{m}_2 = m_2 - \sigma_x^2(\widehat{d}_n - d)^2,$$

the result follows directly from (27.4.1) of Cramér (1946) and Lemma C.1.

(ii) Observe that

$$\widetilde{m}_2^2 = m_2^2 + \sigma_x^4(\widehat{d}_n - d)^4 - 2\sigma_x^2 m_2(\widehat{d}_n - d)^2.$$

Since

$$m_2(\widehat{d}_n - d)^2 = \left(\frac{1}{n}\sum_{i=1}^{n}(\varepsilon_i^2 - \mu_2)\right)\left(\widehat{d}_n - d\right)^2 - \bar{\varepsilon}^2\left(\widehat{d}_n - d\right)^2 + \mu_2\left(\widehat{d}_n - d\right)^2,$$

we have by (B.8), (B.9) and Lemma C.1 that

$$\mathrm{E}\left\{m_2(\widehat{d}_n - d)^2\right\} = \frac{\mu_2^2}{n\sigma_x^2} + O(n^{-2}).$$

The result now follows directly from (27.4.2) of Cramér (1946) and Lemma C.1. □

Corollary C.1. *Assume that* $\max_{1\le i\le n}|x_i - \bar{x}_n| = O(1)$ *as* $n\to\infty$ *and that* ε_1 *has sufficiently many finite moments. Then*

(i) $\mathrm{E}\left(\widetilde{m}_2 - \mu_2\right) = -\dfrac{2\mu_2}{n}$,

(ii) $\mathrm{E}\left\{\left(\widetilde{m}_2 - \mu_2\right)^2\right\} = \dfrac{\mu_4 - \mu_2^2}{n} + O(n^{-2})$.

Proof. Follows directly from Lemma C.5. □

Lemma C.6. *Assume that* $\max_{1\le i\le n}|x_i - \bar{x}_n| = O(1)$ *as* $n\to\infty$ *and that* ε_1 *has sufficiently many finite moments. Then*

$$\mathrm{E}(\widetilde{m}_3) = \mu_3 - \frac{6\mu_3}{n} + O(n^{-2}).$$

Proof. Observe that

$$\begin{aligned}
\widetilde{m}_3 &= \frac{1}{n}\sum_{i=1}^{n}\left\{(\varepsilon_i - \bar{\varepsilon}) - \widetilde{x}_i(\widehat{d}_n - d)\right\}^3 \\
&= \frac{1}{n}\sum_{i=1}^{n}\left\{(\varepsilon_i - \bar{\varepsilon})^3 - 3(\varepsilon_i - \bar{\varepsilon})^2\widetilde{x}_i(\widehat{d}_n - d) + 3(\varepsilon_i - \bar{\varepsilon})\widetilde{x}_i^2(\widehat{d}_n - d)^2 - \widetilde{x}_i^3(\widehat{d}_n - d)^3\right\} \\
&= m_3 - \frac{3}{n}\sum_{i=1}^{n}\widetilde{x}_i(\varepsilon_i - \bar{\varepsilon})^2\left(\widehat{d}_n - d\right) + \frac{3}{n}\sum_{i=1}^{n}\widetilde{x}_i^2(\varepsilon_i - \bar{\varepsilon})\left(\widehat{d}_n - d\right)^2 + \gamma_x\sigma_x^3\left(\widehat{d}_n - d\right)^3.
\end{aligned}$$

By (B.7) and (B.8) it follows that

$$\begin{aligned}
&\mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i(\varepsilon_i - \bar{\varepsilon})^2\left(\widehat{d}_n - d\right)\right\} \\
&= \mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i\left(\varepsilon_i^2 - \mu_2\right)\left(\widehat{d}_n - d\right)\right\} - 2\sigma_x^2\,\mathrm{E}\left\{\bar{\varepsilon}\left(\widehat{d}_n - d\right)^2\right\}
\end{aligned}$$

$$= \frac{1}{\sigma_x^2} \mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i\left(\varepsilon_i^2-\mu_2\right)\right)\left(\frac{1}{n}\sum_{j=1}^n \widetilde{x}_j\varepsilon_j\right)\right\} - \frac{2}{\sigma_x^2}\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^n \varepsilon_i\right)\left(\frac{1}{n}\sum_{j=1}^n \widetilde{x}_j\varepsilon_j\right)^2\right\}$$
$$= \frac{1}{n\sigma_x^2}\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i^2\right)\mathrm{E}\left\{\left(\varepsilon_1^2-\mu_2\right)\varepsilon_1\right\} - \frac{2}{n^2\sigma_x^2}\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i^2\right)\mathrm{E}\left(\varepsilon_1^3\right)$$
$$= \frac{\mu_3}{n} + O(n^{-2}),$$

and

$$\mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i^2(\varepsilon_i-\bar{\varepsilon})\left(\widehat{d}_n-d\right)^2\right\}$$
$$= \frac{1}{\sigma_x^2}\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i^2\varepsilon_i\right)\left(\frac{1}{n}\sum_{j=1}^n \widetilde{x}_j\varepsilon_j\right)^2\right\} - \frac{1}{\sigma_x^2}\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i^2\right)\mathrm{E}\left\{\bar{\varepsilon}\left(\frac{1}{n}\sum_{j=1}^n \widetilde{x}_j\varepsilon_j\right)^2\right\}$$
$$= \frac{1}{n^2\sigma_x^2}\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i^4\right)\mathrm{E}\left(\varepsilon^3\right) - \left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i^2\right)\mathrm{E}\left(\varepsilon^3\right)$$
$$= O(n^{-2}).$$

Hence, one may obtain the result by applying (27.5.1) of Cramér (1946) and Lemma C.1. □

Lemma C.7. *Assume that* $\max_{1\le i\le n}|x_i-\bar{x}_n| = O(1)$ *as* $n\to\infty$ *and that* ε_1 *has sufficiently many finite moments. Then*

$$\mathrm{E}\left(\widetilde{m}_4\right) = \mu_4 + \frac{8\mu_4-12\mu_2^2}{n} + O(n^{-2}).$$

Proof. Observe that

$$\begin{aligned}\widetilde{m}_4 &= \frac{1}{n}\sum_{i=1}^n \left\{(\varepsilon_i-\bar{\varepsilon})-\widetilde{x}_i(\widehat{d}_n-d)\right\}^4 \\ &= m_4 - \frac{4}{n}\sum_{i=1}^n \widetilde{x}_i(\varepsilon_i-\bar{\varepsilon})^3\left(\widehat{d}_n-d\right) + \frac{6}{n}\sum_{i=1}^n \widetilde{x}_i^2(\varepsilon_i-\bar{\varepsilon})^2\left(\widehat{d}_n-d\right)^2 - \frac{4}{n}\sum_{i=1}^n \widetilde{x}_i^3(\varepsilon_i-\bar{\varepsilon})\left(\widehat{d}_n-d\right)^3 \\ &\quad + (\kappa_x+3)\sigma_x^4\left(\widehat{d}_n-d\right)^4. \end{aligned} \tag{C.7}$$

Since

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i(\varepsilon_i-\bar{\varepsilon})^3 &= \frac{1}{n}\sum_{i=1}^n \widetilde{x}_i\varepsilon_i^3 - 3\bar{\varepsilon}\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i\varepsilon_i^2 + 3\sigma_x^2\bar{\varepsilon}^2\left(\widehat{d}_n-d\right) \\ &= \frac{1}{n}\sum_{i=1}^n \widetilde{x}_i\left(\varepsilon_i^3-\mu_3\right) - 3\bar{\varepsilon}\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i\left(\varepsilon_i^2-\mu_2\right) + 3\sigma_x^2\bar{\varepsilon}^2\left(\widehat{d}_n-d\right),\end{aligned}$$

it follows by (B.7), (B.9) and (B.31) that

$$\mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i(\varepsilon_i-\bar{\varepsilon})^3\left(\widehat{d}_n-d\right)\right\}$$
$$= \mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i\left(\varepsilon_i^3-\mu_3\right)\right)\left(\widehat{d}_n-d\right)\right\} - 3\,\mathrm{E}\left\{\bar{\varepsilon}\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i\left(\varepsilon_i^2-\mu_2\right)\right)\left(\widehat{d}_n-d\right)\right\} + 3\sigma_x^2\mathrm{E}\left\{\bar{\varepsilon}^2\left(\widehat{d}_n-d\right)^2\right\}$$
$$= \frac{1}{n\sigma_x^2}\left(\frac{1}{n}\sum_{i=1}^n \widetilde{x}_i^2\right)\mathrm{E}\left(\varepsilon_1^4\right) + O(n^{-2}) = \frac{\mu_4}{n} + O(n^{-2}).$$

By (B.8), (B.9), Lemma C.1 and Lemma B.2 one has

$$
\begin{aligned}
&\mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^2(\varepsilon_i-\bar{\varepsilon})^2\left(\widehat{d}_n-d\right)^2\right\}\\
&=\mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^2\varepsilon_i^2\left(\widehat{d}_n-d\right)^2\right\}-2\,\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^2\varepsilon_i\right)\bar{\varepsilon}\left(\widehat{d}_n-d\right)^2\right\}+\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^2\right)\mathrm{E}\left\{\bar{\varepsilon}^2\left(\widehat{d}_n-d\right)^2\right\}\\
&=\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^2\left(\varepsilon_i^2-\mu_2\right)\right)\left(\widehat{d}_n-d\right)^2\right\}+\mu_2\sigma_x^2\,\mathrm{E}\left\{\left(\widehat{d}_n-d\right)^2\right\}+O(n^{-2})\\
&=\frac{\mu_2^2}{n}+O(n^{-2}).
\end{aligned}
$$

Finally, by (B.10),

$$
\begin{aligned}
&\mathrm{E}\left\{\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^3(\varepsilon_i-\bar{\varepsilon})\left(\widehat{d}_n-d\right)^3\right\}\\
&=\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^3\varepsilon_i\right)\left(\widehat{d}_n-d\right)^3\right\}-\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^3\right)\mathrm{E}\left\{\bar{\varepsilon}\left(\widehat{d}_n-d\right)^3\right\}=O(n^{-2}).
\end{aligned}
$$

Now, taking expected values in (C.7), the result follows directly by making use of (27.5.1) of Cramér (1946) and Lemma C.1. □

Lemma C.8. *Assume that* $\max_{1\le i\le n}|x_i-\bar{x}_n|=O(1)$ *as* $n\to\infty$ *and that* ε_1 *has sufficiently many finite moments. Then*

(i) $\mathrm{E}(\widetilde{m}_3\widetilde{m}_2)=\mu_3\mu_2+\dfrac{\mu_5-12\mu_3\mu_2}{n}+O(n^{-2})$,

(ii) $\mathrm{E}(\widetilde{m}_4\widetilde{m}_2)=\mu_4\mu_2+O(n^{-1})$.

Proof.

(i) Expanding $\widetilde{m}_3\widetilde{m}_2$ yields

$$
\begin{aligned}
\widetilde{m}_3\widetilde{m}_2&=\left\{\frac{1}{n^2}\sum_{i,j}\left((\varepsilon_i-\bar{\varepsilon})-\tilde{x}_i(\widehat{d}_n-d)\right)^3\left((\varepsilon_i-\bar{\varepsilon})-\tilde{x}_i(\widehat{d}_n-d)\right)^2\right\}\\
&=m_3m_2-\sigma_x^2m_3\left(\widehat{d}_n-d\right)^2-\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^3\right)m_2\left(\widehat{d}_n-d\right)^3\\
&\quad+\sigma_x^2\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^3\right)\left(\widehat{d}_n-d\right)^5-3\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i(\varepsilon_i-\bar{\varepsilon})^2\right)m_2\left(\widehat{d}_n-d\right)\\
&\quad+3\sigma_x^2\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i(\varepsilon_i-\bar{\varepsilon})^2\right)\left(\widehat{d}_n-d\right)^3+3\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^2(\varepsilon_i-\bar{\varepsilon})\right)m_2\left(\widehat{d}_n-d\right)^2\\
&\quad-3\sigma_x^2\left(\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i^2(\varepsilon_i-\bar{\varepsilon})\right)\left(\widehat{d}_n-d\right)^4.
\end{aligned}
$$

Now, note that by (B.8), (B.11), Lemma C.1 and Lemma B.2 we have that

$$
\begin{aligned}
&\mathrm{E}\left\{m_3\left(\widehat{d}_n-d\right)^2\right\} \\
&=\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_i^3-\mu_3\right)\right)\left(\widehat{d}_n-d\right)^2\right\}+\mu_3\,\mathrm{E}\left\{\left(\widehat{d}_n-d\right)^2\right\} \\
&\quad-3\,\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i^2\right)\bar{\varepsilon}\left(\widehat{d}_n-d\right)^2\right\}+3\,\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i\right)\bar{\varepsilon}^2\left(\widehat{d}_n-d\right)^2\right\}+\mathrm{E}\left\{\bar{\varepsilon}^3\left(\widehat{d}_n-d\right)^2\right\} \\
&=\frac{\mu_3\mu_2}{n\sigma_x^2}+O(n^{-2}).
\end{aligned}
$$

Also, by (B.7), (B.8), (B.11) and Lemma B.2,

$$
\begin{aligned}
&\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i\left(\varepsilon_i^2-\bar{\varepsilon}\right)^2\right)m_2\left(\widehat{d}_n-d\right)\right\} \\
&=\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i\left(\varepsilon_i^2-\mu_2\right)\right)^2\left(\widehat{d}_n-d\right)\right\}-\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i\left(\varepsilon_i^2-\mu_2\right)\right)\bar{\varepsilon}^2\left(\widehat{d}_n-d\right)\right\} \\
&\quad+\mu_2\,\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i\left(\varepsilon_i^2-\mu_2\right)\right)\left(\widehat{d}_n-d\right)\right\}-2\sigma_x^2\,\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\widetilde{x}_i\left(\varepsilon_i^2-\mu_2\right)\right)\bar{\varepsilon}\left(\widehat{d}_n-d\right)^2\right\} \\
&\quad+2\sigma_x^2\,\mathrm{E}\left\{\bar{\varepsilon}^3\left(\widehat{d}_n-d\right)^2\right\}-2\mu_2\sigma_x^2\,\mathrm{E}\left\{\bar{\varepsilon}\left(\widehat{d}_n-d\right)^2\right\} \\
&=\frac{\mu_3\mu_2}{n\sigma_x^2}+O(n^{-2}).
\end{aligned}
$$

By (B.10), (B.11), and Lemma C.1 and Lemma B.2 we have

$$
\begin{aligned}
\mathrm{E}\left\{m_2\left(\widehat{d}_n-d\right)^3\right\}&=\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\left(\varepsilon_i^2-\mu_2\right)\right)\left(\widehat{d}_n-d\right)^3\right\}+\mu_2\,\mathrm{E}\left\{\left(\widehat{d}_n-d\right)^3\right\} \\
&\quad-2\,\mathrm{E}\left\{\left(\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i\right)\bar{\varepsilon}\left(\widehat{d}_n-d\right)^3\right\}+\mathrm{E}\left\{\bar{\varepsilon}^2\left(\widehat{d}_n-d\right)^3\right\}=O(n^{-2}).
\end{aligned}
$$

Similarly one may show that $n^{-1}\sum_{i=1}^{n}\widetilde{x}_i(\varepsilon_i-\bar{\varepsilon})^2(\widehat{d}_n-d)^3$ and $n^{-1}\sum_{i=1}^{n}\widetilde{x}_i^2(\varepsilon_i-\bar{\varepsilon})m_2(\widehat{d}_n-d)^2$ both have expectation $O(n^{-2})$. Finally, noting that by Cauchy-Schwarz any term containing $(\widehat{d}_n-d)^4$ has expectation $O(n^{-2})$, we have

$$
\begin{aligned}
\mathrm{E}\left(\widetilde{m}_3\widetilde{m}_2\right)&=\mathrm{E}\left(m_3m_2\right)-\frac{4\mu_3\mu_2}{n}+O(n^{-2}) \\
&=\mu_3\mu_2+\frac{\mu_5-12\mu_3\mu_2}{n}+O(n^{-2}),
\end{aligned}
$$

where we used Lemma A.4 in the final step.

(ii) By the Cauchy-Schwarz inequality and Lemma C.4,

$$
\left|\mathrm{E}\left\{\left(\widetilde{m}_4-\mu_4\right)\left(\widetilde{m}_2-\mu_2\right)\right\}\right|\le\sqrt{\mathrm{E}\left\{\left(\widetilde{m}_4-\mu_4\right)^2\right\}\mathrm{E}\left\{\left(\widetilde{m}_2-\mu_2\right)^2\right\}}=O(n^{-1}),
$$

which, together with Leamms C.5 and C.7, proves the result. □

Corollary C.2. *Assume that* $\max_{1\le i\le n}|x_i-\bar{x}_n| = O(1)$ *as* $n\to\infty$ *and that* ε_1 *has sufficiently many finite moments. Then*

$$\text{(i)} \qquad \mathrm{E}\left\{\left(\widetilde{m}_3-\mu_3\right)\left(\widetilde{m}_2-\mu_2\right)\right\} = \frac{\mu_5-4\mu_3\mu_2}{n}+O(n^{-2}).$$

$$\text{(ii)} \qquad \mathrm{E}\left\{\left(\widetilde{m}_4-\mu_4\right)\left(\widetilde{m}_2-\mu_2\right)\right\} = O(n^{-1}).$$

Proof. Follows directly from Lemmas C.6, C.7 and C.8. □

C.3 Moments of sample cumulants of residuals

Lemma C.9. *Suppose that* $\max_{1\le i\le n}|x_i-\bar{x}_n| = O(1)$ *as* $n\to\infty$, *that* $\mathrm{E}(|\varepsilon_1|^k)<\infty$ *for some sufficiently large* k, *and that* $\mu_2>0$. *Then*

$$\mathrm{E}\left(\widetilde{\kappa}'_{3,n}-\kappa'_3\right) = -\frac{1}{8n}\left(12\kappa'_5-15\kappa'_4\kappa'_3+66\kappa'_3\right)+O(n^{-2}),$$

and $\mathrm{E}\{(\widetilde{\kappa}'_{3,n}-\kappa'_3)^4\} = O(n^{-2})$.

Proof. In the context of Corollary 7.1, set

$$V_1=\widetilde{m}_3, \qquad v_1=\mu_3, \qquad V_2=\widetilde{m}_2=\widehat{\sigma}^2_n, \qquad v_2=\mu_2=\sigma^2,$$

and let

$$h(x,y)=\frac{x}{y^{3/2}}.$$

We would now like to apply Corollary 7.1 to the functions $g=h,h^2,h^3,h^4$.

Along the same lines as in Lemma A.1 it may be shown that

$$|h(V_1,V_2)| = |h(\widetilde{m}_3,\widetilde{m}_2)| = \left|\frac{\widetilde{m}_3}{\widetilde{m}_2^{3/2}}\right| \le \sqrt{n}.$$

This implies that

$$|h(V_1,V_2)|^j\le n^{j/2}, \qquad j=1,2,\ldots.$$

Since we need to apply Corollary 7.1 to h^4, we require $\delta=4$. Noting that $k=6\vee(\delta+4)=8$, we must have that

$$\mathrm{E}\left\{(V_1-v_1)^8\right\} = \mathrm{E}\left\{\left(\widetilde{m}_3-\mu_3\right)^8\right\} = O(n^{-4})$$

and

$$\mathrm{E}\left\{(V_2-v_2)^8\right\} = \mathrm{E}\left\{\left(\widetilde{m}_2-\mu_2\right)^8\right\} = O(n^{-4}),$$

which are both confirmed by Lemma C.4 under the assumption of sufficiently many finite moments. Finally, given that $\mu_2>0$, the functions h, h^2, h^3 and h^4 all have bounded derivatives up to the fourth order in an open neighbourhood of (v_1,v_2).

In the context of Corollary 7.1, we have the following constants which we obtain from the results of Lemmas C.6 and C.7 and Corollary C.2:

$$C_1=-6\mu_3, \qquad C_2=-2\mu_2,$$
$$D_2=\mu_4-\mu_2^2 \qquad E_{12}=\mu_5-4\mu_3\mu_2.$$

Also note that (by the same results) item *(d)* of Corollary 7.1 is satisfied.

Now, by Corollary 7.1 we may write

$$\begin{aligned}
\mathrm{E}\left(\widetilde{\kappa}'_{3,n}-\kappa'_3\right) &= \frac{1}{\mu_2^{3/2}}\mathrm{E}\left(\widetilde{m}_3-\mu_3\right)-\frac{3\mu_3}{2\mu_2^{5/2}}\mathrm{E}\left(\widetilde{m}_2-\mu_2\right)+\frac{15\mu_3}{8\mu_2^{7/2}}\mathrm{E}\left\{\left(\widetilde{m}_2-\mu_2\right)^2\right\}\\
&\quad -\frac{3}{2\mu_2^{5/2}}\mathrm{E}\left\{\left(\widetilde{m}_3-\mu_3\right)\left(\widetilde{m}_2-\mu_2\right)\right\}+O(n^{-2})\\
&= \frac{1}{8n}\left(-12\kappa'_5+15\kappa'_4\kappa'_3-66\kappa'_3\right)+O(n^{-2}).
\end{aligned}$$

which proves the first part of the theorem.

By exactly the same arguments as in the proof of Lemma 7.2 it may be shown that $\mathrm{E}\{(\widetilde{\kappa}'_{3,n}-\kappa'_3)^4\}=O(n^{-2})$. □

Lemma C.10. *Suppose that* $\max_{1\le i\le n}|x_i-\bar{x}_n|=O(1)$ *as* $n\to\infty$*, that* $\mathrm{E}(|\varepsilon_1|^k)<\infty$ *for some sufficiently large* k*, and that* $\mu_2>0$*. Then*

$$\mathrm{E}\left(\widetilde{\kappa}'_{4,n}-\kappa'_4\right)=O(n^{-1}) \quad \textit{and} \quad \mathrm{E}\left\{\left(\widetilde{\kappa}'_{4,n}-\kappa'_4\right)^2\right\}=O(n^{-1}).$$

Proof. In the context of Theorem 7.1, set

$$V_1=\widetilde{m}_4, \qquad v_1=\mu_4, \qquad V_2=\widetilde{m}_2=\widehat{\sigma}^2_n, \qquad v_2=\mu_2=\sigma^2,$$

and let

$$h(x,y)=\frac{x}{y^2}.$$

We would like to apply Theorem 7.1 to the functions $g=h,h^2$.

Along the same lines as in Lemma A.1 it may be shown that

$$|h(V_1,V_2)|=|h(\widetilde{m}_4,\widetilde{m}_2)|=\left|\frac{\widetilde{m}_4}{\widetilde{m}_2^2}\right|\le n.$$

This implies that

$$|h(V_1,V_2)|^j\le n^j, \quad j=1,2,\ldots.$$

Since we need to apply Theorem 7.1 to h^2, we require $\delta=4$. Noting that $k=4\vee(\delta+4)=8$, we must have that

$$\mathrm{E}\left\{(V_1-v_1)^8\right\}=\mathrm{E}\left\{\left(\widetilde{m}_4-\mu_4\right)^8\right\}=O(n^{-4})$$

and

$$\mathrm{E}\left\{(V_2-v_2)^8\right\}=\mathrm{E}\left\{\left(\widetilde{m}_2-\mu_2\right)^8\right\}=O(n^{-4}),$$

which are both confirmed by Lemma C.4 under the assumption of sufficiently many finite moments. Finally, given that $\mu_2>0$, the functions h and h^2 have bounded derivatives up to the second order in an open neighbourhood of (v_1,v_2).

Now, by Theorem 7.1 we may write

$$\mathrm{E}\left(\widetilde{\kappa}'_{4,n}-\kappa'_4\right)=\frac{1}{\mu_2^2}\mathrm{E}\left(\widetilde{m}_4-\mu_4\right)-\frac{2\mu_4}{\mu_2^3}\mathrm{E}\left(\widetilde{m}_2-\mu_2\right)+O(n^{-1}).$$

By Theorem 7.1 and Lemma C.7 the result follows.

By exactly the same arguments as in the proof of Lemma 7.3 it may be shown that $\mathrm{E}\{(\widetilde{\kappa}'_{4,n} - \kappa'_4)^2\} = O(n^{-1})$. □

Lemma C.11. *Suppose that* $\max_{1 \le i \le n} |x_i - \bar{x}_n| = O(1)$ *as* $n \to \infty$*, that* $\mathrm{E}(|\varepsilon_1|^k) < \infty$ *for some sufficiently large* k*, and that* $\mu_2 > 0$*. Then*

$$\mathrm{E}\left(\widetilde{\kappa}'_{5,n} - \kappa'_5\right) = O(n^{-1}) \quad \textit{and} \quad \mathrm{E}\left\{\left(\widetilde{\kappa}'_{5,n} - \kappa'_5\right)^2\right\} = O(n^{-1}).$$

Proof. Using the results of this appendix, this lemma may be proved by exactly the same arguments as those found in the proof of Lemma 7.4. □

Appendix D

Supporting lemmas

D.1 Lemmas required in Chapter 4

This section contains a number of lemmas which are required to prove the theorems in Chapter 4. The notation used here is defined in that chapter.

Lemma D.1. *From the assumptions of Theorem 4.1 it follows that*

(i) $\mathrm{E}\left\{\widehat{P}_{1,r}(x)\left(\widehat{P}_{1,r}(x)-P_1(x)\right)\right\}=O(r^{-1})$,

(ii) $\mathrm{E}\left\{\widehat{P}_{1,r}(x)\left(\widehat{P}'_{1,r}(x)-P'_1(x)\right)\right\}=O(r^{-1})$,

(iii) $\mathrm{E}\left\{\widehat{P}^{cf}_{2,r}(x)-P^{cf}_2(x)\right\}=O(r^{-1})$,

(iv) $\mathrm{E}\left\{\left(\widehat{P}^{cf}_{2,r}(x)-P^{cf}_2(x)\right)^2\right\}=O(r^{-1})$,

(v) $\mathrm{E}\left\{\left(\widehat{\beta}_r-\beta\right)\widehat{P}_{1,r}(x)\right\}=O(r^{-1})$,

(vi) $\mathrm{E}\left\{\left(\widehat{\beta}_r-\beta\right)\widehat{P}^2_{1,r}(x)\right\}=O(r^{-1})$,

(vii) $\mathrm{E}\left\{\left(\widehat{\beta}_r-\beta\right)^2\widehat{P}_{1,r}(x)\right\}=O(r^{-1})$,

(viii) $\mathrm{E}\left\{\left(\widehat{\beta}_r-\beta\right)^2\widehat{P}^2_{1,r}(x)\right\}=O(r^{-1})$,

(ix) $\mathrm{E}\left\{\left(\widehat{\beta}_r-\beta\right)\widehat{P}^{cf}_{2,r}(x)\right\}=O(r^{-1})$,

(x) $\mathrm{E}\left\{\left(\widehat{\beta}_r-\beta\right)^2\widehat{P}^{cf}_{2,r}(x)\right\}=O(r^{-1})$.

Proof.

(i) It follows from assumption (A4) of Theorem 4.1 and Lemma 7.1 that

$$\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2\right\}=O(r^{-1}). \tag{D.1}$$

Therefore, using assumption (A3) of Theorem 4.1,

$$\mathrm{E}\left\{\widehat{P}_{1,r}(x)\left(\widehat{P}_{1,r}(x)-P_1(x)\right)\right\}=\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2\right\}+P_1(x)\,\mathrm{E}\left(\widehat{P}_{1,r}(x)-P_1(x)\right)=O(r^{-1}).$$

(ii) It follows from assumption (A6) of Theorem 4.1 and Lemma 7.1 that

$$\mathrm{E}\left\{\left(\widehat{P}'_{1,r}(x)-P'_1(x)\right)^2\right\}=O(r^{-1}). \tag{D.2}$$

Then, by the Cauchy-Schwarz inequality,

$$\begin{aligned}\mathrm{E}\left|\left(\widehat{P}_{1,r}(x)-P_1(x)\right)\left(\widehat{P}'_{1,r}(x)-P'_1(x)\right)\right| &\le \sqrt{\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2\right\}\mathrm{E}\left\{\left(\widehat{P}'_{1,r}(x)-P'_1(x)\right)^2\right\}}\\ &= \sqrt{O(r^{-1})O(r^{-1})}\\ &= O(r^{-1}). \end{aligned} \tag{D.3}$$

Hence, from assumption (A3) of Theorem 4.1, we have

$$\begin{aligned}&\mathrm{E}\left\{\widehat{P}_{1,r}(x)\left(\widehat{P}'_{1,r}(x)-P'_1(x)\right)\right\}\\ &=\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)\left(\widehat{P}'_{1,r}(x)-P'_1(x)\right)\right\}+P_1(x)\mathrm{E}\left(\widehat{P}'_{1,r}(x)-P'_1(x)\right)\\ &=O(r^{-1}).\end{aligned}$$

(iii) Notice that it follows from assumption (A3) of Theorem 4.1 and (D.1) that

$$\mathrm{E}\left(\widehat{P}^2_{1,r}(x)-P^2_1(x)\right)=\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2\right\}+2P_1(x)\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)\right\}=O(r^{-1}). \tag{D.4}$$

Now, using the expressions for $\widehat{P}^{cf}_{2,r}$ and P^{cf}_2 given in (3.10), we may use assumption (A7) of Theorem 4.1 and (D.3) to obtain

$$\begin{aligned}&\mathrm{E}\left(\widehat{P}^{cf}_{2,r}(x)-P^{cf}_2(x)\right)\\ &=\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)\widehat{P}'_{1,r}(x)-\tfrac{1}{2}x\widehat{P}^2_{1,r}(x)-\widehat{P}_{2,r}(x)\right)-\left(P_1(x)P'_1(x)-\tfrac{1}{2}xP^2_1(x)-P_2(x)\right)\right\}\\ &=\mathrm{E}\left(\widehat{P}_{1,r}(x)\widehat{P}'_{1,r}(x)-P_1(x)P'_1(x)\right)-\tfrac{1}{2}x\mathrm{E}\left(\widehat{P}^2_{1,r}(x)-P^2_1(x)\right)-\mathrm{E}\left(\widehat{P}_{2,r}(x)-P_2(x)\right)\\ &=O(r^{-1}).\end{aligned}$$

(iv) By the c_r-inequality and assumption (A4) of Theorem 4.1 we have

$$\begin{aligned}\mathrm{E}\left\{\left(\widehat{P}^2_{1,r}(x)-P^2_1(x)\right)^2\right\}&=\mathrm{E}\left\{\left(\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2+2P_1(x)\left(\widehat{P}_{1,r}(x)-P_1(x)\right)\right)^2\right\}\\ &\le 2\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^4\right\}+8P^2_1(x)\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2\right\}\\ &=O(r^{-1}),\end{aligned}$$

and also, from (D.1), (D.2) and assumption (A6) of Theorem 4.1,

$$\begin{aligned}&\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)\widehat{P}'_{1,r}(x)-P_1(x)P'_1(x)\right)^2\right\}\\ &=\mathrm{E}\left\{\left(\left(\widehat{P}_{1,r}(x)-P_1(x)\right)\left(\widehat{P}'_{1,r}(x)-P'_1(x)\right)+P_1(x)\left(\widehat{P}'_{1,r}(x)-P'_1(x)\right)\right.\right.\\ &\quad\left.\left.-P'_1(x)\left(\widehat{P}_{1,r}(x)-P_1(x)\right)\right)^2\right\}\end{aligned}$$

$$
\begin{aligned}
&\leq 2\,\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2\left(\widehat{P}'_{1,r}(x)-P'_1(x)\right)^2\right\}+4P_1^2(x)\,\mathrm{E}\left\{\left(\widehat{P}'_{1,r}(x)-P'_1(x)\right)^2\right\}\\
&\quad+4\left(P'_1(x)\right)^2\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2\right\}\\
&\leq 2\sqrt{\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^4\right\}\mathrm{E}\left\{\left(\widehat{P}'_{1,r}(x)-P'_1(x)\right)^4\right\}}+O(r^{-1})\\
&\leq 2\sqrt{O(r^{-2})O(r^{-2})}+O(r^{-1})\\
&=O(r^{-1}).
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
&\mathrm{E}\left\{\left(\widehat{P}_2^{cf}(x)-P_2^{cf}(x)\right)^2\right\}\\
&=\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)\widehat{P}'_{1,r}(x)-\tfrac{1}{2}x\widehat{P}^2_{1,r}(x)-\widehat{P}_{2,r}(x)-P_1(x)P'_1(x)+\tfrac{1}{2}xP_1^2(x)+P_2(x)\right)^2\right\}\\
&=\mathrm{E}\left\{\left(\left\{\widehat{P}_{1,r}(x)\widehat{P}'_{1,r}(x)-P_1(x)P'_1(x)\right\}-\tfrac{1}{2}x\left\{\widehat{P}^2_{1,r}(x)-P_1^2(x)\right\}-\{\widehat{P}_{2,r}(x)-P_2(x)\}\right)^2\right\}\\
&\leq 2\,\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)\widehat{P}'_{1,r}(x)-P_1(x)P'_1(x)\right)^2\right\}+x^2\,\mathrm{E}\left\{\left(\widehat{P}^2_{1,r}(x)-P_1^2(x)\right)^2\right\}\\
&\quad+4\,\mathrm{E}\left\{\left(\widehat{P}_{2,r}(x)-P_2(x)\right)^2\right\}\\
&=O(r^{-1}). \qquad\qquad (D.5)
\end{aligned}
$$

(v) We have from Lemma 7.1 and assumption (A2) of Theorem 4.1 that

$$\mathrm{E}\left\{\left(\widehat{\beta}_r-\beta\right)^2\right\}=O(r^{-1}). \qquad (D.6)$$

Then, by the Cauchy-Schwarz inequality and (D.1), it follows that

$$
\begin{aligned}
\mathrm{E}\left|\left(\widehat{\beta}_r-\beta\right)\left(\widehat{P}_{1,r}(x)-P_1(x)\right)\right| &\leq \sqrt{\mathrm{E}\left\{\left(\widehat{\beta}_r-\beta\right)^2\right\}\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2\right\}}\\
&=\sqrt{O(r^{-1})O(r^{-1})}\\
&=O(r^{-1}).
\end{aligned}
$$

Consequently, using assumption (A1) of Theorem 4.1, we obtain

$$\mathrm{E}\left(\left(\widehat{\beta}_r-\beta\right)\widehat{P}_{1,r}(x)\right)=\mathrm{E}\left\{\left(\widehat{\beta}_r-\beta\right)\left(\widehat{P}_{1,r}(x)-P_1(x)\right)\right\}+P_1(x)\,\mathrm{E}\left(\widehat{\beta}_r-\beta\right)=O(r^{-1}).$$

(vi) By the Cauchy-Schwarz inequality and assumptions (A2) and (A4) of Theorem 4.1 it follows that

$$
\begin{aligned}
\mathrm{E}\left|\left(\widehat{\beta}_r-\beta\right)\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2\right| &\leq \sqrt{\mathrm{E}\left\{\left(\widehat{\beta}_r-\beta\right)^2\right\}\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^4\right\}}\\
&=\sqrt{O(r^{-1})O(r^{-2})}\\
&=O(r^{-3/2}).
\end{aligned}
$$

This, together with (v) of this lemma and assumptions (A2) and (A4) of Theorem 4.1, implies that

$$\begin{aligned}
&\mathrm{E}\left((\widehat{\beta}_r-\beta)\widehat{P}_{1,r}^2(x)\right)\\
&=\mathrm{E}\left\{(\widehat{\beta}_r-\beta)\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2\right\}-P_1^2(x)\mathrm{E}\left(\widehat{\beta}_r-\beta\right)+2P_1(x)\mathrm{E}\left\{(\widehat{\beta}_r-\beta)\widehat{P}_{1,r}(x)\right\}\\
&=O(r^{-1}).
\end{aligned}$$

(vii) Assumption (A2) of Theorem 4.1 and (D.1) imply that

$$\begin{aligned}
\mathrm{E}\left|(\widehat{\beta}_r-\beta)^2\left(\widehat{P}_{1,r}(x)-P_1(x)\right)\right| &\le \sqrt{\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^4\right\}\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2\right\}}\\
&=\sqrt{O(r^{-2})O(r^{-1})}\\
&=O(r^{-3/2}).
\end{aligned}$$

Therefore, by (D.6),

$$\mathrm{E}\left((\widehat{\beta}_r-\beta)^2\widehat{P}_{1,r}(x)\right)=\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\left(\widehat{P}_{1,r}(x)-P_1(x)\right)\right\}+P_1(x)\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\right\}=O(r^{-1}).$$

(viii) Since it follows from assumptions (A2) and (A4) of Theorem 4.1 that

$$\begin{aligned}
\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2\right\} &\le \sqrt{\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^4\right\}\mathrm{E}\left\{\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^4\right\}}\\
&=\sqrt{O(r^{-2})O(r^{-2})}\\
&=O(r^{-2}),
\end{aligned}$$

we have by (vii) of this lemma and (D.6) that

$$\begin{aligned}
&\mathrm{E}\left((\widehat{\beta}_r-\beta)^2\widehat{P}_{1,r}^2(x)\right)\\
&=\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\left(\widehat{P}_{1,r}(x)-P_1(x)\right)^2\right\}+2P_1(x)\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\widehat{P}_{1,r}(x)\right\}-P_1^2(x)\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\right\}\\
&=O(r^{-1}).
\end{aligned}$$

(ix) By (iv) of this lemma and (D.6) we may show that

$$\begin{aligned}
\mathrm{E}\left|(\widehat{\beta}_r-\beta)\left(\widehat{P}_{2,r}^{cf}(x)-P_2^{cf}(x)\right)\right| &\le \sqrt{\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\right\}\mathrm{E}\left\{\left(\widehat{P}_{2,r}^{cf}(x)-P_2^{cf}(x)\right)^2\right\}}\\
&=\sqrt{O(r^{-1})O(r^{-1})}\\
&=O(r^{-1}),
\end{aligned}$$

and consequently, by using assumption (A1) of Theorem 4.1,

$$\mathrm{E}\left((\widehat{\beta}_r-\beta)\widehat{P}_{2,r}^{cf}(x)\right)=\mathrm{E}\left\{(\widehat{\beta}_r-\beta)\left(\widehat{P}_{2,r}^{cf}(x)-P_2^{cf}(x)\right)\right\}+P_2^{cf}(x)\mathrm{E}\left(\widehat{\beta}_r-\beta\right)=O(r^{-1}).$$

(x) Assumption (A2) of Theorem 4.1 and (iv) of this lemma imply that

$$\begin{aligned}
\mathrm{E}\left|(\widehat{\beta}_r-\beta)^2\left(\widehat{P}^{cf}_{2,r}(x)-P^{cf}_2(x)\right)\right| &\le \sqrt{\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^4\right\}\mathrm{E}\left\{\left(\widehat{P}^{cf}_{2,r}(x)-P^{cf}_2(x)\right)^2\right\}}\\
&=\sqrt{O(r^{-2})O(r^{-1})}\\
&=O(r^{-3/2}).
\end{aligned}$$

We therefore have from (D.6) that

$$\begin{aligned}
\mathrm{E}\left((\widehat{\beta}_r-\beta)^2\widehat{P}^{cf}_{2,r}(x)\right) &= \mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\left(\widehat{P}^{cf}_{2,r}(x)-P^{cf}_2(x)\right)\right\}+P^{cf}_2(x)\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\right\}\\
&=O(r^{-1}).
\end{aligned}$$

□

Lemma D.2. *From the assumptions of Theorem 4.1 it follows that*

(i) $\mathrm{E}\left\{(\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,\alpha}-z_\alpha)^2\right\}=O(r^{-1}+m^{-1}),$

(ii) $\mathrm{E}\left\{\left|\widehat{\beta}_r\beta^{-1}\widetilde{\xi}_{m,r,\alpha}-z_\alpha\right|^3\right\}=O(r^{-3/2}+m^{-3/2}).$

Proof.

(i) By the c_r-inequality we have

$$\begin{aligned}
&\mathrm{E}\left\{\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha}-z_\alpha\right)^2\right\}\\
&=\mathrm{E}\left\{\left(\frac{(\widehat{\beta}_r-\beta)}{\beta}\widetilde{\xi}_{m,r,\alpha}+(\widetilde{\xi}_{m,r,\alpha}-z_\alpha)\right)^2\right\}\\
&\le\frac{2}{\beta^2}\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2(\widetilde{\xi}_{m,r,\alpha})^2\right\}+2\mathrm{E}\left\{(\widetilde{\xi}_{m,r,\alpha}-z_\alpha)^2\right\}\\
&=\frac{2}{\beta^2}\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\left(z_\alpha+m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha)+m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right)^2\right\}\\
&\quad+2\mathrm{E}\left\{\left(m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha)+m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right)^2\right\}\\
&\le\frac{4z_\alpha^2}{\beta^2}\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\right\}+\frac{8}{\beta^2}m^{-1}\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\left(\widehat{P}^{cf}_{1,r}(z_\alpha)\right)^2\right\}+\frac{8}{\beta^2}m^{-2}\mathrm{E}\left\{(\widehat{\beta}_r-\beta)^2\left(\widehat{P}^{cf}_{2,r}(z_\alpha)\right)^2\right\}\\
&\quad+4m^{-1}\mathrm{E}\left\{\left(\widehat{P}^{cf}_{1,r}(z_\alpha)\right)^2\right\}+4m^{-2}\mathrm{E}\left\{\left(\widehat{P}^{cf}_{2,r}(z_\alpha)\right)^2\right\}.
\end{aligned}$$

Under the assumption of Theorem 4.1 that $\mathbf{W}_1$ has a sufficiently large number of finite moments, all the above expected values will be finite. It therefore follows from (D.6) that

$$\mathrm{E}\left\{\left(\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha}-z_\alpha\right)^2\right\}=O(r^{-1}+m^{-1}).$$

(ii) By the c_r-inequality we have

$$
\begin{aligned}
&\mathrm{E}\left(\left|\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha} - z_\alpha\right|^3\right)\\
&= \mathrm{E}\left(\left|\frac{(\widehat{\beta}_r - \beta)}{\beta}\widetilde{\xi}_{m,r,\alpha} + (\widetilde{\xi}_{m,r,\alpha} - z_\alpha)\right|^3\right)\\
&\le \frac{4}{\beta^3}\mathrm{E}\left(\left|\widehat{\beta}_r - \beta\right|^3 \left|\widetilde{\xi}_{m,r,\alpha}\right|^3\right) + 4\,\mathrm{E}\left(\left|\widetilde{\xi}_{m,r,\alpha} - z_\alpha\right|^3\right)\\
&= \frac{4}{\beta^3}\mathrm{E}\left(\left|\widehat{\beta}_r - \beta\right|^3 \left|z_\alpha + m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right|^3\right)\\
&\quad + 4\,\mathrm{E}\left(\left|m^{-1/2}\widehat{P}^{cf}_{1,r}(z_\alpha) + m^{-1}\widehat{P}^{cf}_{2,r}(z_\alpha)\right|^3\right)\\
&\le \frac{16|z_\alpha|^3}{\beta^3}\mathrm{E}\left(\left|\widehat{\beta}_r - \beta\right|^3\right) + \frac{64}{\beta^3}m^{-3/2}\,\mathrm{E}\left(\left|\widehat{\beta}_r - \beta\right|^3 \left|\widehat{P}^{cf}_{1,r}\right|^3\right) + \frac{64}{\beta^3}m^{-3}\,\mathrm{E}\left(\left|\widehat{\beta}_r - \beta\right|^3 \left|\widehat{P}^{cf}_{2,r}\right|^3\right)\\
&\quad + 16m^{-3/2}\,\mathrm{E}\left(\left|\widehat{P}^{cf}_{1,r}\right|^3\right) + 16m^{-3}\,\mathrm{E}\left(\left|\widehat{P}^{cf}_{2,r}\right|^3\right).
\end{aligned}
$$

By assumption (A2) of Theorem 4.1 it follows from Lemma 7.1 that

$$\mathrm{E}\left\{\left|\widehat{\beta}_r - \beta\right|^3\right\} = O(r^{-3/2}).$$

Under the assumption of Theorem 4.1 that $\mathbf{W}_1$ has a sufficiently large number of finite moments, all the above expected values will be finite. It therefore follows that

$$\mathrm{E}\left(\left|\frac{\widehat{\beta}_r}{\beta}\widetilde{\xi}_{m,r,\alpha} - z_\alpha\right|^3\right) = O(r^{-3/2} + m^{-3/2}). \qquad \square$$

D.2 Lemmas required in Chapter 5

This section contains a number of lemmas which are required to prove the theorems in Chapter 5. The notation used here is defined in that chapter.

Lemma D.3. *From the assumptions of Theorem 5.1 it follows that*

(i) $\mathrm{E}\left\{\widehat{Q}_{1,r}(x)\left(\widehat{Q}_{1,r}(x) - Q_1(x)\right)\right\} = O(r^{-1}),$

(ii) $\mathrm{E}\left\{\widehat{Q}_{2,r}(x)\left(\widehat{Q}_{1,r}(x) - Q_1(x)\right)\right\} = O(r^{-1}),$

(iii) $\mathrm{E}\left\{\widehat{Q}_{1,r}(x)\left(\widehat{Q}'_{1,r}(x) - Q'_1(x)\right)\right\} = O(r^{-1}),$

(iv) $\mathrm{E}\left\{\widehat{Q}_{1,r}(x)\left(\widehat{Q}''_{1,r}(x) - Q''_1(x)\right)\right\} = O(r^{-1}),$

(v) $\mathrm{E}\left\{\widehat{Q}_{1,r}(x)\left(\widehat{Q}'_{2,r}(x) - Q'_2(x)\right)\right\} = O(r^{-1}),$

(vi) $\mathrm{E}\left\{\widehat{Q}^2_{1,r}(x)\left(\widehat{Q}_{1,r}(x) - Q_1(x)\right)\right\} = O(r^{-1}),$

(vii) $\mathrm{E}\left\{\widehat{Q}^2_{1,r}(x)\left(\widehat{Q}'_{1,r}(x) - Q'_1(x)\right)\right\} = O(r^{-1}),$

(viii) $\mathrm{E}\left\{\widehat{Q}^2_{1,r}(x)\left(\widehat{Q}''_{1,r}(x)-Q''_1(x)\right)\right\}=O(r^{-1})$,

(ix) $\mathrm{E}\left\{\widehat{Q}_{1,r}(x)\widehat{Q}'_{1,r}(x)\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\right\}=O(r^{-1})$,

Proof.

(i) It follows from assumption (B2) of Theorem 5.1 and Lemma 7.1 that

$$\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^2\right\}=O(r^{-1}). \tag{D.7}$$

Therefore, using assumption (B1) of Theorem 5.1, we have

$$\mathrm{E}\left\{\widehat{Q}_{1,r}(x)\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\right\}=\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^2\right\}+Q_1(x)\mathrm{E}\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)=O(r^{-1}).$$

(ii) By the Cauchy-Schwarz inequality, it follows from assumption (B5) of Theorem 5.1 and (D.7) that

$$\begin{aligned}\mathrm{E}\left|\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\left(\widehat{Q}_{2,r}(x)-Q_2(x)\right)\right| &\le \sqrt{\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^2\right\}\mathrm{E}\left\{\left(\widehat{Q}_{2,r}(x)-Q_2(x)\right)^2\right\}}\\ &=\sqrt{O(r^{-1})O(r^{-1})}\\ &=O(r^{-1}).\end{aligned}$$

Then, from assumption (B1) of Theorem 5.1, we have

$$\begin{aligned}&\mathrm{E}\left\{\widehat{Q}_{2,r}(x)\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\right\}\\ &=\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\left(\widehat{Q}_{2,r}(x)-Q_2(x)\right)\right\}+Q_2(x)\mathrm{E}\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\\ &=O(r^{-1}).\end{aligned}$$

(iii) It follows from assumption (B3) of Theorem 5.1 and (D.7) that

$$\begin{aligned}\mathrm{E}\left|\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\left(\widehat{Q}'_{1,r}(x)-Q'_1(x)\right)\right| &\le \sqrt{\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^2\right\}\mathrm{E}\left\{\left(\widehat{Q}'_{1,r}(x)-Q'_1(x)\right)^2\right\}}\\ &=\sqrt{O(r^{-1})O(r^{-1})}\\ &=O(r^{-1}).\end{aligned} \tag{D.8}$$

Hence, also by assumption (B3) of Theorem 5.1, we have

$$\begin{aligned}&\mathrm{E}\left\{\widehat{Q}_{1,r}(x)\left(\widehat{Q}'_{1,r}(x)-Q'_1(x)\right)\right\}\\ &=\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\left(\widehat{Q}'_{1,r}(x)-Q'_1(x)\right)\right\}+Q_1(x)\mathrm{E}\left(\widehat{Q}'_{1,r}(x)-Q'_1(x)\right)\\ &=O(r^{-1}).\end{aligned}$$

(iv) It follows from assumption (B4) of Theorem 5.1 and (D.7) that

$$\begin{aligned}\mathrm{E}\left|\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\left(\widehat{Q}''_{1,r}(x)-Q''_1(x)\right)\right| &\le \sqrt{\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^2\right\}\mathrm{E}\left\{\left(\widehat{Q}''_{1,r}(x)-Q''_1(x)\right)^2\right\}}\\ &=\sqrt{O(r^{-1})O(r^{-1})}\\ &=O(r^{-1}).\end{aligned}$$

Hence, also by assumption (B4) of Theorem 5.1, we have

$$\begin{aligned}
&\mathrm{E}\left\{\widehat{Q}_{1,r}(x)\left(\widehat{Q}''_{1,r}(x)-Q''_1(x)\right)\right\}\\
&=\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\left(\widehat{Q}''_{1,r}(x)-Q''_1(x)\right)\right\}+Q_1(x)\mathrm{E}\left(\widehat{Q}''_{1,r}(x)-Q''_1(x)\right)\\
&=O(r^{-1}).
\end{aligned}$$

(v) By assumption (B6) of Theorem 5.1 and (D.7) we have that

$$\begin{aligned}
\mathrm{E}\left|\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\left(\widehat{Q}'_{2,r}(x)-Q'_2(x)\right)\right| &\le \sqrt{\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^2\right\}\mathrm{E}\left\{\left(\widehat{Q}'_{2,r}(x)-Q'_2(x)\right)^2\right\}}\\
&=\sqrt{O(r^{-1})O(r^{-1})}\\
&=O(r^{-1}).
\end{aligned}$$

By the same assumptions of Theorem 5.1 it follows that

$$\begin{aligned}
&\mathrm{E}\left\{\widehat{Q}_{1,r}(x)\left(\widehat{Q}'_{2,r}(x)-Q'_2(x)\right)\right\}\\
&=\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\left(\widehat{Q}'_{2,r}(x)-Q'_2(x)\right)\right\}+Q_1(x)\mathrm{E}\left(\widehat{Q}'_{2,r}(x)-Q'_2(x)\right)\\
&=O(r^{-1}).
\end{aligned}$$

(vi) Applying Lemma 7.1 to assumption (B2) of Theorem 5.1 we obtain

$$\left|\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^3\right\}\right|\le \mathrm{E}\left\{\left|\widehat{Q}_{1,r}(x)-Q_1(x)\right|^3\right\}=O(r^{-3/2}).$$

Now, noting that

$$\begin{aligned}
&\mathrm{E}\left\{\widehat{Q}^2_{1,r}(x)\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\right\}\\
&=\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^3\right\}+2Q_1(x)\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^2\right\}+Q_1^2(x)\mathrm{E}\left\{\widehat{Q}_{1,r}(x)-Q_1(x)\right\},
\end{aligned}$$

we have from (D.7) and assumption (B1) of Theorem 5.1 that

$$\mathrm{E}\left\{\widehat{Q}^2_{1,r}(x)\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\right\}=O(r^{-1}).$$

(vii) Notice that assumptions (B2) and (B3) of Theorem 5.1 imply that

$$\begin{aligned}
&\left|\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^2\left(\widehat{Q}'_{1,r}(x)-Q'_1(x)\right)\right\}\right|\\
&\le\sqrt{\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^4\right\}\mathrm{E}\left\{\left(\widehat{Q}'_{1,r}(x)-Q'_1(x)\right)^2\right\}}\\
&=\sqrt{O(r^{-2})O(r^{-1})}\\
&=O(r^{-3/2}). \qquad \text{(D.9)}
\end{aligned}$$

Now since we may write

$$\begin{aligned}
\mathrm{E}\left\{\widehat{Q}^2_{1,r}(x)\left(\widehat{Q}'_{1,r}(x)-Q'_1(x)\right)\right\}=&\ \mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^2\left(\widehat{Q}'_{1,r}(x)-Q'_1(x)\right)\right\}\\
&+2Q_1(x)\mathrm{E}\left\{\widehat{Q}_{1,r}(x)\left(\widehat{Q}'_{1,r}(x)-Q'_1(x)\right)\right\}\\
&-Q_1^2(x)\mathrm{E}\left\{\widehat{Q}'_{1,r}(x)-Q'_1(x)\right\},
\end{aligned}$$

it follows from result (iii) of this lemma and assumption (B3) of Theorem 5.1 that

$$\mathrm{E}\left\{\widehat{Q}^2_{1,r}(x)\left(\widehat{Q}'_{1,r}(x)-Q'_1(x)\right)\right\}=O(r^{-1}).$$

(viii) Notice that assumptions (B2) and (B4) of Theorem 5.1 imply that

$$\begin{aligned}
&\left|\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^2\left(\widehat{Q}''_{1,r}(x)-Q''_1(x)\right)\right\}\right| \\
&\le\sqrt{\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^4\right\}\mathrm{E}\left\{\left(\widehat{Q}''_{1,r}(x)-Q''_1(x)\right)^2\right\}} \\
&=\sqrt{O(r^{-2})O(r^{-1})} \\
&=O(r^{-3/2}).
\end{aligned}$$

Now since we may write

$$\begin{aligned}
\mathrm{E}\left\{\widehat{Q}^2_{1,r}(x)\left(\widehat{Q}''_{1,r}(x)-Q''_1(x)\right)\right\}=\;&\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^2\left(\widehat{Q}''_{1,r}(x)-Q''_1(x)\right)\right\} \\
&+2Q_1(x)\,\mathrm{E}\left\{\widehat{Q}_{1,r}(x)\left(\widehat{Q}''_{1,r}(x)-Q''_1(x)\right)\right\} \\
&-Q^2_1(x)\,\mathrm{E}\left\{\widehat{Q}''_{1,r}(x)-Q''_1(x)\right\},
\end{aligned}$$

it follows from result (iv) of this lemma and assumption (B4) of Theorem 5.1 that

$$\mathrm{E}\left\{\widehat{Q}^2_{1,r}(x)\left(\widehat{Q}''_{1,r}(x)-Q''_1(x)\right)\right\}=O(r^{-1}).$$

(ix) Note that

$$\begin{aligned}
&\mathrm{E}\left\{\widehat{Q}_{1,r}(x)\widehat{Q}'_{1,r}(x)\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\right\} \\
&=\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^2\left(\widehat{Q}'_{1,r}(x)-Q'_1(x)\right)\right\}+Q'_1(x)\,\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)^2\right\} \\
&\quad+Q_1(x)\,\mathrm{E}\left\{\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\left(\widehat{Q}'_{1,r}(x)-Q'_1(x)\right)\right\}+Q_1(x)Q'_1(x)\,\mathrm{E}\left\{\widehat{Q}_{1,r}(x)-Q_1(x)\right\}.
\end{aligned}$$

Therefore, by (D.7), (D.8), (D.9), and assumption (B4) of Theorem 5.1 we have that

$$\mathrm{E}\left\{\widehat{Q}_{1,r}(x)\widehat{Q}'_{1,r}(x)\left(\widehat{Q}_{1,r}(x)-Q_1(x)\right)\right\}=O(r^{-1}).$$

□

Lemma D.4. *From the assumptions of Theorem 5.1 it follows that*

(i) $\mathrm{E}\left\{\left(\widetilde{\eta}_{m,r,\alpha}-x\right)^2\right\}=O(m^{-1}),$

(ii) $\mathrm{E}\left\{\left|\widetilde{\eta}_{m,r,\alpha}-x\right|^3\right\}=O(m^{-3/2}),$

(iii) $\mathrm{E}\left\{\left(\widetilde{\eta}_{m,r,\alpha}-x\right)^4\right\}=O(m^{-2}).$

Proof.

(i) By the c_r-inequality we have

$$\begin{aligned}
&\mathrm{E}\left\{\left(\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right)^2\right\}\\
&=\mathrm{E}\left\{\left(m^{-1/2}\widehat{Q}^{cf}_{1,r}(z_\alpha)+m^{-1}\widehat{Q}^{cf}_{2,r}(z_\alpha)+m^{-3/2}\widehat{Q}^{cf}_{3,r}(z_\alpha)\right)^2\right\}\\
&\leq 2m^{-1}\,\mathrm{E}\left\{\left(\widehat{Q}^{cf}_{1,r}(z_\alpha)\right)^2\right\}+4m^{-2}\,\mathrm{E}\left\{\left(\widehat{Q}^{cf}_{2,r}(z_\alpha)\right)^2\right\}+4m^{-3}\,\mathrm{E}\left\{\left(\widehat{Q}^{cf}_{3,r}(z_\alpha)\right)^2\right\}.
\end{aligned}$$

Under the assumption of Theorem 5.1 that $\mathbf{W}_1$ has a sufficiently large number of finite moments, all the above expected values will be finite. It therefore follows that

$$\mathrm{E}\left\{\left(\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right)^2\right\}=O(m^{-1}).$$

(ii) By the c_r-inequality we have

$$\begin{aligned}
&\mathrm{E}\left\{\left|\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right|^3\right\}\\
&=\mathrm{E}\left\{\left|m^{-1/2}\widehat{Q}^{cf}_{1,r}(z_\alpha)+m^{-1}\widehat{Q}^{cf}_{2,r}(z_\alpha)+m^{-3/2}\widehat{Q}^{cf}_{3,r}(z_\alpha)\right|^3\right\}\\
&\leq 4m^{-3/2}\,\mathrm{E}\left\{\left|\widehat{Q}^{cf}_{1,r}(z_\alpha)\right|^3\right\}+16m^{-3}\,\mathrm{E}\left\{\left|\widehat{Q}^{cf}_{2,r}(z_\alpha)\right|^3\right\}+16m^{-9/2}\,\mathrm{E}\left\{\left|\widehat{Q}^{cf}_{3,r}(z_\alpha)\right|^3\right\}.
\end{aligned}$$

Under the assumption of Theorem 5.1 that $\mathbf{W}_1$ has a sufficiently large number of finite moments, all the above expected values will be finite. It therefore follows that

$$\mathrm{E}\left\{\left|\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right|^3\right\}=O(m^{-3/2}).$$

(iii) By the c_r-inequality we have

$$\begin{aligned}
&\mathrm{E}\left\{\left(\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right)^4\right\}\\
&=\mathrm{E}\left\{\left(m^{-1/2}\widehat{Q}^{cf}_{1,r}(z_\alpha)+m^{-1}\widehat{Q}^{cf}_{2,r}(z_\alpha)+m^{-3/2}\widehat{Q}^{cf}_{3,r}(z_\alpha)\right)^4\right\}\\
&\leq 8m^{-2}\,\mathrm{E}\left\{\left(\widehat{Q}^{cf}_{1,r}(z_\alpha)\right)^4\right\}+64m^{-4}\,\mathrm{E}\left\{\left(\widehat{Q}^{cf}_{2,r}(z_\alpha)\right)^4\right\}+64m^{-6}\,\mathrm{E}\left\{\left(\widehat{Q}^{cf}_{3,r}(z_\alpha)\right)^4\right\}.
\end{aligned}$$

Under the assumption of Theorem 5.1 that $\mathbf{W}_1$ has a sufficiently large number of finite moments, all the above expected values will be finite. It therefore follows that

$$\mathrm{E}\left\{\left(\widetilde{\eta}_{m,r,\alpha}-z_\alpha\right)^4\right\}=O(m^{-2}).\qquad\square$$

Bibliography

Beran, R. (1987). Prepivoting to reduce level error of confidence sets, *Biometrika* **74**: 151–173.

Bhattacharya, R. N. and Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion, *Ann. Statist.* **6**: 434–451.

Bickel, P. J. and Freedman, D. A. (1981). Some asymptotic theory for the bootstrap, *Ann. Statist.* **9**: 1196–1217.

Carpenter, J. and Bithell, J. (2000). Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians., *Stat. Med.* **19**: 1141–1164.

Chang, J. and Hall, P. (2015). Double-bootstrap methods that use a single double-bootstrap simulation, *Biometrika* **102**: 203–214.

Cheung, K. Y. and Lee, S. M. S. (2005). Variance estimation for sample quantiles using the m out of n bootstrap, *Ann. Inst. Statist. Math.* **57**: 279–290.

Chibishov, D. M. (1972). An asymptotic expansion for the distribution of a statistic admitting an asymptotic expansion, *Theor. Probab. Appl.* **17**: 620–630.

Chibishov, D. M. (1973). An asymptotic expansion for the distribution of sums of a special form with an application to minimum-contrast estimates, *Theor. Probab. Appl.* **18**: 649–661.

Chung, K.-H. and Lee, S. M. S. (2001). Optimal bootstrap sample size in construction of percentile confidence bounds, *Scand. J. Stat.* **28**: 225–239.

Chung, K. L. (1946). The approximate distribution of Student's statistic, *Ann. Math. Statist.* **17**: 442–465.

Cook, M. B. (1951). Bi-variate k-statistics and cumulants of their joint sampling distribution, *Biometrika* **38**: 179–195.

Cramér, H. (1928). On the composition of elementary errors, *Skand. Aktuarietidskr.* **11**: 13–74, 141–180.

Cramér, H. (1946). *Mathematical Methods of Statistics*, Princeton University Press, Princeton.

Davison, A. C. and Hinkley, D. V. (1997). *Bootstrap Methods and their Applications*, Cambridge.

DiCiccio, T. J. and Romano, J. P. (1988). A review of bootstrap confidence intervals (with discussions), *J. R. Statist. Soc. B* **50**: 338–354.

Dressel, P. L. (1940). Statistical seminvariants and their setimates with particular emphasis on their relation to algebraic invariants, *Ann. Math. Statist.* **11**: 33–57.

Edgeworth, F. Y. (1905). The law of error, *Proc. Cambridge Philos. Soc.* **20**: 36–65.

Efron, B. (1979). Bootstrap methods: another look at the jackknife, *Ann. Statist.* **7**: 1–26.

Efron, B. (1981). Nonparametric standard errors and confidence intervals, *Can. J. of Stat.* **9**: 139–158.

Efron, B. (1987). Better bootstrap confidence intervals, *J. Stat. Plan. Infer.* **82**: 171–185.

Efron, B. and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*, Chapman & Hall/CRC.

Finner, H. (1992). A generalization of Hölder's inequality and some probability inequalities, *Ann. Probab.* **20**: 1893–1901.

Finner, H. and Dickhaus, T. (2010). Edgeworth expansions and rates of convergence for normalized sums: Chung's 1946 method revisited, *Stat. Probabil. Lett.* **80**: 1875–1880.

Fisher, R. A. (1930). Moments and product moments of sampling distributions, *Proceedings of the London Mathematical Society*, pp. 199–238.

Giacomini, R., Politis, D. N. and White, H. (2013). A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators, *Econometric Theory* **29**: 567–589.

Giné, E., Götze, F. and Mason, D. (1997). When is the Student t-statistic asymptotically standard normal?, *Ann. Probab.* **25**: 1514–1531.

Hall, P. (1986). On the bootstrap and confidence intervals, *Ann. Stat.* **14**: 1431–1452.

Hall, P. (1987). Edgeworth expansion for Student's t statistic under minimal moment conditions, *Ann. Probab.* **15**: 920–931.

Hall, P. (1988). Theoretical comparison of bootstrap confidence intervals, *Ann. Statist.* **16**: 927–953.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*, Springer, New York.

Hall, P. and Martin, M. A. (1988). On bootstrap resampling and iteration, *Biometrika* **75**: 661–671.

Hall, P. and Wang, Q. (2004). Exact convergence rate and leading term in central limit theorem for Student's t statistc, *Ann. Probab.* **32**: 1419–1437.

Hsu, P. L. (1945). The approximate distributions of the mean and variance of a sample of independent variables, *Ann. Math. Statist.* **16**: 1–29.

Hurt, J. (1976). Asymptotic expansions of functions of statistics, *Apl. Mat.* **21**: 444–456.

Janson, S. (2009). Probability asymptotics: notes on notation. Accessed 05/11/2014.
URL: *http://arxiv.org/abs/1108.3924*

Kendall, M. G. (1940). The derivation of multivariate sampling formulae from univariate formulae by symbolic operation, *Ann. Eugen.* **10**: 392–402.

Loève, M. (1977). *Probability Theory, Vol. 1*, fourth edn, Springer-Verlag, New York.

Loh, W. Y. (1987). Calibrating confidence coefficients, *J. Amer. Statist. Assoc.* **82**: 155–162.

Loh, W. Y. (1988). Discussion of "Theoretical comparison of bootstrap confidence intervals" by P. Hall, *Ann. Statist.* **16**: 972–976.

Loh, W. Y. (1991). Bootstrap calibration for confidence interval construction and selection, *Statist. Sinica* **1**: 479–495.

Ogasawara, H. (2006). Asymptotic expansion of the sample correlation coefficient under nonnormality, *Comput. Stat. Data Anal.* **50**: 891–910.

Sargan, J. D. (1975). Gram-Charlier approximation applied to t ratios of k-class estimators, *Econometrica* **43**: 327–346.

Sargan, J. D. (1976). Econometric estimators and the edgeworth approximation, *Econometrica* **44**: 421–448.

Schenker, N. (1985). Qualms about bootstrap confidence intervals, *Am. Stat.* **80**: 360–361.

Shao, J. (1994). Bootstrap sample size in nonregular cases, *Proc. Amer. Math. Soc.* **122**: 1251–1262.

Shao, J. and Tu, D. (1995). *The Jackknife and Bootstrap*, Springer.

Stuart, A. and Ord, J. K. (1994). *Kendall's Advanced Theory of Statistics*, Vol. 1, John Wiley & Sons.

Swanepoel, J. W. H. (1986). A note on proving that the (modified) bootstrap works, *Commun. Stat. – Theor. M.* **15**: 3193–3203.

von Bahr, B. and Esseen, C.-G. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 \le r \le 2$, *Ann. Math. Statist.* **36**: 299–303.

Wallace, D. L. (1958). Asymptotic approximations to distributions, *Ann. Math. Statist.* **29**: 635–654.