Exploring Functional Asymptotic Confidence Intervals for a Population Mean

by

Ekaterina Tuzov

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of

Master of Science

Department of Statistics
University of Manitoba
Winnipeg

Abstract

We take a Student process that is based on independent copies of a random variable X and has trajectories in the function space $D[0, 1]$. As a consequence of a functional central limit theorem for this process, with X in the domain of attraction of the normal law, we consider convergence in distribution of several functionals of this process and derive respective asymptotic confidence intervals for the mean of X. We explore the expected lengths and finite-sample coverage probabilities of these confidence intervals and the one obtained from the asymptotic normality of the Student t-statistic, thus concluding some alternatives to the latter confidence interval that are shorter and/or have at least as high coverage probabilities.

Acknowledgments

Foremost, I would like to express my deep gratitude to my advisor, Dr. Yuliya V. Martsynyuk, for all the efforts and hard work she invested in this thesis. Her guidance, patience, and knowledge were invaluable to me.

Besides my advisor, I would like to extend my gratitude to other members of my examining committee, Dr. Abba B. Gumel and Dr. Liqun Wang, for their insightful input. I must also thank Dave Gabrielson for teaching me to understand, and even like, LaTeX.

I greatly appreciate the generous financial support I received from my advisor and from the Department of Statistics of the University of Manitoba. Many thanks to the Faculty of Graduate Studies of the University of Manitoba for the Student Travel Award and the GETS program funds.

Last but not least, my friends and family, your encouragement and cheerfulness kept me going when my legs were tired. I could have never gotten as far without you.

Dedication

To caffeine, without which the pursuit of knowledge would be futile.

Contents

List of Tables

List of Figures

Notations and Abbreviations

$:=$	Equality by definition
$\Longleftrightarrow$	If and only if
$[x]$	Greatest integer that does not exceed x (floor function)
$\mathbb{R}$	The space of real numbers
$\mathbb{1}_A$	Indicator function of set A
$D[0,1]$	The space of real-valued functions on $[0,1]$ that are right-continuous and have left-hand limits
$E(X)$	Expected value of random variable X
$Var(X)$	Variance of random variable X
$X \stackrel{\mathcal{D}}{=} Y$	Random variables X and Y have the same distribution
$\stackrel{\mathcal{D}}{\rightarrow}$	Convergence in distribution
$\mathrm{Exp}(\lambda)$	Exponential distribution with mean $1/\lambda$
$N(\mu,\sigma^2)$	Normal distribution with mean μ and variance σ^2
$\Phi(x)$	Standard normal cumulative distribution function
$z_{\alpha/2}$	$1-\alpha/2$ quantile of the standard normal distribution, that is, $\Phi(z_{\alpha/2}) = 1-\alpha/2$, for $\alpha \in (0,1)$
$\{W(t),\, 0 \le t \le 1\}$	Standard real-valued Wiener process (Brownian motion) on $[0,1]$
i.i.d.	Independent and identically distributed
DAN	Domain of attraction of the normal law
CLT	Central limit theorem
FCLT	Functional central limit theorem
FACI	Functional asymptotic confidence interval

Frequently Used Definitions

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

$$s_n := \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2}{n-1}}$$

$$T_n(X_1, \ldots, X_n) := \frac{\sum_{i=1}^{n} X_i}{s_n\sqrt{n}} \qquad \text{(Student } t\text{-statistic)}$$

$$T_n^t(X_1, \ldots, X_n) := \frac{\sum_{i=1}^{[nt]} X_i}{s_n\sqrt{n}}, \quad 0 \le t \le 1 \qquad \text{(Student process)}$$

$$I_0 := \left[\overline{X}_n - \frac{z_{\alpha/2}\, s_n}{\sqrt{n}}, \overline{X}_n + \frac{z_{\alpha/2}\, s_n}{\sqrt{n}}\right]$$

$$r_i := \frac{E(\text{length of } I_i)}{E(\text{length of } I_0)}, \quad i = \overline{1,5}$$

(Expected lengths ratio of confidence intervals I_i and I_0)

$$\widehat{r}_i := \frac{\sum_{k=1}^{10{,}000}(\text{length of } I_i \text{ for sample } k)/10,000}{\sum_{k=1}^{10{,}000}(\text{length of } I_0 \text{ for sample } k)/10,000}, \quad i = \overline{1,5}$$

(Empirical expected lengths ratio of confidence intervals I_i and I_0)

$$\widehat{CP}_i := \frac{\sum_{k=1}^{10{,}000} \mathbb{1}_{\{\mu \in I_i, \text{ for sample } k\}}}{10,000}, \quad i = \overline{0,5} \qquad \text{(Empirical coverage probability of } I_i\text{)}$$

$$G_m(\mu) := \int_0^1 \left(T_n^t(X_1 - \mu, \ldots, X_n - \mu)\right)^m dt = \frac{1}{n}\sum_{k=1}^{n-1}\left(\frac{\sum_{i=1}^{k}(X_i - \mu)}{s_n\sqrt{n}}\right)^m$$

Chapter 1

Introduction

Let $\{X, X_1, X_2, \dots\}$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean $\mu = E(X)$. Consider the Student t-statistic

$$T_n(X_1, \dots, X_n) := \frac{\sum_{i=1}^n X_i}{s_n \sqrt{n}}, \tag{1.1}$$

where

$$s_n := \sqrt{\frac{\sum_{i=1}^n (X_i - \overline{X}_n)^2}{n-1}} \quad \text{and} \quad \overline{X}_n := \frac{\sum_{i=1}^n X_i}{n}.$$

Gosset (1908) (also known as "Student") concluded the exact distribution of $T_n(X_1 - \mu, \dots, X_n - \mu)/\sqrt{n-1}$ for a random sample drawn from a normally distributed population with unknown mean μ and variance and used it to derive confidence intervals for μ for small samples.

Giné et al. (1997) showed that if $\{X, X_1, X_2, \dots\}$ are i.i.d. random variables and μ is a constant, then

$$\begin{aligned} &T_n(X_1 - \mu, \dots, X_n - \mu) \xrightarrow[n\to\infty]{\mathcal{D}} N(0,1) \quad \Longleftrightarrow \\ &X \in \text{DAN} \text{ and } E(X) = \mu, \end{aligned} \tag{1.2}$$

where DAN stands for the domain of attraction of the normal law, and X is said to belong to DAN if there exist constants a_n and $b_n > 0$ such that

$$\frac{\sum_{i=1}^{n} X_i - a_n}{b_n} \xrightarrow[n\to\infty]{\mathcal{D}} N(0,1). \tag{1.3}$$

Remark 1. If $X \in \text{DAN}$, then $E|X|^\nu < \infty$ for all $\nu \in (0,2)$, and a_n can be taken as $nE(X)$, while $b_n = \sqrt{n}\ell_X(n)$, where $\ell_X(n)$ is a slowly varying function at infinity defined by the distribution of X, that is $\ell_X(az)/\ell_X(z) \to 1$, as $z \to \infty$, for any $a > 0$. Also, $\ell_X(n) = \sqrt{Var(X)} > 0$, if $Var(X) < \infty$, and $\ell_X(n)$ is a nondecreasing function that converges to ∞, as $n \to \infty$, if $Var(X) = \infty$. There are various useful characterizations of DAN. For example, due to Lévy (1937), we have the following one:

$$X \in \text{DAN} \quad \Longleftrightarrow \quad \lim_{x\to\infty} \frac{x^2 P(|X| > x)}{E(X^2 \mathbb{1}_{\{|X|\le x\}})} = 0. \tag{1.4}$$

Remark 2. If $0 < Var(X) < \infty$, the classical central limit theorem (CLT) can be used to show that condition (1.3) is satisfied, and therefore $X \in \text{DAN}$. Thus, all the distributions with finite positive variances are in DAN. As to some examples of the distributions in DAN that have infinite variances, a typical one is the Pareto(1,2) distribution with the probability density function

$$f(x) = \begin{cases} \dfrac{2}{x^3}, & x \ge 1, \\ 0, & \text{otherwise.} \end{cases} \tag{1.5}$$

Indeed, since for $X \overset{\mathcal{D}}{=}$ Pareto(1,2) we have

$$\lim_{x\to\infty} \frac{x^2\, P(|X|>x)}{E(X^2 \mathbb{1}_{\{|X|\le x\}})} = \lim_{x\to\infty} \frac{x^2 \displaystyle\int_x^{\infty} \frac{2}{y^3}\, dy}{\displaystyle\int_1^{x} y^2 \frac{2}{y^3}\, dy} = \lim_{x\to\infty} \frac{-x^2 \cdot y^{-2}\Big|_x^{\infty}}{2\ln y\Big|_1^{x}} = \lim_{x\to\infty} \frac{1}{2\ln x} = 0,$$

then $X \in$ DAN in view of (1.4). Moreover, it was shown in Example 1 of Martsynyuk (2013) that the corresponding b_n in (1.3) can be taken as $\sqrt{n\log n}$. Additionally, one can consider the distribution that is a discrete version of Pareto(1,2) and has the following probability mass function:

$$g(x) = \frac{c}{x^3}, \quad x = 1, 2, \ldots \quad , \tag{1.6}$$

where $c = (\sum_{x=1}^{\infty} x^{-3})^{-1} \approx 0.8319$ (cf. p. 811 of Abramowitz and Stegun (1964) for example). We will refer to it as the discrete Pareto distribution hereafter. Clearly, for a random variable X from this distribution, $E(X) < \infty$, while $Var(X) = \infty$. Using (1.4), we establish that $X \in$ DAN:

$$0 \le \frac{x^2\, P(|X|>x)}{E(X^2 \mathbb{1}_{\{|X|\le x\}})} = \frac{x^2 \sum_{y=[x]+1}^{\infty} \dfrac{c}{y^3}}{\sum_{y=1}^{[x]} y^2 \dfrac{c}{y^3}} = \frac{x^2 \sum_{y=1}^{\infty} ([x]+y)^{-3}}{\sum_{y=1}^{[x]} y^{-1}}$$

$$\le \frac{x^2 \int_0^{\infty} ([x]+y)^{-3}\, dy}{\int_1^{[x]} y^{-1}\, dy} = \frac{x^2\, [x]^{-2}\, \dfrac{1}{2}}{\ln [x]} \xrightarrow[x\to\infty]{} 0.$$

Coming back to (1.2), we note that its $\Leftarrow$ part can be viewed as an extension of the Studentized classical CLT to DAN. It can be used to construct the following completely data-based $1-\alpha$ asymptotic confidence interval for a typically unknown μ:

$$I_0 := \left[\overline{X}_n - \frac{z_{\alpha/2}\, s_n}{\sqrt{n}}, \overline{X}_n + \frac{z_{\alpha/2}\, s_n}{\sqrt{n}}\right], \tag{1.7}$$

where $\alpha \in (0,1)$ is fixed and $z_{\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution, that is, $P(|N(0,1)| > z_{\alpha/2}) = \alpha$.

In view of the Student t-statistic $T_n(X_1, \ldots, X_n)$ of (1.1), one can define a Student process in $D[0,1]$, the space of real-valued functions on $[0,1]$ that are right-continuous and have left-hand limits, as follows:

$$T_n^t(X_1, \ldots, X_n) := \frac{\sum_{i=1}^{[nt]} X_i}{s_n\sqrt{n}}, \quad 0 \le t \le 1, \tag{1.8}$$

where $\sum_{i=1}^{0} X_i := 0$. Equivalently,

$$T_n^t(X_1, \ldots, X_n) = \begin{cases} 0, & 0 \le t < \frac{1}{n}, \\ \dfrac{X_1}{s_n\sqrt{n}}, & \frac{1}{n} \le t < \frac{2}{n}, \\ \quad\vdots & \\ \dfrac{X_1 + X_2 + \cdots + X_{n-1}}{s_n\sqrt{n}}, & \frac{n-1}{n} \le t < 1, \\ \dfrac{X_1 + X_2 + \cdots + X_n}{s_n\sqrt{n}}, & t = 1. \end{cases} \tag{1.9}$$

In other words, $T_n^t(X_1, \ldots, X_n)$ is a random step function on $[0,1]$ that coincides at $t=1$ with the Student t-statistic.

Csörgő et al. (2003), among other things, concluded a special convergence, known as a functional central limit theorem (FCLT), for the so-called self-normalized partial sums process and hence also for the Student process of (1.8). Let $\{W(t), 0 \le t \le 1\}$ denote a standard Wiener process (Brownian motion). A simpler form of this FCLT for $T_n^t(X_1, \dots, X_n)$ reads as follows:

$$X \in \mathrm{DAN} \quad \text{and} \quad E(X) = \mu \quad \Longleftrightarrow$$

$$h(T_n^t(X_1 - \mu, \dots, X_n - \mu)) \xrightarrow[n\to\infty]{\mathcal{D}} h(W(t))$$

$$\text{for all functionals } h : D[0,1] \to \mathbb{R} \text{ that are } \mathscr{D}\text{-measurable and } \rho\text{-continuous,} \tag{1.10}$$

where $\mathscr{D}$ is the σ-field of subsets of $D[0,1]$ that is generated by the finite-dimensional subsets of $D[0,1]$ and ρ is the sup-norm metric in $D[0,1]$.

Remark 3. We recall that a functional $h : D[0,1] \to \mathbb{R}$ is called sup-norm continuous if for any $f_0(t) \in D[0,1]$ and $\varepsilon > 0$, there exists $\delta > 0$, such that for any $f(t) \in D[0,1]$ satisfying $\sup_{0\le t\le 1} |f(t) - f_0(t)| < \delta$ we have $|h(f(t)) - h(f_0(t))| < \varepsilon$. The notion of $\mathscr{D}$-measurability for a functional is beyond the scope of the present work. We just note that most functionals on $D[0,1]$, including the ones studied here (cf. (1.11)–(1.15)), are $\mathscr{D}$-measurable.

The FCLT of (1.10) for the Student process contains the CLT of (1.2) for the Student t-statistic. Indeed, reading the convergence in (1.10) in the special case of the ρ-continuous functional $h(\cdot)$ such that $h(f(t)) = f(1)$, for any function $f(t) \in D[0,1]$, we have $T_n^1(X_1 - \mu, \dots, X_n - \mu) \xrightarrow{\mathcal{D}} W(1)$, or, equivalently, $T_n(X_1 - \mu, \dots, X_n - \mu) \xrightarrow{\mathcal{D}} N(0,1)$, as $n \to \infty$.

Martsynyuk (2009a) considered a special analog of the Student process of (1.8) for independent, but not necessarily identically distributed, random variables with a common mean μ that either satisfy the so-called Lindeberg's condition or are symmetric around μ. Using this analog, Martsynyuk (2009a) refined previously known FCLT's for Studentized partial sums processes that are based on such random variables, thus establishing completely data-based versions of these FCLT's. The work in Martsynyuk (2009a) was largely motivated by the need to improve on the applicability of previous FCLT's and to construct completely data-based asymptotic confidence intervals for μ from FCLT's, for which the term functional asymptotic confidence intervals (FACI's) was given in this work. Using the newly obtained FCLT's, Martsynyuk (2009a) constructed three such FACI's for μ, by considering convergence in distribution of three special functionals of the Student process: $\sup_{0\le t\le 1}|\cdot|$, $\int_0^1(\cdot)\,dt$, and the functional that returns the value of a function at a fixed point $t_0 \in (0,1]$. The studies of Martsynyuk (2009a) had been motivated in part by the problems that arose in Martsynyuk (2009b) in connection to establishing FACI's for the slope in the context of linear errors-in-variables models.

Martsynyuk (2009a) posed an open problem of investigating the individual and comparative performance of the three obtained FACI's for μ. Inspired by this problem, in this thesis, we construct several FACI's for an unknown mean μ of a population from DAN and study their comparative performance.

We consider several special cases of the convergence in distribution in (1.10), where the functional $h(\cdot)$ is replaced with one of the functionals $h_i(\cdot)$, $i=\overline{1,5}$, such that for any function $f(t) \in D[0,1]$,

$$h_1(f(t)) = f(t_0), \quad \text{for a fixed } t_0 \in (0,1], \tag{1.11}$$

$$h_2(f(t)) = \sup_{0 \le t \le 1} |f(t)|, \tag{1.12}$$

$$h_3(f(t)) = \sup_{0 \le t \le 1} f(t), \tag{1.13}$$

$$h_4(f(t)) = \int_0^1 f^m(t)\, dt, \quad \text{for } m = 1, 2, 3, 4, \text{and } 8, \tag{1.14}$$

and

$$h_5(f(t)) = a_1 f(1) + a_2 \int_0^1 f(t)\, dt, \quad \text{for some constants } a_1, a_2 \in \mathbb{R}. \tag{1.15}$$

One can verify that all these functionals are sup-norm continuous. As mentioned before, the functionals $h_1(\cdot)$, $h_2(\cdot)$, and $h_4(\cdot)$ with $m = 1$ were used in Martsynyuk (2009a) to construct FACI's for a common mean of independent, but not necessarily identically distributed, random variables.

We note in passing that Erdős and Kac (1946) established convergence in distribution as in (1.10), with the partial sums process $\{\sum_{i=1}^{[nt]} X_i/\sqrt{n}, 0 \le t \le 1\}$ replacing $T_n^t(X_1 - \mu, \ldots, X_n - \mu)$, for the functionals $h_2(\cdot)$, $h_3(\cdot)$, $h_4(\cdot)$ with $m = 2$, and $\int_0^1 |\cdot|\, dt$ on $D[0, 1]$, assuming that $X_1, X_2, \ldots$ are i.i.d. random variables with mean zero and variance one. They also concluded the analytic forms or the Laplace transforms of the corresponding limiting distributions, and noted that the i.i.d. random variable condition for these results can be replaced by assuming only that X_i's are such that the CLT is applicable. The significance of the work of Erdős and Kac (1946) is that the authors established examples of one of the first invariance principles, the laws that do not depend on the underlying distribution structure, or on the distribution of X_i's in this case. Moreover, the idea of the invariance inspired their proofs. They wrote:

> "The proofs of all these theorems follow the same pattern. It is first proved that the limiting distribution exists and is independent of the distribution of the X's; then the distribution of the X's is chosen conveniently so that the limiting distribution can be calculated explicitly."

We use a similar invariance approach when tabulating quantiles of $\int_0^1 W^m(t)\,dt$ for $m = 3, 4$, and 8 in section 2.4.

In Chapter 2, for each of the functionals h_i in (1.11)–(1.15), we first consider the convergence in distribution in the FCLT of (1.10) with $h = h_i$ and use it to construct a FACI for μ. Then, we evaluate the performance of the FACI by comparing its expected length and finite-sample coverage probability to those of the commonly used asymptotic confidence interval I_0 of (1.7) that is based on the asymptotic normality of the Student t-statistic. The FACI's are presented in the same order as the functionals h_i, so that simpler FACI's are studied first.

In Chapter 3, we review the performances of the FACI's obtained in Chapter 2 and conclude which of the FACI's present reasonable alternatives to, and overall improvement upon, the interval I_0.

Following Chapter 3, we include Appendix A with a somewhat lengthy proof of an auxiliary result, and Appendix B with the R syntax of all the simulation results presented throughout the thesis.

Chapter 2

Main results

Let $\{X, X_1, X_2, \dots\}$ be a sequence of i.i.d. random variables in DAN with a population mean $\mu = E(X)$. In this chapter, as a consequence of the FCLT of (1.10) for the Student process $\{T_n^t(X_1 - \mu, \dots, X_n - \mu),\ 0 \le t \le 1\}$ of (1.8), we consider the following convergence in distribution:

$$h_i(T_n^t(X_1 - \mu, \dots, X_n - \mu)) \xrightarrow[n\to\infty]{\mathcal{D}} h_i(W(t)), \tag{2.1}$$

where $h_i : D[0,1] \to \mathbb{R}$, $i = \overline{1,5}$, are the sup-norm continuous functionals in (1.11) – (1.15) and $\{W(t),\ 0 \le t \le 1\}$ is a standard Wiener process. Based on (2.1), for each h_i, we construct a FACI for μ, denoted by I_i hereafter, by using already available quantiles, or ones that we tabulate ourselves, of the limiting random variable $h_i(W(t))$ and solving certain inequalities for μ.

The goodness of each FACI is assessed by how its expected length and finite-sample coverage probability compare to those of the commonly used asymptotic confidence interval I_0 of (1.7) that is based on the asymptotic normality of the Student t-statistic.

The finite-sample coverage probabilities of I_i's and I_0 are approximated by their respective empirical coverage probabilities. For this purpose, we take 10,000 samples of size n from a distribution of X, and for each sample compute the bounds of I_i and those of I_0. Then, we calculate the empirical coverage probability of I_i,

$$\widehat{CP}_i := \frac{\sum_{k=1}^{10{,}000} \mathbb{1}_{\{\mu \in I_i, \text{ for sample } k\}}}{10,000}, \quad i = \overline{1,5}, \tag{2.2}$$

and that of I_0, namely,

$$\widehat{CP}_0 := \frac{\sum_{k=1}^{10{,}000} \mathbb{1}_{\{\mu \in I_0, \text{ for sample } k\}}}{10,000}, \tag{2.3}$$

from the same 10,000 samples, where $\mathbb{1}_A$ is the indicator function of set A. Five different distributions are considered for $X:\ N(0,1)$, Exp(1), Poisson(3), Pareto(1,2), and the discrete Pareto as given in (1.6). These distributions vary with respect to their means, skewness, and variances. We also use four sample sizes, $n = 50, 100, 500,$ and 1000, and consider three confidence levels, $1-\alpha = 0.9, 0.95,$ and 0.98, which are often used in practice. Overall, we calculate the empirical coverage probabilities in 60 different scenarios, taking 10,000 samples for each, and report the corresponding values of the pair $(\widehat{CP}_i, \widehat{CP}_0)$ in tables.

To facilitate the comparison of the expected lengths of the intervals I_i to I_0, we study the ratio

$$r_i := \frac{E(\text{length of } I_i)}{E(\text{length of } I_0)}, \tag{2.4}$$

$i = \overline{1,5}$. Except for r_1, r_4 that corresponds to h_4 of (1.14) with $m = 1$, and r_5, theoretical evaluations of r_i's do not appear to be feasible. Therefore, for each of the aforementioned 60 combinations of the distribution of X, sample size, and confidence level, we use the same 10,000 samples that are generated for computing $\widehat{CP_i}$ and approximate r_i with the ratio of the empirical expected lengths of I_i and I_0 as follows:

$$\widehat{r_i} := \frac{\sum_{k=1}^{10,000}(\text{length of } I_i \text{ for sample } k)/10,000}{\sum_{k=1}^{10,000}(\text{length of } I_0 \text{ for sample } k)/10,000}. \tag{2.5}$$

Values of $\widehat{r_i}$ are reported in the same table as those of $(\widehat{CP_i}, \widehat{CP_0})$.

The syntax for all the presented simulations is found in Appendix B.

2.1 Functional Asymptotic Confidence Interval (FACI) based on $h_1(\cdot)$

Let t_0 be any fixed number from $(0, 1]$. From (2.1) with $h_1(\cdot)$ of (1.11), we have the following convergence in distribution:

$$T_n^{t_0}(X_1 - \mu, \ldots, X_n - \mu) \xrightarrow[n\to\infty]{\mathcal{D}} W(t_0). \tag{2.6}$$

Due to the known property of the Wiener process that $W(t_0) \stackrel{\mathcal{D}}{=} N(0, t_0)$, (2.6) is equivalent to:

$$\frac{\sum_{i=1}^{[nt_0]}(X_i - \mu)}{s_n\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{D}} N(0, t_0). \tag{2.7}$$

Set $0 < \alpha < 1$ and let $P(|N(0, 1)| > z_{\alpha/2}) = \alpha$. It follows from (2.7) that

$$P\left(-z_{\alpha/2}\sqrt{t_0} \le \frac{\sum_{i=1}^{[nt_0]}(X_i - \mu)}{s_n\sqrt{n}} \le z_{\alpha/2}\sqrt{t_0}\right) \xrightarrow[n\to\infty]{} 1 - \alpha, \tag{2.8}$$

or, equivalently, that

$$P\left(\frac{\sum_{i=1}^{[nt_0]} X_i - z_{\alpha/2} s_n \sqrt{nt_0}}{[nt_0]} \leq \mu \leq \frac{\sum_{i=1}^{[nt_0]} X_i + z_{\alpha/2} s_n \sqrt{nt_0}}{[nt_0]}\right) \xrightarrow[n\to\infty]{} 1-\alpha. \quad (2.9)$$

Thus, we obtain the following $1-\alpha$ FACI for μ:

$$I_1 := \left[\frac{\sum_{i=1}^{[nt_0]} X_i - z_{\alpha/2}\, s_n \sqrt{nt_0}}{[nt_0]}, \frac{\sum_{i=1}^{[nt_0]} X_i + z_{\alpha/2}\, s_n \sqrt{nt_0}}{[nt_0]}\right]. \quad (2.10)$$

In view of Theorem 9.3.2 in Casella and Berger (2002) and the limiting distribution $N(0, t_0)$ in (2.7), selecting symmetric cut-off points $-z_{\alpha/2}\sqrt{t_0}$ and $z_{\alpha/2}\sqrt{t_0}$ in (2.8) leads to the shortest FACI.

Evaluation of I_1

The center of the interval I_1 equals $\sum_{i=1}^{[nt_0]} X_i/[nt_0]$, which is an unbiased estimator of μ. The ratio of the expected lengths of I_1 and I_0 is:

$$r_1 = \frac{2z_{\alpha/2}E(s_n)\sqrt{nt_0}/[nt_0]}{2z_{\alpha/2}E(s_n)/\sqrt{n}} = \frac{1}{\sqrt{t_0}}\frac{nt_0}{[nt_0]} \geq 1. \quad (2.11)$$

We see that $\lim_{n\to\infty} r_1 = 1/\sqrt{t_0}$ and that r_1 is shortest when $t_0 = 1$, i.e., when I_1 coincides with I_0 of (1.7).

It can be shown theoretically that the finite-sample coverage probability of I_1 is greater or equal to that of I_0 when $X_i \stackrel{\mathcal{D}}{=} N(0,1)$, $1 \leq i \leq n$. However, it is challenging to compare the two probabilities for the other four distributions.

Therefore, we calculate the empirical coverage probabilities $\widehat{CP_1}$ of I_1 and $\widehat{CP_0}$ of I_0, as in (2.2) and (2.3), respectively, for the example of $t_0 = 0.9$.

		$\mathbf{1-\alpha}$		
Distribution	$\boldsymbol{n}$	0.9	0.95	0.98
$N(0,1)$	50	(0.892,0.893)	(0.943,0.942)	(0.977,0.976)
	100	(0.899,0.897)	(0.944,0.944)	(0.980,0.979)
	500	(0.902,0.904)	(0.952,0.951)	(0.979,0.979)
	1000	(0.901,0.900)	(0.947,0.947)	(0.980,0.980)
Exp(1)	50	(0.879,0.879)	(0.930,0.928)	(0.962,0.961)
	100	(0.892,0.893)	(0.942,0.941)	(0.970,0.968)
	500	(0.896,0.897)	(0.948,0.949)	(0.977,0.980)
	1000	(0.900,0.902)	(0.952,0.951)	(0.977,0.976)
Pareto(1,2)	50	(0.761,0.755)	(0.819,0.806)	(0.859,0.850)
	100	(0.790,0.788)	(0.843,0.836)	(0.888,0.878)
	500	(0.834,0.830)	(0.891,0.886)	(0.931,0.922)
	1000	(0.843,0.841)	(0.898,0.896)	(0.938,0.935)
Poisson(3)	50	(0.893,0.892)	(0.941,0.944)	(0.975,0.976)
	100	(0.895,0.892)	(0.950,0.949)	(0.978,0.980)
	500	(0.901,0.898)	(0.949,0.951)	(0.978,0.978)
	1000	(0.901,0.896)	(0.949,0.949)	(0.981,0.982)
Discrete Pareto	50	(0.753,0.750)	(0.796,0.787)	(0.836,0.828)
	100	(0.784,0.780)	(0.837,0.830)	(0.876,0.867)
	500	(0.842,0.833)	(0.880,0.875)	(0.921,0.914)
	1000	(0.847,0.841)	(0.900,0.892)	(0.933,0.930)

Table 2.1: Empirical coverage probabilities: $(\widehat{CP_1}, \widehat{CP_0})$, $t_0 = 0.9$

The difference $\widehat{CP_1} - \widehat{CP_0}$ is mostly positive and ranges from -0.3% to 1.3%. For $t_0 = 0.9$ and the sample sizes $n = 50, 100, 500,$ and 1000, we have r_1=1.0541, that is I_1 is 1.0541 times longer than I_0.

2.2 FACI based on $h_2(\cdot)$

It follows from (2.1) with $h_2(\cdot)$ as in (1.12) that

$$\sup_{0\le t\le 1}\left|T_n^t(X_1-\mu,\ldots,X_n-\mu)\right| \xrightarrow[n\to\infty]{\mathcal{D}} \sup_{0\le t\le 1}|W(t)|, \tag{2.12}$$

which, in view of (1.9), is equivalent to

$$\max_{1\le k\le n}\frac{\left|\sum_{i=1}^{k}(X_i-\mu)\right|}{s_n\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{D}} \sup_{0\le t\le 1}|W(t)|. \tag{2.13}$$

We set $0<\alpha<1$ and define b such that $P(\sup_{0\le t\le 1}|W(t)|\le b)=1-\alpha$. Tabulated values of b can be found in Csörgő and Horváth (1985). It follows from (2.13) that

$$P\left(\max_{1\le k\le n}\frac{\left|\sum_{i=1}^{k}(X_i-\mu)\right|}{s_n\sqrt{n}}\le b\right) \xrightarrow[n\to\infty]{} 1-\alpha. \tag{2.14}$$

Having (2.14) and that

$$\begin{aligned}
P\left(\max_{1\le k\le n}\frac{\left|\sum_{i=1}^{k}(X_i-\mu)\right|}{s_n\sqrt{n}}\le b\right) &= P\left(\bigcap_{k=1}^{n}\left\{\frac{\left|\sum_{i=1}^{k}(X_i-\mu)\right|}{s_n\sqrt{n}}\le b\right\}\right)\\
=&P\left(\bigcap_{k=1}^{n}\left\{-b\le\frac{\sum_{i=1}^{k}X_i-k\mu}{s_n\sqrt{n}}\le b\right\}\right) \qquad (2.15)\\
=&P\left(\bigcap_{k=1}^{n}\left\{\frac{\sum_{i=1}^{k}X_i-b\,s_n\sqrt{n}}{k}\le\mu\le\frac{\sum_{i=1}^{k}X_i+b\,s_n\sqrt{n}}{k}\right\}\right)
\end{aligned}$$

leads to the following $1-\alpha$ FACI for μ:

$$\bigcap_{k=1}^{n}\left[\frac{\sum_{i=1}^{k}X_i-b\,s_n\sqrt{n}}{k},\frac{\sum_{i=1}^{k}X_i+b\,s_n\sqrt{n}}{k}\right]. \tag{2.16}$$

An alternative formulation of the FACI in (2.16) is given by

$$I_2 := \left[\max_{1 \le k \le n} \left(\frac{\sum_{i=1}^{k} X_i - b\, s_n \sqrt{n}}{k} \right), \min_{1 \le k \le n} \left(\frac{\sum_{i=1}^{k} X_i + b\, s_n \sqrt{n}}{k} \right) \right]. \quad (2.17)$$

Evaluation of I_2

The bounds of the FACI I_2 are random variables whose exact distributions are difficult to obtain. After being unable to solve this problem on the example of normally distributed X_i's, we also believe that the exact distributions of interest are likely dependent on the underlying distribution of X_i. Thus, we failed to obtain the closed form expression for the expected length of I_2 by direct calculations. We also do not know whether the expected length of I_2 can be successfully computed by some other methods, such as conditioning for example. Consequentially, we compare the expected lengths of I_2 and I_0 numerically.

Simulated values of $\widehat{r_2}$ as in (2.5), the ratio of the empirical expected lengths of I_2 and I_0, and values of $\widehat{CP_2}$ and $\widehat{CP_0}$ as in (2.2) and (2.3), the empirical coverage probabilities of I_2 and I_0, are summarized in the following table. These variables are evaluated in 60 scenarios with various sample sizes, confidence levels, and distributions, as discussed on page 10, where each time the value of $\widehat{r_2}$ is reported in front of the $(\widehat{CP_2}, \widehat{CP_0})$ values.

Distribution	$\boldsymbol{n}$	$\boldsymbol{1-\alpha}$ 0.9	0.95	0.98
$N(0,1)$	50	1.086 (0.912,0.893)	1.070 (0.955,0.943)	1.056 (0.981,0.976)
	100	1.074 (0.913,0.902)	1.059 (0.955,0.948)	1.048 (0.982,0.980)
	500	1.053 (0.909,0.901)	1.042 (0.953,0.950)	1.035 (0.983,0.981)
	1000	1.048 (0.906,0.903)	1.038 (0.953,0.952)	1.032 (0.979,0.978)
Exp(1)	50	1.093 (0.896,0.877)	1.074 (0.937,0.926)	1.060 (0.967,0.962)
	100	1.077 (0.905,0.892)	1.062 (0.946,0.936)	1.050 (0.975,0.973)
	500	1.056 (0.910,0.903)	1.046 (0.950,0.948)	1.036 (0.979,0.977)
	1000	1.051 (0.904,0.896)	1.040 (0.949,0.946)	1.033 (0.980,0.979)
Pareto(1,2)	50	1.114 (0.773,0.758)	1.093 (0.824,0.810)	1.070 (0.862,0.852)
	100	1.105 (0.808,0.791)	1.083 (0.854,0.843)	1.067 (0.885,0.876)
	500	1.089 (0.844,0.829)	1.073 (0.894,0.887)	1.057 (0.925,0.924)
	1000	1.087 (0.855,0.843)	1.067 (0.902,0.896)	1.056 (0.935,0.932)
Poisson(3)	50	1.087 (0.912,0.893)	1.070 (0.953,0.942)	1.055 (0.978,0.973)
	100	1.075 (0.912,0.897)	1.058 (0.954,0.947)	1.047 (0.981,0.978)
	500	1.053 (0.903,0.893)	1.043 (0.955,0.953)	1.035 (0.981,0.982)
	1000	1.050 (0.909,0.904)	1.039 (0.951,0.948)	1.032 (0.981,0.980)
Discrete Pareto	50	1.124 (0.762,0.751)	1.094 (0.796,0.781)	1.074 (0.839,0.829)
	100	1.113 (0.784,0.770)	1.089 (0.833,0.825)	1.069 (0.863,0.861)
	500	1.092 (0.838,0.823)	1.073 (0.884,0.876)	1.057 (0.924,0.922)
	1000	1.088 (0.859,0.845)	1.068 (0.894,0.887)	1.055 (0.935,0.931)

Table 2.2: Empirical expected lengths ratio and coverage probabilities: $\widehat{r}_2\,(\widehat{CP}_2, \widehat{CP}_0)$

The $\widehat{r}_2$ values in Table 2.2 vary from 1.032 to 1.124, indicating that I_2 is longer than I_0 on average. We note that the $\widehat{r}_2$'s are higher for X_i's from the Pareto(1,2) and discrete Pareto distributions. This suggests that the infinite variance and/or skewness of such X_i's amplifies the expected length of I_2 more than it amplifies the expected length of I_0. We also observe that for each of the distributions in Table 2.2, as n or $1-\alpha$ increases, $\widehat{r}_2$ decreases. Finally, the empirical coverage probabilities of I_2 are, except for one case, higher than those of I_0 by 0.1%–1.9%, with larger differences seen for smaller $1-\alpha$ levels and smaller sample sizes.

2.3 FACI based on $h_3(\cdot)$

We now consider (2.1) with $h_3(\cdot)$ of (1.13):

$$\sup_{0\le t\le 1} T_n^t(X_1-\mu,\ldots,X_n-\mu) \xrightarrow[n\to\infty]{\mathcal{D}} \sup_{0\le t\le 1} W(t).$$

Via the definition of the Student process in (1.9), this is equivalent to

$$\max_{1\le k\le n} \frac{\sum_{i=1}^{k}(X_i-\mu)}{s_n\sqrt{n}} \xrightarrow[n\to\infty]{\mathcal{D}} \sup_{0\le t\le 1} W(t). \tag{2.18}$$

The cumulative distribution function of $\sup_{0\le t\le 1} W(t)$ is well-known to be related to that of $N(0,1)$, denoted by $\Phi(\cdot)$, as follows:

$$P\left(\sup_{0\le t\le 1} W(t)\le y\right) = \begin{cases} 2\Phi(y)-1, & y\ge 0,\\ 0, & \text{otherwise.}\end{cases} \tag{2.19}$$

Let $0<\alpha<1$, and define $a,b\ge 0$ such that $P\left(a\le \sup_{0\le t\le 1} W(t)\le b\right)=1-\alpha$. In view of (2.19), the relationship between a and b is

$$b=\Phi^{-1}\left(\frac{1-\alpha}{2}+\Phi(a)\right). \tag{2.20}$$

It follows from (2.18) that

$$P\left(a\le \max_{1\le k\le n}\frac{\sum_{i=1}^{k}(X_i-\mu)}{s_n\sqrt{n}}\le b\right) \xrightarrow[n\to\infty]{} 1-\alpha, \tag{2.21}$$

where

$$
\begin{aligned}
&P\left(a \le \max_{1\le k\le n} \frac{\sum_{i=1}^{k}(X_i-\mu)}{s_n\sqrt{n}} \le b\right)\\
&= P\left(\left\{a \le \max_{1\le k\le n} \frac{\sum_{i=1}^{k}(X_i-\mu)}{s_n\sqrt{n}}\right\} \bigcap \left\{\max_{1\le k\le n} \frac{\sum_{i=1}^{k}(X_i-\mu)}{s_n\sqrt{n}} \le b\right\}\right)\\
&= P\left(\left(\bigcup_{k=1}^{n}\left\{a \le \frac{\sum_{i=1}^{k}(X_i-\mu)}{s_n\sqrt{n}}\right\}\right) \bigcap \left(\bigcap_{k=1}^{n}\left\{\frac{\sum_{i=1}^{k}(X_i-\mu)}{s_n\sqrt{n}} \le b\right\}\right)\right)\\
&= P\left(\left(\bigcup_{k=1}^{n}\left\{\mu \le \frac{\sum_{i=1}^{k}X_i - as_n\sqrt{n}}{k}\right\}\right) \bigcap \left(\bigcap_{k=1}^{n}\left\{\frac{\sum_{i=1}^{k}X_i - bs_n\sqrt{n}}{k} \le \mu\right\}\right)\right)\\
&= P\left(\left\{\mu \le \max_{1\le k\le n} \frac{\sum_{i=1}^{k}X_i - as_n\sqrt{n}}{k}\right\} \bigcap \left\{\max_{1\le k\le n} \frac{\sum_{i=1}^{k}X_i - bs_n\sqrt{n}}{k} \le \mu\right\}\right)\\
&= P\left(\max_{1\le k\le n} \frac{\sum_{i=1}^{k}X_i - bs_n\sqrt{n}}{k} \le \mu \le \max_{1\le k\le n} \frac{\sum_{i=1}^{k}X_i - as_n\sqrt{n}}{k}\right).
\end{aligned}
\tag{2.22}
$$

Hence, we obtain the following $1-\alpha$ FACI for μ from (2.21):

$$
I_3 := \left[\max_{1\le k\le n}\left(\frac{\sum_{i=1}^{k}X_i - b\,s_n\sqrt{n}}{k}\right), \max_{1\le k\le n}\left(\frac{\sum_{i=1}^{k}X_i - a\,s_n\sqrt{n}}{k}\right)\right]. \tag{2.23}
$$

Evaluation of I_3

The interval I_3 has a similar form to that of I_2 given in (2.17). On page 15, we discussed the difficulty of obtaining a closed form expression for the expected length of I_2. Due to the same reasons, we are also unable to find a closed form expression for the expected length of I_3. Therefore, we compare the expected lengths of

I_3 and I_0 numerically, via $\widehat{r}_3$ of (2.5). To do so, we first need to select a and b according to (2.20), by choosing a first for example. With many possibilities for a, we would naturally prefer to choose a such that the expected length of I_3 is minimized. However, without having a closed form expression for the expected length of I_3, we can only estimate the desired value of a numerically as follows (the corresponding syntax is provided in part B.4 of Appendix B). We take 10,000 random samples of size n from a distribution of X_i's and calculate the average length of I_3 for each of 50 different values of a that form a uniform subdivision of the interval $(0, \Phi^{-1}(\frac{1+\alpha}{2}))$ (the respective values of b are calculated from (2.20)). Then we select the value of a which produces the shortest average length of I_3. Such a value of a turns out not to be the same for different distributions of X_i's ($N(0,1)$, Exp(1), Poisson(3), Pareto(1,2) and discrete Pareto), confidence levels ($1-\alpha = 0.9, 0.95$, and 0.98), and sample sizes ($n = 50, 100, 500$, and 1000).

In the following table, we present the values of $\widehat{r}_3$ and $(\widehat{CP}_3, \widehat{CP}_0)$, where $\widehat{CP}_3$ and $\widehat{CP}_0$ are as in (2.2) and (2.3), respectively.

Distribution	$\boldsymbol{n}$	$\mathbf{1-\alpha}$ 0.9	0.95	0.98
$N(0,1)$	50	1.416 (0.833,0.892)	1.439 (0.882,0.946)	1.376 (0.900,0.975)
	100	1.560 (0.852,0.900)	1.681 (0.904,0.947)	1.720 (0.923,0.978)
	500	1.778 (0.881,0.898)	2.205 (0.929,0.948)	2.656 (0.960,0.979)
	1000	1.850 (0.884,0.899)	2.341 (0.935,0.951)	3.176 (0.966,0.982)
Exp(1)	50	1.576 (0.820,0.879)	1.555 (0.859,0.934)	1.485 (0.880,0.965)
	100	1.781 (0.843,0.891)	1.855 (0.884,0.943)	1.817 (0.907,0.973)
	500	2.112 (0.877,0.899)	2.632 (0.918,0.948)	2.995 (0.948,0.977)
	1000	2.154 (0.879,0.895)	2.844 (0.931,0.952)	3.638 (0.958,0.979)
Pareto(1,2)	50	1.598 (0.728,0.759)	1.507 (0.761,0.805)	1.358 (0.789,0.854)
	100	1.866 (0.745,0.783)	1.812 (0.797,0.833)	1.636 (0.830,0.879)
	500	2.578 (0.794,0.829)	2.441 (0.846,0.882)	2.616 (0.888,0.918)
	1000	2.748 (0.805,0.844)	2.776 (0.858,0.897)	3.104 (0.900,0.935)
Poisson(3)	50	1.494 (0.821,0.897)	1.493 (0.871,0.940)	1.403 (0.872,0.974)
	100	1.634 (0.859,0.899)	1.739 (0.909,0.944)	1.747 (0.910,0.977)
	500	1.861 (0.876,0.900)	2.325 (0.923,0.953)	2.841 (0.961,0.980)
	1000	1.907 (0.879,0.895)	2.507 (0.930,0.950)	3.350 (0.958,0.981)
Discrete Pareto	50	1.614 (0.720,0.752)	1.505 (0.746,0.780)	1.383 (0.761,0.831)
	100	1.923 (0.753,0.785)	1.833 (0.793,0.827)	1.670 (0.813,0.865)
	500	2.442 (0.791,0.825)	2.731 (0.841,0.874)	2.569 (0.880,0.917)
	1000	2.848 (0.809,0.840)	2.845 (0.850,0.887)	3.020 (0.902,0.931)

Table 2.3: Empirical expected lengths ratio and coverage probabilities: $\widehat{r}_3\,(\widehat{CP}_3, \widehat{CP}_0)$

The values of $\widehat{r}_3$ in Table 2.3 range from 1.358 to 3.638. There is a lot of variability in the $\widehat{r}_3$ values, more than for any other FACI we explore in this thesis. We observe that as the sample size increases, the ratio $\widehat{r}_3$ also increases. Also, the empirical coverage probabilities of I_3 are lower than those of I_0 by 1.5% to 10.2%. Overall, I_3 is not a very desirable confidence interval in terms of its expected length and coverage probability.

2.4 FACI's based on $h_4(\cdot)$

Convergence in (2.1) with $h_4(\cdot)$ of (1.14) reads as

$$\int_0^1 (T_n^t(X_1-\mu,\ldots,X_n-\mu))^m\,dt \xrightarrow[n\to\infty]{\mathcal{D}} \int_0^1 W^m(t)\,dt, \quad \text{for } m=1,2,3,4, \text{and } 8, \tag{2.24}$$

or, equivalently, via (1.9), as

$$G_m(\mu) := \frac{1}{n}\sum_{k=1}^{n-1}\left(\frac{\sum_{i=1}^{k}(X_i-\mu)}{s_n\sqrt{n}}\right)^m \xrightarrow[n\to\infty]{\mathcal{D}} \int_0^1 W^m(t)\,dt, \quad \text{for } m=1,2,3,4, \text{and } 8. \tag{2.25}$$

We note that for $m=3,4,$ and 8, the exact or tabulated distribution function of the limiting random variable $\int_0^1 W^m(t)\,dt$ seems to be unavailable in the literature. Inspired by the invariance approach of Erdős and Kac (1946) (cf. p. 8), we use (2.25) and estimate quantiles of the distribution of $\int_0^1 W^m(t)\,dt$ by those of the empirical distribution of $G_m(\mu)$, which is based on 500,000 simulated values of the random variable $G_m(\mu)$, where each value is computed from 10,000 independent X_i's having Bernoulli(0.5) distribution. We note that the quantiles of the empirical distribution of $G_m(\mu)$ are calculated with the *quantile* function in R. The tabulated quantiles of $\int_0^1 W^m(t)\,dt$ for $m=3,4$ and 8 are presented respectively in Tables 2.7, 2.9, and 2.11 of this section, while the syntax for their computation is provided in part B.3 of Appendix B.

Set $0 < \alpha < 1$ and a and b such that $P\left(a \leq \int_0^1 W^m(t)\,dt \ \leq b\right) = 1 - \alpha$. It follows from (2.25) that

$$P\left(a \leq G_m(\mu) \leq b\right) \xrightarrow[n\to\infty]{} 1 - \alpha. \tag{2.26}$$

The derivations of FACI's for μ from (2.26) with different values of m follow similar lines and depend on whether m is even or odd, as summarized next.

Case I: odd m ($m = 1$ and 3)

Constructing a $1-\alpha$ FACI for μ from (2.26) amounts to finding $\{\mu : a \leq G_m(\mu) \leq b\}$. As a random function of μ, $G_m(\mu)$ is a polynomial of power m and therefore has m roots. Disregarding the event when all X_i's are equal to μ, we notice that

$$\frac{d}{d\mu}G_m(\mu) = \frac{m}{n(s_n\sqrt{n})^m}\sum_{k=1}^{n-1}(-k)\left(\sum_{i=1}^{k}(X_i - \mu)\right)^{m-1} < 0, \tag{2.27}$$

indicating that the polynomial $G_m(\mu)$ is decreasing and that the equations $G_m(\mu) = a$ and $G_m(\mu) = b$ have exactly one real solution each. Therefore, the lower and upper bounds of the FACI we get from (2.26) are the real-valued solutions of $G_m(\mu) = b$ and $G_m(\mu) = a$, respectively, as illustrated in the next figure on the example of $m = 3$.

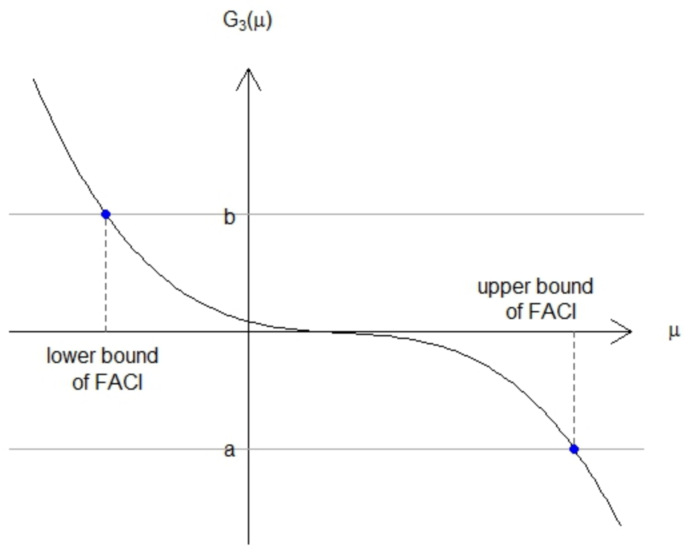

Figure 2.1: Illustration of FACI construction from (2.26) with $m = 3$

For m= 1 and 3, we choose "equal tail" cut-off points a and b in (2.26) such that $P(\int_0^1 W^m(t)\,dt \le a) = P(b \le \int_0^1 W^m(t)\,dt) = \alpha/2$. This choice of a and b is shown to result in the shortest FACI when $m = 1$, but for $m = 3$, these cut-off values are chosen simply for convenience.

Case II: even m (m=2, 4, and 8)

When m is even, $G_m(\mu)$ of (2.25) and $\int_0^1 W^m(t)\,dt$ are nonnegative random variables, and by choosing $a = 0$ in (2.26), we have:

$$P\left(G_m(\mu) \le b\right) \xrightarrow[n\to\infty]{} 1 - \alpha. \tag{2.28}$$

It is possible, but cumbersome, to work with $a > 0$ and evaluate numerically the choice of such a that leads to a shortest FACI for μ. Thus, we do not include any of our work in this regard here.

To construct a $1-\alpha$ FACI for μ from (2.28), we need to find $\{\mu : G_m(\mu) \le b\}$. On disregarding the event when all X_i's are equal to μ, we have

$$\frac{d^2}{d\mu^2} G_m(\mu) = \frac{m(m-1)}{n(s_n\sqrt{n})^m} \sum_{k=1}^{n-1} k^2 \left(\sum_{i=1}^{k} (X_i - \mu) \right)^{m-2} > 0, \tag{2.29}$$

which indicates that the polynomial $G_m(\mu)$ has a convex shape, and thus the equation $G_m(\mu) = b$ has at most two real solutions. Hence, we get the FACI for μ with the lower and upper bounds that are respectively the smaller and larger real solutions of $G_m(\mu) = b$, as illustrated in the figure below for the case of $m = 4$.

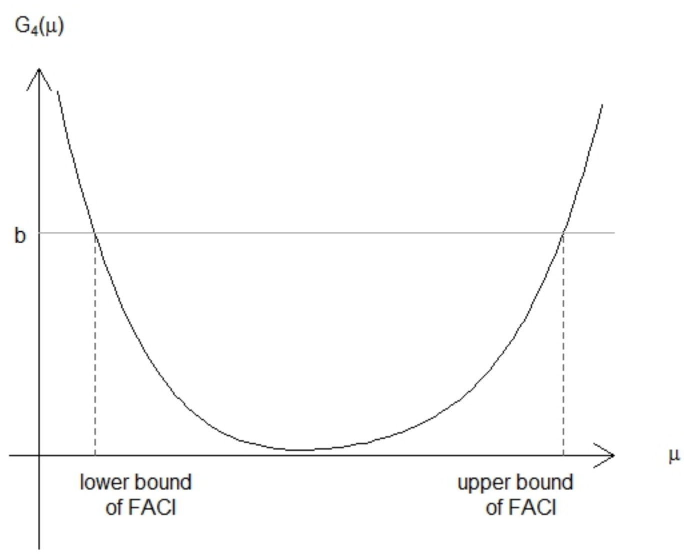

Figure 2.2: Illustration of FACI construction from (2.28) with $m = 4$

2.4.1 $m = 1$

By using integration by parts and Theorem 5.1 with $g(x) = t - x$ in Chapter 8 of Karlin and Taylor (1998), it can be shown that $\int_0^1 W(t)\,dt \overset{\mathcal{D}}{=} N(0, 1/3)$. Therefore, (2.26) with $m = 1$ and $b = -a$ reads as

$$P\left(\frac{-z_{\alpha/2}}{\sqrt{3}} \leq \frac{1}{n}\sum_{k=1}^{n-1} \frac{\sum_{i=1}^{k}(X_i - \mu)}{s_n\sqrt{n}} \leq \frac{z_{\alpha/2}}{\sqrt{3}}\right) \xrightarrow[n\to\infty]{} 1 - \alpha, \tag{2.30}$$

or, equivalently, as

$$P\left(\frac{2}{n-1}\left(\sum_{k=1}^{n-1}\sum_{i=1}^{k}\frac{X_i}{n} - z_{\alpha/2}\frac{s_n\sqrt{n}}{\sqrt{3}}\right) \leq \mu \leq \frac{2}{n-1}\left(\sum_{k=1}^{n-1}\sum_{i=1}^{k}\frac{X_i}{n} + z_{\alpha/2}\frac{s_n\sqrt{n}}{\sqrt{3}}\right)\right) \to 1-\alpha, \tag{2.31}$$

as $n \to \infty$, where $P(|N(0,1)| > z_{\alpha/2}) = \alpha$. Thus, we obtain the following $1 - \alpha$ FACI for μ:

$$I_4 := \left[\frac{2}{n-1}\left(\sum_{k=1}^{n-1}\sum_{i=1}^{k}\frac{X_i}{n} - z_{\alpha/2}\frac{s_n\sqrt{n}}{\sqrt{3}}\right), \frac{2}{n-1}\left(\sum_{k=1}^{n-1}\sum_{i=1}^{k}\frac{X_i}{n} + z_{\alpha/2}\frac{s_n\sqrt{n}}{\sqrt{3}}\right)\right]. \tag{2.32}$$

We note in passing that (2.26) with $m = 1$ leads to the FACI for μ with a length of $4(b-a)\dfrac{s_n\sqrt{n}}{n-1}$. In view of Theorem 9.3.2 in Casella and Berger (2002) and having an $N(0, 1/3)$ limiting distribution for $G_1(\mu)$ in (2.25), this length is the shortest when $b = -a = z_{\alpha/2}/\sqrt{3}$.

Evaluation of I_4, $m = 1$

At the center of the interval we have

$$\frac{2}{n-1}\sum_{k=1}^{n-1}\sum_{i=1}^{k}\frac{X_i}{n},$$

which is an unbiased estimator of the population mean μ. The ratio r_4 of (2.4) of the expected lengths of I_4 and I_0 is

$$r_4 = \frac{\dfrac{4\, z_{\alpha/2}}{n-1} \cdot \dfrac{E(s_n)\sqrt{n}}{\sqrt{3}}}{2\, z_{\alpha/2}\dfrac{E(s_n)}{\sqrt{n}}} = \frac{2n}{\sqrt{3}(n-1)}. \tag{2.33}$$

It is a monotonically decreasing function of n, and $\lim_{n\to\infty} r_4 \approx 1.155$. We conclude that I_4 is longer than I_0 on average, by a factor of approximately 1.155 for large n.

In the following table, we present finite-sample values of r_4 and $(\widehat{CP}_4, \widehat{CP}_0)$, the empirical coverage probabilities of I_4 and I_0.

Distribution	$\boldsymbol{n}$	$\boldsymbol{r_4}$	$\boldsymbol{1-\alpha}$ 0.9	0.95	0.98
$N(0,1)$	50	1.178266	(0.901,0.892)	(0.951,0.949)	(0.977,0.974)
	100	1.166364	(0.902,0.900)	(0.949,0.946)	(0.978,0.979)
	500	1.157015	(0.895,0.897)	(0.952,0.951)	(0.980,0.980)
	1000	1.155856	(0.899,0.898)	(0.951,0.947)	(0.980,0.977)
Exp(1)	50	1.178266	(0.881,0.877)	(0.936,0.930)	(0.969,0.959)
	100	1.166364	(0.899,0.888)	(0.942,0.936)	(0.975,0.972)
	500	1.157015	(0.897,0.898)	(0.950,0.949)	(0.976,0.977)
	1000	1.155856	(0.896,0.896)	(0.950,0.951)	(0.982,0.978)
Pareto(1,2)	50	1.178266	(0.779,0.761)	(0.841,0.816)	(0.885,0.857)
	100	1.166364	(0.807,0.791)	(0.856,0.829)	(0.904,0.881)
	500	1.157015	(0.846,0.835)	(0.894,0.882)	(0.933,0.918)
	1000	1.155856	(0.859,0.851)	(0.907,0.893)	(0.945,0.935)
Poisson(3)	50	1.178266	(0.898,0.891)	(0.950,0.946)	(0.978,0.975)
	100	1.166364	(0.898,0.893)	(0.948,0.946)	(0.979,0.976)
	500	1.157015	(0.899,0.899)	(0.951,0.951)	(0.982,0.982)
	1000	1.155856	(0.900,0.901)	(0.949,0.951)	(0.978,0.979)
Discrete Pareto	50	1.178266	(0.765,0.751)	(0.817,0.790)	(0.855,0.825)
	100	1.166364	(0.791,0.777)	(0.845,0.821)	(0.893,0.869)
	500	1.157015	(0.837,0.826)	(0.895,0.877)	(0.935,0.920)
	1000	1.155856	(0.848,0.841)	(0.899,0.888)	(0.944,0.932)

Table 2.4: Expected lengths ratio and empirical coverage probabilities: r_4 and $(\widehat{CP}_4, \widehat{CP}_0)$, $m = 1$

From Table 2.4, we see from the r_4 values that I_4 is 1.156 to 1.178 times longer than I_0. The coverage probability of I_4 is almost always higher than that of I_0. The difference $\widehat{CP}_4 - \widehat{CP}_0$ is more significant for samples from the two Pareto distributions, where it ranges from 0.7% to 3%, with the larger values observed for smaller samples and higher confidence levels.

2.4.2 $\boldsymbol{m = 2}$

When $m = 2$, (2.28) reads as

$$P\left(\frac{1}{n}\sum_{k=1}^{n-1}\left(\frac{\sum_{i=1}^{k}(X_i-\mu)}{s_n\sqrt{n}}\right)^2 \le b\right) \xrightarrow[n\to\infty]{} 1-\alpha. \tag{2.34}$$

Tabulated values of b can be found in Csörgő and Horváth (1981). For the probability in (2.34), we have

$$\begin{aligned}
&P\left(\frac{1}{n}\sum_{k=1}^{n-1}\left(\frac{\sum_{i=1}^{k}(X_i-\mu)}{s_n\sqrt{n}}\right)^2 \le b\right)\\
=&P\left(\sum_{k=1}^{n-1}\left[\left(\sum_{i=1}^{k}X_i\right)^2 - 2k\mu\sum_{i=1}^{k}X_i + k^2\mu^2\right] \le b\,n^2{s_n}^2\right)\\
=&P\left(c_n\,\mu^2 - 2\mu\sum_{k=1}^{n-1}k\sum_{i=1}^{k}X_i\; + \sum_{k=1}^{n-1}\left(\sum_{i=1}^{k}X_i\right)^2 - b\,n^2{s_n}^2 \le 0\right)\\
=&P\left(\left\{\mu \ge \frac{\sum_{k=1}^{n-1}k\sum_{i=1}^{k}X_i - \sqrt{(\sum_{k=1}^{n-1}k\sum_{i=1}^{k}X_i)^2 - c_n\left(\sum_{k=1}^{n-1}(\sum_{i=1}^{k}X_i)^2 - b\,n^2{s_n}^2\right)}}{c_n}\right\}\right.\\
&\left.\bigcap\left\{\mu \le \frac{\sum_{k=1}^{n-1}k\sum_{i=1}^{k}X_i + \sqrt{(\sum_{k=1}^{n-1}k\sum_{i=1}^{k}X_i)^2 - c_n\left(\sum_{k=1}^{n-1}(\sum_{i=1}^{k}X_i)^2 - b\,n^2{s_n}^2\right)}}{c_n}\right\}\right),
\end{aligned} \tag{2.35}$$

where $c_n := \sum_{k=1}^{n-1} k^2 = \dfrac{n(n-1)(2n-1)}{6}$. Hence, we obtain the following $1-\alpha$ FACI for μ from (2.34):

$$I_4 := \left[\frac{\sum_{k=1}^{n-1} k \sum_{i=1}^{k} X_i \mp \sqrt{(\sum_{k=1}^{n-1} k \sum_{i=1}^{k} X_i)^2 - c_n \left(\sum_{k=1}^{n-1} (\sum_{i=1}^{k} X_i)^2 - b\, n^2 {s_n}^2 \right)}}{c_n} \right]. \tag{2.36}$$

Evaluation of I_4, $m = 2$

The center of I_4,

$$\frac{\sum_{k=1}^{n-1} k \sum_{i=1}^{k} X_i}{c_n},$$

is an unbiased estimator for μ. The length of I_4 is a random variable given by

$$\frac{2}{c_n} \sqrt{\left(\sum_{k=1}^{n-1} k \sum_{i=1}^{k} X_i \right)^2 - c_n \left(\sum_{k=1}^{n-1} \left(\sum_{i=1}^{k} X_i \right)^2 - b\, n^2 {s_n}^2 \right)}. \tag{2.37}$$

Theoretical evaluation of the expectation of (2.37) is complicated by the presence of the square root. Therefore, the expectation of (2.37) is approximated numerically, and we study $\widehat{r_4}$ of (2.5) instead of r_4 of (2.4).

The values of $\widehat{r_4}$ and those of the empirical coverage probabilities $\widehat{CP_4}$ and $\widehat{CP_0}$ are summarized next. They are evaluated in 60 scenarios of different sample sizes, confidence levels, and distributions, as discussed on page 10.

Distribution	$\boldsymbol{n}$	$\boldsymbol{1-\alpha}$ 0.9	0.95	0.98
$N(0,1)$	50	1.119 (0.893,0.887)	1.120 (0.948,0.946)	1.119 (0.977,0.977)
	100	1.110 (0.900,0.899)	1.110 (0.951,0.945)	1.110 (0.980,0.978)
	500	1.103 (0.896,0.901)	1.103 (0.948,0.949)	1.104 (0.979,0.978)
	1000	1.102 (0.895,0.899)	1.103 (0.950,0.953)	1.102 (0.979,0.981)
Exp(1)	50	1.119 (0.882,0.874)	1.119 (0.935,0.928)	1.119 (0.969,0.963)
	100	1.109 (0.897,0.894)	1.110 (0.942,0.938)	1.110 (0.973,0.971)
	500	1.103 (0.904,0.899)	1.103 (0.950,0.945)	1.103 (0.981,0.979)
	1000	1.102 (0.895,0.895)	1.102 (0.947,0.946)	1.102 (0.980,0.978)
Pareto(1,2)	50	1.120 (0.775,0.758)	1.119 (0.825,0.807)	1.119 (0.871,0.849)
	100	1.111 (0.800,0.790)	1.110 (0.854,0.840)	1.110 (0.897,0.882)
	500	1.104 (0.845,0.837)	1.104 (0.893,0.885)	1.103 (0.938,0.926)
	1000	1.102 (0.849,0.846)	1.103 (0.911,0.900)	1.103 (0.940,0.933)
Poisson(3)	50	1.119 (0.900,0.895)	1.119 (0.946,0.946)	1.119 (0.977,0.975)
	100	1.110 (0.895,0.893)	1.111 (0.951,0.952)	1.110 (0.977,0.973)
	500	1.103 (0.904,0.907)	1.103 (0.950,0.949)	1.103 (0.980,0.981)
	1000	1.101 (0.898,0.900)	1.102 (0.945,0.948)	1.103 (0.984,0.985)
Discrete Pareto	50	1.120 (0.762,0.750)	1.120 (0.813,0.791)	1.119 (0.847,0.823)
	100	1.110 (0.788,0.776)	1.110 (0.842,0.825)	1.110 (0.881,0.866)
	500	1.103 (0.841,0.834)	1.104 (0.887,0.876)	1.103 (0.932,0.920)
	1000	1.102 (0.842,0.838)	1.103 (0.901,0.895)	1.102 (0.940,0.930)

Table 2.5: Empirical expected lengths ratio and coverage probabilities: $\widehat{r}_4\,(\widehat{CP}_4, \widehat{CP}_0)$, $m=2$

The values of $\widehat{r}_4$ in Table 2.5 range from 1.101 to 1.120, indicating that I_4 is longer than I_0 on average. The $\widehat{r}_4$ values are very similar for the five distributions of X_i. We observe that as the sample size n increases, $\widehat{r}_4$ decreases. The difference $\widehat{CP}_4 - \widehat{CP}_0$ deviates nonsignificantly around 0, with the exception of the samples from both Pareto distributions, where this difference is always positive and more significant, reaching a high of 2.4% for one of the small samples.

While we are unable to derive the expected length of I_4, we can get a closed form expression for the expectation of the squared length of I_4, provided we assume additionally that $Var(X) = \sigma^2 < \infty$. Therefore, in addition to the numerical comparison of the expected lengths of I_4 and I_0 via $\widehat{r_4}$ in Table 2.5, we compare I_4 to I_0 by evaluating the ratio

$$r_{4SQ} := \frac{E[(\text{length of } I_4)^2]}{E[(\text{length of } I_0)^2]}. \tag{2.38}$$

The expectation of the squared length of I_4 is computed in Lemma A.1 of Appendix A, and the expected squared length of I_0 is $4z_{\alpha/2}^2\sigma^2/n$. Thus,

$$r_{4SQ} = \frac{\frac{12\sigma^2}{5}\frac{(10b-1)n^2+n+2}{n(n-1)(2n-1)}}{\frac{4z_{\alpha/2}^2\sigma^2}{n}} = \frac{3}{5}\cdot\frac{(10b-1)\,n^2+n+2}{(n-1)(2n-1)z_{\alpha/2}^2} \underset{n\to\infty}{\longrightarrow} \frac{3(10b-1)}{10z_{\alpha/2}^2}. \tag{2.39}$$

Using the normal table for $z_{\alpha/2}$ and the tabulated b values in Csörgő and Horváth (1981), we obtain the following values of $\lim_{n\to\infty} r_{4SQ}$ for different confidence levels $1-\alpha$:

$1-\alpha$	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	0.98	0.99
r_{4SQ}	1.253	1.246	1.232	1.230	1.223	1.220	1.215	1.216	1.214	1.214	1.214	1.216

Table 2.6: Tabulated values of $\lim_{n\to\infty} r_{4SQ}$

For the confidence levels of $1-\alpha = 0.9, 0.95$, and 0.98 that are considered throughout this presentation, $E[(\text{length of } I_4)^2]$ is greater than $E[(\text{length of } I_0)^2]$ by a factor of approximately 1.214, when n is large.

2.4.3 $m = 3$

First, we generate the empirical distribution function of the limiting random variable $\int_0^1 W^3(t)\,dt$ in (2.25), according to the scheme described on page 21, and use it to compute approximate values of a and b such that $P(\int_0^1 W^3(t)\,dt \le a) = P(\int_0^1 W^3(t)\,dt \ge b) = \alpha/2$, for different values of α.

$\boldsymbol{1-\alpha}$	0.8	0.85	0.9	0.95	0.98	0.99
$\boldsymbol{a}$	-0.862	-1.161	-1.658	-2.672	-4.315	-5.708
$\boldsymbol{b}$	0.873	1.173	1.667	2.684	4.3	5.834

Table 2.7: Empirical quantiles of $\int_0^1 W^3(t)\,dt$

Now, as was explained on page 22, the lower and upper bounds of the FACI from (2.26), which is denoted by I_4 from now on, are the real-valued solutions of $G_3(\mu) = b$ and $G_3(\mu) = a$, respectively. To get these solutions, we expand $G_3(\mu)$ as follows:

$$G_3(\mu) = \frac{1}{n}\sum_{k=1}^{n-1}\left(\frac{\sum_{i=1}^{k}(X_i-\mu)}{s_n\sqrt{n}}\right)^3 =$$

$$-\frac{\sum_{k=1}^{n-1}k^3}{s_n^3\,n^{5/2}}\mu^3+3\frac{\sum_{k=1}^{n-1}k^2\sum_{i=1}^{k}X_i}{s_n^3\,n^{5/2}}\mu^2-3\frac{\sum_{k=1}^{n-1}k\left(\sum_{i=1}^{k}X_i\right)^2}{s_n^3\,n^{5/2}}\mu+\frac{\sum_{k=1}^{n-1}\left(\sum_{i=1}^{k}X_i\right)^3}{s_n^3\,n^{5/2}}, \tag{2.40}$$

and then evaluate the real solutions of $G_3(\mu) = b$ and $G_3(\mu) = a$ numerically using the *solve* function in R.

Evaluation of I_4, $m = 3$

Since we solve for the bounds of I_4 numerically, there is no closed form expression for them. Thus, deriving the expected length of I_4 is not feasible, and we must turn to $\widehat{r}_4$ as in (2.5). We present the values of $\widehat{r}_4$ and those of the empirical coverage probabilities $(\widehat{CP}_4, \widehat{CP}_0)$ of I_4 and I_0 in the following table.

		$1-\alpha$		
Distribution	n	0.9	0.95	0.98
$N(0,1)$	50	1.097 (0.892,0.893)	1.097 (0.949,0.946)	1.094 (0.976,0.974)
	100	1.089 (0.895,0.894)	1.089 (0.951,0.945)	1.086 (0.978,0.978)
	500	1.083 (0.897,0.898)	1.082 (0.947,0.949)	1.080 (0.981,0.978)
	1000	1.082 (0.898,0.902)	1.082 (0.956,0.955)	1.080 (0.982,0.982)
Exp(1)	50	1.097 (0.885,0.882)	1.096 (0.931,0.927)	1.094 (0.965,0.960)
	100	1.089 (0.889,0.887)	1.087 (0.940,0.940)	1.086 (0.974,0.972)
	500	1.083 (0.894,0.897)	1.083 (0.948,0.949)	1.080 (0.980,0.978)
	1000	1.082 (0.893,0.893)	1.081 (0.947,0.946)	1.079 (0.980,0.981)
Pareto(1,2)	50	1.097 (0.770,0.758)	1.097 (0.830,0.816)	1.094 (0.873,0.850)
	100	1.088 (0.799,0.791)	1.088 (0.850,0.840)	1.086 (0.887,0.875)
	500	1.084 (0.842,0.836)	1.082 (0.893,0.882)	1.080 (0.931,0.922)
	1000	1.083 (0.850,0.846)	1.081 (0.903,0.896)	1.080 (0.939,0.933)
Poisson(3)	50	1.097 (0.897,0.888)	1.096 (0.948,0.944)	1.094 (0.978,0.975)
	100	1.089 (0.898,0.897)	1.088 (0.950,0.949)	1.086 (0.978,0.977)
	500	1.084 (0.905,0.904)	1.082 (0.949,0.949)	1.080 (0.979,0.979)
	1000	1.083 (0.899,0.901)	1.081 (0.950,0.951)	1.079 (0.980,0.980)
Discrete Pareto	50	1.098 (0.753,0.748)	1.096 (0.806,0.789)	1.093 (0.851,0.833)
	100	1.088 (0.782,0.777)	1.088 (0.833,0.820)	1.085 (0.883,0.872)
	500	1.083 (0.830,0.824)	1.082 (0.886,0.872)	1.080 (0.927,0.919)
	1000	1.083 (0.842,0.837)	1.081 (0.896,0.888)	1.079 (0.940,0.932)

Table 2.8: Empirical expected lengths ratio and coverage probabilities: $\widehat{r}_4\,(\widehat{CP}_4, \widehat{CP}_0)$, $m = 3$

The values of $\widehat{r}_4$ in Table 2.8 range from 1.079 to 1.098, indicating that I_4 is longer than I_0 on average. The $\widehat{r}_4$ values are similar for all the distributions of X_i, and they decrease as n increases. The empirical coverage probabilities of I_4 are mostly nonsignificantly higher than those of I_0, except in the cases of the two Pareto distributions, where $0.4\% \le \widehat{CP}_4 - \widehat{CP}_0 \le 2.3\%$.

2.4.4 $\boldsymbol{m = 4}$

Similarly to the previous case of $m = 3$, we first tabulate b for the random variable $\int_0^1 W^4(t)\,dt$ along the lines of page 21, where $P(\int_0^1 W^4(t)\,dt \le b) = 1 - \alpha$:

$\boldsymbol{1-\alpha}$	0.8	0.85	0.9	0.95	0.98	0.99
$\boldsymbol{b}$	1.074	1.571	2.485	4.629	8.748	12.79

Table 2.9: Empirical quantiles of $\int_0^1 W^4(t)\,dt$

We recall from the introduction on page 23 that the lower and upper bounds of the FACI corresponding to (2.28) with $m = 4$, which is denoted by I_4, are the smaller and larger real solutions of $G_4(\mu) = b$, respectively, where

$$\begin{aligned} G_4(\mu) &= \frac{1}{n}\sum_{k=1}^{n-1}\left(\frac{\sum_{i=1}^{k}(X_i-\mu)}{s_n\sqrt{n}}\right)^4 \\ &= \frac{\sum_{k=1}^{n-1}k^4}{s_n^4\,n^3}\mu^4 - 4\frac{\sum_{k=1}^{n-1}k^3\sum_{i=1}^{k}X_i}{s_n^4\,n^3}\mu^3 + 6\frac{\sum_{k=1}^{n-1}k^2\left(\sum_{i=1}^{k}X_i\right)^2}{s_n^4\,n^3}\mu^2 \\ &\quad - 4\frac{\sum_{k=1}^{n-1}k\left(\sum_{i=1}^{k}X_i\right)^3}{s_n^4\,n^3}\mu + \frac{\sum_{k=1}^{n-1}\left(\sum_{i=1}^{k}X_i\right)^4}{s_n^4\,n^3}. \end{aligned} \tag{2.41}$$

These solutions are evaluated numerically by using the *solve* function in R.

Evaluation of I_4, $m = 4$

The simulated values of $\widehat{r_4}$ and $\widehat{CP_4}$ for I_4 (m=4), and those of $\widehat{CP_0}$ are summarized next.

Distribution	n	$1-\alpha$		
		0.9	0.95	0.98
$N(0,1)$	50	1.087 (0.896,0.895)	1.083 (0.947,0.944)	1.084 (0.979,0.975)
	100	1.079 (0.898,0.895)	1.077 (0.951,0.947)	1.077 (0.977,0.976)
	500	1.073 (0.902,0.900)	1.071 (0.953,0.952)	1.071 (0.980,0.980)
	1000	1.072 (0.901,0.906)	1.071 (0.952,0.950)	1.070 (0.981,0.982)
Exp(1)	50	1.086 (0.886,0.882)	1.084 (0.935,0.929)	1.085 (0.969,0.966)
	100	1.079 (0.888,0.886)	1.077 (0.941,0.940)	1.077 (0.971,0.969)
	500	1.074 (0.904,0.900)	1.071 (0.945,0.944)	1.072 (0.978,0.977)
	1000	1.072 (0.897,0.896)	1.070 (0.950,0.950)	1.071 (0.978,0.980)
Pareto(1,2)	50	1.088 (0.775,0.762)	1.084 (0.822,0.812)	1.084 (0.865,0.852)
	100	1.080 (0.791,0.779)	1.077 (0.841,0.829)	1.077 (0.891,0.880)
	500	1.074 (0.839,0.831)	1.071 (0.889,0.884)	1.071 (0.934,0.927)
	1000	1.073 (0.860,0.851)	1.071 (0.905,0.899)	1.071 (0.942,0.933)
Poisson(3)	50	1.086 (0.901,0.895)	1.083 (0.945,0.941)	1.083 (0.978,0.976)
	100	1.079 (0.901,0.895)	1.077 (0.948,0.946)	1.077 (0.978,0.977)
	500	1.074 (0.900,0.898)	1.071 (0.951,0.950)	1.071 (0.983,0.983)
	1000	1.072 (0.903,0.899)	1.070 (0.947,0.950)	1.071 (0.980,0.977)
Discrete Pareto	50	1.087 (0.754,0.744)	1.083 (0.807,0.792)	1.084 (0.846,0.829)
	100	1.080 (0.786,0.775)	1.077 (0.839,0.827)	1.077 (0.876,0.865)
	500	1.073 (0.825,0.820)	1.071 (0.882,0.875)	1.072 (0.923,0.916)
	1000	1.073 (0.848,0.842)	1.070 (0.902,0.898)	1.070 (0.936,0.929)

Table 2.10: Empirical expected lengths ratio and coverage probabilities: $\widehat{r_4}\,(\widehat{CP_4}, \widehat{CP_0})$, $m = 4$

From Table 2.10, we see that I_4 is longer than I_0 on average, since $1.070 \leq \widehat{r_4} \leq 1.088$. The values of $\widehat{r_4}$ are similar for all the distributions of X_i's and decrease as n increases. As before, the $\widehat{CP_4}$ values are mostly higher than those of $\widehat{CP_0}$, particularly for the Pareto distributions, although the difference $\widehat{CP_4} - \widehat{CP_0}$ is more modest than for I_4 with a smaller m.

Looking back at the values of r_4 for $m = 1$ in (2.33), $\widehat{r_4}$ for $m = 2$ in Table 2.5, $\widehat{r_4}$ for $m = 3$ in Table 2.8, and $\widehat{r_4}$ for $m = 4$ in Table 2.10, we notice that the average length of I_4 relative to that of I_0 seems to decrease as m increases. Therefore, we continue to increase m to see if this trend continues. The next FACI presented here is constructed from (2.28) with $m = 8$.

2.4.5 $\boldsymbol{m = 8}$

Just like in the case of $m = 4$, we first generate the empirical distribution function of $\int_0^1 W^8(t)\,dt$ and use it to tabulate approximate values of b, where $P\left(\int_0^1 W^8(t)\,dt \ \le b\right) = 1 - \alpha$, for several values of $1 - \alpha$:

$\boldsymbol{1-\alpha}$	0.8	0.85	0.9	0.95	0.98	0.99
$\boldsymbol{b}$	3.301	6.832	16.234	53.536	181.368	383.002

Table 2.11: Empirical quantiles of $\int_0^1 W^8(t)\,dt$

Then, expanding $G_8(\mu)$ of (2.25) as

$$G_8(\mu) = \frac{1}{n}\sum_{k=1}^{n-1}\left(\frac{\sum_{i=1}^{k}(X_i-\mu)}{s_n\sqrt{n}}\right)^8 = \frac{\sum_{k=1}^{n-1}k^8}{s_n^8\,n^5}\mu^8 - 8\frac{\sum_{k=1}^{n-1}k^7\sum_{i=1}^{k}X_i}{s_n^8\,n^5}\mu^7$$

$$+ 28\frac{\sum_{k=1}^{n-1}k^6\left(\sum_{i=1}^{k}X_i\right)^2}{s_n^8\,n^5}\mu^6 - 56\frac{\sum_{k=1}^{n-1}k^5\left(\sum_{i=1}^{k}X_i\right)^3}{s_n^8\,n^5}\mu^5$$

$$+ 70\frac{\sum_{k=1}^{n-1}k^4\left(\sum_{i=1}^{k}X_i\right)^4}{s_n^8\,n^5}\mu^4 - 56\frac{\sum_{k=1}^{n-1}k^3\left(\sum_{i=1}^{k}X_i\right)^5}{s_n^8\,n^5}\mu^3$$

$$+28\frac{\sum_{k=1}^{n-1}k^2\left(\sum_{i=1}^{k}X_i\right)^6}{s_n^8\,n^5}\mu^2-8\frac{\sum_{k=1}^{n-1}k\left(\sum_{i=1}^{k}X_i\right)^7}{s_n^8\,n^5}\mu+\frac{\sum_{k=1}^{n-1}\left(\sum_{i=1}^{k}X_i\right)^8}{s_n^8\,n^5}, \tag{2.42}$$

we find the real solutions of $G_8(\mu)=b$ numerically, using the *solve* function in R, thus obtaining the bounds of the FACI for μ based on (2.28) with $m=8$.

Evaluation of I_4, m=8

In the following table, we present simulated values for the empirical coverage probabilities $\widehat{CP_4}$ of (2.2) and $\widehat{CP_0}$ of (2.3), and for the ratio $\widehat{r_4}$ of (2.5).

		$1-\alpha$		
Distribution	n	0.9	0.95	0.98
$N(0,1)$	50	1.063 (0.896,0.894)	1.058 (0.948,0.947)	1.054 (0.976,0.974)
	100	1.055 (0.905,0.900)	1.050 (0.949,0.950)	1.048 (0.980,0.980)
	500	1.050 (0.901,0.902)	1.046 (0.948,0.947)	1.043 (0.979,0.981)
	1000	1.051 (0.905,0.901)	1.045 (0.950,0.952)	1.042 (0.978,0.978)
Exp(1)	50	1.063 (0.880,0.875)	1.058 (0.931,0.928)	1.054 (0.966,0.963)
	100	1.056 (0.892,0.889)	1.051 (0.940,0.937)	1.048 (0.974,0.973)
	500	1.050 (0.899,0.900)	1.046 (0.948,0.946)	1.043 (0.979,0.977)
	1000	1.049 (0.900,0.901)	1.045 (0.952,0.952)	1.042 (0.978,0.978)
Pareto(1,2)	50	1.066 (0.769,0.757)	1.059 (0.819,0.810)	1.056 (0.860,0.851)
	100	1.061 (0.795,0.786)	1.053 (0.841,0.835)	1.049 (0.887,0.879)
	500	1.052 (0.841,0.837)	1.048 (0.890,0.884)	1.046 (0.928,0.923)
	1000	1.054 (0.851,0.852)	1.047 (0.896,0.895)	1.043 (0.938,0.934)
Poisson(3)	50	1.063 (0.891,0.889)	1.057 (0.948,0.945)	1.055 (0.977,0.976)
	100	1.056 (0.896,0.895)	1.051 (0.949,0.952)	1.048 (0.981,0.979)
	500	1.049 (0.898,0.897)	1.046 (0.947,0.946)	1.043 (0.980,0.978)
	1000	1.049 (0.898,0.898)	1.045 (0.948,0.947)	1.042 (0.980,0.980)
Discrete Pareto	50	1.069 (0.753,0.744)	1.059 (0.797,0.785)	1.056 (0.839,0.827)
	100	1.060 (0.783,0.778)	1.053 (0.837,0.827)	1.049 (0.868,0.863)
	500	1.055 (0.829,0.824)	1.049 (0.881,0.880)	1.043 (0.920,0.915)
	1000	1.052 (0.842,0.835)	1.047 (0.900,0.896)	1.043 (0.932,0.927)

Table 2.12: Empirical expected lengths ratio and coverage probabilities: $\widehat{r_4}\,(\widehat{CP_4},\widehat{CP_0})$, $m=8$

The $\widehat{r_4}$ values in Table 2.12 range from 1.042 to 1.069. They are smaller than the corresponding values of $\widehat{r_4}$ in Table 2.10 ($m = 4$), thus supporting the trend that the relative average length of I_4 becomes shorter as m increases. For each of the five distributions in Table 2.12, as the sample size n or the confidence level $1 - \alpha$ increases, $\widehat{r_4}$ decreases. Since $\widehat{r_4}$ is higher than 1, we conclude that, on average, I_4 is longer than I_0. The empirical coverage probabilities of I_4 are mostly nonsignificantly higher than those of I_0.

Studying (2.25) with even m that are larger than 8 becomes problematic already for $m = 10$. This is because the coefficients of the polynomial $G_m(\mu)$ become enormous, especially for large samples, which leads to the numerical roots of $G_m(\mu) = b$ being inaccurate.

2.5 FACI based on $h_5(\cdot)$

Finally, we consider (2.1) with $h_5(\cdot)$, where the functional $h_5(\cdot)$ in (1.15) is a linear combination of the projection functional $h_1(\cdot)$ with $t_0 = 1$ and the integral functional $h_4(\cdot)$ with $m = 1$. We have

$$h_5(T_n^t(X_1 - \mu, \ldots, X_n - \mu)) = \frac{a_1 \sum_{i=1}^n (X_i - \mu)}{s_n \sqrt{n}} + \frac{a_2}{n} \sum_{k=1}^{n-1} \frac{\sum_{i=1}^k (X_i - \mu)}{s_n \sqrt{n}}$$

$$\xrightarrow[n\to\infty]{\mathcal{D}} a_1 \, W(1) + a_2 \int_0^1 W(t)\, dt, \qquad (2.43)$$

where a_1 and a_2 are fixed real numbers. Notice that when $a_1 = 1$ and $a_2 = 0$, (2.43) is equivalent to (1.2), the asymptotic normality of the Student t-statistic, while when $a_1 = 0$ and $a_2 = 1$, (2.43) coincides with (2.25) with $m = 1$.

To find the limiting distribution in (2.43), we first use the integration by parts formula as in (5.20) of Chapter 8 of Karlin and Taylor (1998) and conclude that

$$a_1 W(1) = \int_0^1 a_1 \, dW(x)$$

and

$$a_2 \int_0^1 W(t) \, dt = \int_0^1 a_2(1-x) \, dW(x).$$

Consequentially, by Theorem 5.1 in Chapter 8 of Karlin and Taylor (1998) with $f(x) = a_1$ and $g(x) = a_2(1-x)$, the random variables $a_1 W(1)$ and $a_2 \int_0^1 W(t) \, dt$ have a bivariate normal distribution with the covariance matrix

$$\begin{pmatrix} \int_0^1 {a_1}^2 \, dx & \int_0^1 a_1 a_2 (1-x) \, dx \\ \int_0^1 a_1 a_2 (1-x) \, dx & \int_0^1 {a_2}^2 (1-x)^2 \, dx \end{pmatrix} = \begin{pmatrix} {a_1}^2 & \dfrac{a_1 a_2}{2} \\ \dfrac{a_1 a_2}{2} & \dfrac{{a_2}^2}{3} \end{pmatrix}. \tag{2.44}$$

Therefore,

$$a_1 \, W(1) + a_2 \int_0^1 W(t) \, dt \overset{\mathcal{D}}{=} N(0, {a_1}^2 + {a_2}^2/3 + a_1 a_2). \tag{2.45}$$

Set $0 < \alpha < 1$, and define a and b such that

$$P(a \le N(0, {a_1}^2 + {a_2}^2/3 + a_1 a_2) \le b) = 1 - \alpha. \tag{2.46}$$

It follows from (2.43), (2.45) and (2.46) that

$$P\left(a \le \frac{a_1 \sum_{i=1}^n (X_i - \mu)}{s_n \sqrt{n}} + \frac{a_2}{n} \sum_{k=1}^{n-1} \frac{\sum_{i=1}^k (X_i - \mu)}{s_n \sqrt{n}} \le b \right) \underset{n \to \infty}{\longrightarrow} 1 - \alpha. \tag{2.47}$$

For the probability in (2.47), we have

$$P\left(a \le \frac{a_1 \sum_{i=1}^n (X_i-\mu)}{s_n\sqrt{n}} + \frac{a_2}{n}\sum_{k=1}^{n-1}\frac{\sum_{i=1}^k (X_i-\mu)}{s_n\sqrt{n}} \le b\right)$$

$$= P\left(a\, s_n n^{3/2} \le a_1\, n \sum_{i=1}^n (X_i-\mu) + a_2 \sum_{k=1}^{n-1}\sum_{i=1}^k (X_i-\mu) \le b\, s_n n^{3/2}\right)$$

$$= P\left(a\, s_n n^{3/2} \le a_1\, n \sum_{i=1}^n (X_i-\mu) + a_2 \sum_{i=1}^{n-1} (X_i-\mu)(n-i) \le b\, s_n n^{3/2}\right)$$

$$= P\left(a\, s_n n^{3/2} \le \sum_{i=1}^n (X_i-\mu)(a_1\, n + a_2(n-i)) \le b\, s_n n^{3/2}\right)$$

$$= P\left(\frac{-a\, s_n n^{3/2} + \sum_{i=1}^n X_i(a_1\, n + a_2(n-i))}{a_1 n^2 + a_2\dfrac{n(n-1)}{2}} \ge \mu \ge \frac{-b\, s_n n^{3/2} + \sum_{i=1}^n X_i(a_1\, n + a_2(n-i))}{a_1 n^2 + a_2\dfrac{n(n-1)}{2}}\right), \tag{2.48}$$

where, in the last step, we assumed that $a_1 + a_2\frac{n-1}{2n} > 0$. Hence, we obtain the following $1-\alpha$ FACI for μ:

$$I_5 := \left[\frac{\sum_{i=1}^n X_i(a_1 n + a_2(n-i)) - b s_n n^{3/2}}{a_1 n^2 + a_2\dfrac{n(n-1)}{2}}, \frac{\sum_{i=1}^n X_i(a_1 n + a_2(n-i)) - a s_n n^{3/2}}{a_1 n^2 + a_2\dfrac{n(n-1)}{2}}\right]. \tag{2.49}$$

Evaluation of I_5

The center of I_5,

$$\frac{\sum_{i=1}^n X_i(a_1 n + a_2(n-i))}{a_1 n^2 + a_2\dfrac{n(n-1)}{2}},$$

is an unbiased estimator for μ. The length of I_5,

$$\frac{(b-a)s_n n^{3/2}}{a_1 n^2 + a_2 \dfrac{n(n-1)}{2}}, \tag{2.50}$$

is minimized when $b=-a$, as a result of Theorem 9.3.2 in Casella and Berger (2002). This implies that $-a = b = z_{\alpha/2}\sqrt{{a_1}^2 + {a_2}^2/3 + a_1 a_2}$ (cf. (2.46)). Plugging these values into (2.50) gives

$$\text{Length}(I_5) = \frac{2z_{\alpha/2}s_n}{\sqrt{n}} \frac{\sqrt{{a_1}^2 + \dfrac{{a_2}^2}{3} + a_1 a_2}}{a_1 + a_2\left(\dfrac{1}{2} - \dfrac{1}{2n}\right)} = \text{Length}(I_0)\frac{\sqrt{{a_1}^2 + \dfrac{{a_2}^2}{3} + a_1 a_2}}{a_1 + a_2\left(\dfrac{1}{2} - \dfrac{1}{2n}\right)}, \tag{2.51}$$

where I_0 is as in (1.7). Hence, the ratio of the expected lengths of I_5 and I_0 is:

$$r_5 = \frac{\sqrt{{a_1}^2 + \dfrac{{a_2}^2}{3} + a_1 a_2}}{a_1 + a_2\left(\dfrac{1}{2} - \dfrac{1}{2n}\right)}. \tag{2.52}$$

Let a_1 be fixed. To find the values of a_2 that minimize r_5, we examine $\dfrac{d\,r_5}{d\,a_2}$, where

$$\frac{d\,r_5}{d\,a_2} = \frac{\dfrac{\left(\dfrac{2}{3}a_2 + a_1\right)\left(a_1 + a_2\left(\dfrac{1}{2} - \dfrac{1}{2n}\right)\right)}{2\sqrt{{a_1}^2 + \dfrac{{a_2}^2}{3} + a_1 a_2}} - \sqrt{{a_1}^2 + \dfrac{{a_2}^2}{3} + a_1 a_2}\left(\dfrac{1}{2} - \dfrac{1}{2n}\right)}{\left[a_1 + a_2\left(\dfrac{1}{2} - \dfrac{1}{2n}\right)\right]^2}$$

$$= \frac{\dfrac{1}{2}\sqrt{{a_1}^2 + \dfrac{{a_2}^2}{3} + a_1 a_2}}{\left[a_1 + a_2\left(\dfrac{1}{2} - \dfrac{1}{2n}\right)\right]^2}\left[\frac{\left(\dfrac{2}{3}a_2 + a_1\right)\left(a_1 + a_2\left(\dfrac{1}{2} - \dfrac{1}{2n}\right)\right)}{{a_1}^2 + \dfrac{{a_2}^2}{3} + a_1 a_2} - \left(1 - \frac{1}{n}\right)\right]$$

$$= \frac{1}{2\left[a_1 + a_2\left(\frac{1}{2} - \frac{1}{2n}\right)\right]^2} \frac{\frac{a_1}{n}\left(a_2\left(\frac{n+3}{6}\right) + a_1\right)}{\sqrt{{a_1}^2 + \frac{{a_2}^2}{3} + a_1 a_2}}. \tag{2.53}$$

Without loss of generality, we assume that $a_1 > 0$. Then, solving $d\,r_5/d\,a_2 = 0$ for a_2, we get

$$a_2 = -\frac{6a_1}{n+3}, \tag{2.54}$$

and since $d\,r_5/d\,a_2$ is negative when $a_2 < -6a_1/(n+3)$ and positive when $a_2 > -6a_1/(n+3)$, r_5 has a local minimum at $a_2 = -6a_1/(n+3)$. Plugging this value of a_2 back into r_5 of (2.52) gives

$$\frac{\sqrt{{a_1}^2 + \frac{36{a_1}^2}{3(n+3)^2} - \frac{6{a_1}^2}{n+3}}}{a_1 - \frac{6a_1}{n+3}\left(\frac{1}{2} - \frac{1}{2n}\right)} = \frac{n}{\sqrt{n^2+3}} < 1\,. \tag{2.55}$$

Thus, for a fixed $a_1 > 0$ and a_2 chosen according to (2.54), the length of I_5 of (2.49) is shorter than that of I_0. For large n, the two intervals have nearly the same length.

For example, when $a_1 = 1$ and $a_2 = -6/(n+3)$, the values of r_5 as well as the empirical coverage probabilities $\widehat{CP}_5$ of I_5 and $\widehat{CP}_0$ of I_0 are provided next.

Distribution	n	r_5	$1-\alpha$ 0.9	0.95	0.98
$N(0,1)$	50	0.9994005	(0.896,0.896)	(0.943,0.942)	(0.977,0.977)
	100	0.99985	(0.896,0.897)	(0.946,0.946)	(0.981,0.981)
	500	0.999994	(0.896,0.896)	(0.947,0.947)	(0.978,0.978)
	1000	0.9999985	(0.895,0.895)	(0.952,0.952)	(0.981,0.981)
Exp(1)	50	0.9994005	(0.883,0.883)	(0.925,0.926)	(0.959,0.958)
	100	0.99985	(0.898,0.898)	(0.937,0.938)	(0.969,0.969)
	500	0.999994	(0.895,0.895)	(0.951,0.952)	(0.978,0.978)
	1000	0.9999985	(0.896,0.896)	(0.951,0.951)	(0.978,0.978)
Pareto(1,2)	50	0.9994005	(0.761,0.761)	(0.809,0.809)	(0.851,0.852)
	100	0.99985	(0.787,0.788)	(0.838,0.839)	(0.886,0.886)
	500	0.999994	(0.836,0.836)	(0.888,0.887)	(0.921,0.921)
	1000	0.9999985	(0.842,0.842)	(0.895,0.895)	(0.937,0.937)
Poisson(3)	50	0.9994005	(0.890,0.891)	(0.941,0.940)	(0.974,0.975)
	100	0.99985	(0.901,0.899)	(0.947,0.947)	(0.975,0.975)
	500	0.999994	(0.899,0.899)	(0.947,0.947)	(0.980,0.980)
	1000	0.9999985	(0.901,0.900)	(0.953,0.953)	(0.982,0.982)
Discrete Pareto	50	0.9994005	(0.743,0.746)	(0.797,0.795)	(0.831,0.831)
	100	0.99985	(0.773,0.772)	(0.832,0.832)	(0.867,0.869)
	500	0.999994	(0.826,0.827)	(0.875,0.875)	(0.919,0.919)
	1000	0.9999985	(0.842,0.842)	(0.890,0.890)	(0.926,0.926)

Table 2.13: Expected lengths ratio and empirical coverage probabilities: r_5 and $(\widehat{CP}_5, \widehat{CP}_0)$

In Table 2.13, the values of $\widehat{CP}_5$ are nearly identical to those of $\widehat{CP}_0$, and the ratio r_5 of the expected lengths is smaller but close to 1. The similarity between the two confidence intervals is not surprising. After all, the functional $h_5(\cdot)$ of (1.15) that was used to derive I_5 is simply $h_1(\cdot) - \frac{6}{n+3} h_4(\cdot)$, where $t_0 = 1$ in $h_1(\cdot)$ and $m = 1$ in $h_4(\cdot)$. That is $h_5(\cdot)$ is almost the functional $h_1(\cdot)$ with $t_0 = 1$ that was used to derive I_0, as the weight $-6/(n+3)$ assigned to $h_4(\cdot)$ is small.

Even though I_5 does not offer significant improvement over I_0, it demonstrates that FACI's that are shorter than I_0 and have coverage probabilities nearly equal to that of I_0 do exist. Perhaps other functionals or (not necessarily linear) combinations of functionals that were not considered here could yield more striking examples of such FACI's.

Chapter 3

Conclusions

In the previous chapter, we derived FACI's I_1 to I_5 for the mean μ of a population in DAN and compared their expected lengths and empirical finite-sample coverage probabilities to those of the classical asymptotic confidence interval I_0 of (1.7). We note that these FACI's have simple enough forms that do not aggravate the duration of their numerical computation. Now, we review the performances of all the obtained FACI's and conclude which ones present reasonable alternatives to, and overall improvement upon, I_0.

FACI with a shorter expected length and nearly equal coverage probability

The FACI I_5 of (2.49) with $a_1 = 1$ and $a_2 = -6/(n+3)$ is $\sqrt{n^2+3}/n$ times shorter than I_0 and has nearly the same coverage as I_0 (see Table 2.13). While the expected length of I_5 is not significantly shorter than that of I_0, I_5 does improve overall upon the interval I_0.

FACI's with longer expected lengths and higher coverage probabilities

We found that I_1 of (2.10) with $t_0 = 0.9$, I_2 of (2.17), and I_4, which is based on (2.25) and studied for $m = 1, 2, 3, 4$, and 8 in section 2.4, are all longer than I_0 on average, but their empirical coverage probabilities $\widehat{CP_i}$ are mostly higher than the empirical coverage probability $\widehat{CP_0}$ of I_0. Based on the corresponding values of $(\widehat{CP_i}, \widehat{CP_0})$, r_i of (2.4) and $\widehat{r_i}$ of (2.5) that were obtained in Chapter 2, we present a summary table with ranges of each FACI's $\widehat{CP_i} - \widehat{CP_0}$ values and averages of the values of their $\widehat{r_i}$ (or r_i instead, when available). Since the $\widehat{CP_i} - \widehat{CP_0}$ values were previously observed to be higher for samples from the Pareto(1,2) and discrete Pareto distributions, we summarize the results for these two distributions separately from the other distributions considered in Chapter 2 ($N(0, 1)$, Exp(1), and Poisson(3)).

FACI		I_1 (t_0=0.9)	I_2	I_4 (m=1)	I_4 (m=2)	I_4 (m=3)	I_4 (m=4)	I_4 (m=8)
samples from $N(0,1)$, Exp(1) and Poisson(3)	$\widehat{CP_i}-\widehat{CP_0}$ range (in %)	-0.3%–0.5%	-0.1%–1.9%	-0.2%–1.1%	-0.5%–0.8%	-0.4%–0.9%	-0.5%–0.6%	-0.3%–0.5%
	average r_i or $\widehat{r_i}$	1.0541*	1.055	1.164*	1.109	1.087	1.076	1.050
samples from Pareto(1,2) and discrete Pareto	$\widehat{CP_i}-\widehat{CP_0}$ range (in %)	0.2%–1.3%	0.1%–1.7%	0.7%–3.0%	0.3%–2.4%	0.4%–2.3%	0.4%–1.7%	-0.1%–1.2%
	average r_i or $\widehat{r_i}$	1.0541*	1.082	1.164*	1.109	1.087	1.077	1.053

Table 3.1: Summary of (empirical) expected lengths ratios and empirical coverage probabilities for I_1, I_2, and I_4 (* marks r_i value)

The FACI's I_1 with $t_0 = 0.9$, I_2, and I_4 present alternatives to I_0 when having higher frequency of "catching" the real value of μ is more important than having a narrower interval. However, neither of the FACI's in Table 3.1 is clearly better than the others when taking both the expected length and coverage probability into consideration. For example, the highest $\widehat{CP_i} - \widehat{CP_0}$ values for the two Pareto distributions are observed for I_4 with $m = 1$, and as m increases, improvement in the coverage probability of I_4 over that of I_0 becomes smaller. On the other hand, the ratio of the expected lengths of I_4 and I_0 gets better (decreases), as m increases. Thus, choosing the "best" FACI among I_1, I_2, and I_4 depends on the importance that the user assigns to having a shorter expected length versus having a higher coverage probability.

FACI with a longer expected length and lower coverage probability

I_3 of (2.23), which is based on the supremum functional, is much longer than I_0 compared to our other FACI's. As opposed to other I_i's, the empirical coverage probability of I_3 is always lower than that of I_0, by 1.5%–10.2%. This makes I_3 the least desirable among our FACI's.

All in all, in this thesis, we hope to have demonstrated a good potential of the FCLT based FACI's for the mean μ of a population in DAN. With countless choices for the functional $h(\cdot)$ in (1.10), this may inspire some to construct new FACI's for μ with properties that would be better than those of the FACI's I_1 to I_5 and that would further improve on those of the commonly used asymptotic confidence interval I_0 of (1.7), which follows from the asymptotic normality of the Student t-statistic.

Appendix A

Deriving the expectation of the squared length of I_4 of (2.36)

Lemma A.1. Let $X_1, \ldots, X_n$ be i.i.d. random variables with mean μ and finite variance $\sigma^2 > 0$. Then, for the random interval I_4 of (2.36),

$$E[(\text{length of } I_4)^2] = \frac{12\sigma^2}{5} \frac{(10b-1)n^2 + n + 2}{n(n-1)(2n-1)}. \tag{A.1}$$

Proof. The following known summation results will be used throughout the proof:

$$\sum_{i=1}^{n-1} i = \frac{n(n-1)}{2}, \tag{A.2}$$

$$\sum_{i=1}^{n-1} i^2 = \frac{n(n-1)(2n-1)}{6}, \tag{A.3}$$

$$\sum_{i=1}^{n-1} i^3 = \left(\frac{n(n-1)}{2}\right)^2, \tag{A.4}$$

and

$$\sum_{i=1}^{n-1} i^4 = \frac{n(n-1)(2n-1)(3n^2-3n-1)}{30}. \tag{A.5}$$

Now, we recall that

$$(\text{length of } I_4)^2 = \frac{4}{{c_n}^2}\left[\left(\sum_{k=1}^{n-1} k \sum_{i=1}^{k} X_i\right)^2 - c_n\left(\sum_{k=1}^{n-1}\left(\sum_{i=1}^{k} X_i\right)^2 - b\, n^2 {s_n}^2\right)\right], \tag{A.6}$$

where $c_n := \sum_{k=1}^{n-1} k^2 = \dfrac{n(n-1)(2n-1)}{6}$.

First, we evaluate $E\left(\sum_{k=1}^{n-1} k \sum_{i=1}^{k} X_i\right)$ and $Var\left[\sum_{k=1}^{n-1} k \sum_{i=1}^{k} X_i\right]$ as follows. In view of (A.3),

$$E\left(\sum_{k=1}^{n-1} k \sum_{i=1}^{k} X_i\right) = \sum_{k=1}^{n-1} k \sum_{i=1}^{k} E(X_i) = \sum_{k=1}^{n-1} k \sum_{i=1}^{k} \mu = \sum_{k=1}^{n-1} k^2 \mu = \mu c_n, \tag{A.7}$$

while, from the independence of $X_1, \ldots, X_n$ and (A.2)–(A.5),

$$\begin{aligned}
& Var\left[\sum_{k=1}^{n-1} k \sum_{i=1}^{k} X_i\right] \\
&= Var\left[X_1 + 2(X_1+X_2) + \cdots + (n-1)(X_1+X_2+\cdots+X_{n-1})\right] \\
&= Var\left[X_1 \sum_{k=1}^{n-1} k + X_2 \sum_{k=2}^{n-1} k + \cdots + X_{n-1} \sum_{k=n-1}^{n-1} k\right] \\
&= \sigma^2\left(\sum_{k=1}^{n-1} k\right)^2 + \sigma^2\left(\sum_{k=2}^{n-1} k\right)^2 + \cdots + \sigma^2\left(\sum_{k=n-1}^{n-1} k\right)^2
\end{aligned}$$

$$= \sigma^2 \sum_{k=1}^{n-1} \left(\frac{n(n-1)}{2} - \frac{k(k-1)}{2} \right)^2$$

$$= \sigma^2 \sum_{k=1}^{n-1} \left[\left(\frac{n(n-1)}{2} \right)^2 - \frac{n(n-1)k(k-1)}{2} + \left(\frac{k(k-1)}{2} \right)^2 \right]$$

$$= \sigma^2 \left[\left(\frac{n(n-1)}{2} \right)^2 (n-1) - \frac{n(n-1)}{2} \sum_{k=1}^{n-1} (k^2 - k) + \frac{1}{4} \sum_{k=1}^{n-1} (k^4 - 2k^3 + k^2) \right]$$

$$= \sigma^2 \left[\left(\frac{n(n-1)}{2} \right)^2 (n-1) - \left(\frac{n(n-1)}{2} \right)^2 \frac{2n-1}{3} + \left(\frac{n(n-1)}{2} \right)^2 \right.$$

$$+ \frac{1}{4} \frac{n(n-1)(2n-1)}{6} \cdot \frac{3n^2-3n-1}{5} - \frac{1}{2} \left(\frac{n(n-1)}{2} \right)^2$$

$$\left. + \frac{1}{4} \frac{n(n-1)(2n-1)}{6} \right]$$

$$= \sigma^2 \left[\left(\frac{n(n-1)}{2} \right)^2 \left(n - 1 - \frac{2n-1}{3} + 1 - \frac{1}{2} \right) \right.$$

$$\left. + \frac{1}{4} \cdot \frac{n(n-1)(2n-1)}{6} \left(\frac{3n^2-3n-1}{5} + 1 \right) \right]$$

$$= \sigma^2 \left[\left(\frac{n(n-1)}{2} \right)^2 \frac{2n-1}{6} + \frac{1}{4} \cdot \frac{n(n-1)(2n-1)}{6} \left(\frac{3n^2-3n-1}{5} + 1 \right) \right]$$

$$= \frac{\sigma^2 c_n}{4} \left[n(n-1) + \frac{3n^2-3n-1}{5} + 1 \right]$$

$$= \frac{\sigma^2 c_n}{5} (2n^2 - 2n + 1). \tag{A.8}$$

Using (A.7) and (A.8), we have

$$E\left[\left(\sum_{k=1}^{n-1} k \sum_{i=1}^{k} X_i\right)^2\right] = Var\left[\sum_{k=1}^{n-1} k \sum_{i=1}^{k} X_i\right] + \left[E\left(\sum_{k=1}^{n-1} k \sum_{i=1}^{k} X_i\right)\right]^2$$
$$= \frac{\sigma^2 c_n}{5}(2n^2 - 2n + 1) + (\mu c_n)^2. \tag{A.9}$$

Clearly, by the independence of $X_1, \ldots, X_n$,

$$E\left(\sum_{k=1}^{n-1}\left(\sum_{i=1}^{k} X_i\right)^2\right) = \sum_{k=1}^{n-1} E\left(\sum_{i=1}^{k} X_i\right)^2 = \sum_{k=1}^{n-1}\left[Var\left(\sum_{i=1}^{k} X_i\right) + \left(E\left(\sum_{i=1}^{k} X_i\right)\right)^2\right]$$
$$= \sum_{k=1}^{n-1}\left[k\sigma^2 + (k\mu)^2\right] = \sigma^2 \frac{n(n-1)}{2} + \mu^2 c_n. \tag{A.10}$$

Finally, upon combining (A.9) and (A.10), we get

$E[(\text{length of } I_4)^2]$

$$= \frac{4}{{c_n}^2} E\left[\left(\sum_{k=1}^{n-1} k \sum_{i=1}^{k} X_i\right)^2 - c_n\left(\sum_{k=1}^{n-1}\left(\sum_{i=1}^{k} X_i\right)^2 - b\, n^2 {s_n}^2\right)\right]$$
$$= \frac{4}{{c_n}^2}\left[\frac{\sigma^2 c_n}{5}(2n^2 - 2n + 1) + (\mu c_n)^2 - c_n\left(\sigma^2 \frac{n(n-1)}{2} + \mu^2 c_n\right) + b c_n n^2 \sigma^2\right]$$
$$= \frac{4\sigma^2}{c_n}\left[\frac{2n^2 - 2n + 1}{5} - \frac{n(n-1)}{2} + bn^2\right]$$
$$= \frac{4\sigma^2}{c_n}\left[\frac{(10b-1)n^2 + n + 2}{10}\right] = \frac{12\sigma^2}{5}\frac{(10b-1)n^2 + n + 2}{n(n-1)(2n-1)}. \tag{A.11}$$

$\square$

Appendix B

R Code

B.1 Functions for computing FACI bounds

```
################################################################################
# These functions generate a sample x of size n from a specified distribution.
# They also compute the bounds of a specific FACI and of I_0 for that sample.
# They return the lengths of both confidence intervals and indicators of whether
# these intervals contain the real mu.
################################################################################

# Generate a Pareto(1,2) sample
rpareto=function(n){sqrt(1/(1-runif(n)))}

# Generate a discrete Pareto sample of size n
# discard samples where all data enteries are equal (this situation has a non
# negligable probability for n=50)
c=1/1.20205690315959428540 # as computed in: Handbook of mathematical functions
    with formulas-Abramowitz
ddpareto=function(x){c/x^3} # pmf of discrete pareto for x=1,2,3,...
dpareto.pmf=ddpareto(seq(1:100000)) #Store pmf values to avoid recomputing them for
     each sample
rdpareto=function(n)
{allEqual=1
 while(allEqual==1)
 {x=sample(1:100000,size=n,replace=T,prob=dpareto.pmf)
  allEqual=all(x==x[1])}
 return(x)}

x.sample.function=function(n,dist){
  if (dist=="normal"){x=rnorm(n);mean=0}
  if (dist=="exp"){x=rexp(n,1);mean=1}
  if (dist=="pareto"){x=rpareto(n);mean=2}
  if (dist=="poisson"){x=rpois(n,3);mean=3}
  if (dist=="dpareto"){x=rdpareto(n); mean=c*pi^2/6}
  return(c(x,mean))}


#    I_0   (STUDENT STATISTIC)
```

```
# -------------------------

CLT=function(x,n,s,alpha,mean)
  {CLT.l=mean(x)-qnorm(1-(alpha)/2)*s/sqrt(n) # Calculate the lower bound of I_0
  CLT.u=mean(x)+qnorm(1-(alpha)/2)*s/sqrt(n) # Calculate the upper bound of I_0
  CLT.length=CLT.u-CLT.l
  if((CLT.l<=mean)&&(CLT.u>=mean)) {in.CLT=1} else in.CLT=0
  return(c(CLT.length,in.CLT))}


#   I_1 with t0=0.9
# -------------------------

projection=function(dist,n,alpha)
  {x.sample=x.sample.function(n,dist)
  x=x.sample[1:n]
  s=sd(x)
  mean=x.sample[n+1]

  z=qnorm(1-(alpha)/2)
  t0=0.9
  x.partial=x[1:(n*t0)]

  upper=(sum(x.partial)+z*s*sqrt(n*t0) )/floor(n*t0)
  lower=(sum(x.partial)-z*s*sqrt(n*t0) )/floor(n*t0)
  length=upper-lower
  if (length<=0) {stop("length can't be negative or 0")}
  if((lower<=mean)&&(upper>=mean)) {in.interval=1} else in.interval=0

  return(c(length,in.interval,CLT(x,n,s,alpha,mean)))}


#   I_2  (ABSOLUTE SUPREMUM)
# -------------------------

abs.sup=function(dist,n,alpha)
  {b.s=c(1.645,1.78,1.96,2.24,2.575,2.81)
  alphas=c(0.2,0.15,0.1,0.05,0.02,0.01)
  values=matrix(c(alphas,b.s),length(alphas),2,byrow=F)
  if (alpha %in% alphas == 0) {stop("Choose a different alpha!")}
  b=values[values[,1]==alpha,2] # select a value of b according to alpha

  x.sample=x.sample.function(n,dist)
  x=x.sample[1:n]
  s=sd(x)
  mean=x.sample[n+1]

  k=seq(1:n)
  lower=max((cumsum(x)-b*s*sqrt(n))/k)  # Calculate the lower bound of the FACI
  upper=min((cumsum(x)+b*s*sqrt(n))/k)  # Calculate the upper bound of the FACI
  length=upper-lower
  if (length<=0) {stop("length can't be negative or 0")}
  if((lower<=mean)&&(upper>=mean)) {in.interval=1} else in.interval=0

  return(c(length,in.interval,CLT(x,n,s,alpha,mean)))}


#   I_3  (SUPREMUM)
# -------------------------

sup=function(dist,n,alpha)
```

```
  {if (dist=="normal")
    {a.s=c(0.0955, 0.1056, 0.1131 ,0.1131 ,0.0414 ,0.0502 ,0.0577, 0.0589 ,0.0070
        ,0.0160, 0.0231, 0.0236)
    values=matrix(c(n.vec,alpha.vec,a.s),12,3)
    a=values[values[,1]==n & values[,2]==alpha,3]}

  if (dist=="exp")
    {a.s=c(0.0905 ,0.1030, 0.1156 ,0.1156, 0.0376, 0.0489 ,0.0577, 0.0589 ,0.0035,
       0.0140 ,0.0226 ,0.0236)
    values=matrix(c(n.vec,alpha.vec,a.s),12,3)
    a=values[values[,1]==n & values[,2]==alpha,3]}

  if (dist=="pareto")
    {a.s=c(0.0754, 0.0930 ,0.1131 ,0.1131, 0.0263, 0.0389, 0.0539 ,0.0564 ,0.0000
        ,0.0075, 0.0206 ,0.0221)
    values=matrix(c(n.vec,alpha.vec,a.s),12,3)
    a=values[values[,1]==n & values[,2]==alpha,3]}

  if (dist=="poisson")
    {a.s=c(0.0930, 0.1056 ,0.1131 ,0.1131 ,0.0401, 0.0502 ,0.0577 ,0.0589, 0.0055,
       0.0155 ,0.0231, 0.0236)
    values=matrix(c(n.vec,alpha.vec,a.s),12,3)
    a=values[values[,1]==n & values[,2]==alpha,3]}

  if (dist=="dpareto")
    {a.s=c( 0.0754, 0.0905, 0.1106 ,0.1131, 0.0238, 0.0389,0.0539, 0.0564, 0.0000,
       0.0065, 0.0201, 0.0221)
    values=matrix(c(n.vec,alpha.vec,a.s),12,3)
    a=values[values[,1]==n & values[,2]==alpha,3]}

  max.a=qnorm(1-(1-alpha)/2)
  b=qnorm((1-alpha)/2+pnorm(a))

  x.sample=x.sample.function(n,dist)
  x=x.sample[1:n]
  s=sd(x)
  mean=x.sample[n+1]

  k=seq(1:n)
  lower=max((cumsum(x)-b*s*sqrt(n))/k)
  upper=max((cumsum(x)-a*s*sqrt(n))/k)

  length=upper-lower
  if (length<=0)  {stop("length can't be negative")}
  if((lower<=mean)&&(upper>=mean)) {in.interval=1} else in.interval=0

  return(c(length,in.interval,CLT(x,n,s,alpha,mean)))}


#   I_4 (INTEGRAL with m=1)
# ------------------------------

int.1=function(dist,n,alpha)
  {b=qnorm(1-alpha/2) # select a value of b according to alpha

  x.sample=x.sample.function(n,dist)
  x=x.sample[1:n]
  s=sd(x)
  mean=x.sample[n+1]

  sum.sum.x=sum(cumsum(x[1:(n-1)]))
```

```
  lower=2/(n-1)*(sum.sum.x/n-b*s*sqrt(n)/sqrt(3)) # Calculate the lower bound of
      the FACI
  upper=2/(n-1)*(sum.sum.x/n+b*s*sqrt(n)/sqrt(3)) # Calculate the upper bound of
      the FACI
  length=upper-lower
  if (length<=0) {stop("length can't be negative or 0")}
  if((lower<=mean)&&(upper>=mean)) {in.interval=1} else in.interval=0

  return(c(length,in.interval,CLT(x,n,s,alpha,mean)))}


#   I_4 (INTEGRAL with m=2)
# ------------------------------

int.2=function(dist,n,alpha)
  {b.s=c(0.765,0.94,1.195,1.655,2.29,2.79)
  alphas=c(0.2,0.15,0.1,0.05,0.02,0.01)
  values=matrix(c(alphas,b.s),length(alphas),2,byrow=F)
  if (alpha %in% alphas == 0) {stop("Choose a different alpha!")}
  b=values[values[,1]==alpha,2] # select a value of b according to alpha

  x.sample=x.sample.function(n,dist)
  x=x.sample[1:n]
  s=sd(x)
  mean=x.sample[n+1]

  sum.x=cumsum(x[1:(n-1)])
  k=seq(1,n-1)
  sum.x.k=sum(k*sum.x)
  sum.x.sq=sum((sum.x)^2)
  c=(n*(n-1)*(2*n-1))/6

  lower=(sum.x.k-sqrt(((sum.x.k)^2)-c*(sum.x.sq-b*n^2*s^2)))/c # Calculate the
      lower bound of the FACI
  upper=(sum.x.k+sqrt(((sum.x.k)^2)-c*(sum.x.sq-b*n^2*s^2)))/c # Calculate the
      upper bound of the FACI
  length=upper-lower
  if (length<=0) {stop("length can't be negative or 0")}
  if((lower<=mean)&&(upper>=mean)) {in.interval=1} else in.interval=0

  return(c(length,in.interval,CLT(x,n,s,alpha,mean)))}


#  I_4 (INTEGRAL with m=3)
# ------------------------------

int.3=function(dist,n,alpha)
  {b.s=c(-0.862, -1.161, -1.658, -2.672, -4.315, -5.708, 0.873, 1.173, 1.667,
      2.684, 4.3, 5.834)
  alphas=c(0.2,0.15,0.1,0.05,0.02,0.01)
  values=matrix(c(alphas,b.s),length(alphas),3,byrow=F)
  if (alpha %in% alphas == 0) {stop("Choose a different alpha!")}
  a=values[values[,1]==alpha,2] # select a value of a according to alpha
  b=values[values[,1]==alpha,3] # select a value of b according to alpha

  x.sample=x.sample.function(n,dist)
  x=x.sample[1:n]
  s=sd(x)
  mean=x.sample[n+1]

  sum.x=cumsum(x[1:(n-1)])
```

```
  k=seq(1,n-1)
  c=s^3*n^(5/2)

  a3=-sum(k^3)/c
  a2=3*sum(k^2*sum.x)/c
  a1=-3*sum(k*sum.x^2)/c
  a0=sum(sum.x^3)/c

  roots.a=solve(polynomial(c(a0-a,a1,a2,a3)))
  upper=as.double(roots.a[Im(roots.a)==0]) # Calculate the upper bound of the FACI
  roots.b=solve(polynomial(c(a0-b,a1,a2,a3)))
  lower=as.double(roots.b[Im(roots.b)==0]) # Calculate the lower bound of the FACI
  length=upper-lower
  if (length<=0) {stop("length can't be negative or 0")}
  if((lower<=mean)&&(upper>=mean)) {in.interval=1} else in.interval=0

  return(c(length,in.interval,CLT(x,n,s,alpha,mean)))}


#   I_4 (INTEGRAL with m=4)
# ------------------------------

int.4=function(dist,n,alpha)
  {b.s=c(1.074,1.571,2.485,4.629,8.748,12.79)
  alphas=c(0.2,0.15,0.1,0.05,0.02,0.01)
  values=matrix(c(alphas,b.s),length(alphas),2,byrow=F)
  if (alpha %in% alphas == 0) {stop("Choose a different alpha!")}
  b=values[values[,1]==alpha,2] # select a value of b according to alpha

  x.sample=x.sample.function(n,dist)
  x=x.sample[1:n]
  s=sd(x)
  mean=x.sample[n+1]

  sum.x=cumsum(x[1:(n-1)])
  k=seq(1,n-1)
  c=s^4*n^3

  a4=sum(k^4)/c
  a3=-4*sum(k^3*sum.x)/c
  a2=6*sum(k^2*sum.x^2)/c
  a1=-4*sum(k*sum.x^3)/c
  a0=sum(sum.x^4)/c-b

  roots=solve(polynomial(c(a0,a1,a2,a3,a4)))
  real.roots=as.double(roots[Im(roots)==0]) # Find bounds of FACI (real roots of
      polynomial)
  lower=min(real.roots)
  upper=max(real.roots)
  length=upper-lower
  if (length<=0) {stop("length should not be negative or 0")}
  if((lower<=mean)&&(upper>=mean)) {in.interval=1} else in.interval=0

  return(c(length,in.interval,CLT(x,n,s,alpha,mean)))}


#   I_4 (INTEGRAL with m=8)
# ------------------------------

int.8=function(dist,n,alpha)
  {b.s=c(3.301,6.832,16.234,53.536,181.368,383.002)
```

```
  alphas=c(0.2,0.15,0.1,0.05,0.02,0.01)
  values=matrix(c(alphas,b.s),length(alphas),2,byrow=F)
  if (alpha %in% alphas == 0) {stop("Choose a different alpha!")}
  b=values[values[,1]==alpha,2] # select a value of b according to alpha

  x.sample=x.sample.function(n,dist)
  x=x.sample[1:n]
  s=sd(x)
  mean=x.sample[n+1]

  sum.x=cumsum(x[1:(n-1)])
  k=seq(1,n-1)
  c=s^8*n^5

  a8=sum(k^8)/c
  a7=-8*sum(k^7*sum.x)/c
  a6=28*sum(k^6*sum.x^2)/c
  a5=-56*sum(k^5*sum.x^3)/c
  a4=70*sum(k^4*sum.x^4)/c
  a3=-56*sum(k^3*sum.x^5)/c
  a2=28*sum(k^2*sum.x^6)/c
  a1=-8*sum(k*sum.x^7)/c
  a0=sum(sum.x^8)/c-b

  roots=solve(polynomial(c(a0,a1,a2,a3,a4,a5,a6,a7,a8)))
  real.roots=as.double(roots[Im(roots)==0]) # Find bounds of FACI (real roots of
      polynomial)
  lower=min(real.roots)
  upper=max(real.roots)
  length=upper-lower
  if (length<=0) {stop("length can't be negative or 0")}
  if((lower<=mean)&&(upper>=mean)) {in.interval=1} else in.interval=0

  return(c(length,in.interval,CLT(x,n,s,alpha,mean)))}


#   I_5 (LINEAR COMBINATION OF FUNCTIONALS)
# ----------------------------------------

w1.int1=function(dist,n,alpha)
  {x.sample=x.sample.function(n,dist)
  x=x.sample[1:n]
  s=sd(x)
  mean=x.sample[n+1]
  i=seq(1,n)

  a.1=1
  a.2=-6*a.1/(n+3)
  a=qnorm(alpha/2,mean=0,sd=sqrt(a.1^2+a.2^2/3+a.1*a.2)) # select "equal tail" cut-
      off points
  b=-a

  lower=(sum(x*(a.1*n+a.2*(n-i)))-b*s*n^(3/2))/(a.1*n^2+a.2*(n*(n-1))/2) #
      Calculate the lower bound of the FACI
  upper=(sum(x*(a.1*n+a.2*(n-i)))-a*s*n^(3/2))/(a.1*n^2+a.2*(n*(n-1))/2) #
      Calculate the upper bound of the FACI
  length=upper-lower
  if (length<=0) {stop("length can't be negative or 0")}
  if((lower<=mean)&&(upper>=mean)) {in.interval=1} else in.interval=0

  return(c(length,in.interval,CLT(x,n,s,alpha,mean)))}
```

B.2 Computing empirical expected lengths and coverage probabilities of FACI's

```
################################################################################
# This code repeats a FACI computation function for 10,000 samples of size n from
# a specified distribution and confidence level. Then it  returns
# (Mean length of FACI)/(Mean length of I_0) and  coverage probabilities
# for the FACI and for I_0.
# This is repeated for 60 different combinations of distribution, alpha, and n.
#
# Run "Functions for computing FACI bounds" first!
################################################################################

library(xtable)
library(polynom)

reps=10000
alphas=c(0.1,0.05,0.02)
n.values=c(50,100,500,1000)

n.vec=rep(n.values,3)
alpha.vec=c(rep(alphas[1],4),rep(alphas[2],4),rep(alphas[3],4))
args=matrix(c(n.vec,alpha.vec),12,2) # Create a matrix of arguments for the
    replication function.

# Choose the m value for the simulations.
m=1 # If you change m, re-run the following "if" lines.
if(exists("faci.fun",mode="function")){rm(faci.fun)} # safety measure: remove
    previously selected faci.fun
if (m=="abs.sup"){faci.fun=abs.sup ; seed=25}
if (m=="sup"){faci.fun=sup ; seed=-171}
if (m=="combination"){faci.fun=w1.int1; seed=10}
if (m=="projection"){faci.fun=projection; seed=399}
if (m==1){faci.fun=int.1;seed=1}
if (m==2){faci.fun=int.2;seed=223}
if (m==3){faci.fun=int.3;seed=35}
if (m==4){faci.fun=int.4;seed=149}
if (m==8){faci.fun=int.8;seed=80}

# The order of the distributions in the next two lines must match
dists=c("normal","exp","pareto","poisson","dpareto")
dists.names=c("$N(0,1)$","Exp(1)","Pareto(1,2)","Poisson(3)","Discrete Pareto")

set.seed(seed)

for (j in 1:5){
  rep.faci=function(arg)
    {sample=replicate(reps,faci.fun(dist=dists[j],n=arg[1],alpha=arg[2]))
      if (m=="combination"| m==1 | m=="projection")
        {return(paste(" (", formatC(sum(sample[2,])/reps,format="f",digits=3), ",",
             formatC(sum(sample[4,])/reps,format="f",digits=3),")" ,sep=""))}
      else
        {return(paste(formatC(mean(sample[1,])/mean(sample[3,]),format="f",digits
            =3) ,
                " (", formatC(sum(sample[2,])/reps,format="f",digits=3), ",",
                    formatC(sum(sample[4,])/reps,format="f",digits=3),")" ,sep=""))
                    }}
```

```
  # Apply the replication function to different alpha and n values and put the
      results in a LaTeX table.
  res.mat=matrix(apply(args,1,rep.faci),nrow=4,ncol=3,F)
  distribution=c("",dists.names[j],"","")
  results=cbind(distribution,n.values,res.mat)
  colnames(results)=c("Distribution","n",alphas[1],alphas[2],alphas[3])
  print(xtable(results), include.rownames=FALSE,sanitize.text.function=function(x){
      x})

  # Remove results from workspace to guarantee the next code doesn't accidentally
      use these results.
  rm(results,res.mat,distribution,rep.faci)}
```

B.3 Computing empirical quantiles of $\int_0^1 W^m(t)\,dt$

```
# Calculation of empirical quantiles of int_0^1(W(t))^m dt, m=3,4,8.

reps=500000
n=10000
alpha=c(0.2,0.15,0.1,0.05,0.02,0.01)
m=3 # Run this code for m=3,4, and 8
set.seed(m)

simulate=function(n,m)
  {p=0.5
  x=rbinom(n,1,p)
  s=sd(x)
  (1/n)*sum((cumsum(x[1:(n-1)]-p)/(s*sqrt(n)))^m)}

lim.var=replicate(reps,simulate(n,m))

###############################################################################
###################                 Even power m          #####################
###############################################################################

if (m%%2 == 0)
  {hist(lim.var, prob=T,breaks=m*100, xlim=c(0,quantile(lim.var,.99,type=8)),xlab="
      ",main="")
  # Sample Quantiles in Statistical Packages
  # Author(s): Rob J. Hyndman and Yanan Fan
  # Recommend using the type=8 in the quantile function
  round(t(as.matrix(quantile(lim.var,1-alpha,type=8))),3)}

###############################################################################
###################                 Odd power m           #####################
###############################################################################

if (m%%2 == 1)
  {hist(lim.var, prob=T,breaks=m*100, xlim=c(quantile(lim.var,.01,type=8),quantile(
      lim.var,.98,type=8)),xlab="",main="")
  # Sample Quantiles in Statistical Packages
  # Author(s): Rob J. Hyndman and Yanan Fan
  # Recommend using the type=8 in the quantile function
  round(t(as.matrix(quantile(lim.var,c(alpha/2,1-alpha/2),type=8), nrows=2,byrow=T)
      ),3)}
```

B.4 Finding minimizing a's for I_3 of (2.23)

```
#############################################################################
# This code checks the average FACI length based on the supremum functional for
# different values of a.
# The objective is to find an optimal a for the simulations.
#
# Run "Functions for computing FACI bounds" first!
#############################################################################

reps=10000
l=51 # number of a's to consider plus 1
alphas=c(0.1,0.05,0.02)
n.values=c(50,100,500,1000)

sup.a.sim=function(arg)
  {n=as(arg[1],"numeric")
  alpha=as(arg[2],"numeric")
  dist=arg[3]
  max.a=qnorm(1-(1-alpha)/2)
  a=seq(0,max.a,length.out=l)
  b=qnorm((1-alpha)/2+pnorm(a))

  lengths=function(n,alpha,a,b)
    {x.sample=x.sample.function(n,dist)
    x=x.sample[1:n]
    s=sd(x)
    k=seq(1:n)

    lower=NULL
    upper=NULL
    for (i in 1:(l-1)) #Ignore last a and b as they produce infinite values
      {lower[i]=max((cumsum(x)-b[i]*s*sqrt(n))/k)
      upper[i]=max((cumsum(x)-a[i]*s*sqrt(n))/k)}
    length=upper-lower
    return(length)}

  results=replicate(reps,lengths(n,alpha,a,b))
  min.a=a[which.min(rowMeans(results))]
  plot(a[1:(l-1)],rowMeans(results),xlab=paste("min a=",round(min.a,4)),ylab="
      Average FACI length",main=paste(arg[3],", alpha=",alpha,", n=",n))
  abline(v=min.a,lty="dotdash")
  return(round(min.a,4))}


set.seed(8238)

##########################          Normal        #####################

par(mfrow=c(3, 4))

n.vec=rep(n.values,3)
alpha.vec=c(rep(alphas[1],4),rep(alphas[2],4),rep(alphas[3],4))
dist.vec=c(rep("normal",12))
args=matrix(c(n.vec,alpha.vec,dist.vec),12,3) # Create a matrix of arguments.
apply(args,1,sup.a.sim)

##########################     Exponential     #####################

par(mfrow=c(3, 4))
```

```r
dist.vec=c(rep("exp",12))
args=matrix(c(n.vec,alpha.vec,dist.vec),12,3) # Create a matrix of arguments.
apply(args,1,sup.a.sim)

##########################    Pareto    #####################

par(mfrow=c(3, 4))

dist.vec=c(rep("pareto",12))
args=matrix(c(n.vec,alpha.vec,dist.vec),12,3) # Create a matrix of arguments.
apply(args,1,sup.a.sim)

##########################    Poisson    #####################

par(mfrow=c(3, 4))

dist.vec=c(rep("poisson",12))
args=matrix(c(n.vec,alpha.vec,dist.vec),12,3) # Create a matrix of arguments.
apply(args,1,sup.a.sim)

######################### Discrete pareto     #####################

par(mfrow=c(3, 4))

dist.vec=c(rep("dpareto",12))
args=matrix(c(n.vec,alpha.vec,dist.vec),12,3) # Create a matrix of arguments.
apply(args,1,sup.a.sim)
```

Bibliography

Abramowitz, M. and I. A. Stegun (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, Volume 55 of *National Bureau of Standards Applied Mathematics Series*. U.S. Government Printing Office, Washington, D.C. (Cited on page 3.)

Casella, G. and R. L. Berger (2002). *Statistical Inference* (2nd ed.). Wadsworth. (Cited on pages 12, 25 and 41.)

Csörgő, M., B. Szyszkowicz, and Q. Wang (2003). Donsker's theorem for self-normalized partial sums processes. *Ann. Probab. 31*(3), 1228–1240. (Cited on page 4.)

Csörgő, S. and L. Horváth (1981). On the Koziol-Green model for random censorship. *Biometrika 68*(2), 391–401. (Cited on pages 28 and 31.)

Csörgő, S. and L. Horváth (1985). The baboons come down from the trees quite normally. In *Mathematical statistics and applications, Vol. B (Bad Tatzmannsdorf, 1983)*, pp. 95–106. Dordrecht: Reidel. (Cited on page 14.)

Erdős, P. and M. Kac (1946). On certain limit theorems of the theory of probability. *Bull. Amer. Math. Soc. 52*, 292–302. (Cited on pages 7 and 21.)

Giné, E., F. Götze, and D. M. Mason (1997). When is the Student t-statistic asymptotically standard normal? *Ann. Probab. 25*(3), 1514–1531. (Cited on page 1.)

Gosset, W. S. (1908). The probable error of a mean. *Biometrika 6*(1), 1–25. (Cited on page 1.)

Karlin, S. and H. M. Taylor (1998). *An introduction to stochastic modeling* (Third ed.). San Diego, CA: Academic Press Inc. (Cited on pages 25 and 39.)

Lévy, P. (1937). *Théorie de l'addition des variables aléatoires.* Gauthier-Villars. (Cited on page 2.)

Martsynyuk, Y. V. (2009a). Functional asymptotic confidence intervals for a common mean of independent random variables. *Electron. J. Stat. 3*, 25–40. (Cited on pages 5, 6 and 7.)

Martsynyuk, Y. V. (2009b). Functional asymptotic confidence intervals for the slope in linear error-in-variables models. *Acta Math. Hungar. 123*(1-2), 133–168. (Cited on page 6.)

Martsynyuk, Y. V. (2013). On the generalized domain of attraction of the multivariate normal law and asymptotic normality of the multivariate Student t-statistic. *J. Multivariate Anal. 114*, 402–411. (Cited on page 3.)