



**HAL**  
open science

# On the Dependence between Quantiles and Dispersion Estimators

Marcel Bräutigam, Marie Kratz

► **To cite this version:**

Marcel Bräutigam, Marie Kratz. On the Dependence between Quantiles and Dispersion Estimators. 2018. hal-02296832

**HAL Id: hal-02296832**

**<https://hal.science/hal-02296832>**

Preprint submitted on 25 Sep 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**ESSEC**  
BUSINESS SCHOOL

*The pioneering spirit*

# ON THE DEPENDENCE BETWEEN QUANTILES AND DISPERSION ESTIMATORS

---

RESEARCH CENTER

MARCEL BRÄUTIGAM, MARIE KRATZ

ESSEC WORKING PAPER 1807

DECEMBER 2018



# On the Dependence between Quantiles and Dispersion Estimators

Marcel Bräutigam<sup>† ‡ \*</sup> and Marie Kratz<sup>† \*</sup>

<sup>†</sup> ESSEC Business School Paris, CREAR

<sup>‡</sup> Sorbone University, LPSM

\* LabEx MME-DII

## Abstract

In this study, we derive the joint asymptotic distributions of functionals of quantile estimators (the non-parametric sample quantile and the parametric location-scale quantile) and functionals of measure of dispersion estimators (the sample standard deviation, sample mean absolute deviation, sample median absolute deviation) - assuming an underlying identically and independently distributed sample. Additionally, for location-scale distributions, we show that asymptotic correlations of such functionals do not depend on the mean and variance parameter of the distribution. Further, we compare the impact of the choice of the quantile estimator (sample quantile vs. parametric location-scale quantile) in terms of speed of convergence of the asymptotic covariance and correlations respectively. As application, we show in simulations a good finite sample performance of the asymptotics. Further, we show how the theoretical dependence results can be applied to the most well-known risk measures (Value-at-Risk, Expected Shortfall, expectile). Finally, we relate the theoretical results to empirical findings in the literature of the dependence between risk measure prediction (on historical samples) and the estimated volatility.

2010 AMS classification: 60F05; 62H10; 62H20; 62P05; 62P20; 91B30

JEL classification: C13; C14; C30; C58; C69; G32

*Keywords:* asymptotic distribution; sample quantile; measure of dispersion; non-linear dependence; VaR; ES; correlation;

# Contents

<b>1</b>	<b>Introduction and Notation</b>	<b>4</b>
<b>2</b>	<b>Asymptotic Joint Properties of Quantile and Dispersion Estimators for iid rv's</b>	<b>9</b>
2.1	Historical Estimation . . . . .	9
2.2	Location-Scale Quantile . . . . .	11
2.3	Specific Cases and Methods . . . . .	14
2.4	The Effect of Sample Size in Estimation . . . . .	15
<b>3</b>	<b>Case of Location-Scale Distributions</b>	<b>18</b>
3.1	Historical Estimation . . . . .	19
3.2	Location-scale Quantile . . . . .	24
3.3	The Impact of the Choice of the Quantile Estimator for Elliptical Distributions . . . . .	28
3.3.1	Gaussian Distribution . . . . .	28
3.3.2	Student Distribution . . . . .	30
<b>4</b>	<b>Application to Quantitative Risk Management</b>	<b>34</b>
4.1	Estimation with Various Risk and Dispersion Measures . . . . .	35
4.1.1	Dispersion Measures . . . . .	35
4.1.2	Risk Measures . . . . .	36
4.1.3	Implications of the Theoretical Results . . . . .	37
4.2	Example in Finance: Explaining Procyclicality in Risk Measurements . . . . .	42
4.2.1	Gaussian Distribution . . . . .	43
4.2.2	Student Distribution . . . . .	46
4.2.3	Impact of Using Longer Samples . . . . .	50
4.2.4	Comparison with Empirical Work . . . . .	52
<b>5</b>	<b>Concluding Remarks: Implications for Use in Practice</b>	<b>55</b>
	<b>Appendix A Proofs of Section 2</b>	<b>59</b>
A.1	Proofs of Subsection 2.1 . . . . .	59

A.1.1	Bahadur’s Method . . . . .	59
A.1.2	Taylor’s Method . . . . .	65
A.2	Proofs of Subsection 2.2 . . . . .	67
A.3	Proofs of Subsection 2.3 . . . . .	70
A.4	Proof of Subsection 2.4 . . . . .	71
<b>Appendix B</b>	<b>Proofs of Section 3</b>	<b>73</b>
B.1	Proofs of Subsection 3.1 . . . . .	73
B.2	Proofs of Subsection 3.2 . . . . .	77
<b>Appendix C</b>	<b>Proofs of Section 4</b>	<b>79</b>
<b>Appendix D</b>	<b>Miscellanea</b>	<b>86</b>
D.1	Extensions of Theorem 1 . . . . .	86
D.2	Tables of Finite Sample Size Simulation Study . . . . .	87

# 1 Introduction and Notation

In this paper we analyse the asymptotic dependence properties of functionals of quantile estimators with functionals of measure of dispersion estimators for underlying identically and independently distributed (iid) samples. By measures of dispersion we mean well-known quantities as the variance or standard deviation, but also less frequently used ones as, for example, the mean or median absolute deviation - we introduce and formalise those concepts in the next subsection.

One motivation to look at this dependence comes from the empirical study in [9]. Therein the correlation between a log-ratio of sample quantiles with the sample standard deviation is measured empirically using log-returns from different stock indices. Note that looking at a functional of the sample quantile (as e.g. the logarithm) is more general than considering only the linear correlation between sample quantile and sample standard deviation. Another reason of considering those functionals is that the results directly relate to risk measure estimation (and their dependence with volatility measures, represented in this case by measures of dispersion): The sample quantile can be seen as a Value-at-Risk (VaR) estimator and the functional framework allows us to extend the results also to Expected Shortfall (ES).

In this spirit the results presented in the paper treat in a broad way the joint distribution and linear correlation of general functionals of quantile estimators and functionals of measure of dispersion estimators. The reason why we consider also other measures of dispersion than the classical sample variance is that using this latter brings asymptotic constraints (such as the existence of the fourth moment of the underlying distribution) which can be relaxed when using other measures like the mean absolute deviation or median absolute deviation. It is quite important in practice, as pointed out in Section 4 in the context of quantitative risk management (QRM) and risk measures: We illustrate how results on the asymptotic dependence between VaR, ES or the expectile with these measures of dispersion follow from the results presented. Also, coming back to the first motivation of this research, theoretical results for the dependence between the log-ratio of sample quantiles and the sample standard deviation are presented - as counterpart to the empirical estimation in [9]. We also link these theoretical results to empirical findings in [43, 44].

In the existing literature there exist works on the asymptotic joint distribution of the sample mean and the sample quantile ([29], [17] for two different approaches). Still, we only found one paper, [8], which looks as a biproduct at the joint distribution of a vector of sample quantiles with the  $r$ -th absolute sample moment in the case of an underlying Gaussian distribution with known mean. Although their framework seems to include the joint distribution of a sample quantile with the sample variance (second absolute sample moment) as well as with the sample mean absolute deviation (first absolute sample moment), note that the work in [8] differs as it assumes the mean to be known and a Gaussian underlying distribution. Moreover, in the case of the joint distribution of the sample quantile with the sample median absolute deviation, to the best of our knowledge, no result exists.

The structure of the paper is as follows. We present in Section 2 the main result about the asymptotic joint distribution and dependence between functionals of quantile estimators and functionals of three different measures of dispersion, in the following order: sample variance, sample mean absolute deviation and sample median absolute deviation. First, we consider the sample quantile as quantile estimator. We show its joint asymptotic distribution with the above mentioned sample measures of dispersion using their existing Bahadur representations. Those results presented are valid for all distributions (which fulfil certain moment and smoothness conditions). Next, we consider the parametric location-scale quantile estimator, which means to restrict ourselves to the class of location-scale distributions, an important class in applications. Hence, we subsequently present the asymptotic joint distributions for functionals of the three different measure of dispersion estimators considered, but now with functionals of the

location-scale quantile estimator. We conclude this section by analysing the effect of the sample size, first when the quantile estimator and the measure of dispersion estimator each have different sample sizes (in an asymptotic sense). Second, we evaluate the finite sample approximation of the asymptotics: In a simulation study we compare the sample correlation between quantile estimators and measure of dispersion estimators (each on a finite sample) to the theoretical asymptotic correlation - considering elliptical distributions with light and heavy tails, respectively.

To compare the two quantile estimators considered, we focus in Section 3 on location-scale distributions as this is a broad group of distributions often used in financial modelling and from which the theoretical question arised. In Subsections 3.1 and 3.2 we present the asymptotic covariances and correlations for functionals of the sample quantile and location-scale quantile respectively with each of the three measures of dispersion. Further, we provide all the explicit asymptotic covariances and correlations in the cases of a Gaussian and a Student distribution. As described, one has two different alternatives of using (consistent) estimators to estimate the underlying true quantile. The parametric location-scale quantile is especially of interest as this estimator has a higher speed of convergence than the sample quantile. Still, such a parametric estimator needs to assume a specific underlying distribution which might add some model risk in practice. With regard to this, we present an important result in Subsection 3.3. Therein we show the asymptotic difference in the dependence structure (for each corresponding measure of dispersion) when using the one or the other quantile estimator for the two main examples considered, the Gaussian and the Student distribution.

In Section 4, we apply the results from the previous sections in the context of quantitative risk management. In a first part, we make the link to risk measures apparent. Taking as example the three most popular risk measures in quantitative risk management, Value-at-Risk, Expected Shortfall and expectile, we show their asymptotic dependence with each of the three measure of dispersion estimators. Given the constant debate about the properties and advantages of these different risk measures (see e.g. [10], [15], and the references therein), this adds another layer to the discussion. Again, in the cases of an underlying Gaussian or Student distribution, we provide the explicit formulae and also plots of the asymptotic correlation in these cases. In the second part, we look at a specific problem in finance: Therein we consider the non-linear dependence looking at the linear correlation between log-ratios of Value-at-Risk estimators (i.e. sample quantiles) and the sample standard deviation (i.e. a non-linear dependence between the sample quantile and the sample standard deviation). This was computed empirically in [9] to assess the pro-cyclicality of risk measurement using historical data from different stock indices. We also relate this dependence between sample quantile and sample standard deviation to the empirical work in [43, 44]. We then provide theoretical values in the setup given in [9] when assuming underlying iid models. Additionally to the asymptotic correlation between the sample quantile and sample standard deviation, we consider this correlation when using the different measures of dispersion introduced as well as with the parametric location-scale quantile instead of the sample quantile. Finally, we show how these theoretical results scale when using longer samples (in an asymptotic meaning that we will explain) and compare finite sample empirical results, in an identical setting as in [9], with the theoretical asymptotics. We conclude in Section 5 discussing implications of the different results for their use in practice, also giving a concise comparison of the theoretical conditions on the underlying distribution (of the asymptotic results) depending on the different quantile and dispersion estimators used.

All proofs are developed in the Appendix, so that the series of results appears clearly in the main body of the paper to offer a statistical tool set.

Let us end with a remark about the structure. The new and general results about the joint asymptotics of the sample quantile (location-scale quantile) with different measures of dispersion are in Section 2 and hold for any distribution (in particular for the case of location-scale distributions) - with certain

restrictions on its smoothness and existence of moments. Then we apply those results to the specific cases of location-scale distributions in Section 3 and each time explicitly consider the three different cases of our measure of dispersion separately with examples, although they can be derived from the main theorems in Section 2. We are aware it extends the paper, but think it may be a useful and good reference in applications with specific results being explicitly available.

## Notation

Let us start with common mathematical notation. The rounded-up and rounded-off integer-parts of a real number  $x \in \mathbb{R}$  are denoted, respectively, as

$$\lceil x \rceil = \min \{m \in \mathbb{Z} : m \geq x\} \quad \text{and} \quad \lfloor x \rfloor = \max \{m \in \mathbb{Z} : m \leq x\}.$$

The signum function is denoted by  $\text{sgn}$  and defined, as usual, by

$$\text{sgn}(x) := \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

Further,  $\Gamma(\cdot)$  is the Gamma function; we will need it either for integers or half-integers, i.e.

$$\Gamma(x) := \begin{cases} (x-1)! & \text{for integers } x > 0, \\ \sqrt{\pi} \frac{(2x-2)!!}{2^{\frac{2x-1}{2}}} & \text{for half-integers } x, \text{ i.e. odd integer-multiples of } \frac{1}{2}, \end{cases}$$

where  $!$  and  $!!$  denote the factorial and double-factorial function, respectively. Also, we denote by  $\xrightarrow{d}, \xrightarrow{P}$  the convergence in distribution and in probability, respectively.

Focusing on statistical quantities, let us work with a sample of size  $n$ ,  $(X_1, \dots, X_n)$ , with parent random variable (rv)  $X$ , parent cumulative distribution function (cdf)  $F_X$ , (and, given they exist,) probability density function (pdf)  $f_X$ , mean  $\mu$ , variance  $\sigma^2$ , and quantile of order  $p$  as  $q_X(p) := \inf\{x \in \mathbb{R} : p \leq F_X(x)\}$ . We denote its ordered sample by  $X_{(1)} \leq \dots \leq X_{(n)}$ . Whenever it exists, we introduce the standardised version (with mean 0 and variance 1) of  $X$ ,  $Y := \frac{X-\mu}{\sigma}$ , and correspondingly the cdf, pdf and quantile of order  $p$  as  $F_Y, f_Y$  and  $q_Y(p)$ . In the special case of the standard normal distribution  $\mathcal{N}(0, 1)$ , we use the standard notation  $\Phi, \phi, \Phi^{-1}(p)$  for the cdf, pdf and quantile of order  $p$ , respectively. We use the symbol  $\sim$  for ‘distributed as’, i.e.  $Y \sim \mathcal{N}(0, 1)$  means that  $Y$  is  $\mathcal{N}(0, 1)$ -distributed.

In this paper, we focus on the following five estimators. First, we consider the three estimators of the dispersion: (i) the sample variance  $\hat{\sigma}_n^2$ , (ii) the sample mean absolute deviation around the sample mean (MAD)  $\hat{\theta}_n$ , and (iii) the sample median absolute deviation around the sample median (MedianAD)  $\hat{\xi}_n$ :

- (i)  $\hat{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ , where  $\bar{X}_n = \frac{1}{n} \sum_{j=1}^n X_j$ ,
- (ii)  $\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n |X_i - \bar{X}_n|$ ,
- (iii)  $\hat{\xi}_n = \frac{1}{2}(W_{(\lfloor \frac{n+1}{2} \rfloor)} + W_{(\lfloor \frac{n+2}{2} \rfloor)})$ , where  $W_i = |X_i - \hat{\nu}_n|, i = 1, \dots, n$  (and its ordered sample  $W_{(1)} \leq \dots \leq W_{(n)}$ ), and  $\hat{\nu}_n = \frac{1}{2}(X_{(\lfloor \frac{n+1}{2} \rfloor)} + X_{(\lfloor \frac{n+2}{2} \rfloor)})$  (sample median of the original sample).

Note that, for the sample MAD, we could have taken a factor of  $\frac{1}{n-1}$  instead of  $\frac{1}{n}$ , and this does not matter asymptotically. We chose the latter to be in line with the literature (see e.g. [20],[34],[39]), although we think, conceptually,  $\frac{1}{n-1}$  does make more sense.



To unify the notation of sample variance and the sample MAD, one can use the notion of the  $r$ -th absolute centred sample moment, i.e.

$$\hat{m}(X, n, r) := \frac{1}{n} \sum_{i=1}^n |X_i - \bar{X}_n|^r, \quad \text{with } r = 1, 2, \quad (1)$$

so that  $\hat{\sigma}_n^2 = \frac{n-1}{n} \hat{m}(X, n, 2)$  and  $\hat{\theta}_n = \hat{m}(X, n, 1)$ . To have a framework that incorporates the three measures of dispersion, we introduce the notation  $D_{i,n}$ , for  $i = 1, 2, 3$ , with  $\hat{D}_{1,n} = \hat{\sigma}_n^2$  or  $\sqrt{\hat{D}_{1,n}} = \hat{\sigma}_n$  the sample variance or sample standard deviation respectively, with  $\hat{D}_{2,n} = \hat{\theta}_n$  the sample MAD, and  $\hat{D}_{3,n} = \hat{\xi}_n$  the sample MedianAD. Correspondingly, the theoretical counterparts are denoted as  $D_i$ ,  $i = 1, 2, 3$  with  $D_1 = \sigma^2$ ,  $\sqrt{D_1} = \sigma$ ,  $D_2 = \theta$ ,  $D_3 = \xi$ .

Then, we consider two estimators of the quantile of order  $p$  of the distribution:

- (iv) the sample quantile  $q_n(p) = X_{(\lceil np \rceil)}$ ,
- (v) the quantile estimated parametrically by its location-scale model whenever  $F_X$  belongs to the location-scale family of distributions:  $q_{n,\hat{\mu},\hat{\sigma}}(p) = \hat{\mu}_n + \hat{\sigma}_n q_Y(p)$ , where  $\hat{\mu}_n$  and  $\hat{\sigma}_n$  are any estimators of the mean  $\mu$  and standard deviation  $\sigma$ . Here we choose them to be the sample mean  $\bar{X}_n$  and the square root of the sample variance  $\hat{\sigma}_n$  (by abuse of notation) respectively. When assuming  $\mu$  to be known, we write  $q_{n,\hat{\sigma}}(p) = \mu + \hat{\sigma}_n q_Y(p)$ .

Recall that the location-scale family of distributions  $\mathcal{F}$  is the class of distributions such that

$$\text{if } F \in \mathcal{F}, \text{ then for any } a \in \mathbb{R}, 0 < b < \infty, G(x) := F(ax + b) \in \mathcal{F}. \quad (2)$$

Again, to have a unified notation, we introduce the quantile estimator  $\hat{q}_n$  that should represent either the sample quantile  $q_n$ , or the location-scale quantile models  $q_{n,\hat{\mu},\hat{\sigma}}$  (or  $q_{n,\hat{\sigma}}$  for  $\mu$  known).

In addition, to be consistent in the notation with related results in the literature, we generalise a notation used in [6, 17]: First recall that the quantile can be written as a minimiser (see [25])

$$q_X(p) = \operatorname{argmin}_{a \in \mathbb{R}} \mathbb{E}[L_p(X - a)],$$

for the so called quantile loss function  $L_p(x) := (p - \mathbf{1}_{(x \leq 0)})x$ . Hence, one can consider the case of  $a = q_X(p)$  and write

$$\mathbb{E}[L_p(X - q_X(p))] = p (\mathbb{E}[X] - \mathbb{E}[X | X \leq q_X(p)])$$

called ‘minimized expected quantile loss function’ in [6], and denoted by  $\tau(p)$  in [17], where it appeared as part of the asymptotic covariance of the sample quantile with the sample mean.

So, assuming that the underlying rv  $X$  has finite moments up to order  $l$ , and that  $\eta$  is a continuous real-valued function, we use in this paper the following generalization: for  $1 \leq k \leq l$  and  $p \in (0, 1)$ ,

$$\tau_k(\eta(X), p) = (1 - p) \left( \mathbb{E}[\eta^k(X) | X > q_X(p)] - \mathbb{E}[\eta^k(X)] \right), \quad (3)$$

$$= p(1 - p) \left( \mathbb{E}[\eta^k(X) | X > q_X(p)] - \mathbb{E}[\eta^k(X) | X < q_X(p)] \right), \quad (4)$$

where we provided (4) to point out that this quantity involves the truncated moments of both tails. We abbreviate, when  $\eta$  is the identity function,  $\tau_k(X, p)$  as  $\tau_k(p)$ .

Further, note that for the results we present, we need different conditions on the underlying random variable  $X$  depending on the choice of quantile estimator and measure of dispersion estimator. Although not always needed for the theoretical results (if interested in minimal requirements, see Table 9 in Section 4), we will assume for simplicity that

(C<sub>1</sub>)  $F_X$  is absolutely continuous and strictly monotonically increasing,

i.e.  $X$  has a pdf  $f_X(x)$  such that, both,  $0 < f_X(x) < \infty$  and  $F_X(q_X(p)) = p$  almost everywhere. Two other conditions, which do not need to hold for all  $x \in \mathbb{R}$  (but in given intervals), are

(C<sub>2</sub>)  $F_X$  is Hölder-continuous,

(C<sub>3</sub>) the second derivative of  $F_X$  exists and is bounded

Also, we denote necessary moment conditions, for an integer  $k > 0$ , as follows

(M<sub>k</sub>)  $\mathbb{E}[X^{2k}] < \infty$  (i.e. finite 2k-th moment)

and will usually need in this study (M<sub>1</sub>) (i.e. existence of second moment) or (M<sub>2</sub>) (existence of the fourth moment).

Finally, to have results as general as possible (in view of applications), all along the paper we will consider functionals  $h_1, h_2$  of the estimators that we assume to be continuous real-valued functions with existing derivatives denoted by  $h'_1$  and  $h'_2$  respectively. Note that in fact, to apply the Delta method, it suffices for the derivatives to exist only at one point - the point where they are evaluated at. We will omit recalling it in the conditions of the results.

To end this section, we summarise for convenience the notation of the different statistical quantities with their corresponding estimators in Table 1.

Table 1: Notation of statistical quantities and their (possibly various) estimators used in this paper

Statistical quantities		Corresponding Estimators	
mean	$\mu$	sample mean	$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$
variance	$\sigma^2$	sample variance (unknown $\mu$ )	$\hat{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$
		(known $\mu$ )	$\tilde{\sigma}_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$
mean absolute deviation (MAD)	$\theta = \mathbb{E}[ X - \mu ]$	Sample MAD (unknown $\mu$ )	$\hat{\theta}_n = \frac{1}{n} \sum_{i=1}^n  X_i - \bar{X}_n $
		(known $\mu$ )	$\tilde{\theta}_n = \frac{1}{n} \sum_{i=1}^n  X_i - \mu $
r-th centred/central moment	$\mu_r = \mathbb{E}[(X - \mu)^r]$		
r-th absolute centred moment (measure of dispersion)	$m(X, r) = \mathbb{E}[ X - \mu ^r]$	Sample measure of dispersion (unknown $\mu$ )	$\hat{m}(X, n, r) = \frac{1}{n} \sum_{i=1}^n  X_i - \bar{X}_n ^r$
		Sample measure of dispersion (known $\mu$ )	$\tilde{m}(X, n, r) = \frac{1}{n} \sum_{i=1}^n  X_i - \mu ^r$
cdf	$F_X(x)$	empirical cdf	$F_n(x) = F_{n,X}(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{(X_i \leq x)}$
pdf	$f_X(x)$		
quantile of order p	$q_X(p) = F_X^{-1}(p)$	sample quantile	$q_n(p) = X_{(\lceil np \rceil)}$
		parametric location-scale quantile (unknown $\mu$ )	$q_{n,\hat{\mu},\hat{\sigma}}(p) = \hat{\mu}_n + \hat{\sigma}_n q_Y(p)$
		parametric location-scale quantile (known $\mu$ )	$q_{n,\hat{\sigma}}(p) = \mu + \hat{\sigma}_n q_Y(p)$
median	$\nu = q_X(1/2)$	sample median	$\hat{\nu}_n = \frac{1}{2} (X_{(\lfloor \frac{n+1}{2} \rfloor)} + X_{(\lfloor \frac{n+2}{2} \rfloor)})$
median absolute deviation (MedianAD)	$\xi = q_{ X-\nu }(1/2)$ where $F_{ X-\nu }(x) = F_X(\nu+x) - F_X(\nu-x)$	sample MedianAD	$\hat{\xi}_n = \frac{1}{2} (W_{(\lfloor \frac{n+1}{2} \rfloor)} + W_{(\lfloor \frac{n+2}{2} \rfloor)})$

## 2 Asymptotic Joint Properties of Quantile and Dispersion Estimators for iid rv's

Having fixed the general notation, let us present our main results. We divide them into two subsections according to the quantile estimator we used. We start with the asymptotic joint distributions when using a sample quantile. Subsequently, we present the analogous results when using a location-scale quantile. As mentioned in the introduction, the results with the sample quantile hold for any iid distributions while the ones with the location-scale quantile only for location-scale distributions (both being subject to some smoothness and moment conditions).

### 2.1 Historical Estimation

We present two results. The first one for the relation of functionals of the sample quantile with functionals of either the sample MAD or the sample variance in a unified setting. The second for functionals of the sample quantile with functionals of the sample MedianAD.

**Theorem 1** *Consider an iid sample with parent rv  $X$  having mean  $\mu$ , variance  $\sigma^2$ . Assume conditions  $(C_1), (C_3)$  in a neighbourhood of  $q_X(p)$  and  $(M_r)$  for  $r = 1, 2$  respectively as well as  $(C_2)$  at  $\mu$  for  $r = 1$ . Then the joint behaviour of the functionals  $h_1$  of the sample quantile  $q_n(p)$  (for  $p \in (0, 1)$ ) and  $h_2$  of the sample measure of dispersion  $\hat{m}(X, n, r)$  (defined in (1)), is asymptotically normal:*

$$\sqrt{n} \begin{pmatrix} h_1(q_n(p)) - h_1(q_X(p)) \\ h_2(\hat{m}(X, n, r)) - h_2(m(X, r)) \end{pmatrix} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Sigma^{(r)}), \quad (5)$$

where the asymptotic covariance matrix  $\Sigma^{(r)} = (\Sigma_{ij}^{(r)}, 1 \leq i, j \leq 2)$  satisfies

$$\Sigma_{11}^{(r)} = \frac{p(1-p)}{f_X^2(q_X(p))} (h_1'(q_X(p)))^2; \quad \Sigma_{22}^{(r)} = (h_2'(m(X, r)))^2 \text{Var}(|X - \mu|^r + (2-r)(2F_X(\mu) - 1)X); \quad (6)$$

$$\Sigma_{12}^{(r)} = \Sigma_{21}^{(r)} = h_1'(q_X(p)) h_2'(m(X, r)) \times \frac{\tau_r(|X - \mu|, p) + (2-r)(2F_X(\mu) - 1)\tau_1(p)}{f_X(q_X(p))}, \quad (7)$$

where  $\tau_r$  is defined in (3).

The asymptotic correlation between the functional  $h_1$  of the sample quantile and the functional  $h_2$  of the measure of dispersion is - up to its sign  $a_{\pm}$  - the same whatever the choice of  $h_1, h_2$ :

$$\lim_{n \rightarrow \infty} \text{Cor}(h_1(q_n(p)), h_2(\hat{m}(X, n, r))) = a_{\pm} \times \frac{\tau_r(|X - \mu|, p) + (2-r)(2F_X(\mu) - 1)\tau_1(p)}{\sqrt{p(1-p) \text{Var}(|X - \mu|^r + (2-r)(2F_X(\mu) - 1)X)}}, \quad (8)$$

where  $a_{\pm} = \text{sgn}(h_1'(q_X(p)) \times h_2'(m(X, r)))$ .

**Remark 2** Notice that in the case the functionals  $h_1, h_2$  are such that  $h_1'(q_X(p)) h_2'(m(X, r)) = 0$ , we have asymptotic linear independence: Equations (7) and (8) will equal zero (as  $\text{sgn}(0) = 0$  by definition). This remark holds for all other results involving the functionals  $h_1, h_2$  too, but we will not mention it each time.

Note that apart from choosing the dependence with the sample variance ( $r = 2$ ) or sample MAD ( $r = 1$ ), the choice of  $r$  also has an impact on the required existence of moments of  $X$ . Indeed, with  $r = 2$ , we require the existence of the fourth moment of  $X$ , while, with  $r = 1$ , only a finite second moment.

Further, note that we could have formulated Theorem 1 in a more general form using a function  $h(x, y) = \begin{pmatrix} h_1(x, y) \\ h_2(x, y) \end{pmatrix}$  or considering a vector-valued version of the theorem (vector of sample quantiles), but preferred for readability the above presentation (see Appendix D for such extensions).

The case of the dependence between functionals of the sample quantile  $q_n(p)$  and MedianAD  $\hat{\xi}_n$  cannot be included in the same mathematical framework and needs to be formulated separately. The different dependent structure appears clearly in the expressions of the covariance and correlation (e.g. involving maxima - something we did not have in Theorem 1).

**Proposition 3** Consider an iid sample with parent rv  $X$  with, if defined, mean  $\mu$  and variance  $\sigma^2$ . Under conditions  $(C_1)$  and  $(C_3)$  in the neighbourhoods of  $q_X(p), \nu, \nu \pm \xi$ , the joint behaviour of the functionals  $h_1$  of the sample quantile  $q_n(p)$  (for  $p \in (0, 1)$ ) and  $h_2$  of the sample MedianAD  $\hat{\xi}_n$  (defined in Table 1) is asymptotically normal:

$$\sqrt{n} \begin{pmatrix} h_1(q_n(p)) - h_1(q_X(p)) \\ h_2(\hat{\xi}_n) - h_2(\xi) \end{pmatrix} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Gamma),$$

where the asymptotic covariance matrix  $\Gamma = (\Gamma_{ij}, 1 \leq i, j \leq 2)$  satisfies

$$\Gamma_{11} = \frac{p(1-p)}{f_X^2(q_X(p))} (h_1'(q_X(p)))^2; \quad \Gamma_{22} = \frac{1 + \gamma/f_X^2(\nu)}{4(f_X(\nu + \xi) + f_X(\nu - \xi))^2} (h_2'(\xi))^2; \quad (9)$$

$$\Gamma_{12} = \Gamma_{21} = h_1'(q_X(p)) h_2'(\xi) \times \frac{-\max(0, F_X(\nu + \xi) - \max(F_X(\nu - \xi), p)) + \frac{1-p}{2} + \frac{f_X(\nu + \xi) - f_X(\nu - \xi)}{f_X(\nu)} \max\left(-\frac{p}{2}, \frac{p-1}{2}\right)}{f_X(q_X(p)) (f_X(\nu + \xi) + f_X(\nu - \xi))} \quad (10)$$

where  $\gamma = (f_X(\nu + \xi) - f_X(\nu - \xi)) f_X(\nu) (f_X(\nu + \xi) - f_X(\nu - \xi) - 4) (1 - F_X(\nu - \xi) - F_X(\nu + \xi))$  and  $\nu$  denotes the median of the sample (see Table 1).

The asymptotic correlation between the two functionals is - up to its sign  $a_{\pm}$  - the same whatever the choice of  $h_1, h_2$ :  $\lim_{n \rightarrow \infty} \text{Cor} \left( h_1(q_n(p)), h_2(\hat{\xi}_n) \right) =$

$$a_{\pm} \times \frac{-\max(0, F_X(\nu + \xi) - \max(F_X(\nu - \xi), p)) + \frac{1-p}{2} + \frac{f_X(\nu + \xi) - f_X(\nu - \xi)}{f_X(\nu)} \max\left(-\frac{p}{2}, \frac{p-1}{2}\right)}{\sqrt{\frac{p(1-p)}{4}} \sqrt{1 + \frac{\gamma}{f_X^2(\nu)}}} \quad (11)$$

where  $a_{\pm} := \text{sgn}(h'_1(q_X(p)) h'_2(\xi))$ .

Two remarks with respect to the result presented. First, note that the asymptotic dependence with the sample MedianAD does not even require a finite mean. Second, for symmetric distributions, it holds that  $f_X(\nu + \xi) = f_X(\nu - \xi)$  and  $\gamma = 0$ , so the expressions of  $\Gamma$  and (11) simplify a lot.

## 2.2 Location-Scale Quantile

As a comparison to using historical estimation via sample quantiles (denoted by  $q_n$ ), let us estimate the quantile via the known analytical formula for the quantile of the model when considering a given location-scale distribution with unknown but finite mean  $\mu$  and variance  $\sigma^2$  as defined in equation (2). Consequently, we can write the quantile of order  $p$  in such cases as

$$q_X(p) = \mu + \sigma q_Y(p), \quad (12)$$

where  $Y$  is the corresponding rv with standardised distribution with mean 0 and variance 1.

Hence, if we estimate  $\mu$  by the sample mean  $\bar{X}_n$  and  $\sigma$  by the square-root of the sample variance,  $\hat{\sigma}_n$ , the quantile estimator based on (12) can be written as

$$\hat{q}_{n, \hat{\mu}, \hat{\sigma}}(p) = \bar{X}_n + \hat{\sigma}_n q_Y(p). \quad (13)$$

In the case  $\mu$  is known, this reduces to  $q_{n, \hat{\sigma}}(p) = \mu + \hat{\sigma}_n q_Y(p)$ , and studying the dependence with the dispersion measure estimators will be simpler.

We keep the same structure as in Subsection 2.1: First, we present a unified result for the dependence of (here) the location-scale quantile with the sample variance or sample MAD. Then, we present the corresponding result when using the sample MedianAD. Since we distinguish the cases  $\mu$  unknown from known, we always start with the more general case ( $\mu$  unknown), and, as a Corollary, present the case with known mean  $\mu$ .

Let us start with presenting the analogon of Theorem 1 for functionals of the location-scale quantile estimator.

**Proposition 4** Consider an iid sample with parent rv  $X$  having mean  $\mu$  and variance  $\sigma^2$ . Assume conditions  $(C_1), (M_2)$ , as well as  $(C_2)$  at  $\mu$  for  $r = 1$ . Then, taking  $r = 1, 2$ , the joint behaviour of the functionals  $h_1$  of the quantile  $q_{n, \hat{\mu}, \hat{\sigma}}(p)$  from a location-scale model (for  $p \in (0, 1)$ ) and  $h_2$  of the measure of dispersion  $\hat{m}(X, n, r)$  (defined in Table 1) is asymptotically normal:

$$\sqrt{n} \begin{pmatrix} h_1(q_{n, \hat{\mu}, \hat{\sigma}}(p)) - h_1(q_X(p)) \\ h_2(\hat{m}(X, n, r)) - h_2(m(X, r)) \end{pmatrix} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Lambda^{(r)}), \quad (14)$$

where the asymptotic covariance matrix  $\Lambda^{(r)} = (\Lambda_{ij}^{(r)}, 1 \leq i, j \leq 2)$  satisfies

$$\Lambda_{11}^{(r)} = \sigma^2 (h_1'(q_X(p)))^2 \left( 1 + q_Y(p) \left( q_Y(p) (\mathbb{E}[Y^4] - 1) / 4 + \mathbb{E}[Y^3] \right) \right); \quad (15)$$

$$\Lambda_{22}^{(r)} = (h_2'(m(X, r)))^2 \text{Var} \left( |X - \mu|^r + (2 - r)(2F_X(\mu) - 1)X \right); \quad (16)$$

$$\Lambda_{12}^{(r)} = \Lambda_{21}^{(r)} = \sigma^{r+1} h_1'(q_X(p)) h_2'(m(X, r)) \times \quad (17)$$

$$\left( \mathbb{E}[Y^{r+1}] + (2 - r)(2F_Y(0) - 1 - 2\mathbb{E}[|Y|^{r+1} \mathbf{1}_{(Y < 0)}]) + \frac{q_Y(p)}{2} \left( \mathbb{E}[|Y|^{r+2}] - \mathbb{E}[|Y|^r] + (2 - r)(2F_Y(0) - 1) \mathbb{E}[Y^3] \right) \right).$$

The asymptotic correlation between the functional  $h_1$  of the sample quantile and the functional  $h_2$  of the measure of dispersion is - up to its sign  $a_{\pm}$  - the same whatever the choice of  $h_1, h_2$ :

$$\lim_{n \rightarrow \infty} \text{Cor} (h_1(q_{n, \hat{\mu}, \hat{\sigma}}(p)), h_2(\hat{m}(X, n, r))) = a_{\pm} \times \quad (18)$$

$$\frac{\mathbb{E}[Y^{r+1}] + (2 - r)(2F_Y(0) - 1 - 2\mathbb{E}[|Y|^{r+1} \mathbf{1}_{(Y < 0)}]) + \frac{q_Y(p)}{2} \left( \mathbb{E}[|Y|^{r+2}] - \mathbb{E}[|Y|^r] + (2 - r)(2F_Y(0) - 1) \mathbb{E}[Y^3] \right)}{\sqrt{\left( 1 + q_Y(p) \left( q_Y(p) \frac{\mathbb{E}[Y^4] - 1}{4} + \mathbb{E}[Y^3] \right) \right) \text{Var} \left( |Y|^r + (2 - r)(2F_Y(0) - 1)Y \right)}}$$

where  $a_{\pm} = \text{sgn}(h_1'(q_X(p)) \times h_2'(m(X, r)))$ . Hence, the asymptotic correlation does not depend on the mean  $\mu$  and the variance  $\sigma^2$  of the underlying location-scale distribution, and, as before, is - up to its sign - independent of the specific choice of  $h_1, h_2$ .

Further, if we assume the location-scale distribution to be symmetric, then the asymptotic correlation between the location-scale quantile  $q_{n, \hat{\mu}, \hat{\sigma}}$  and the measure of dispersion estimator can be proved to have its minimum absolute correlation (value of 0) for  $p = 0.5$  and to be (point-)symmetric with respect to  $p = 0.5$ .

Note that using the location-scale quantile model implies assuming the existence of a finite fourth moment with the sample variance and with the sample MAD - this is in contrast to the historical estimation with the sample quantile.

Further, while equations (17) and (18) for covariance and correlation simplify when either having a symmetric distribution or with zero third central moment  $\mu_3$  (e.g. for elliptical distributions), it is also worth looking specifically at the case when the mean of the distribution is known, thus estimating the quantile as  $q_{n, \hat{\sigma}} = \mu + q_Y(p) \hat{\sigma}_n$ :

**Corollary 5** Under the same conditions and notations as in Proposition 4, we can say that the asymptotic covariance of the functional  $h_1$  of the quantile  $q_{n, \hat{\sigma}}$  from a location-scale distribution with known mean and the functional  $h_2$  of the measure of dispersion  $\hat{m}(X, n, r)$ , is, for  $p \in (0, 1)$ , asymptotically normal:

$$\sqrt{n} \begin{pmatrix} h_1(q_{n, \hat{\sigma}}(p)) - h_1(q_X(p)) \\ h_2(\hat{m}(X, n, r)) - h_2(m(X, r)) \end{pmatrix} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \tilde{\Lambda}^{(r)}), \quad (19)$$

where the asymptotic covariance matrix  $\tilde{\Lambda}^{(r)} = (\tilde{\Lambda}_{ij}^{(r)}, 1 \leq i, j \leq 2)$  satisfies

$$\tilde{\Lambda}_{11}^{(r)} = \sigma^2 (h_1'(q_X(p)))^2 q_Y^2(p) \frac{\mathbb{E}[Y^4] - 1}{4}; \quad \tilde{\Lambda}_{22}^{(r)} = (h_2'(m(X, r)))^2 \text{Var} \left( |X - \mu|^r + (2 - r)(2F_X(\mu) - 1)X \right); \quad (20)$$

$$\tilde{\Lambda}_{12}^{(r)} = \tilde{\Lambda}_{21}^{(r)} = \sigma^{r+1} \frac{q_Y(p)}{2} \left( \mathbb{E}[|Y|^{r+2}] - \mathbb{E}[|Y|^r] + (2 - r)(2F_Y(0) - 1) \mathbb{E}[Y^3] \right) h_1'(q_X(p)) h_2'(m(X, r)). \quad (21)$$

The asymptotic correlation remains independent - up to its sign - of the specific choice of  $h_1, h_2$ :

$$\lim_{n \rightarrow \infty} \text{Cor} (h_1(q_{n,\hat{\sigma}}(p)), h_2(\hat{m}(X, n, r))) = \text{sgn} (h'_1(q_X(p)) h'_2(m(X, r))(p - F_Y(0))) \times \frac{\mathbb{E}[|Y|^{r+2}] - \mathbb{E}[|Y|^r] + (2-r)(2F_Y(0) - 1) \mathbb{E}[Y^3]}{\sqrt{(\mathbb{E}[Y^4] - 1) \text{Var}(|Y|^r + (2-r)(2F_Y(0) - 1)Y)}}. \quad (22)$$

Note, that in this case the correlation is (up to sign changes) constant in  $p$ .

We now consider the joint asymptotics of functionals of the location-scale quantile estimator with functionals of the sample MedianAD.

**Proposition 6** Consider an iid sample with parent rv  $X$  from a location-scale distribution having mean  $\mu$  and variance  $\sigma^2$ . Under  $(C_1)$ ,  $(M_2)$  and  $(C_3)$  in the neighbourhoods of  $\nu, \nu \pm \xi$ , the joint behaviour of the functionals  $h_1$  of the quantile  $q_{n,\hat{\mu},\hat{\sigma}}(p)$  from a location-scale model (for  $p \in (0, 1)$ ) and  $h_2$  of the sample MedianAD  $\hat{\xi}_n$  (defined in Table 1) is asymptotically normal:

$$\sqrt{n} \begin{pmatrix} h_1(q_{n,\hat{\mu},\hat{\sigma}}(p)) - h_1(q_X(p)) \\ h_2(\hat{\xi}_n) - h_2(\xi) \end{pmatrix} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Pi), \quad (23)$$

where the asymptotic covariance matrix  $\Pi = (\Pi_{ij}, 1 \leq i, j \leq 2)$  satisfies

$$\Pi_{11} = \sigma^2 (h'_1(q_X(p)))^2 \left( 1 + q_Y(p) \left( q_Y(p) (\mathbb{E}[Y^4] - 1)/4 + \mathbb{E}[Y^3] \right) \right); \quad (24)$$

$$\Pi_{22} = \frac{1 + \gamma/f_X^2(\nu)}{4(f_X(\nu + \xi) + f_X(\nu - \xi))^2} (h'_2(m(X, r)))^2; \quad (25)$$

$$\begin{aligned} \Pi_{12} = \Pi_{21} = & \frac{h'_1(q_X(p)) h'_2(\xi) \sigma^2}{2 \left( f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) + f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right) \right)} \times \left( -\mathbb{E} \left[ (Y^2 q_Y(p) + 2Y) \mathbf{I}_{\left(\frac{\nu-\xi-\mu}{\sigma} < Y \leq \frac{\nu+\xi-\mu}{\sigma}\right)} \right] \right. \\ & \left. + \frac{f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right)}{f_Y\left(\frac{\nu-\mu}{\sigma}\right)} \mathbb{E} \left[ (Y^2 q_Y(p) + 2Y) \mathbf{I}_{(Y \leq \frac{\nu-\mu}{\sigma})} \right] + \frac{q_Y(p)}{2} \left( 1 - \frac{f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right)}{f_Y\left(\frac{\nu-\mu}{\sigma}\right)} \right) \right). \end{aligned} \quad (26)$$

The asymptotic correlation remains independent - up to its sign - of the specific choice of  $h_1, h_2$ :

$$\lim_{n \rightarrow \infty} \text{Cor} (h_1(q_{n,\hat{\mu},\hat{\sigma}}(p)), h_2(\hat{\xi}_n)) = \text{sgn}(h'_1(q_X(p)) h'_2(\xi)) \times \quad (27)$$

$$\frac{-\mathbb{E} \left[ (Y^2 q_Y(p) + 2Y) \mathbf{I}_{\left(\frac{\nu-\xi-\mu}{\sigma} < Y \leq \frac{\nu+\xi-\mu}{\sigma}\right)} \right] + \frac{f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right)}{f_Y\left(\frac{\nu-\mu}{\sigma}\right)} \mathbb{E} \left[ (Y^2 q_Y(p) + 2Y) \mathbf{I}_{(Y \leq \frac{\nu-\mu}{\sigma})} \right] + \frac{q_Y(p)}{2} \left( 1 - \frac{f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right)}{f_Y\left(\frac{\nu-\mu}{\sigma}\right)} \right)}{\sqrt{1 + \frac{(f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right))^2}{f_Y^2\left(\frac{\nu-\mu}{\sigma}\right)} - \frac{4(f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right))}{f_Y\left(\frac{\nu-\mu}{\sigma}\right)} \left( 1 - F_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right) - F_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) \right) \times \sqrt{1 + q_Y^2(p) \frac{\mathbb{E}[Y^4] - 1}{4} + q_Y \mathbb{E}[Y^3]}}}$$

In the case where the mean of the distribution is known, we have:

**Corollary 7** Under the same conditions and notations as in Proposition 6, the joint behaviour of the functional  $h_1$  of the quantile  $q_{n,\hat{\sigma}}$  from a location-scale distribution with known mean and the functional  $h_2$  of our sample MedianAD  $\hat{\xi}_n$ , is, for  $p \in (0, 1)$ , asymptotically normal:

$$\sqrt{n} \begin{pmatrix} h_1(q_{n,\hat{\sigma}}(p)) - h_1(q_X(p)) \\ h_2(\hat{\xi}_n) - h_2(\xi) \end{pmatrix} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \tilde{\Pi}), \quad (28)$$

where the asymptotic covariance matrix  $\tilde{\Pi} = (\tilde{\Pi}_{ij}, 1 \leq i, j \leq 2)$  satisfies

$$\tilde{\Pi}_{11} = \sigma^2 (h'_1(q_X(p)))^2 q_Y^2(p) (\mathbb{E}[Y^4] - 1)/4; \quad (29)$$

$$\tilde{\Pi}_{22} = \frac{1 + \gamma/f_X^2(\nu)}{4(f_X(\nu + \xi) + f_X(\nu - \xi))^2} (h'_2(m(X, r)))^2; \quad (30)$$

$$\tilde{\Pi}_{12} = \tilde{\Pi}_{21} = \frac{\sigma^2 q_Y(p)}{2 \left( f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) + f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right) \right)} \times h'_1(q_X(p)) h'_2(\xi) \times \quad (31)$$

$$\left( -\mathbb{E} \left[ Y^2 \mathbf{I}_{\left(\frac{\nu-\xi-\mu}{\sigma} < Y \leq \frac{\nu+\xi-\mu}{\sigma}\right)} \right] + \frac{f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right)}{f_Y\left(\frac{\nu-\mu}{\sigma}\right)} \mathbb{E} \left[ Y^2 \mathbf{I}_{(Y \leq \frac{\nu-\mu}{\sigma})} \right] + \frac{1}{2} \left( 1 - \frac{f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right)}{f_Y\left(\frac{\nu-\mu}{\sigma}\right)} \right) \right).$$

The asymptotic correlation remains, as before, independent - up to its sign - of the specific choice of  $h_1, h_2$ :

$$\lim_{n \rightarrow \infty} \text{Cor} \left( h_1(q_{n,\hat{\sigma}}(p)), h_2(\hat{\xi}_n) \right) = \text{sgn} \left( (p - F_Y(0)) \times h'_1(q_X(p)) h'_2(\xi) \right) \times \quad (32)$$

$$\frac{-\mathbb{E} \left[ Y^2 \mathbf{I}_{\left(\frac{\nu-\xi-\mu}{\sigma} < Y \leq \frac{\nu+\xi-\mu}{\sigma}\right)} \right] + \frac{f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right)}{f_Y\left(\frac{\nu-\mu}{\sigma}\right)} \mathbb{E} \left[ Y^2 \mathbf{I}_{(Y \leq \frac{\nu-\mu}{\sigma})} \right] + \frac{1}{2} \left( 1 - \frac{f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right)}{f_Y\left(\frac{\nu-\mu}{\sigma}\right)} \right)}{\sqrt{1 + \frac{\left( f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right) \right)^2}{f_Y^2\left(\frac{\nu-\mu}{\sigma}\right)} - \frac{4 \left( f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right) \right)}{f_Y\left(\frac{\nu-\mu}{\sigma}\right)} \left( 1 - F_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right) - F_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) \right) \times \sqrt{\frac{\mathbb{E}[Y^4] - 1}{4}}}}$$

While in the case of the asymptotics of the sample MedianAD with the sample quantile we did not even required a finite mean of the underlying distribution, we need in this case a finite fourth moment, as we are considering the location-scale quantile.

### 2.3 Specific Cases and Methods

We briefly want to address two points. First, recall that the results about the asymptotic joint distribution so far have been presented in a functional form, i.e. for arbitrary functions  $h_1, h_2$  (fulfilling, as mentioned, the conditions of the Delta method, i.e. to be continuous and have existing derivatives at the evaluated points). Here we turn to specific choices of functions  $h_1, h_2$  and show how they are related. Second, we briefly present and comment on the methods and ideas needed in the proofs given in the Appendix A.

**Common structure of cases considered** - Clearly, in applications one considers specific choices of  $h_1, h_2$ . For example the covariance/correlation with the quantile estimator itself, as well as with its logarithm, is of interest in finance. The following proposition shows how such cases are related to each other.

**Proposition 8** Consider an iid sample with parent rv  $X$  with, if defined, mean  $\mu$  and variance  $\sigma^2$ . Denote by  $\hat{q}_n$  a given quantile estimator (either  $q_n(p), q_{n,\hat{\mu},\hat{\sigma}}(p)$  or  $q_{n,\hat{\sigma}}(p)$ ) and by  $\hat{D}_{i,n}$  a measure of dispersion estimator ( $i = 1, 2, 3$ ). Assume sufficient conditions on  $X$  to be given (i.e. depending on the chosen estimators, some of  $(C_1) - (C_3), (M_k)$ ). Then, the asymptotic covariance and correlation respectively of different functionals of these quantities are related to each other in the following way:

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\hat{q}_n, \sqrt{n}\hat{D}_{i,n}) &= \frac{1}{2q_X(p)} \times \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\hat{q}_n^2, \sqrt{n}\hat{D}_{i,n}) = 2D_i \times \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n}\hat{q}_n, \sqrt{n}\sqrt{\hat{D}_{i,n}} \right) \\ &= q_X(p) \times \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \log|\hat{q}_n|, \sqrt{n}\hat{D}_{i,n}) = q_X(p) \sqrt{D_i} \times \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} \log|\hat{q}_n|, \sqrt{n}\sqrt{\hat{D}_{i,n}} \right), \end{aligned}$$



and for the asymptotic correlations it holds that

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cor}(\hat{q}_n, \hat{D}_{i,n}) &= \lim_{n \rightarrow \infty} \text{Cor}\left(\hat{q}_n, \sqrt{\hat{D}_{i,n}}\right) = \text{sgn}(p - F_X(0)) \times \lim_{n \rightarrow \infty} \text{Cor}(\hat{q}_n^2, \hat{D}_{i,n}) \\ &= \text{sgn}(p - F_X(0)) \times \lim_{n \rightarrow \infty} \text{Cor}(\log|\hat{q}_n|, \hat{D}_{i,n}) = \text{sgn}(p - F_X(0)) \times \lim_{n \rightarrow \infty} \text{Cor}\left(\log|\hat{q}_n|, \sqrt{\hat{D}_{i,n}}\right). \end{aligned}$$

Hence, in the following we will always provide results for the asymptotic covariance/correlation between the quantile estimator  $\hat{q}_n$  and the measure of dispersion estimator  $\hat{D}_i$  itself, as the other cases can be deduced from this proposition. Note also that the asymptotic linear correlation does not change (up to its sign) under non-linear transformation - although the linear correlation is known to be invariant only under linear transformations. This is due to the fact that, asymptotically, all terms of higher order than the linear approximation in the Taylor series are negligible.

**Remark 9** *In some cases, e.g. when using the logarithm, some care has to be taken when applying the Delta method, as the conditions will not always be satisfied: For  $p = F_X(0)$ , the quantity  $1/q_X(p)$  is not defined. In such cases, if the left-sided and right-sided limit (for the asymptotic covariance and variance respectively) coincide, we simply set the value at the point itself, by continuity of the limit, to be the left-sided limit.*

**Methods -** Let us now comment on the proofs, given in detail in the Appendix A, of the main results presented.

For the proofs involving the sample quantile, the main approach in proving Theorem 1 and Proposition 3 relies on the Bahadur representation of the sample quantile ([5]). Note that already in [29] the asymptotic distribution of sample quantile and sample mean was computed using this representation. This was subsequently used in [8] (see Theorem A.1) to obtain the asymptotic distribution of sample quantiles with  $r$ -th absolute sample moments, in the case of an iid Gaussian sample with known mean  $\mu$ . We will show that the extension of their result to the general iid case is straightforward. Then, when considering the sample variance and sample MAD, we extend the corresponding results to the case of an unknown, hence estimated, mean for any underlying iid distribution (with respective moment and smoothness conditions). Analogously, we show further that it is also possible to use the Bahadur representation for the sample MedianAD to obtain Proposition 3. The proof is presented separately as it cannot be included in the above framework.

Note that an alternative approach is possible to prove Theorem 1; it is based on a Taylor expansion of the sample variance (or sample MAD respectively) and the sample quantile, and has more restrictive smoothness conditions on the underlying distribution. We present it in the Appendix too, as it is a natural extension and adaption of the techniques developed in [17] by Ferguson, who considered the asymptotic joint distribution of the sample mean with the sample quantile.

In contrast to the work with the sample quantile the proofs involving the parametric location-scale quantile estimator are direct computations. Thus, no specific comments on the procedure are given here.

## 2.4 The Effect of Sample Size in Estimation

We conclude this section by analysing the two impacts of the sample size  $n$ . The first one is of asymptotic nature and simply considers the situation when using different sample sizes (in an asymptotic sense)

for the measure of dispersion estimator and the quantile estimator respectively. Second, we evaluate empirically how well the finite sample results approximate the theoretical asymptotics.

**Asymptotic Scaling Law** - We want to understand how the dependence effects of the quantile estimator with the measures of dispersion estimator (sample MAD, sample standard deviation, sample MedianAD) depend on the chosen (asymptotic) sample size  $n$ .

Instead of looking separately at all the different cases when using either the historical estimation of the quantile or the location scale model, and one of the three measures of dispersion, we consider a unified approach. This is motivated by the fact that we are more interested in seeing how the values change with different sample sizes, than presenting formulae for each of the sub-cases, which, anyway, can be deduced from the results. Further, we will observe the same scaling property for each measure of dispersion; thus, a unified approach seems appropriate.

We consider sample sizes, say  $n_v = vn$  and  $n_w = wn$  for integers  $v, w > 0$ , so that they are asymptotically multiples of each other, *i.e.*  $\lim_{n \rightarrow \infty} n_v/n_w = v/w$ . This way we can introduce ‘different’ sample sizes into an asymptotic framework. With the above notations, we can show the following result.

**Theorem 10 (Asymptotic Scaling Law)** *Let  $v, w$  be positive integers and consider an iid sample with parent rv  $X$  with, if defined, mean  $\mu$  and variance  $\sigma^2$ . Under appropriate moment and continuity conditions for  $X$  (*i.e.*  $(C_1) - (C_3), (M_k)$ , depending on the estimators), the asymptotic covariance between functionals of a quantile estimator with sample size  $vn$ ,  $h_1(\hat{q}_{vn})$ , and functionals of the measure of dispersion estimator for  $i \in \{1, 2, 3\}$  with sample size  $wn$ ,  $h_2(\hat{D}_{i,wn})$ , satisfies*

$$\lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n}h_1(\hat{q}_{vn}), \sqrt{n}h_2(\hat{D}_{i,wn}) \right) = \frac{1}{\max(v, w)} \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n}h_1(\hat{q}_n), \sqrt{n}h_2(\hat{D}_{i,n}) \right). \quad (33)$$

Accordingly, one can show for the asymptotic correlation that

$$\lim_{n \rightarrow \infty} \text{Cor} \left( h_1(\hat{q}_{vn}), h_2(\hat{D}_{i,wn}) \right) = \sqrt{\frac{\min(v, w)}{\max(v, w)}} \lim_{n \rightarrow \infty} \text{Cor} \left( h_1(\hat{q}_n), h_2(\hat{D}_{i,n}) \right). \quad (34)$$

Note that the conditions of applicability (*i.e.* moment and continuity conditions on rv  $X$ ) of this proposition depend on the chosen estimators ( $q_{n,t}$  or  $q_{n,\hat{\mu},\hat{\sigma}}, q_{n,\hat{\sigma},t}$  as well as  $\hat{D}_{i,n,t}$  for  $i = 1, 2, 3$ ) and are the same as in the corresponding cases for equal sample sizes  $n$ . An application of this result will be discussed in Section 4.

**Finite sample performance** - So far, we have presented asymptotic results for the joint distribution of the quantile and dispersion measure estimators, given a sample of size  $n$ , with  $n \rightarrow \infty$ .

In practice, we need to estimate the quantile and measure of dispersion estimators on a finite sample of size  $n$ , as well their corresponding covariance and correlation. It means we need a time-series of quantile estimates and measure of dispersion estimates which we use to then empirically compute the sample covariance and correlation.

Thus, by  $q_{n,t}$ ,  $q_{n,\hat{\mu},\hat{\sigma},t}$ , and  $q_{n,\hat{\sigma},t}$  respectively, we denote the quantile estimators (depending on the chosen estimation method) at time  $t$  over the last  $n$  observations before time  $t$ . Accordingly we introduce the time  $t$  for our different measures of dispersion. For notational convenience, we use a framework that incorporates all different cases (already introduced in Section 1): Denote the quantile estimator simply by  $\hat{q}_{n,t}$  for a quantile estimated at time  $t$  over a sample of size  $n$  (it could be either the sample quantile

$q_{n,t}$  or the location-scale quantile model  $q_{n,\hat{\mu},\hat{\sigma},t}$ ,  $q_{n,\hat{\sigma},t}$  respectively). By  $\hat{D}_{i,n,t}$ , for  $i = 1, 2, 3$ , we denote the measure of dispersion of sample size  $n$  (and by  $D_i$  its theoretical counterpart), referring with  $\hat{D}_{1,n,t} = \hat{\sigma}_{n,t}^2$  or  $\sqrt{\hat{D}_{1,n,t}} = \hat{\sigma}_{n,t}$  to the sample variance or sample standard deviation respectively, with  $\hat{D}_{2,n,t} = \hat{\theta}_{n,t}$  to the sample MAD, and with  $\hat{D}_{3,n,t} = \hat{\xi}_{n,t}$  to the sample MedianAD.

To assess the finite sample performance, we conduct a simulation study in the following way: We simulate an iid sample with mean  $\mu = 0$  from three different distributions each: Either a Gaussian distribution or Student distributions with 3 and 5 degrees of freedom, respectively. The sample is of varying size  $N$ . The overall sample size is determined by the fact that we use different sample sizes  $n$  for the estimation of either the quantile or the dispersion measure, with  $n = 126, 252, 504, 1008$  (being multiples or fractions of one year of data, i.e. 252 data points), and different lengths of time-series  $l$ . In each of the cases, the overall sample size needed is  $N = nl$ . We compute the time series of quantile estimates  $\hat{q}_{n,t}(p)$  on disjoint samples for quantiles of order  $p = 0.95$  and  $p = 0.99$  each, and accordingly the time series of measure of dispersion estimates  $\hat{D}_{i,n,t}$  too. Recalling the correlation of interest,  $\text{Cor}(\hat{q}_{n,t}(p), \hat{D}_{i,n,t})$ , we then estimate the linear Pearson correlation using these two time series of  $l$  estimates. This procedure is repeated 1'000-fold in each case. Then, we report the averages of the 1'000-fold repetition with, into brackets, the corresponding empirical 95% confidence interval. Further, we provide as benchmark the theoretical asymptotic value of the correlation in the last column. Also, we provide confidence intervals for the sample Pearson linear correlation coefficient (using the classical variance-stabilizing Fisher transform of the correlation coefficient for a bivariate normal distribution to compute the confidence intervals -see the original paper [18] or e.g. a standard encyclopedia entry [37]). Note that those confidence interval values have to be considered with care. Recall that the bivariate normality of the quantile estimator and measure of dispersion estimator holds asymptotically. Hence, it is not clear if for the sample sizes  $n$  considered, we can assume bivariate normality (this could be tested). Still, we provide those theoretical confidence intervals as approximate guidance.

The full results of the simulation study are available in the Appendix D.2. In Table 2 we focus on the approximation of the joint asymptotic correlation as a function of the sample size  $n$ , the different dispersion estimators and the three different distributions considered. Thus, we only consider the sample quantile (and not the location-scale quantile) and fix the length of the sample correlation time series to  $l = 50$  (from the simulations in Appendix D.2, one can see that such a time series is long enough for a good estimation of the correlation). Also, here we focus on  $p = 0.95$ . Clearly, for a higher quantile, as e.g.  $p = 0.99$ , using the same sample size  $n$  for the estimation of the quantile and the sample correlation will be less precise.

The explicit expressions in the case of a Gaussian or Student distribution of the asymptotic correlation used to calculate the theoretical values in Table 2 will be shown in Section 3 (but could be already derived from the theorems presented). Similarly, in Section 3 we will see that the correlation results for location-scale distributions are independent of its parameters, hence the specification of  $\mu, \sigma^2$  is not needed.

We recall that when working with the sample standard deviation, the existence of the fourth moment is a necessary condition, also when working with the location-scale quantile. Thus, as they do not exist for a Student distribution with 3 degrees of freedom, we simply write 'NA' as theoretical value instead.

Let us look at the results in Table 2. First we consider the Gaussian case. For the three dispersion measures, the behaviour is equal: We see that a sample size of  $n = 126$  suffices to estimate on average the asymptotic correlation well enough. Also, the theoretical confidence intervals coming from a sample correlation of size  $l = 50$  are captured well by the empirical confidence intervals. Moving to the Student(5) distribution, the picture changes a bit. The sample correlation with the sample variance does

Table 2: Average values from a 1'000-fold repetition. Comparing the sample correlation of the sample measure of dispersion with the sample quantile, as a function of the sample size  $n$  on which the quantile is estimated (fixed length  $l = 50$  of the time-series used to estimate the correlation). Underlying samples are simulated from a Gaussian, Student(5) and Student(3) distributions. Average empirical values are written first (with empirical 95% confidence interval in brackets). The corresponding theoretical values, mean and 95% confidence interval, are provided in the last column. We consider the threshold  $p = 0.95$ .

$p = 0.95$	$n = 126$	$n = 252$	$n = 504$	$n = 1008$	theoretical value
<b>Gaussian distr.</b>					
$\widehat{\text{Cor}}(\hat{\sigma}_n^2, \hat{q}_n(p))$	55 (33,71)	55 (34,73)	55 (34,73)	55 (34,71)	55 (32,72)
$\widehat{\text{Cor}}(\hat{\theta}_n, \hat{q}_n(p))$	48 (26,66)	48 (26,69)	48 (25,69)	48 (26,66)	48 (23,67)
$\widehat{\text{Cor}}(\hat{\xi}_n, \hat{q}_n(p))$	23 (-4,48)	23 (-3,48)	23 (-4,49)	23 (-4,45)	23 (-5,48)
<b>Student(5) distr.</b>					
$\widehat{\text{Cor}}(\hat{\sigma}_n^2, \hat{q}_n(p))$	51 (19,75)	49 (19,71)	47 (19,68)	46 (20,67)	43 (17,63)
$\widehat{\text{Cor}}(\hat{\theta}_n, \hat{q}_n(p))$	50 (27,71)	50 (27,69)	50 (27,70)	51 (27,69)	51 (27,69)
$\widehat{\text{Cor}}(\hat{\xi}_n, \hat{q}_n(p))$	23 (-6,50)	23 (-6,47)	23 (-6,47)	23 (-6,48)	23 (-5,48)
<b>Student(3) distr.</b>					
$\widehat{\text{Cor}}(\hat{\sigma}_n^2, \hat{q}_n(p))$	25 (-8,55)	22 (-9,52)	19 (-9,47)	17 (-14,44)	NA
$\widehat{\text{Cor}}(\hat{\theta}_n, \hat{q}_n(p))$	48 (21,68)	47 (23,67)	47 (20,68)	47 (23,67)	48 (23,67)
$\widehat{\text{Cor}}(\hat{\xi}_n, \hat{q}_n(p))$	23 (-4,49)	22 (-7,48)	22 (-7,47)	23 (-7,49)	23 (-5,48)

not estimate on average accurately the theoretical value. For increasing  $n$ , it approaches the theoretical value. This can be explained by the fact that the theoretical correlation values come from the underlying asymptotic bivariate normal distribution. Hence, for a small  $n$  the corresponding sample quantities are not yet bivariate normally distributed and one would need a larger sample for this. Correspondingly, the confidence intervals (whose theoretical values, again, are based on a bivariate normal sample) are not that exact. This different behaviour is not observed for the MAD or MedianAD with more accurate results, as in the Gaussian case. While the average with the MAD is slightly (one percent point) below the theoretical value for most values of  $n$  (which is acceptable), it equals the theoretical value exactly in the case of the MedianAD. In both cases, the sample confidence intervals correspond quite well to the theoretical ones, potentially indicating that the sample quantities converge faster to a bivariate normal distribution. The case of the Student(3) distribution is as the Student(5) distribution, the only difference being that the theoretical correlation values with the sample variance are not defined (as the fourth moment does not exist). However, the behaviour of the sample correlation with the MAD and MedianAD does not change.

### 3 Case of Location-Scale Distributions

We presented in Section 2 the general results about the joint asymptotic distributions for the functionals of measure of dispersion estimators with functionals of either the sample quantile or the location-scale quantile estimator. While the results for the former hold for all underlying distributions (with some smoothness conditions and restrictions on the existence of its moments), the location-scale quantile

can be built only for location-scale distributions. Thus, to be able to compare the behaviour with both quantile estimators, we will assume in this section that the rv  $X$  comes from a location-scale distribution  $F_X$  with finite mean  $\mu$  and finite variance  $\sigma^2$ . Such distributions appear often in financial applications. Standard examples of location-scale distributions are the elliptical distributions such as the Gaussian and the Student. For instance, if  $X \sim \mathcal{N}(\mu, \sigma^2)$ , then

$$q_X(p) = \mu + \sigma \Phi^{-1}(p),$$

whereas, if  $X$  is Student distributed rv (also referred to as  $t$ -distributed or Student  $t$ -distributed) with  $\nu$  degrees of freedom, mean  $\mu$  and variance  $\sigma^2$ , then

$$q_X(p) = \mu + \sigma q_Y(p) = \mu + \frac{\sigma}{\sqrt{\frac{\nu}{\nu-2}}} q_{\tilde{Y}}(p)$$

where  $Y \sim t(0, \frac{\nu-2}{\nu}, \nu)$  and the standard Student rv  $\tilde{Y} \sim t(0, 1, \nu)$ . Note that we follow the usual convention to denote a Student rv with mean  $\mu$  and variance  $\sigma^2$  by  $X \sim t(\mu, \sigma^2(\nu - 2)/\nu, \nu)$ , the second parameter being the scale parameter (and not the variance).

Consequently, we have two different ways of estimating the quantile, either via the sample quantile, or via the given quantile relation for location-scale distributions. Although both estimators converge to the same quantity, the theoretical quantile, they do not have the same rate of convergence. This has some impact on their asymptotic covariance and correlation with the corresponding measure of dispersion (sample variance, sample MAD or sample MedianAD), thus in practice too. Using the location-scale quantile, we obtain a better rate of convergence than with the historical estimation, as expected. Hence our interest of investigating this second way of estimation, as it provides a benchmark in comparison to the convergence with the historical estimation. Still, the location-scale quantile is not often used in practice, as one does not generally know the underlying distribution, so using it could add model risk to the estimation.

In the following, we consider both estimators of the quantile, starting with the historical one, then using the location-scale formula. In each case, we look at the covariance and correlation with the sample variance, sample MAD and sample MedianAD separately and apply it to our two exemplary distributions, the Gaussian and the Student distributions. In the examples it would suffice to consider either the case of the estimated quantile or its logarithm (recall Proposition 8) but we provide both for the convenience of the reader. In the third and last part, we bring together the results of the historical estimation and the usage of the location-scale quantile model. Using the location-scale quantile as benchmark, we can state the difference in speed of convergence when using the two estimators by considering the ratio of speeds of convergence (again separately for the three measures of dispersion considered in this paper). We do so by looking at the two main examples used in this paper, the Gaussian and the Student distributions.

### 3.1 Historical Estimation

We start with the version of Theorem 1 using the sample quantile for location-scale distributions, then consider separately the cases of sample variance and sample MAD. For location-scale distributions, using the notations above, Theorem 1 simplifies as follows:

**Proposition 11** *Consider an iid sample with parent rv  $X$  from a location-scale distribution having mean  $\mu$ , variance  $\sigma^2$ , and corresponding rv  $Y$  with standardised distribution. Under the conditions  $(C_1), (C_3)$  in a neighbourhood of  $q_X(p)$ , for  $r = 1, 2$  ( $M_r$ ) respectively and, additionally,  $(C_2)$  at  $\mu$ , the joint behaviour of the functionals  $h_1$  of the sample quantile  $q_n(p)$  (for  $p \in (0, 1)$ ) and  $h_2$  of the measure of dispersion  $\hat{m}(X, n, r)$  (defined in (1)) is asymptotically normal (via Theorem 1), satisfying*

(5) where the asymptotic covariance  $\Sigma_{12}^{(r)} = \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}h_1(q_n(p)), \sqrt{n}h_2(\hat{m}(X, n, r)))$  simplifies to:

$$\Sigma_{12}^{(r)} = \frac{\sigma^{r+1} \left( \tau_r(|Y|, p) + (2-r)(2F_Y(0) - 1)\tau_1(Y, p) \right)}{f_Y(q_Y(p))} h_1'(q_X(p)) h_2'(m(X, r)), \quad (35)$$

and the asymptotic correlation to:

$$\lim_{n \rightarrow \infty} \text{Cor}(h_1(q_n(p)), h_2(\hat{m}(X, n, r))) = a_{\pm} \times \frac{\tau_r(|Y|, p) + (2-r)(2F_Y(0) - 1)\tau_1(Y, p)}{\sqrt{p(1-p) \text{Var}(|Y|^r + (2-r)(2F_Y(0) - 1)Y)}} \quad (36)$$

with  $a_{\pm} = \text{sgn}(h_1'(q_X(p)) \times h_2'(m(X, r)))$ , and  $\tau_r$  defined in (3).

Further, if we assume the location-scale distribution to be symmetric, then the asymptotic correlation between the sample quantile and the measure of dispersion estimator can be proved to have its minimum absolute correlation (value of 0) for  $p = 0.5$  and to be (point-)symmetric with respect to  $p = 0.5$ .

Note that the asymptotic correlation (36) is not only independent - up to the sign - of the specific choice of  $h_1, h_2$ , but also does not depend on the mean  $\mu$  and the variance  $\sigma^2$  of the underlying location-scale distribution.

We now look at the specific choices of  $r = 2$  (sample variance) and  $r = 1$  (MAD) separately, to obtain specific explicit expressions for the covariance  $\Sigma_{12}^{(r)}$  or corresponding correlation.

**Dependence with Sample Variance** - When choosing  $r = 2$  in Proposition 11, we are in the case of dependence with the sample variance, which requires the existence of the fourth moment of the underlying distribution.

**Corollary 12** Consider an iid sample with parent rv  $X$  following a location-scale distribution under conditions  $(C_1), (C_3)$  in a neighbourhood of  $q_X(p)$  and  $(M_2)$ . Let  $Y$  be the corresponding rv with standardised distribution. Then the asymptotic covariance of the sample quantile and the sample variance satisfies, for  $p \in (0, 1)$ ,

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_n(p), \sqrt{n}\hat{\sigma}_n^2) = \frac{p\sigma^3(1 - \mathbb{E}[Y^2|Y \leq q_Y(p)])}{f_Y(q_Y(p))}, \quad (37)$$

from which we deduce the asymptotic correlation:

$$\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\sigma}_n^2) = \sqrt{\frac{p}{1-p}} \times \frac{1 - \mathbb{E}[Y^2|Y \leq q_Y(p)]}{\sqrt{\mathbb{E}[Y^4] - 1}}, \quad (38)$$

which can be rewritten to better show the dependence of the truncated moments on both, the upper and lower tails, as

$$\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\sigma}_n^2) = \frac{\sqrt{p(1-p)} (\mathbb{E}[Y^2|Y > q_Y(p)] - \mathbb{E}[Y^2|Y \leq q_Y(p)])}{\sqrt{\mathbb{E}[Y^4] - 1}}. \quad (39)$$

As an application of Corollary 12, let us consider the two elliptical examples of the Gaussian and Student distributions, for which we can get explicit expressions. The conditions of Corollary 12 are clearly met: Both distributions are absolutely continuous and strictly monotonically increasing, i.e.  $(C_1)$ , and have bounded second derivative, i.e.  $(C_3)$ . For the Gaussian distribution the condition  $(M_2)$  is always satisfied whereas for the Student distribution we need a degree of freedom  $\nu > 4$ . For  $p \in (0, 1)$ , we have:

**Example 13** (i) *In the case of a Gaussian distribution with mean  $\mu$  and variance  $\sigma^2$ :*

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{\sigma}_n^2) = \sigma^3 \Phi^{-1}(p) = (\mu + \sigma \Phi^{-1}(p)) \times \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \log|q_n(p)|, \sqrt{n} \hat{\sigma}_n^2), \quad (40)$$

whereas all the correlations are identical - up to the sign- and equal to:

$$\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\sigma}_n^2) = \frac{\phi(\Phi^{-1}(p)) \Phi^{-1}(p)}{\sqrt{2p(1-p)}} = \text{sgn} \left( p - \Phi \left( \frac{-\mu}{\sigma} \right) \right) \times \lim_{n \rightarrow \infty} \text{Cor}(\log|q_n(p)|, \hat{\sigma}_n^2). \quad (41)$$

(ii) *In the case of a Student distribution with mean  $\mu$ , variance  $\sigma^2$  and  $\nu > 4$  degrees of freedom,*

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{\sigma}_n^2) = \sigma^3 \sqrt{\frac{\nu-2}{\nu}} q_{\tilde{Y}}(p) \left( 1 + \frac{q_{\tilde{Y}}^2(p)}{\nu} \right) = \sqrt{\frac{\nu-2}{\nu}} \left( \sqrt{\frac{\nu}{\nu-2}} \mu + \sigma q_{\tilde{Y}}(p) \right) \times \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \log(|q_n(p)|), \sqrt{n} \hat{\sigma}_n^2), \quad (42)$$

whereas, as before, the correlations are up to their sign all identical in the different cases:

$$\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\sigma}_n^2) = \frac{f_{\tilde{Y}}(q_{\tilde{Y}}(p)) q_{\tilde{Y}}(p) \left( 1 + \frac{q_{\tilde{Y}}^2(p)}{\nu} \right)}{\sqrt{\frac{\nu-1}{\nu-4}} 2p(1-p)} = \text{sgn} \left( p - F_{\tilde{Y}} \left( \frac{-\mu}{\sigma} \sqrt{\frac{\nu}{\nu-2}} \right) \right) \times \lim_{n \rightarrow \infty} \text{Cor}(\log|q_n(p)|, \hat{\sigma}_n^2) \quad (43)$$

As expected, letting  $\nu \rightarrow \infty$  gives back the results for the Gaussian distribution given in (i).

**Dependence with Sample Mean Absolute Deviation** - We have seen, when using the sample variance as measure of dispersion, that we are restricted for the results of asymptotic normality to distributions with finite fourth moment. If we choose  $r = 1$  in Proposition 11, *i.e.* when considering the dependence between functionals of the sample quantile and functionals of the sample MAD, it allows us to relax this strong moment condition (only a finite second moment is required). Such a less restrictive condition is needed for most financial applications, when using for instance Student distributions with  $2 < \nu < 4$ .

**Corollary 14** *Consider an iid sample with parent rv  $X$  following a location-scale distribution continuous at  $\mu$  and fulfilling conditions  $(C_1)$ ,  $(C_2)$  at  $\mu$ ,  $(C_3)$  in a neighbourhood of  $q_X(p)$  and  $(M_1)$ . Then, the asymptotic covariance of the sample quantile and sample MAD satisfies, for  $p \in (0, 1)$ ,*

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{\theta}_n) = \frac{\sigma^2 p}{f_Y(q_Y(p))} \left( \theta/\sigma - \mathbb{E}[|Y| | Y \leq q_Y(p)] - (2F_Y(0) - 1) \mathbb{E}[Y | Y \leq q_Y(p)] \right) \quad (44)$$

from which we deduce the asymptotic correlation:

$$\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\theta}_n) = \frac{\theta/\sigma - \mathbb{E}[|Y| | Y \leq q_Y(p)] - (2F_Y(0) - 1) \mathbb{E}[Y | Y \leq q_Y(p)]}{\sqrt{p(1-p) \text{Var}(|Y| + (2F_Y(0) - 1)Y)}}. \quad (45)$$

Let us quickly look at how these formulae simplify for a Gaussian and a Student distributions (since both distributions are symmetric around the mean). For this, recall that  $\text{Var}(|X - \mu|) = \sigma^2 - \theta^2$  where  $\theta$  is defined in Table 1 and the values of  $\theta$  are tabulated for many of the well-known distributions. Note that we present the expressions only for  $p \geq 0.5$ , the case  $p < 0.5$  being deduced by symmetry. For  $p < 0.5$ ,

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{\theta}_n) = - \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(1-p), \sqrt{n} \hat{\theta}_n)$$

or, in the case involving the logarithm (and assuming  $\mu = 0$ ),

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \log|q_n(p)|, \sqrt{n} \hat{\theta}_n) = \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \log|q_n(1-p)|, \sqrt{n} \hat{\theta}_n),$$

and accordingly for the correlations. Also, we already discussed before Example 13 why conditions  $(C_1)$  and  $(C_3)$  are met for the Gaussian and Student distributions. Further, for the dependence with the sample MAD, we only require  $(M_1)$ , i.e. finite variance, as mentioned before. While this is fulfilled for the Gaussian distribution, we need  $\nu > 2$  for the Student case. Thus, we only need to argue why  $(C_2)$  is fulfilled: As both distributions are continuously differentiable, they are also Hölder continuous.

**Example 15** (i) *In the case of a Gaussian distribution we have, for  $p \geq 0.5$ ,*

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{\theta}_n) = \frac{\sigma^2}{\phi(\Phi^{-1}(p))} \left( \phi(\Phi^{-1}(p)) - 1 - p \sqrt{\frac{2}{\pi}} \right) = (\mu + \sigma \Phi^{-1}(p)) \times \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \log|q_n(p)|, \sqrt{n} \hat{\theta}_n), \quad (46)$$

whereas the correlation is -up to the sign - identical in the two different cases

$$\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\theta}_n) = \text{sgn} \left( p - F_Y \left( \frac{-\mu}{\sigma} \right) \right) \times \lim_{n \rightarrow \infty} \text{Cor}(\log|q_n(p)|, \hat{\theta}_n) = \frac{\phi(\Phi^{-1}(p)) - (1-p) \sqrt{2/\pi}}{\sqrt{p(1-p)} \sqrt{1-2/\pi}}. \quad (47)$$

(ii) *For the case of a Student distribution with  $\nu > 2$  degrees of freedom, for  $p \geq 0.5$ ,*

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{\theta}_n) = \frac{\sigma^2(\nu-2)}{f_{\tilde{Y}}(q_{\tilde{Y}}(p))} \left( \frac{f_{\tilde{Y}}(q_{\tilde{Y}}(p))}{\nu-1} \left( 1 + \frac{q_{\tilde{Y}}^2(p)}{\nu} \right) - \frac{\Gamma(\frac{\nu-1}{2})}{\sqrt{\pi\nu} \Gamma(\frac{\nu}{2})} (1-p) \right) \quad (48)$$

and for the case involving the logarithm,

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \log|q_n(p)|, \sqrt{n} \hat{\theta}_n) = \frac{\sigma^2(\nu-2) \left( \frac{1}{\nu-1} \left( 1 + \frac{q_{\tilde{Y}}^2(p)}{\nu} \right) - \frac{\Gamma(\frac{\nu-1}{2})}{\sqrt{\pi\nu} \Gamma(\frac{\nu}{2})} \frac{(1-p)}{f_{\tilde{Y}}(q_{\tilde{Y}}(p))} \right)}{\mu + \sigma \sqrt{\frac{\nu-2}{\nu}} q_{\tilde{Y}}(p)} \quad (49)$$

$$\begin{aligned} \text{and } \lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\theta}_n) &= \text{sgn} \left( p - F_{\tilde{Y}} \left( \frac{-\mu}{\sigma} \sqrt{\frac{\nu}{\nu-2}} \right) \right) \times \lim_{n \rightarrow \infty} \text{Cor}(\log|q_n(p)|, \hat{\theta}_n) \\ &= \frac{\frac{\sqrt{\nu(\nu-2)}}{\nu-1} f_{\tilde{Y}}(q_{\tilde{Y}}(p)) \left( 1 + \frac{q_{\tilde{Y}}^2(p)}{\nu} \right) - (1-p) \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)}}{\sqrt{p(1-p)} \sqrt{1 - \frac{\nu-2}{\pi} \frac{\Gamma^2(\frac{\nu-1}{2})}{\Gamma^2(\nu/2)}}}. \end{aligned} \quad (50)$$

As expected, letting  $\nu \rightarrow \infty$  provides the results for the Gaussian distribution given in (i). This might be not as obvious as in Example 13, and one might need to recall the asymptotic property of the Gamma function  $\lim_{n \rightarrow \infty} \frac{\Gamma(n+\alpha)}{\Gamma(n)n^\alpha} = 1$  that we need to use here with  $n = \nu\alpha$  and  $\alpha = 1/2$ .

**Dependence with Sample Median Absolute Deviation** - As in Section 2, we treat the case of dependence with the sample MedianAD separately. We start by presenting the dependence of functionals of the sample quantile and the sample MedianAD in the case of general location-scale distributions.



**Proposition 16** Consider an iid sample with parent rv  $X$  from a location-scale distribution fulfilling conditions  $(M_1)$ ,  $(C_1)$  and  $(C_3)$  in neighbourhoods of  $q_X(p)$ ,  $\nu$ ,  $\nu \pm \xi$ . Then, the joint behaviour of the functionals  $h_1$  of the sample quantile  $q_n(p)$  (for  $p \in (0, 1)$ ) and  $h_2$  of the sample MedianAD  $\hat{\xi}_n$  (defined in Table 1) is asymptotically normal (via Proposition 3) and (10) rewrites as:

$$\lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} h_1(q_n(p)), \sqrt{n} h_2(\hat{\xi}_n) \right) = h'_1(q_X(p)) h'_2(\xi) \times \frac{\sigma^2}{f_Y(q_Y(p)) \left( f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) + f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right) \right)} \times \left( \frac{\left( f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right) \right) \max\left(-\frac{p}{2}, \frac{p-1}{2}\right)}{f_Y\left(\frac{\nu-\mu}{\sigma}\right)} - \max\left(0, F_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - \max\left(F_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right), p\right)\right) + \frac{1-p}{2} \right). \quad (51)$$

The asymptotic correlation is - up to the sign - independent of the specific choice of  $h_1, h_2$ . Introducing

$$\tilde{\gamma} = \left( f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right) \right) \left( f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\mu}{\sigma}\right) \right) \left( 1 - F_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right) - F_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) \right),$$

and with  $a_{\pm} = \text{sgn}(h'_1(q_X(p)) h'_2(\xi))$ , we have  $\lim_{n \rightarrow \infty} \text{Cor} \left( h_1(q_n(p)), h_2(\hat{\xi}_n) \right) = a_{\pm} \times$

$$\frac{2 \left( \frac{f_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - f_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right)}{f_Y\left(\frac{\nu-\mu}{\sigma}\right)} \max\left(-\frac{p}{2}, \frac{p-1}{2}\right) + \frac{1-p}{2} - \max\left(0, F_Y\left(\frac{\nu+\xi-\mu}{\sigma}\right) - \max\left(F_Y\left(\frac{\nu-\xi-\mu}{\sigma}\right), p\right)\right) \right)}{\sqrt{p(1-p)} \sqrt{1 + \frac{\tilde{\gamma}}{f_Y^2\left(\frac{\nu-\mu}{\sigma}\right)}}}. \quad (52)$$

We observe that, in the case of the sample MedianAD, the formulae for asymptotic covariance and correlation do not simplify for general location-scale distributions (compare (51), (52) with those for general distributions in Proposition 3), contrarily to the case of the sample variance and the sample MAD. Nevertheless, when assuming additionally the distribution to be symmetric, which covers a big part of the location-scale distributions (as e.g. the Gaussian and Student distributions), those expressions simplify a lot, as can be seen in the next proposition.

**Proposition 17** Consider an iid sample with parent rv  $X$  from a symmetric location-scale distribution having mean  $\mu$ , variance  $\sigma^2$  and fulfilling conditions  $(M_1)$ ,  $(C_1)$  and  $(C_3)$  in neighbourhoods of  $q_X(p)$ ,  $\nu$ ,  $\nu \pm \xi$ . Then, the joint behaviour of the functionals  $h_1$  of the sample quantile  $q_n(p)$  (for  $p \in (0, 1)$ ) and  $h_2$  of the sample MedianAD  $\hat{\xi}_n$  (defined in Table 1) is asymptotically normal (via Proposition 3) and (10) rewrites as:

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(q_n(p)), \sqrt{n} h_2(\hat{\xi}_n)) = \frac{\sigma^2 (1-p - 2 \max(0, 3/4 - \max(1/4, p)))}{4 f_Y(q_Y(p)) f_Y(q_Y(3/4))} h'_1(q_X(p)) h'_2(\xi). \quad (53)$$

The asymptotic correlation, independent - up to its sign - of the specific choice of  $h_1, h_2$ , satisfies:

$$\lim_{n \rightarrow \infty} \text{Cor}(h_1(q_n(p)), h_2(\hat{\xi}_n)) = \frac{1-p - 2 \max(3/4 - \max(1/4, p), 0)}{\sqrt{p(1-p)}} \times \text{sgn}(h'_1(q_X(p)) h'_2(\xi)) \quad (54)$$

and does not depend on the underlying symmetric location-scale distribution at all.

Also, the asymptotic correlation between the sample quantile and the sample MedianAD can be proved to have its minimum absolute correlation (value of 0) for  $p = 0.5$  and to be (point-)symmetric with respect to  $p = 0.5$ .

The fact that the correlation above is completely independent of the underlying distribution is something to point out! As an application of Proposition 17, let us consider the two elliptical examples of

the Gaussian and Student distributions, for which we can get explicit expressions. Again, we already discussed before Example 13, why conditions  $(C_1)$ ,  $(C_3)$  are met in general for the Gaussian and Student distribution. Because of the location-scale distributions, we need condition  $(M_1)$  (finite variance) to be fulfilled, thus need  $\nu > 2$  in the case of the Student distribution (although the asymptotics with the sample MedianAD do not need a finite variance of the underlying distribution in general). As we already know that the asymptotic correlation will be the same for the two distributions, we present the example in a more compact form. We can show that, for  $p \in (0, 1)$ , we have

**Example 18**

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{\xi}_n) &= \frac{\sigma^2}{4} (1 - p - 2 \max(3/4 - \max(1/4, p), 0)) \times \begin{cases} \frac{1}{\phi(\Phi^{-1}(p))\phi(\Phi^{-1}(3/4))} & \text{if } X \sim \mathcal{N}(\mu, \sigma^2) \\ f_{\hat{Y}}(q_{\hat{Y}}(p))f_{\hat{Y}}(q_{\hat{Y}}(3/4))^{-\frac{\nu}{\nu-2}} & \text{if } X \sim t(\mu, (\nu-2)\sigma^2/\nu, \nu) \end{cases} \\ &= \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \log|q_n(p)|, \sqrt{n} \hat{\xi}_n) \times \begin{cases} \mu + \sigma\Phi^{-1}(p) & \text{if } X \sim \mathcal{N}(\mu, \sigma^2), \\ \mu + \frac{\sigma q_{\hat{Y}}}{\sqrt{\frac{\nu}{\nu-2}}} & \text{if } X \sim t(\mu, (\nu-2)\sigma^2/\nu, \nu). \end{cases} \end{aligned} \quad (55)$$

As expected, if we let  $\nu \rightarrow \infty$  in the Student case, we get back the results for the covariance of the Gaussian distribution. For both distributions, the correlations are identical - up to its sign when considering the log-quantile:

$$\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\xi}_n) = a_X \times \lim_{n \rightarrow \infty} \text{Cor}(\log|q_n(p)|, \hat{\xi}_n) = \frac{1 - p - 2 \max(3/4 - \max(1/4, p), 0)}{\sqrt{p(1-p)}}, \quad (56)$$

$$\text{where } a_X = \begin{cases} \text{sgn}(p - \Phi(-\mu/\sigma)) & \text{if } X \sim \mathcal{N}(\mu, \sigma^2), \\ \text{sgn}\left(p - F_{\hat{Y}}\left(\frac{-\mu}{\sigma} \sqrt{\frac{\nu}{\nu-2}}\right)\right) & \text{if } X \sim t(\mu, (\nu-2)\sigma^2/\nu, \nu). \end{cases}$$

### 3.2 Location-scale Quantile

Recall that, already in Subsection 2.2, we presented results for the location-scale quantile, namely the unified result for the dependence of the location-scale quantile with the sample variance or sample MAD (Proposition 4, Corollary 5) as well as the corresponding results when using the sample MedianAD (Proposition 6, Corollary 7).

Hence, while keeping the same structure as in Subsection 3.1, the presentation differs a bit: We directly start by looking separately at the covariance and correlation of the location-scale quantile with the sample variance ( $r = 2$ ) and sample MAD ( $r = 1$ ). In the case of the sample MedianAD we present the case for symmetric location-scale distributions. Since we distinguish the cases  $\mu$  unknown from known, we present the more general case first ( $\mu$  unknown), and then comment on the case with known mean  $\mu$ . As before, we present explicit expressions for the covariance and correlation in the case of the Gaussian and Student distributions for all three cases of measures of dispersion. Last but not least, recall that using the asymptotics with the location-scale quantile model implies assuming the existence of a finite fourth moment - this is in contrast to the historical estimation with the sample quantile (where we do not need a finite fourth moment when considering the dependence with the sample MAD or MedianAD).

**Dependence with Sample Variance** - Let us start with the dependence of the sample variance and a location-scale quantile model and state the result that follows from Proposition 4 when choosing  $r = 2$ .

**Proposition 19** Consider an iid sample with parent rv  $X$  following a location-scale distribution. Then, under  $(C_1)$ ,  $(M_2)$ , the asymptotic covariance and correlation of the location-scale quantile  $q_{n, \hat{\mu}, \hat{\sigma}}$  and

sample variance  $\hat{\sigma}_n^2$  satisfy, respectively for  $p \in (0, 1)$ ,

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_{n, \hat{\mu}, \hat{\sigma}}(p), \sqrt{n} \hat{\sigma}_n^2) = \mu_3 + q_Y(p) \frac{\mu_4 - \sigma^4}{2\sigma} \quad (57)$$

$$\text{and } \lim_{n \rightarrow \infty} \text{Cor}(q_{n, \hat{\mu}, \hat{\sigma}}(p), \hat{\sigma}_n^2) = \frac{\text{sgn}(p - F_Y(0))}{\sqrt{1 + \frac{4(1 + \mathbb{E}[Y^3]q_Y(p))}{(\mathbb{E}[Y^4] - 1)q_Y^2(p)}}} \times \left(1 + \frac{2\mathbb{E}[Y^3]}{(\mathbb{E}[Y^4] - 1)q_Y(p)}\right). \quad (58)$$

Note that those expressions (57) and (58) simplify quite a lot when assuming zero third centred moment ( $\mu_3 = 0 = \mathbb{E}[Y^3]$ ), as e.g. for elliptical distributions, or when the mean is known, giving in this latter case that the asymptotic correlation of the considered estimators is always 1 (up to its sign), irrespective of the underlying parameters or the distribution itself; indeed, in such a case (58) becomes  $\lim_{n \rightarrow \infty} \text{Cor}(q_{n, \hat{\sigma}}(p), \hat{\sigma}_n^2) = \text{sgn}(p - F_Y(0))$ .

Applying Proposition 19 to the cases of a Gaussian and a Student distributions, we need those distributions to fulfil  $(C_1)$  (fulfilled, as discussed for Example 13) and  $(M_2)$  (because of the location-scale quantile), hence need  $\nu > 4$  for the Student distribution. As before, we assume e.g. that the mean  $\mu$  is known, which provides:

**Example 20** (i) For the Gaussian distribution  $\mathcal{N}(\mu, \sigma^2)$ :

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_{n, \hat{\sigma}}(p), \sqrt{n} \hat{\sigma}_n^2) = \sigma^3 \Phi^{-1}(p) = (\mu + \sigma \Phi^{-1}(p)) \times \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \log|q_{n, \hat{\sigma}}(p)|, \sqrt{n} \hat{\sigma}_n^2), \quad (59)$$

$$\lim_{n \rightarrow \infty} \text{Cor}(q_{n, \hat{\sigma}}(p), \hat{\sigma}_n^2) = \text{sgn}(p - 1/2) = \text{sgn}(p - \Phi(\mu/\sigma)) \times \lim_{n \rightarrow \infty} \text{Cor}(\log|q_{n, \hat{\sigma}}(p)|, \hat{\sigma}_n^2). \quad (60)$$

(ii) For the Student distribution  $t(\mu, (\nu - 2)\sigma^2/\nu, \nu)$  with  $\nu > 4$ ,

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_{n, \hat{\sigma}}(p), \sqrt{n} \hat{\sigma}_n^2) = \sigma^3 q_{\tilde{Y}}(p) \frac{\nu - 1}{\nu - 4} \sqrt{\frac{\nu - 2}{\nu}} \quad (61)$$

$$= \left(\mu + \sigma q_{\tilde{Y}}(p) \sqrt{1 - 2/\nu}\right) \times \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \log|q_{n, \hat{\sigma}}(p)|, \sqrt{n} \hat{\sigma}_n^2),$$

$$\lim_{n \rightarrow \infty} \text{Cor}(q_{n, \hat{\sigma}}(p), \hat{\sigma}_n^2) = \text{sgn}(p - 1/2) = \text{sgn}\left(p - F_{\tilde{Y}}\left(\frac{-\mu}{\sigma} \sqrt{\frac{\nu}{\nu - 2}}\right)\right) \times \lim_{n \rightarrow \infty} \text{Cor}(\log|q_{n, \hat{\sigma}}(p)|, \hat{\sigma}_n^2). \quad (62)$$

While the correlations are already the same - up to a sign - for the Gaussian and Student distributions, we can also check that taking  $\nu \rightarrow \infty$  in (ii) gives back the result for the asymptotic covariance in the Gaussian case.

**Dependence with Sample Mean Absolute Deviation** - We continue with the dependence of functionals of the sample MAD and a location-scale quantile model - which corresponds to the case when choosing  $r = 1$  in Proposition 4.

**Proposition 21** Consider an iid sample with parent rv  $X$  following a location-scale distribution fulfilling conditions  $(C_1)$ ,  $(M_2)$  and  $(C_2)$  at  $\mu$ . Then, the asymptotic covariance and correlation, respectively, of

the following functionals of the location-scale quantile  $q_{n,\hat{\mu},\hat{\sigma}}$  and sample MAD  $\hat{\theta}_n$  satisfy, for  $p \in (0, 1)$ ,

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} q_{n,\hat{\mu},\hat{\sigma}}(p), \sqrt{n} \hat{\theta}_n \right) &= \sigma^2 \left( 2F_Y(0) - 2\mathbb{E}[Y^2 \mathbf{I}_{(Y<0)}] + \frac{q_Y(p)}{2} \left( \mathbb{E}[|Y|^3] - \theta/\sigma + (2F_Y(0) - 1) \mathbb{E}[Y^3] \right) \right), \\ \lim_{n \rightarrow \infty} \text{Cor} \left( q_{n,\hat{\mu},\hat{\sigma}}(p), \hat{\theta}_n \right) &= \frac{2F_Y(0) - 2\mathbb{E}[Y^2 \mathbf{I}_{(Y<0)}] + \frac{q_Y(p)}{2} \left( \mathbb{E}[|Y|^3] - \theta/\sigma + (2F_Y(0) - 1) \mathbb{E}[Y^3] \right)}{\sqrt{\left(1 + q_Y^2(p) \frac{\mathbb{E}[Y^4]-1}{4} + q_Y(p) \mathbb{E}[Y^3]\right) \text{Var}(|Y| + (2F_Y(0) - 1)Y)}}. \end{aligned}$$

Once again, those expressions clearly simplify when assuming zero third centred moment ( $\mu_3 = 0 = \mathbb{E}[Y^3]$ ), as e.g. for elliptical distributions, or when the mean is known. For instance, in this latter case, we obtain:

$$\lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} q_{n,\hat{\sigma}}(p), \sqrt{n} \hat{\theta}_n \right) = \sigma^2 \frac{q_Y(p)}{2} \left( \mathbb{E}[|Y|^3] - \theta/\sigma + (2F_Y(0) - 1) \mathbb{E}[Y^3] \right), \quad (63)$$

$$\text{and } \lim_{n \rightarrow \infty} \text{Cor} \left( q_{n,\hat{\sigma}}(p), \hat{\theta}_n \right) = \text{sgn}(p - F_Y(0)) \times \frac{\mathbb{E}[|Y|^3] - \theta/\sigma + (2F_Y(0) - 1) \mathbb{E}[Y^3]}{\sqrt{(\mathbb{E}[Y^4] - 1) \text{Var}(|Y| + (2F_Y(0) - 1)Y)}}. \quad (64)$$

Revisiting the two examples of the Gaussian and Student distributions, the same remarks as for Example 15 apply here: The case  $p < 0.5$  is deduced by symmetry, and we need to use the property of the Gamma function to recover the Gaussian expression in the case  $\nu \rightarrow \infty$ . Further, we discussed already for example 15 that  $(C_1)$  and  $(C_2)$  are met; contrary to the case with the sample quantile, the condition  $\nu > 4$  cannot be dropped because of the use of a location-scale quantile. The results of Proposition 21 become, assuming e.g. that the mean  $\mu$  is known:

**Example 22** For  $p \geq 0.5$ , we have:

(i) For the Gaussian distribution  $\mathcal{N}(\mu, \sigma^2)$

$$\lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} q_{n,\hat{\sigma}}(p), \sqrt{n} \hat{\theta}_n \right) = \frac{\sigma^2 \Phi^{-1}(p)}{\sqrt{2\pi}} = (\mu + \sigma \Phi^{-1}(p)) \times \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} \log|q_{n,\hat{\sigma}}(p)|, \sqrt{n} \hat{\theta}_n \right) \quad (65)$$

$$\text{and } \lim_{n \rightarrow \infty} \text{Cor} \left( q_{n,\hat{\sigma}}(p), \hat{\theta}_n \right) = \frac{\text{sgn}(p - 1/2)}{\sqrt{\pi - 2}} = \text{sgn}(p - \Phi(\mu/\sigma)) \times \lim_{n \rightarrow \infty} \text{Cor} \left( \log|q_{n,\hat{\sigma}}(p)|, \hat{\theta}_n \right). \quad (66)$$

(ii) For the Student distribution  $t(\mu, (\nu - 2)\sigma^2/\nu, \nu)$  with  $\nu > 4$  degrees of freedom:

$$\lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} q_{n,\hat{\sigma}}(p), \sqrt{n} \hat{\theta}_n \right) = \frac{\sigma^2(\nu - 1)(\nu - 2)\Gamma(\frac{\nu-1}{2})q_{\tilde{Y}}(p)}{2(\nu - 3)\sqrt{\nu\pi}\Gamma(\nu/2)} = (\mu + \sigma q_{\tilde{Y}}(p)\sqrt{1 - 2/\nu}) \times \text{Cov} \left( \sqrt{n} \log|q_{n,\hat{\sigma}}(p)|, \sqrt{n} \hat{\theta}_n \right), \quad (67)$$

$$\text{and } \lim_{n \rightarrow \infty} \text{Cor} \left( q_{n,\hat{\sigma}}(p), \hat{\theta}_n \right) = \frac{\text{sgn}(p - 1/2)\sqrt{(\nu - 1)(\nu - 2)}}{(\nu - 3)\sqrt{\frac{\pi\Gamma^2(\nu/2)}{\Gamma^2(\frac{\nu-1}{2})} - (\nu - 2)}\sqrt{\frac{2}{\nu-4}}} = \text{sgn} \left( p - F_{\tilde{Y}} \left( \frac{-\mu}{\sigma} \sqrt{\frac{\nu}{\nu - 2}} \right) \right) \times \lim_{n \rightarrow \infty} \text{Cor} \left( \log|q_{n,\hat{\sigma}}(p)|, \hat{\theta}_n \right). \quad (68)$$

**Dependence with Sample Median Absolute Deviation -** As mentioned, we already presented the asymptotic distribution of the location-scale quantile with the MedianAD (Proposition 6, Corollary 7). Therein we saw that the general expressions for the covariance and correlation are very long and tedious. Thus, we consider separately the case of symmetric location-scale distributions, where expressions simplify: As before, we then apply those results to the case of the Gaussian and Student distributions.

**Proposition 23** Consider an iid sample with parent rv  $X$  from a symmetric location-scale distribution having mean  $\mu$ , variance  $\sigma^2$ . Then, under  $(C_1)$ ,  $(M_2)$  and  $(C_3)$  in neighbourhoods of  $\nu, \nu \pm \xi$ , the joint behaviour of the functionals  $h_1$  of the quantile from a location-scale model  $q_{n,\hat{\mu},\hat{\sigma}}(p), q_{n,\hat{\sigma}}(p)$  respectively (for  $p \in (0, 1)$ ) and  $h_2$  of the sample MedianAD  $\hat{\xi}_n$  (defined in Table 1) is asymptotically normal and we have:

(i) when  $\mu$  is unknown

$$\lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} h_1(q_{n,\hat{\mu},\hat{\sigma}}(p)), \sqrt{n} h_2(\hat{\xi}_n) \right) = \frac{\sigma^2 h'_1(q_X(p)) h'_2(\xi)}{4 f_Y(q_Y(3/4))} \times q_Y(p) \left( \frac{1}{2} - 2 \mathbb{E}[Y^2 \mathbf{I}_{(0 < Y \leq q_Y(3/4))}] \right), \quad (69)$$

$$\lim_{n \rightarrow \infty} \text{Cor} \left( h_1(q_{n,\hat{\mu},\hat{\sigma}}(p)), h_2(\hat{\xi}_n) \right) = \text{sgn} \left( h'_1(q_X(p)) h'_2(\xi) \right) \times \frac{q_Y(p)/2 - 2 \mathbb{E}[q_Y(p) Y^2 \mathbf{I}_{(0 < Y \leq q_Y(3/4))}]}{\sqrt{1 + q_Y^2(p) \frac{\mathbb{E}[Y^4] - 1}{4}}}, \quad (70)$$

(ii) when  $\mu$  is known

$$\lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} h_1(q_{n,\hat{\sigma}}(p)), \sqrt{n} h_2(\hat{\xi}_n) \right) = \frac{\sigma^2 h'_1(q_X(p)) h'_2(\xi)}{4 f_Y(q_Y(3/4))} \times q_Y(p) \left( \frac{1}{2} - 2 \mathbb{E}[Y^2 \mathbf{I}_{(0 < Y \leq q_Y(3/4))}] \right), \quad (71)$$

$$\lim_{n \rightarrow \infty} \text{Cor} \left( h_1(q_{n,\hat{\sigma}}(p)), h_2(\hat{\xi}_n) \right) = \text{sgn} \left( h'_1(q_X(p)) h'_2(\xi) \right) \times \text{sgn} \left( p - \frac{1}{2} \right) \times \frac{1 - 4 \mathbb{E}[Y^2 \mathbf{I}_{(0 < Y \leq q_Y(3/4))}]}{\sqrt{\mathbb{E}[Y^4] - 1}}. \quad (72)$$

We observe once again that the asymptotic correlation does not depend on the mean  $\mu$  and the variance  $\sigma^2$  of the underlying location-scale distribution, even if the two expressions differ depending on if  $\mu$  is known or not. Further, the correlation is also constant (in  $p$ ) up to its sign.

Let us go back to our main examples of elliptical distributions, the Gaussian and Student distributions, to see how those results apply, when assuming e.g. that the mean  $\mu$  is known. We already discussed for Example 18 that  $(C_1)$  and  $(C_3)$  are met. Again, as we are working with the location-scale quantile the condition  $\nu > 4$  is necessary. We also get back from the Student case for  $\nu \rightarrow \infty$  the Gaussian expressions.

**Example 24** (i) For the Gaussian distribution  $\mathcal{N}(\mu, \sigma^2)$ , we have:

$$\lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} q_{n,\hat{\sigma}}(p), \sqrt{n} \hat{\xi}_n \right) = \frac{\sigma^2 \Phi^{-1}(3/4)}{2} \Phi^{-1}(p) = (\mu + \sigma q_Y(p)) \times \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} \log |q_{n,\hat{\sigma}}(p)|, \sqrt{n} \hat{\xi}_n \right), \quad (73)$$

$$\lim_{n \rightarrow \infty} \text{Cor} \left( q_{n,\hat{\sigma}}(p), \hat{\xi}_n \right) = 2\sqrt{2} \Phi^{-1} \left( \frac{3}{4} \right) \phi \left( \Phi^{-1} \left( \frac{3}{4} \right) \right) \text{sgn} \left( p - \frac{1}{2} \right) = \text{sgn} \left( p - \Phi^{-1} \left( \frac{-\mu}{\sigma} \right) \right) \lim_{n \rightarrow \infty} \text{Cor} \left( \log |q_{n,\hat{\sigma}}(p)|, \hat{\xi}_n \right). \quad (74)$$

(ii) For the Student distribution  $t(\mu, \sigma^2(\nu - 2)/\nu, \nu)$  with  $\nu > 4$ ,

$$\lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} q_{n,\hat{\sigma}}(p), \sqrt{n} \hat{\xi}_n \right) = q_{\hat{\nu}}(p) \frac{\sigma^2 q_{\hat{\nu}}(3/4) \left( 1 + \frac{q_{\hat{\nu}}^2(3/4)}{\nu} \right)}{2 \frac{\nu}{\nu-2}} = \left( \mu + \sigma \sqrt{\frac{\nu-2}{\nu}} q_Y(p) \right) \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} \log |q_{n,\hat{\sigma}}(p)|, \sqrt{n} \hat{\xi}_n \right), \quad (75)$$

$$\text{and } \lim_{n \rightarrow \infty} \text{Cor} \left( q_{n, \hat{\sigma}}(p), \hat{\xi}_n \right) = \text{sgn}(p - 1/2) \times \frac{2\sqrt{2(\nu - 4)} f_{\tilde{Y}}(q_{\tilde{Y}}(3/4)) q_{\tilde{Y}}(3/4) \left( 1 + \frac{q_{\tilde{Y}}^2(3/4)}{\nu} \right)}{\sqrt{\nu - 1}} \quad (76)$$

$$= \text{sgn} \left( p - F_{\tilde{Y}} \left( \frac{-\mu}{\sigma} \sqrt{\frac{\nu}{\nu - 2}} \right) \right) \times \lim_{n \rightarrow \infty} \text{Cor} \left( \log |q_{n, \hat{\sigma}}(p)|, \hat{\xi}_n \right). \quad (77)$$

### 3.3 The Impact of the Choice of the Quantile Estimator for Elliptical Distributions

In the previous two subsections 3.1 and 3.2, we have explicitly computed the different asymptotic correlation and covariance values in the Gaussian and Student cases when using either the sample quantile  $q_n$ , or the quantile estimators obtained via the location-scale model  $q_{n, \hat{\sigma}}$  with known  $\mu$ .

We observe that those expressions differ depending on the choice of the quantile estimator. To better point out this difference in convergence speed, we consider the ratios of asymptotic covariances or correlations when using the two different quantile estimators. We will not provide those ratios in the general case, since, from a mathematical perspective, this would be redundant, all the quantities we want to put into relation being readily available from the last two subsections. Instead, we will focus on the analysis of the two examples, the Gaussian and the Student distributions. We provide not only the theoretical expressions, but also plots, to compare the rate of convergence.

As the ratios do not depend on the functionals  $h_1, h_2$  of the corresponding quantities, we can focus on the case where  $h_1, h_2$  are the identity functions to build the ratios of asymptotic covariances and correlations respectively. This will make the results more traceable.

Further, recall that the expressions containing the sample MAD are valid only for  $p \geq 0.5$ , the case of  $p < 0.5$  being deduced by the corresponding symmetry around  $p = 0.5$ . In general, by the symmetry around  $p = 0.5$  in all cases, we will focus on  $p \geq 0.5$ .

For the ease of readability, by abuse of notation, the term ‘sample’ in the context of estimators might be omitted in this section as we will be exclusively referring to sample quantities throughout: We will use variance, MAD and MedianAD synonymously for sample variance, sample MAD and sample MedianAD, respectively.

#### 3.3.1 Gaussian Distribution

In the case of a Gaussian distribution  $\mathcal{N}(\mu, \sigma^2)$  with known  $\mu$ , we first look at the ratio of asymptotic covariances in the three cases (sample variance, sample MAD, sample MedianAD).

For the sample variance one needs equations (40) and (59), for the expressions involving the sample MAD equations (46), (65), and finally for the sample MedianAD equations (55) and (73).

$$\begin{aligned}
\frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{\sigma}_n^2)}{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_{n,\hat{\sigma}}(p), \sqrt{n} \hat{\sigma}_n^2)} &= \frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(q_n(p)), \sqrt{n} h_2(\hat{\sigma}_n^2))}{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(q_{n,\hat{\sigma}}(p)), \sqrt{n} h_2(\hat{\sigma}_n^2))} = 1, \\
\frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{\theta}_n)}{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_{n,\hat{\sigma}}(p), \sqrt{n} \hat{\theta}_n)} &= \frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(q_n(p)), \sqrt{n} h_2(\hat{\theta}_n))}{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(q_{n,\hat{\sigma}}(p)), \sqrt{n} h_2(\hat{\theta}_n))} = \frac{\sqrt{2\pi} \left( \phi(\Phi^{-1}(p)) - (1-p) \sqrt{\frac{2}{\pi}} \right)}{\phi(\Phi^{-1}(p)) \Phi^{-1}(p)}, \\
\frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{\xi}_n)}{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_{n,\hat{\sigma}}(p), \sqrt{n} \hat{\xi}_n)} &= \frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(q_n(p)), \sqrt{n} h_2(\hat{\xi}_n))}{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(q_{n,\hat{\sigma}}(p)), \sqrt{n} h_2(\hat{\xi}_n))} = \frac{1-p-2 \max(0, 3/4 - \max(1/4, p))}{2\phi(\Phi^{-1}(p)) \Phi^{-1}(p) \Phi^{-1}(3/4) \phi(\Phi^{-1}(3/4))}.
\end{aligned}$$

We have depicted the three cases in the left plot of Figure 1. In the case with the sample variance, the speed of convergence is identical for the covariances for all values of  $p$  (in the case of comparing the sample quantile with the location-scale quantile with known  $\mu$ ), leading to a constant ratio equal to 1. For the sample MAD the behaviour of the ratio of asymptotic covariances is different. First, the ratio is always positive meaning that both covariances have always the same tendencies. Also, it is bigger than 1, i.e. the asymptotic covariance with  $q_n$  is bigger than with  $q_{n,\hat{\sigma}}$  (up to a factor of 2 for  $p = 0.5$ ), except for the following values of  $p$ ,  $p < 0.07$  and  $p > 0.93$ , where the ratio goes below 1 and drops near to 0.6 (i.e. the asymptotic covariance with  $q_{n,\hat{\sigma}}$  is up to 1.5 times bigger than with  $q_n$ ). Recalling Remark 9, the value is not defined for  $p = 0.5$ , thus we simply set the value by continuity. In the case of the sample MedianAD, the ratio of asymptotic covariances is, in contrast to the case with the sample variance or sample MAD, bi-modal (with maxima at  $p = 0.25, 0.75$ ). Otherwise it follows the same tendencies, the values going below 1 (to values as low as 0.2) for boundary values of  $p$ , here  $p < 0.09$  and  $p > 0.91$ . The other values further in the ‘interior’ of  $(0, 1)$  are above 1 and reach values up to 2.72 at their maxima.

Comparing the asymptotic covariance-ratios for the three measures of dispersion, we see that the speed of convergence for  $q_n, q_{n,\hat{\sigma}}$  is similar for  $p$  around their intersection points (i.e. around 0.1 or 0.9). For  $p$  near to the extremes, the variance and MAD seem somewhat similar, while the covariance-ratio for the MedianAD is a lot lower in this cases (approximately half the value of the MAD and one fourth of the variance). For  $p$ -values in the interior of  $(0, 1)$ , the difference between the measures of dispersion increases (the difference between MedianAD and MAD or variance respectively is the highest around  $p = 0.25$  or  $0.75$ , while between MAD and standard deviation for  $p$  around 0.5).

Subsequently, let us turn to the ratio of the correlations in the Gaussian case, using Equations (41, 60), (47, 66) and (56, 74) for the ratios with the sample variance, the sample MAD and the sample MedianAD respectively:

$$\begin{aligned}
\frac{\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\sigma}_n^2)}{\lim_{n \rightarrow \infty} \text{Cor}(q_{n,\hat{\sigma}}(p), \hat{\sigma}_n^2)} &= \frac{\lim_{n \rightarrow \infty} \text{Cor}(h_1(q_n(p)), h_2(\hat{\sigma}_n^2))}{\lim_{n \rightarrow \infty} \text{Cor}(h_1(q_{n,\hat{\sigma}}(p)), h_2(\hat{\sigma}_n^2))} = \frac{|\Phi^{-1}(p)| \phi(\Phi^{-1}(p))}{\sqrt{2p(1-p)}} = \left| \lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\sigma}_n^2) \right|, \\
\frac{\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\theta}_n)}{\lim_{n \rightarrow \infty} \text{Cor}(q_{n,\hat{\sigma}}(p), \hat{\theta}_n)} &= \frac{\lim_{n \rightarrow \infty} \text{Cor}(h_1(q_n(p)), h_2(\hat{\theta}_n))}{\lim_{n \rightarrow \infty} \text{Cor}(h_1(q_{n,\hat{\sigma}}(p)), h_2(\hat{\theta}_n))} = \frac{|\sqrt{\pi} \phi(\Phi^{-1}(p)) - (1-p) \sqrt{2}|}{\sqrt{p(1-p)}} = \sqrt{\pi - 2} \left| \lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\theta}_n) \right|, \\
\frac{\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\xi}_n)}{\lim_{n \rightarrow \infty} \text{Cor}(q_{n,\hat{\sigma}}(p), \hat{\xi}_n)} &= \frac{\lim_{n \rightarrow \infty} \text{Cor}(h_1(q_n(p)), h_2(\hat{\xi}_n))}{\lim_{n \rightarrow \infty} \text{Cor}(h_1(q_{n,\hat{\sigma}}(p)), h_2(\hat{\xi}_n))} = \frac{\sqrt{2} |1-p-2 \max(0, 3/4 - \max(1/4, p))|}{4\Phi^{-1}(3/4) \phi(\Phi^{-1}(3/4)) \sqrt{p(1-p)}}.
\end{aligned}$$

Again, we depict the results in Figure 1 (the middle plot has the same scale as the covariance plot, the right plot an absolute scale): We see that the tendencies are the same for all three measures of dispersion considered: The ratio is symmetric around the axis  $p = 0.5$ , always below 1 meaning that the asymptotic correlation is bigger for  $q_{n,\hat{\sigma}}$  than  $q_n$  and equal to 0 for  $p = 0.5$ . For the sample variance, the result is due to the fact that  $\lim_{n \rightarrow \infty} \text{Cor}(q_{n,\hat{\sigma}}(p), \hat{\sigma}_n^2) = 1$ .

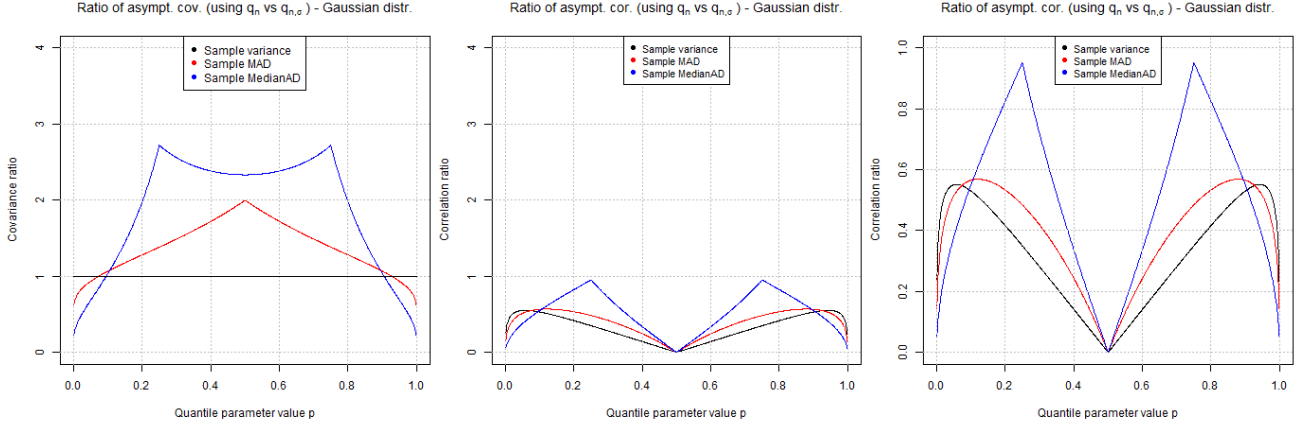


Figure 1: Ratio of asymptotic covariances (left) and correlations (in the middle on same scale as covariance, on the right on absolute scale) between the quantile estimator with the measure of dispersion estimator, when using  $q_n$  versus  $q_{n,\hat{\sigma}}$ , considering three different measures of dispersion each (sample variance, sample MAD, sample MedianAD) in the case of a Gaussian underlying distribution.

For the sample MAD, as noticed, it is structurally similar. The ratio equals  $\sqrt{\pi - 2} \times \lim_{n \rightarrow \infty} \text{Cor} \left( q_n(p), \hat{\theta}_n \right)$ , i.e. it is multiplied by a factor bigger than 1. The overall amplitude is slightly bigger than in the case for the sample variance. With the sample MedianAD, the ratio of asymptotic correlations behaves as for the sample variance and sample MAD, and has a similar shape. But the amplitude is different and is clearly the biggest in this case. Also, from all three measures of dispersion, it is the nearest to 1 at their maxima (value of 0.95 at  $p = 0.25, 0.75$ ), i.e. there the asymptotic correlation is similar for  $q_n$  and  $q_{n,\hat{\sigma}}$ .

Again, we compare the ratios for the three measures of dispersion and look only at the behaviour for  $p \geq 0.5$  as, by symmetry around  $p = 0.5$ , we can observe the same behaviour for values of  $p < 0.5$ . We see that for values of  $p$  around 0.91, the ratios behave similarly in all three cases, while they are most different around  $p=0.75$  (the maximum of the MedianAD correlation-ratio). For  $p > 0.91$ , standard deviation and MAD look very similar, while MedianAD only for values very near to 1.

### 3.3.2 Student Distribution

Let us now consider the case of a Student distribution  $t(\mu, \sigma^2(\nu - 2)/\nu, \nu)$  with known  $\mu$  and  $\nu > 4$  degrees of freedom (to have all quantities well-defined). To compare the behaviour also to the Gaussian distribution (as we know that we get back the Gaussian expressions for  $\nu \rightarrow \infty$ ), we consider three different degrees of freedom in the plots:  $\nu = 5, 10, 40$ .

Referring to Equations (42, 61), (48, 67) and (55, 75) for the three dispersion measures, respectively, we get for the ratio of asymptotic covariances:



$$\begin{aligned}
\frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{\sigma}_n^2)}{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_{n,\hat{\sigma}}(p), \sqrt{n} \hat{\sigma}_n^2)} &= \frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(q_n(p)), \sqrt{n} h_2(\hat{\sigma}_n^2))}{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(q_{n,\hat{\sigma}}(p)), \sqrt{n} h_2(\hat{\sigma}_n^2))} = \frac{\nu - 4}{\nu - 1} \left( 1 + \frac{q_{\hat{Y}}^2(p)}{\nu} \right), \\
\frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{\theta}_n)}{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_{n,\hat{\sigma}}(p), \sqrt{n} \hat{\theta}_n)} &= \frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(q_n(p)), \sqrt{n} h_2(\hat{\theta}_n))}{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(q_{n,\hat{\sigma}}(p)), \sqrt{n} h_2(\hat{\theta}_n))} = \frac{\frac{\sqrt{\nu\pi}}{\nu-1} f_{\hat{Y}}(q_{\hat{Y}}(p)) \left( 1 + \frac{q_{\hat{Y}}^2(p)}{\nu} \right) - (1-p) \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)}}{f_{\hat{Y}}(q_{\hat{Y}}(p)) \frac{q_{\hat{Y}}(p)}{2} \frac{\nu-1}{\nu-3} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)}}, \\
\frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{\xi}_n)}{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_{n,\hat{\sigma}}(p), \sqrt{n} \hat{\xi}_n)} &= \frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(q_n(p)), \sqrt{n} h_2(\hat{\xi}_n))}{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(q_{n,\hat{\sigma}}(p)), \sqrt{n} h_2(\hat{\xi}_n))} = \frac{1 - p - 2 \max(0, 3/4 - \max(1/4, p))}{2 f_{\hat{Y}}(q_{\hat{Y}}(p)) q_{\hat{Y}}(p) f_{\hat{Y}}(q_{\hat{Y}}(3/4)) q_{\hat{Y}}(3/4) (1 + \frac{q_{\hat{Y}}^2(3/4)}{\nu})}.
\end{aligned}$$

Let us have a look at these behaviours in the first row of Figure 2, where the cases of  $\nu = 5, 10, 40$  are presented in separate plots. We first consider the left plot for  $\nu = 5$ . For the sample variance, recalling the case of the ratio for the Gaussian distribution ( $= 1, \forall p \in (0, 1)$ ), it is interesting to see how it behaves for finite  $\nu$ : The ratio values are below 1 for  $p_{thresh} < p < 1 - p_{thresh}$  where  $p_{thresh} \approx 0.01$ . The lowest value is 0.25 meaning that the asymptotic covariance with  $q_{n,\hat{\sigma}}$  is four times as big as with  $q_n$ . For  $p < p_{thresh}, p > 1 - p_{thresh}$  the ratios are above 1 with exploding growth for  $p$  tending to 0 or 1.

In the case of the sample MAD and for  $\nu = 5$ , we see a somewhat similar behaviour as for the asymptotic covariance with the sample variance. The ratio is below 1 for  $0.01 < p < 0.99$  and within this range, also reasonably close to 1, especially for  $p$  near 0.5. As for the sample variance, the ratio seems to explode for  $p$  tending to 0 or 1. With increasing  $\nu$ , the behaviour where ratios are below and above 1 swaps, which we look at in more detail in the second row of Figure 2.

The ratios of asymptotic covariances for the sample MedianAD behave as in the case of the Gaussian distribution (with values being a little bit different for small  $\nu$ ). It means that the ratio of asymptotic covariances is bi-modal and above (below, respectively) 1 for values of  $p$  in (outside of, respectively) the interval (0.07, 0.93). Its maxima are at 0.25 and 0.75 where the ratio reaches a value of 2.71. It is clearly the ratio with the largest amplitude.

Comparing the three measures of dispersion, we see, for  $\nu = 5$ , that their ratio is similar for values around the interval (0.96, 0.99), while very different otherwise (mostly a factor of 2 or more). We also see that for the extreme values of  $p$ , the covariance ratio behaves very similar for the variance and the MAD. Interestingly, the relation between the covariance of the three measures of dispersion changes for increasing  $\nu$ . While they still behave similarly around (0.96, 0.99), we can see that MAD and MedianAD get closer to each other around  $p \approx 0.5$ , whereas in the limit  $p \rightarrow 1$ , the MAD and the variance do not show the same behaviour anymore (they are rather divergent).

In the second row of Figure 2, we present the ratios of the asymptotic covariances again, but in a different way. For each given measure of dispersion we have a plot on its own, showing the cases of different Student distributions as well as the Gaussian distribution in order to compare how big the differences between those two elliptical distributions are. As we already could see from the first row of Figure 2, the difference between the Student and Gaussian cases is clearly visible for the sample variance and the sample MAD - while it is small for the sample MedianAD. With the sample variance, we see that the covariance ratio for values near  $p = 0$  or 1, is clearly distinct from the Gaussian case for small values of  $\nu$ . Even for  $\nu = 40$ , we still see the difference. Also, for increasing  $\nu$ , the value of  $p_{thresh}$  is decreasing (but only minimally), but all the values are getting nearer to 1 (as expected). In the case of the sample MAD, we see that the behaviour of the ratio of covariances is, for small degrees of freedom, quite different from the Gaussian case. The shape of the covariance ratio for  $p$  near 0 or 1 changes from small values for  $\nu$  (ratio above 1) to bigger values of  $\nu$  (ratio below 1). For values in the interior of (0, 1), we see the reverse behaviour. For  $\nu = 5$ , the values are below and, for  $p = 0.5$ , very near to 1, while for increasing  $\nu$ , those values are clearly above 1. For the covariance ratio with the sample MedianAD, we

observe, as already remarked, that for any degree of freedom considered, the values are very close to the Gaussian case.

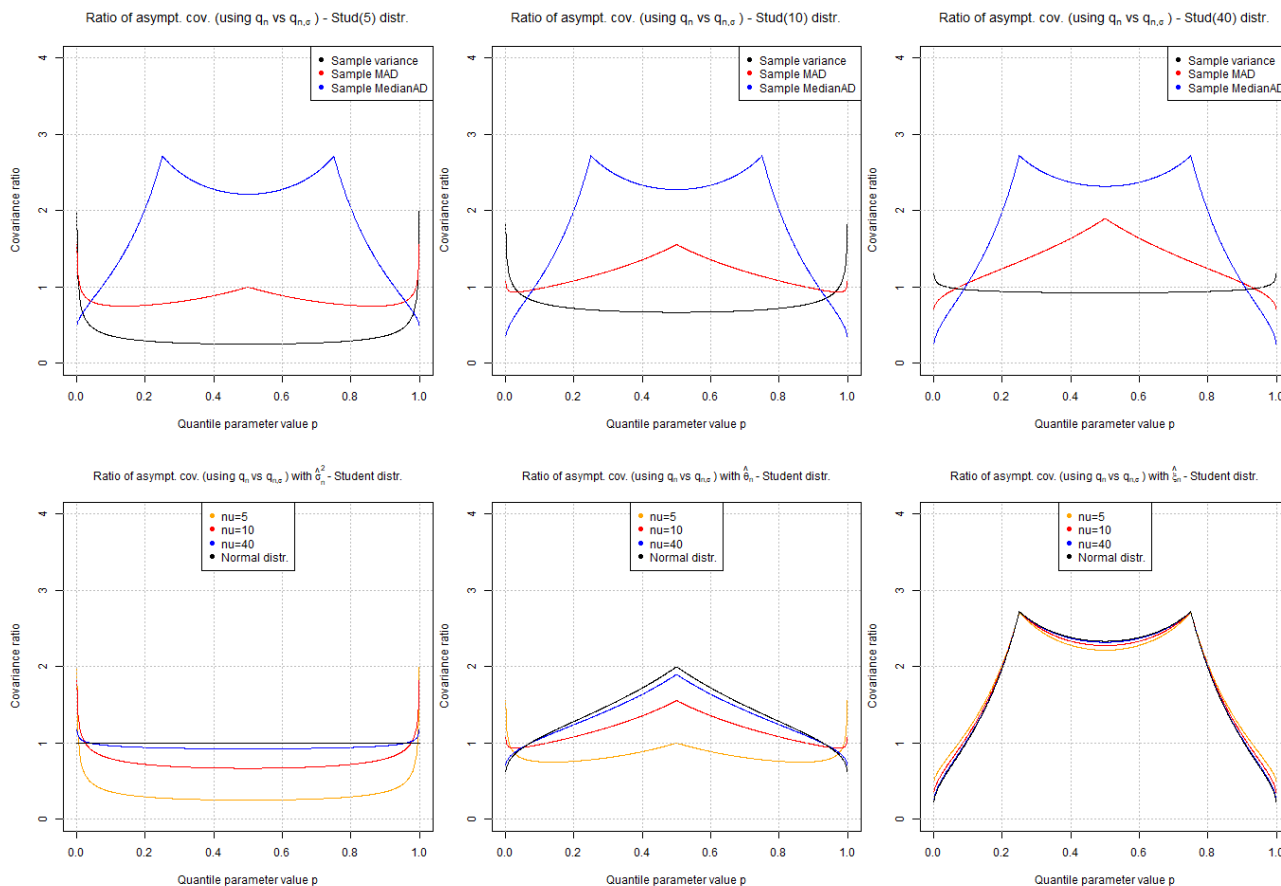


Figure 2: Ratio of asymptotic covariances between the quantile estimator with the measure of dispersion estimator, when using  $q_n$  versus  $q_{n,\hat{\sigma}}$ , considering three different measures of dispersion each (sample variance, sample MAD, sample MedianAD) in the case of a Student distribution. First row: As a function of the measure of dispersion with 5, 10 and 40 degrees of freedom (from left to right). Second row: As a function of the degrees of freedom for each measure of dispersion (sample variance, sample MAD, sample MedianAD -from left to right).

Finally, let us look at the ratio of asymptotic correlations for the Student distribution. Using the pairs of equations (43, 62), (50, 68) and (56, 76) respectively, the ratios are as follows:

$$\begin{aligned}
\frac{\lim_{n \rightarrow \infty} \text{Cor} (q_n(p), \hat{\sigma}_n^2)}{\lim_{n \rightarrow \infty} \text{Cor} (q_{n,\hat{\sigma}}(p), \hat{\sigma}_n^2)} &= \frac{\lim_{n \rightarrow \infty} \text{Cor} (h_1(q_n(p)), h_2(\hat{\sigma}_n^2))}{\lim_{n \rightarrow \infty} \text{Cor} (h_1(q_{n,\hat{\sigma}}(p)), h_2(\hat{\sigma}_n^2))} = \frac{|q_{\hat{Y}}(p)| \left(1 + q_{\hat{Y}}^2(p)/\nu\right) f_{\hat{Y}}(q_{\hat{Y}}(p))}{\sqrt{2 \frac{\nu-1}{\nu-4} p(1-p)}}. \\
\frac{\lim_{n \rightarrow \infty} \text{Cor} (q_n(p), \hat{\theta}_n)}{\lim_{n \rightarrow \infty} \text{Cor} (q_{n,\hat{\sigma}}(p), \hat{\theta}_n)} &= \frac{\lim_{n \rightarrow \infty} \text{Cor} (h_1(q_n(p)), h_2(\hat{\theta}_n))}{\lim_{n \rightarrow \infty} \text{Cor} (h_1(q_{n,\hat{\sigma}}(p)), h_2(\hat{\theta}_n))} \\
&= \left| \frac{\sqrt{\nu\pi}}{\nu-1} f_{\hat{Y}}(q_{\hat{Y}}(p)) \left(1 + \frac{q_{\hat{Y}}^2(p)}{\nu}\right) - (1-p) \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)} \right| \times \frac{\sqrt{2}(\nu-3)\Gamma(\nu/2)}{\sqrt{p(1-p)}\sqrt{(\nu-4)(\nu-1)}\Gamma(\frac{\nu-1}{2})}. \\
\frac{\lim_{n \rightarrow \infty} \text{Cor} (q_n(p), \hat{\xi}_n)}{\lim_{n \rightarrow \infty} \text{Cor} (q_{n,\hat{\sigma}}(p), \hat{\xi}_n)} &= \frac{\lim_{n \rightarrow \infty} \text{Cor} (h_1(q_n(p)), h_2(\hat{\xi}_n))}{\lim_{n \rightarrow \infty} \text{Cor} (h_1(q_{n,\hat{\sigma}}(p)), h_2(\hat{\xi}_n))} \\
&= \frac{|1-p-2\max(0, 3/4-\max(1/4, p))|}{\sqrt{p(1-p)}} \times \frac{\sqrt{2 \frac{\nu-1}{\nu-4}}}{4 f_{\hat{Y}}(q_{\hat{Y}}(3/4)) q_{\hat{Y}}(3/4) \left(1 + q_{\hat{Y}}^2(3/4)/\nu\right)}.
\end{aligned}$$

We compare the three different measures of dispersion in the first row of Figure 3, again, separately for  $\nu = 5, 10, 40$ . The general tendencies for all three measures of dispersion (sample variance, sample MAD, sample MedianAD), even for different degrees of freedom, are as in the Gaussian case: The ratio is symmetric around  $p = 0.5$ , tends to small values for  $p$  tending to 0 or 1 and equals 0 for  $p = 0.5$ . Let us look at the case  $\nu = 5$  (left plot). The correlation-ratio with the sample variance is below 0.5 for all values of  $p$ . Starting from  $p = 0.5$  (with a value of 0), it is increasing in  $p$  until its maximum of 0.43 at 0.99, afterwards it decreases. The MAD behaves as the variance, only that the values increase from 0 at  $p = 0.5$  much faster and reach a maximum of 0.68 at  $p = 0.92$ . Also, the decrease afterwards is quicker, as one can see from the fact that, for  $p \rightarrow 1$ , the values seem lower than for the variance. Again, the MedianAD has the largest amplitude, its maximum of 1.81 is achieved at  $p = 0.75$ , i.e., in contrast to the other measures of dispersion, it has ratio values above 1 (in the interval  $(0.65, 0.91)$ ). As the decrease was the largest compared to the two other measures of dispersion, the increase too. The ratio for  $p \rightarrow 1$  has values around 0.2 meaning that the asymptotic correlation with  $q_{n,\hat{\sigma}}$  is five times bigger than with  $q_n$ .

Comparing the three measures of dispersion, we can make similar comments as in the Gaussian case. The ratio values are similar in a (very small) neighbourhood of  $p = 0.5$  and between the maximum for the MAD and the variance (i.e. in the interval  $(0.96, 0.996)$ ). The biggest difference between the ratios is for the values around the maximum for the MedianAD ratio ( $p = 0.75$ ).

Then, in the second row of Figure 3, we compare the correlation-ratios for each measure of dispersion separately, as a function of the degrees of freedom of the underlying Student distribution. We see a behaviour that is different to the case of the covariance-ratios in the second row of Figure 2. For the sample variance, we can clearly see the different correlation-ratios for the different values of  $\nu$ . Around  $p = 0.5$ , the values are close in the different cases, also somewhat in the very tails (but not the limit). The difference is the largest around the values of the maximum of the Gaussian correlation-ratio (i.e.  $p \approx 0.94$ ). In the case of the sample MAD, we do not have a change in behaviour of the correlation-ratio (i.e. being below or above 1 in dependence of  $\nu$ ), as we had for the covariance-ratio. We see a smooth convergence in  $\nu$  towards the Gaussian case. For most values of  $p$ , there is little difference of the correlation-ratio values as a function of  $\nu$ . Only around the maximum for  $\nu = 5$  ( $p = 0.92$ ), we have a relatively slow convergence in  $\nu$  to the Gaussian case. While for the sample MedianAD covariance-ratio

the difference between the different Student cases and the case of the Gaussian distribution was very small, the difference is pronounced for the correlation ratio: Apart from  $p$  around 0, 0.5 and 1, we can clearly see the difference of the correlation-ratios for  $\nu = 5, 10, 40$  and the Gaussian case. As for the MAD, the difference is the biggest around the correlation-ratio maximum for  $\nu = 5$  ( $p = 0.75$ ). But overall, we have a smooth convergence with increasing degrees of freedom to the Gaussian case.

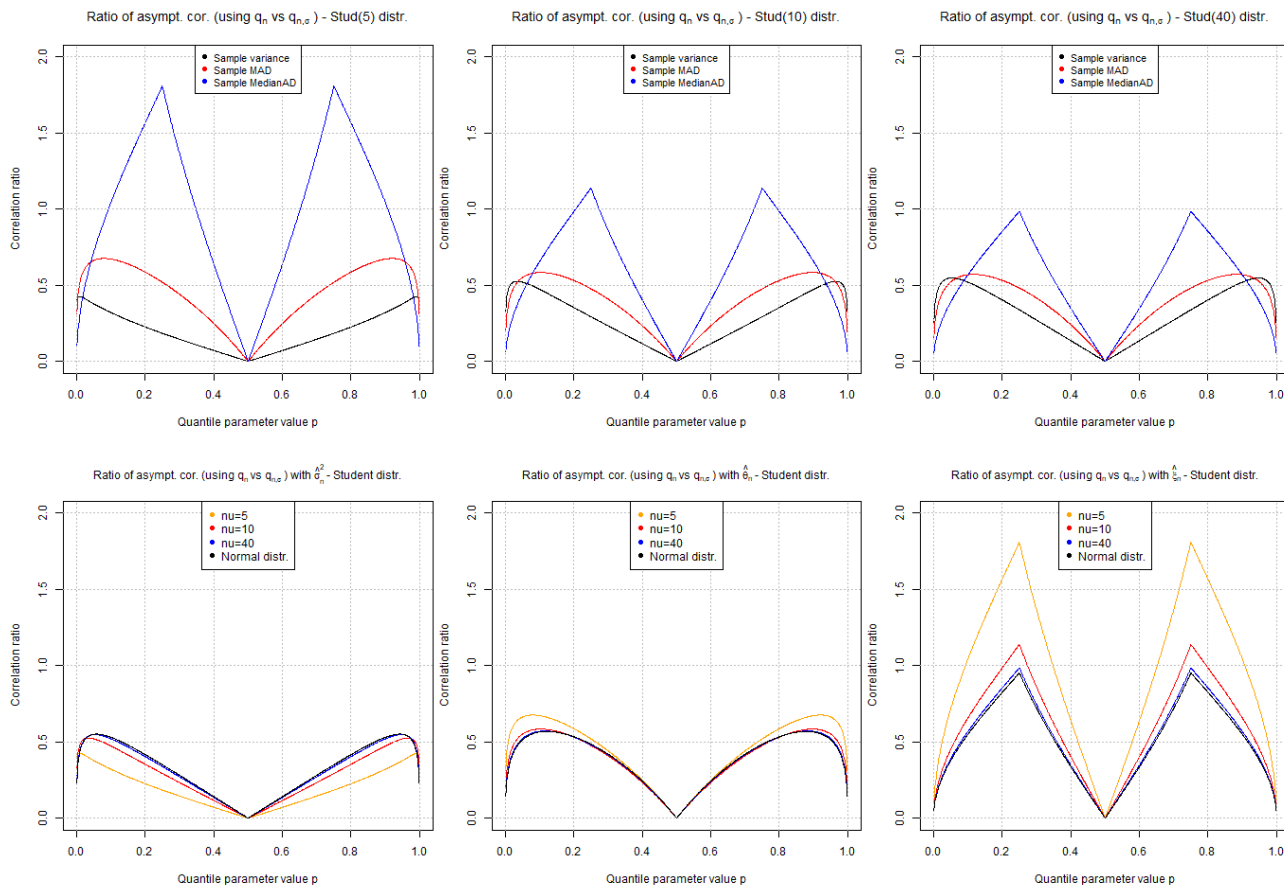


Figure 3: Ratio of asymptotic correlations between the quantile estimator with the measure of dispersion estimator, when using  $q_n$  versus  $q_{n,\hat{\sigma}}$ , considering three different measures of dispersion each (sample variance, sample MAD, sample MedianAD) in the case of a Student distribution. First row: As a function of the measure of dispersion with 5, 10 and 40 degrees of freedom (from left to right). Second row: As a function of the degrees of freedom for each measure of dispersion (sample variance, sample MAD, sample MedianAD -from left to right).

## 4 Application to Quantitative Risk Management

Based on the previous results of this paper, we can now draw a statistical framework for applications in quantitative risk management. We start by giving a brief overview of some well-known risk measures (Value-at-Risk, Expected Shortfall and expectile) and discuss the dispersion measures in the context of quantitative risk management. Using the theoretical asymptotic results from the previous sections, we give corresponding results for the asymptotic dependence between risk measure estimators and measure of dispersion estimators; they might add another layer to the discussion of which risk measure might be best to use (from a theoretical as well as practical point of view). As before, we give explicit formulae in the cases of the Gaussian and Student distribution.

We then turn to an example from finance, which initially raised the general question treated in this study. We relate our theoretical work to empirical findings in the literature as e.g. in [9],[43] (and the related Section 3.6 in Zumbach’s book [44]). To do so, we assume to have an underlying iid distribution and present the explicit asymptotics for the two distributions considered throughout, the Gaussian and Student distributions, using the results from the previous sections. We then look at how the sample size influences the dependence results for this specific example and discuss theoretically the changes observed when considering longer sample sizes in an asymptotic sense. We conclude by discussing the usefulness of these theoretical results for finite sample applications: We compare, for the different risk and dispersion measure estimators, the values provided through the analytical formulae with those obtained via simulated iid samples for the two main elliptical examples of this paper in a setting corresponding to the empirical analysis performed in [9].

## 4.1 Estimation with Various Risk and Dispersion Measures

While we already introduced the dispersion measures and their estimators in Section 1, we have a closer look how they are used in the context of quantitative risk management. Also, we briefly introduce the three risk measures considered and link their estimators to the quantile estimators. Finally, we provide explicit formulae for asymptotic correlations of the estimators of those risk and dispersion measures, and plot them in the case of the Gaussian and Student distribution.

### 4.1.1 Dispersion Measures

When talking about dispersion measures, we focus on those used in this note: The standard deviation (or equivalently the variance), the mean absolute deviation and the median absolute deviation.

While the standard deviation is the most popular measure of spread/dispersion, there are different reasons for considering alternatives in some situations. From a statistical as well as applied point of view, one of them is the theoretical moment condition the sample estimators might require. This we have commented on, throughout the paper, and we summarise it again in Table 3: The variance and MAD only exist for a distribution with finite variance, in contrast to the MedianAD, which exists even for an infinite-mean distribution, and the estimation of such quantities only makes sense if the true parameter exists (in finance, such observations date back to Fama’s analysis of stock-market prices, see [16]). To work with the asymptotic normality of the estimators, we have more restrictive moment conditions: For the sample variance we need the fourth moment to exist, whereas for the sample MAD only the second moment, and for the sample MedianAD no moment condition is necessary. Also, from the so-called Bahadur representations of the sample MAD and sample MedianAD ([4], [31] respectively), one can see that neither of the two estimators is unbiased (although they are asymptotically unbiased, i.e.  $\lim_{n \rightarrow \infty} \mathbb{E}[\hat{\lambda}_n - \lambda] = 0$  for an estimator  $\hat{\lambda}_n$  of the parameter  $\lambda$ ). As the three estimators have non-degenerate asymptotic distribution functions (when the conditions on the underlying distribution are met), they are all three consistent.

Table 3: Overview of moment conditions related to statistical properties of the measure of dispersion estimators

Estimator	Existence of theoretical counterpart	Consistency / Unbiasedness	Moment condition for asymptotic normality
Sample Variance ( $\hat{\sigma}_n^2$ )	$\mathbb{E}[X^2] < \infty$	Yes/Yes	$\mathbb{E}[X^4] < \infty$
Sample MAD ( $\hat{\theta}_n$ )	$\mathbb{E}[ X ] < \infty$	Yes/No	$\mathbb{E}[X^2] < \infty$
Sample MedianAD ( $\hat{\xi}_n$ )	no moment condition	Yes/No	none

Also, clearly, the choice of dispersion measure depends on the objective of the estimation. Without doubt, the sample MedianAD is the most robust estimator with respect to outliers, whereas the sample variance is very sensitive to such and the sample MAD being in between the two (see [20] for someone advocating the MAD; also [28], for a discussion about outlier detection from an applied point of view). For a good overview over advantages and properties of these three different measures of dispersion, see [20] and [34] and the references therein.

The discussion about which measure of dispersion to use, also takes place in areas of QRM. For instance, in [13], this is addressed for the so-called ‘realised volatility’ in financial markets (considering a generalization of the concepts of sample standard deviation and sample MAD), saying that the choice (most often either sample standard deviation or sample MAD) should be made taking into account the existence of the moments of the underlying distribution. For premium calculation in insurance, [14] advocates for the MedianAD instead of the standard deviation (as the risk functionals corresponding to the former are comonotonic additive, whereas the ones corresponding to the latter not). In the context of portfolio optimization, [26] propose a portfolio choice based on the mean absolute deviation instead of the variance minimization as in the classic approach of Markowitz ([30]); but the advantages gained in computational feasibility, so they claim in [41], can be lost again by the higher estimation error. Also, in the context of regression, alternatives to least squares are proposed in the spirit of the dispersion measures we consider: Least absolute deviations regression (see e.g. [7]) trying to remedy the fact that least-squares regression gives single ‘outliers’, in their opinion, a too high weight in the fitting procedure. Or, the least median of squares regression, [38], which has superior ‘breakdown point’ (i.e. ‘the smallest percentage of contaminated data that can cause the estimator to take on arbitrarily large aberrant values’, [38]) in comparison to least squares or least absolute deviation regression. As they use the median of the squares in [38], this is different to the use of the median in our case - but it still illustrates the idea of using the median.

#### 4.1.2 Risk Measures

In contrast to the dispersion measures, risk measures have not appeared explicitly in the previous sections. Nevertheless, one of the most used risk measures, Value-at-Risk (VaR), is simply a quantile at a certain level of the underlying distribution and thus, the results presented directly relate to risk estimation. The VaR for risk management was popularised by JP Morgan in 1996 (see [32]) and is defined as follows: If we assume a loss random variable  $L$  having a continuous, strictly increasing distribution function  $F_L$ , the VaR at level  $\alpha$  of  $L$  is simply the quantile of order  $\alpha$  of  $L$ :

$$\text{VaR}_\alpha(L) = \inf \left\{ x : P[L \leq x] \geq \alpha \right\} = F_L^{-1}(\alpha). \quad (78)$$

The VaR is generally estimated on historical data, using the empirical quantile  $\widehat{\text{VaR}}_n(\alpha) = q_n(\alpha)$  associated to a  $n$ -loss sample  $(L_1, \dots, L_n)$  with  $\alpha \in (0, 1)$  - hence the direct connection to the theoretical

results presented previously.

The use of VaR has been reinforced by regulators even if it has been shown not to be a coherent measure, [3], contrary to Expected Shortfall (ES). ES is defined as follows (e.g. [1]) for a loss random variable  $L$  and a level  $\alpha \in (0, 1)$  :

$$\text{ES}_\alpha(L) = \frac{1}{1-\alpha} \int_\alpha^1 q_L(u) du = \mathbb{E}[L|L \geq q_L(\alpha)]. \quad (79)$$

While the first equality in (79) is the definition of ES, the second holds only if  $L$  is continuous (that we assume throughout the paper). There are different ways of estimating ES, we consider the two most direct ones when using historical estimation. First, simply estimate the quantile by the sample quantile, leading to  $\widetilde{\text{ES}}_n(\alpha) := \frac{1}{1-\alpha} \int_\alpha^1 q_n(u) du$ , which can be seen as the limit ( $k \rightarrow \infty$ ) of the following second ES estimator  $\widehat{\text{ES}}_{n,k}(\alpha) := \frac{1}{k} \sum_{i=1}^k q_n(p_i)$  for  $\alpha = p_1 < p_2 < \dots < p_k < 1$ . As in the context of regulation the level of  $\alpha$  is high, in practice it is sufficient for  $k$  to be small, [12] claim. The question of the appropriateness of the risk measure to use for evaluating the risk of financial institutions has been heavily debated especially after the financial crisis of 2008/2009. For a review of the arguments on this subject, we refer e.g. to [10], [15] and the references therein.

In these discussions, as a third risk measure the expectile is considered (introduced in the context of least-squares estimation by [33], as a risk measure in [27]). It is defined, for a square-integrable random variable and level  $\alpha \in (0, 1)$ , by the following minimiser

$$e_\alpha(L) = \underset{x \in \mathbb{R}}{\text{argmin}} \alpha \mathbb{E}[\max(L - x, 0)^2] + (1 - \alpha) \mathbb{E}[\max(x - L, 0)^2]. \quad (80)$$

While a natural estimator for the expectile is the empirical argmax of (80), there exists another way to define an estimator of  $e_\alpha$ , when recalling the relation between an expectile and quantile, see [42]: Let  $q_X(\alpha)$  be the quantile at level  $\alpha \in (0, 1)$ , then there exists a bijection  $\kappa : (0, 1) \mapsto (0, 1)$  such that  $e_{\kappa(\alpha)}(X) = q_X(\alpha)$  with

$$\kappa(\alpha) = \frac{\alpha q_X(\alpha) - \int_{-\infty}^{q_X(\alpha)} x dF_X(x)}{\mathbb{E}[X] - 2 \int_{-\infty}^{q_X(\alpha)} x dF_X(x) - (1 - 2\alpha) q_X(\alpha)}. \quad (81)$$

Thus, we define a sample estimator for the expectile at level  $\alpha$  to be  $e_n(\alpha) := q_n(\kappa^{-1}(\alpha))$ .

Having presented the three most popular risk measures in quantitative risk management, let us see how they relate to the theoretical findings in this paper.

### 4.1.3 Implications of the Theoretical Results

As mentioned, the consequences for the VaR are direct. Recall that we presented asymptotic distributions for a quantile estimator with any of the three measures of dispersion (Theorem 1, Proposition 3). Hence, if we estimate the VaR at threshold  $p$  as  $\widehat{\text{VaR}}_n(p)$ , the results directly hold for the VaR: We can say how any functional of the VaR estimator  $\widehat{\text{VaR}}_n(p)$  is correlated with any functional of a measure of dispersion estimator.

For the case of the ES, the results with the sample quantile are directly applicable for  $\widetilde{\text{ES}}_n(p)$  (using Fubini theorem and dominated convergence, we can simply interchange limit and integral). Hence, we obtain the following asymptotic dependence between  $\widetilde{\text{ES}}_n(p)$  and either the sample variance or sample MAD:

**Corollary 25** Consider an iid sample with parent rv  $X$  having mean  $\mu$ , variance  $\sigma^2$ . Assume conditions  $(C_1)$ ,  $(C_3)$  in each neighbourhood of  $q_X(u)$ ,  $u \in (p, 1)$ , and  $(M_r)$  for  $r = 1, 2$ , respectively, as well as  $\underline{C}_2$  at  $\mu$  for  $r = 1$ . Then the joint asymptotic behaviour of the historically estimated expected shortfall  $\widetilde{\text{ES}}_n(p)$ , for  $p \in (0, 1)$ , and the measure of dispersion estimator  $\hat{m}(X, n, r)$  is normal with the following asymptotic covariance:

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\widetilde{\text{ES}}_n(p), \sqrt{n}\hat{m}(X, n, r)) &= \frac{1}{1-p} \int_p^1 \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_n(u), \sqrt{n}\hat{m}(X, n, r)) du \\ &= \frac{1}{1-p} \int_p^1 \frac{\tau_r(|X - \mu|, u) + (2-r)(2F_X(\mu) - 1)\tau_1(u)}{f_X(q_X(u))} du, \end{aligned}$$

with  $\tau_r$  defined in (3). Accordingly the asymptotic correlation is:

$$\lim_{n \rightarrow \infty} \text{Cor}(\widetilde{\text{ES}}_n(p), \hat{m}(X, n, r)) = \frac{\int_p^1 \frac{\tau_r(|X - \mu|, u) + (2-r)(2F_X(\mu) - 1)\tau_1(u)}{f_X(q_X(u))} du}{\sqrt{2 \int_p^1 \int_v^1 \frac{v(1-u)}{f_X(q_X(v))f_X(q_X(u))} dudv} \sqrt{\text{Var}(|X - \mu|^r + (2-r)(2F_X(\mu) - 1)X)}}. \quad (82)$$

For the dependence of the ES estimator with the MedianAD estimator, we obtain analogously:

**Corollary 26** Consider an iid sample with parent rv  $X$  with, if defined, mean  $\mu$  and variance  $\sigma^2$ . Under conditions  $(C_1)$  and  $(C_3)$  in the neighbourhoods of  $\nu, \nu \pm \xi$  as well as in each neighbourhood of  $q_X(u)$ ,  $u \in (p, 1)$ , the joint asymptotic behaviour of the historically estimated expected shortfall  $\widetilde{\text{ES}}_n(p)$ , for  $p \in (0, 1)$ , and the sample MedianAD  $\hat{\xi}_n$  (defined in Table 1) is normal with the following asymptotic covariance:

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\widetilde{\text{ES}}_n(p), \sqrt{n}\hat{\xi}_n) &= \frac{1}{1-p} \int_p^1 \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_n(u), \sqrt{n}\hat{\xi}_n) du \\ &= \frac{1}{1-p} \int_p^1 \frac{-\max(0, F_X(\nu + \xi) - \max(F_X(\nu - \xi), u)) + \frac{1-u}{2} + \frac{f_X(\nu + \xi) - f_X(\nu - \xi)}{f_X(\nu)} \max\left(-\frac{p}{2}, \frac{p-1}{2}\right)}{f_X(q_X(u)) (f_X(\nu + \xi) + f_X(\nu - \xi))} du \end{aligned}$$

where  $\nu$  denotes the the median of the sample (see Table 1). The asymptotic correlation equals, if we denote, as before,  $\gamma = (f_X(\nu + \xi) - f_X(\nu - \xi)) f_X(\nu) (f_X(\nu + \xi) - f_X(\nu - \xi) - 4) (1 - F_X(\nu - \xi) - F_X(\nu + \xi))$ :

$$\lim_{n \rightarrow \infty} \text{Cor}(\widetilde{\text{ES}}_n(p), \hat{\xi}_n) = \frac{\int_p^1 \frac{-\max(0, F_X(\nu + \xi) - \max(F_X(\nu - \xi), u)) + \frac{1-u}{2} + \frac{f_X(\nu + \xi) - f_X(\nu - \xi)}{f_X(\nu)} \max\left(-\frac{p}{2}, \frac{p-1}{2}\right)}{f_X(q_X(u)) (f_X(\nu + \xi) + f_X(\nu - \xi))} du}{\sqrt{2 \int_p^1 \int_v^1 \frac{v(1-u)}{f_X(q_X(v))f_X(q_X(u))} dudv} \sqrt{1 + \frac{\gamma}{f_X^2(\nu)}}}. \quad (83)$$

Note, that we can replace  $\widetilde{\text{ES}}_n(p)$  by its sum-approximation  $\widehat{\text{ES}}_{n,k}(p)$  for fixed  $k$ . Then, clearly, the above results in Corollaries 25 and 26 hold with the integrals replaced by the corresponding sums. In such a case, we can also consider the limit when both,  $n, k$ , tend to infinity, giving us as result of this double-limit the asymptotics as stated in the Corollaries. For the expectile, with the definition of the estimator  $e_n(\alpha)$  as sample quantile at level  $\kappa^{-1}(\alpha)$ , we can use the results for the dependence in the case of the VaR and the measure of dispersion estimators (only that we consider the level  $\kappa^{-1}(\alpha)$ ).

We use these results to display the correlation between the estimators of the three risk measures considered and the measure of dispersion estimators  $\hat{\sigma}_n^2, \hat{\theta}_n, \hat{\xi}_n$ . We provide explicit expressions for the Gaussian and Student distribution in Example 27.



### Example 27

Table 4: Asymptotic correlations between the three risk measure estimators and the three measure of dispersion estimator in the case of a Gaussian distribution or Student distribution respectively

Correlation	Gaussian Distribution	Student( $\nu$ ) distribution
...with $\hat{\sigma}_n^2$		
$\lim_{n \rightarrow \infty} \text{Cor}(\text{VaR}_n(p), \hat{\sigma}_n^2)$	$\frac{\phi(\Phi^{-1}(p))\Phi^{-1}(p)}{\sqrt{2p(1-p)}} \quad (84)$	$\frac{f_{\tilde{Y}}(q_{\tilde{Y}}(p))q_{\tilde{Y}}(p)\left(1 + \frac{q_{\tilde{Y}}^2(p)}{\nu}\right)}{\sqrt{\frac{\nu-1}{\nu-4}}2p(1-p)} \quad (85)$
$\lim_{n,k \rightarrow \infty} \text{Cor}(\widehat{\text{ES}}_{n,k}(p), \hat{\sigma}_n^2)$	$\frac{\int_p^1 \Phi^{-1}(u)du}{2\sqrt{\int_p^1 \int_v^1 \frac{v(1-u)}{\phi(\Phi^{-1}(u))\phi(\Phi^{-1}(v))} dudv}} \quad (86)$	$\frac{\int_p^1 q_{\tilde{Y}}(u)\left(1 + \frac{q_{\tilde{Y}}^2(u)}{\nu}\right)du}{2\sqrt{\int_p^1 \int_v^1 \frac{v(1-u)}{f_{\tilde{Y}}(q_{\tilde{Y}}(v))f_{\tilde{Y}}(q_{\tilde{Y}}(u))} dudv}} \quad (87)$
$\lim_{n \rightarrow \infty} \text{Cor}(e_n(p), \hat{\sigma}_n^2)$	$\frac{\phi(\Phi^{-1}(\kappa^{-1}(p)))\Phi^{-1}(\kappa^{-1}(p))}{\sqrt{2\kappa^{-1}(p)(1-\kappa^{-1}(p))}} \quad (88)$	$\frac{f_{\tilde{Y}}(q_{\tilde{Y}}(\kappa^{-1}(p)))q_{\tilde{Y}}(\kappa^{-1}(p))\left(1 + \frac{q_{\tilde{Y}}^2(\kappa^{-1}(p))}{\nu}\right)}{\sqrt{\frac{\nu-1}{\nu-4}}2\kappa^{-1}(p)(1-\kappa^{-1}(p))} \quad (89)$
...with $\hat{\theta}_n$ :		
$\lim_{n \rightarrow \infty} \text{Cor}(\text{VaR}_n(p), \hat{\theta}_n)$	$\frac{\phi(\Phi^{-1}(p)) - (1-p)\sqrt{2/\pi}}{\sqrt{p(1-p)}\sqrt{1-2/\pi}} \quad (90)$	$\frac{\frac{\sqrt{\nu(\nu-2)}}{\nu-1}f_{\tilde{Y}}(q_{\tilde{Y}}(p))\left(1 + \frac{q_{\tilde{Y}}^2(p)}{\nu}\right) - (1-p)\sqrt{\frac{\nu-2}{\pi}}\frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)}}{\sqrt{p(1-p)}\sqrt{1 - \frac{\nu-2}{\pi}\frac{\Gamma^2(\frac{\nu-1}{2})}{\Gamma^2(\nu/2)}}} \quad (91)$
$\lim_{n,k \rightarrow \infty} \text{Cor}(\widehat{\text{ES}}_{n,k}(p), \hat{\theta}_n)$	$\frac{1-p - \int_p^1 \frac{1-u}{\phi(\Phi^{-1}(u))\sqrt{2/\pi}} du}{2\sqrt{\left(\frac{1}{2} - \frac{1}{\pi}\right) \int_p^1 \int_v^1 \frac{v(1-u)}{\phi(\Phi^{-1}(v))\phi(\Phi^{-1}(u))} dudv}} \quad (92)$	$\frac{\int_p^1 \sqrt{\nu-2} \left( \frac{\sqrt{\nu}}{\nu-1} \left( 1 + \frac{q_{\tilde{Y}}^2(u)}{\nu} \right) - \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})} \frac{(1-u)}{\sqrt{\pi} f_{\tilde{Y}}(q_{\tilde{Y}}(u))} \right)}{\sqrt{2 \int_p^1 \int_v^1 \frac{v(1-u)}{f_{\tilde{Y}}(q_{\tilde{Y}}(v))f_{\tilde{Y}}(q_{\tilde{Y}}(u))} dudv} \sqrt{1 - \frac{\nu-2}{\pi} \frac{\Gamma((\nu-1)/2)^2}{\Gamma(\nu/2)^2}} \quad (93)$
$\lim_{n \rightarrow \infty} \text{Cor}(e_n(p), \hat{\theta}_n)$	$\frac{\phi(\Phi^{-1}(\kappa^{-1}(p))) - (1-\kappa^{-1}(p))\sqrt{2/\pi}}{\sqrt{\kappa^{-1}(p)(1-\kappa^{-1}(p))}\sqrt{1-2/\pi}} \quad (94)$	$\frac{\frac{\sqrt{\nu(\nu-2)}}{\nu-1}f_{\tilde{Y}}(q_{\tilde{Y}}(\kappa^{-1}(p)))\left(1 + \frac{q_{\tilde{Y}}^2(\kappa^{-1}(p))}{\nu}\right) - (1-\kappa^{-1}(p))\sqrt{\frac{\nu-2}{\pi}}\frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)}}{\sqrt{\kappa^{-1}(p)(1-\kappa^{-1}(p))}\sqrt{1 - \frac{\nu-2}{\pi}\frac{\Gamma^2(\frac{\nu-1}{2})}{\Gamma^2(\nu/2)}}} \quad (95)$
...with $\hat{\xi}_n$ :		
$\lim_{n \rightarrow \infty} \text{Cor}(\text{VaR}_n(p), \hat{\xi}_n)$	$\frac{1-p - 2 \max(3/4 - \max(1/4, p), 0)}{\sqrt{p(1-p)}} \quad (96)$	$\frac{1-p - 2 \max(3/4 - \max(1/4, p), 0)}{\sqrt{p(1-p)}} \quad (97)$
$\lim_{n,k \rightarrow \infty} \text{Cor}(\widehat{\text{ES}}_{n,k}(p), \hat{\xi}_n)$	$\frac{\int_p^1 \frac{1-u-2 \max(3/4 - \max(1/4, u), 0)}{\phi(\Phi^{-1}(u))} du}{\sqrt{2 \int_p^1 \int_v^1 \frac{v(1-u)}{\phi(\Phi^{-1}(v))\phi(\Phi^{-1}(u))} dudv}} \quad (98)$	$\frac{\int_p^1 \frac{1-u-2 \max(3/4 - \max(1/4, u), 0)}{f_{\tilde{Y}}(q_{\tilde{Y}}(u))} du}{\sqrt{2 \int_p^1 \int_v^1 \frac{v(1-u)}{f_{\tilde{Y}}(q_{\tilde{Y}}(v))f_{\tilde{Y}}(q_{\tilde{Y}}(u))} dudv}} \quad (99)$
$\lim_{n \rightarrow \infty} \text{Cor}(e_n(p), \hat{\xi}_n)$	$\frac{1-\kappa^{-1}(p) - 2 \max(3/4 - \max(1/4, \kappa^{-1}(p)), 0)}{\sqrt{\kappa^{-1}(p)(1-\kappa^{-1}(p))}} \quad (100)$	$\frac{1-\kappa^{-1}(p) - 2 \max(3/4 - \max(1/4, \kappa^{-1}(p)), 0)}{\sqrt{\kappa^{-1}(p)(1-\kappa^{-1}(p))}} \quad (101)$

where the function  $\kappa(p)$  differs for the Gaussian and Student distribution:

$$\kappa(p) = \begin{cases} \frac{\phi(\Phi^{-1}(p)) + p\Phi^{-1}(p)}{2\phi(\Phi^{-1}(p)) - (1-2p)\Phi^{-1}(p)} & \text{if } X \sim \mathcal{N}(\mu, \sigma^2), \\ \frac{\frac{\nu}{\nu-1}f_{\bar{Y}}(q_{\bar{Y}}(p))(1+q_{\bar{Y}}^2(p)/\nu) + pq_{\bar{Y}}(p)}{2\frac{\nu}{\nu-1}f_{\bar{Y}}(q_{\bar{Y}}(p))(1+q_{\bar{Y}}^2(p)/\nu) - (1-2p)q_{\bar{Y}}(p)} & \text{if } X \sim t(\mu, (\nu-2)\sigma^2/\nu, \nu). \end{cases}$$

After having presenting these results, note the following four remarks with respect to Example 27. First, the expressions of the correlation with the ES involve integrals which can be solved analytically. As this makes the formulae very lengthy, we defer these solved integral expressions to the Appendix C. Second, an alternative way of presenting the asymptotic correlations in Equations (84)-(101) would be in terms of the risk measures itself instead of the quantiles and densities. For this, simply recall that e.g. in the Gaussian case  $\text{VaR}_\alpha(Y) = \Phi^{-1}(p)$  and  $\text{ES}_\alpha(Y) = \frac{\phi(\Phi^{-1}(p))}{1-\alpha}$ . Third, the asymptotic correlations in Equations (84)-(101) do not depend on the mean and variance of the underlying distribution (we have seen this being valid for all the asymptotic correlations in the case of location-scale distributions). Fourth, we could use the parametric location-scale quantile as the VaR estimator (for location-scale distributions), applying the results of Section 3, and extend the results to the ES and expectile too.

To conclude this part on the correlation between risk measures and measure of dispersion estimators, we show the asymptotic correlations in the case of the Gaussian and Student(5) distributions in Figure 4. We have separate plots for the asymptotic correlation with each measure of dispersion. In each plot we show the asymptotic correlation with the sample VaR, the sample ES (evaluated in three possible ways: the sum approximation using 4 or 50 summands and the explicit solution of the integral - corresponding to infinite summands) as well as with the expectile. In the first row we show the case of the Gaussian distribution, then, in the second row, we zoom into the tail. The third and fourth row are accordingly for the Student(5) distribution.

Looking at the plots in the first row, we see that, for the Gaussian distribution, we have the same tendencies of the asymptotic correlation (for VaR, ES and expectile respectively) irrespective of the dispersion measure. The behaviour is more similar with the sample variance and sample MAD and a bit different for the sample MedianAD: For  $p \geq 0.5$ , the ES has clearly higher asymptotic correlation than the VaR (except in the tail where they are similar for sample variance and sample MAD, but lower for the ES in the case of the MedianAD). The asymptotic correlation of the expectile is lower than with VaR and ES, except in the tail where it is highest (the point where this behaviour changes is furthest in the tail with the sample variance, and least with the sample MedianAD).

In the case of the Student(5) distribution, we see the same trends as in the Gaussian case with one exception: For  $p \geq 0.5$ , the asymptotic correlation of the sample variance with the ES is always higher than with VaR and with VaR always higher than with the expectile, see the first plot in the third and fourth row (in contrast to the first plot in the first and second row).

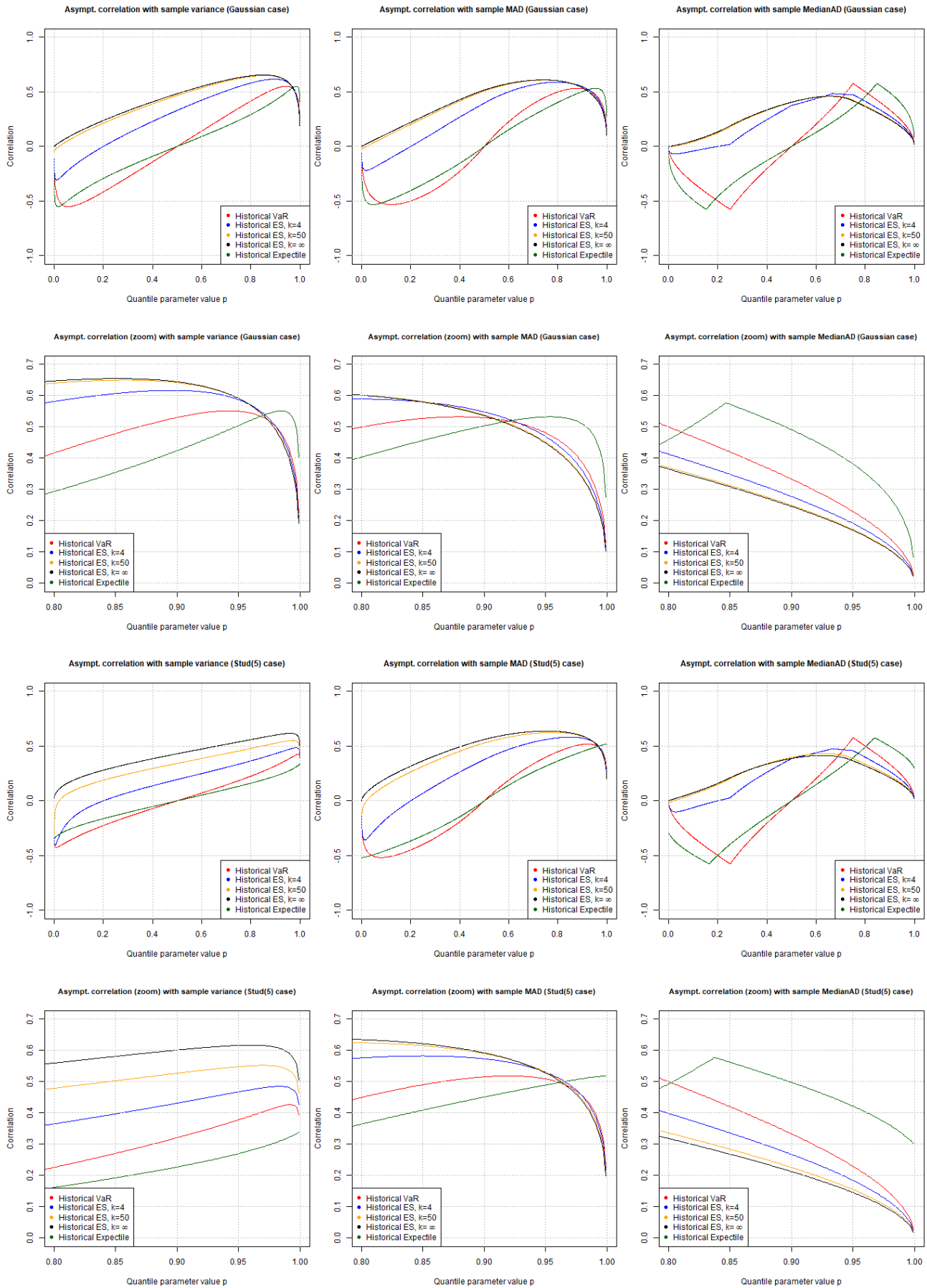


Figure 4: Asymptotic correlations between a risk<sup>41</sup> measure estimator with the measure of dispersion estimator, considering on each plot three different risk measures (VaR, ES, evaluated in 3 possible ways, and expectile). On each row in each plot a different measure of dispersion is considered (from left to right: sample variance, sample MAD, sample MedianAD). First two rows: The case of an underlying Gaussian distribution, third and fourth row: A Student distribution with 5 degrees of freedom.

## 4.2 Example in Finance: Explaining Procyclicality in Risk Measurements

Having made clear the connection of the theoretical results and risk measures in the previous subsection, we can now look at an example in finance and draw a statistical framework in this case. We proceed in different steps. As in the general case for risk and dispersion measures, we relate the theoretical work to empirical findings in the examples we want to consider ([9],[43], Section 3.6 in [44]), and present the theoretical asymptotics - assuming an underlying iid distribution and focussing, as before, on the Gaussian and Student distributions as main examples. Subsequently, we discuss theoretically which changes can be observed in the specific covariances and correlations of risk and dispersion measures, when considering longer sample sizes (in an asymptotic sense we will precise). Finally, we discuss the usefulness of these theoretical results in view of those finite sample applications. We compare, for the different quantile and dispersion estimators considered, the values provided through the analytical formulae with those obtained via simulated iid samples.

We use the time-series notation already introduced in Subsection 2.4.

**Setup -** Considering the example of [9], a measure of interest is the linear correlation of the logarithm of a ratio of sample quantiles (named ‘look-forward ratio’, which measures the predictive quality of the VaR) with the sample standard deviation (used there as a marker of the market state), namely

$$\text{Cor} \left( \log \left| \frac{\hat{q}_{n,t+1y}(p)}{\hat{q}_{n,t}(p)} \right|, \hat{\sigma}_{n,t} \right). \quad (102)$$

This quantity is introduced in [9] to measure the pro-cyclicality of the VaR. Similarly in construction but in another context, Zumbach, in [43]/[44], is interested in the correlation of ‘the realized volatilities with the centred volatility increment’, which in our notation can be translated as

$$\text{Cor} \left( \log \left| \frac{\hat{q}_{a \cdot n, \hat{\sigma}, t+a \cdot 1y}(p)}{\hat{q}_{a \cdot n, \hat{\sigma}, t}(p)} \right|, \hat{\sigma}_{a \cdot n, t+a \cdot 1y} \right), \quad (103)$$

for  $a$  an integer multiple of  $1/252$ . Note that in (102), (103) the quantile estimators in the logarithm are computed on subsequent samples which are disjoint by construction: Choosing the points in time  $t$  and  $t$  plus one year,  $t + 1y$ , in (102) and a sample size of  $n = 252$  (which corresponds to one year of data), the samples do not overlap (and accordingly in (103)).

Note that, in [9], the authors look at this empirical measure of procyclicality/dependence (102) when using data from stock indices, realizations from a GARCH(1,1) model or simulated iid rv’s, while [43] considers currency exchange rates as well as simulated realizations from different stochastic models. Here our goal is to analyse it theoretically when assuming an iid model.

As throughout this paper, in more generality than (102) and (103), we want to consider any quantile estimator  $\hat{q}_{n,t}$  and any measure of dispersion estimator  $\hat{D}_{i,n,t}$ . By abuse of notation we refer in this section to  $\hat{D}_{1,n,t}$  as the sample standard deviation  $\hat{\sigma}_{n,t}$  and not the sample variance (as we only use the former as a marker of the market state).

As mentioned, the sample used for the quantile estimator at time  $t + 1y$ ,  $\hat{q}_{n,t+1y}(p)$ , is disjoint from the sample used at time  $t$  and will thus be uncorrelated with the sample standard deviation  $\hat{\sigma}_{n,t}$ , or more generally the measure of dispersion estimator  $\hat{D}_{i,n,t}$ , at time  $t$ .

Hence, we can summarise our equations of interest as follows (for  $i \in \{1, 2, 3\}$ ):

$$\text{Cov} \left( \log \left| \frac{\hat{q}_{n,t+1y}(p)}{\hat{q}_{n,t}(p)} \right|, \hat{D}_{i,n,t} \right) = - \text{Cov} \left( \log |\hat{q}_{n,t}(p)|, \hat{D}_{i,n,t} \right); \quad \text{Cor} \left( \log \left| \frac{\hat{q}_{n,t+1y}(p)}{\hat{q}_{n,t}(p)} \right|, \hat{D}_{i,n,t} \right) = - \frac{1}{\sqrt{2}} \text{Cor} \left( \log |\hat{q}_{n,t}(p)|, \hat{D}_{i,n,t} \right). \quad (104)$$

Very importantly, note that the quantities computed in [9], [43]/[44] and more generally in any application are for finite sample size (e.g.  $n = 252$  for [9]) while our theoretical results are the asymptotic expressions for equations (104).

We want to compute all asymptotic cases of equations (104) when considering the two standard elliptical cases, namely the Gaussian  $\mathcal{N}(0, \sigma^2)$  and Student  $t(0, \sigma^2(\nu - 2)/\nu, \nu)$  distributions. As our application concerns the modelling of financial log-returns, we can assume  $\mu = 0$  to be known (otherwise one could exploit this for arbitrage). Thus, our location-scale estimator simplifies to  $q_{n,\hat{\sigma},t}(p) = \hat{\sigma}_{n,t} q_Y(p)$ . Hence, the computation is a straightforward application of the results obtained in the previous sections (noting that additionally we have  $\mu = 0$ , such that we can simplify some results further). To enhance comparability, we present the results in a table. Recall, that the expressions involving the sample MAD are always only valid for  $p \geq 0.5$ ; the case for  $p < 0.5$  can be recovered by the corresponding symmetry (in the case involving the logarithm of the absolute value, this is the symmetry around the  $p = 0.5$ -axis), e.g.  $\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \log|\hat{q}_{n,t}(p)|, \sqrt{n}\hat{\theta}_{n,t}) = \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \log|\hat{q}_{n,t}(1-p)|, \sqrt{n}\hat{\theta}_{n,t})$ . Besides, we plot the covariances and correlations using the different measures of dispersion as a function of the quantile parameter  $p$ .

#### 4.2.1 Gaussian Distribution

Before presenting and comparing the different covariances and correlations in Table 5, let us refer to the equations needed to derive the expressions.

For the covariances with the sample quantile and the parametric location-scale quantile, we use equations (40) and (59) for the sample standard deviation, (46) and (65) for the sample MAD, and (55) and (73) for the sample MedianAD. When turning to the case of the asymptotic correlations, we use equations (41), (60) in the case of the sample standard deviation, (47), (66) when computing the correlation with the sample MAD, and (56), (74) for the case with the sample MedianAD.

Table 5: All different cases considered for the covariance (first three rows), then the correlation (last three rows), between the log-ratio of quantile estimators and three measure of dispersion estimators  $\hat{D}_{i,n,t}$  ( $i = 1, 2, 3$ ) each, for a Gaussian distribution.

	$\text{Cov} \left( \log \left  \frac{q_{n,t+1y}(p)}{q_{n,t}(p)} \right , \hat{D}_{i,n,t} \right)$	$\text{Cov} \left( \log \left  \frac{q_{n,\hat{\sigma},t+1y}(p)}{q_{n,\hat{\sigma},t}(p)} \right , \hat{D}_{i,n,t} \right)$
$\hat{D}_{1,n,t} = \text{sample s.d. } \hat{\sigma}_{n,t}$	$-\frac{\sigma}{2}$	$-\frac{\sigma}{2}$
$\hat{D}_{2,n,t} = \text{sample MAD } \hat{\theta}_{n,t}$	$-\frac{\sigma \left( \phi(\Phi^{-1}(p)) - (1-p)\sqrt{\frac{2}{\pi}} \right)}{\phi(\Phi^{-1}(p)) \Phi^{-1}(p)}$	$-\frac{\sigma}{\sqrt{2\pi}}$
$\hat{D}_{3,n,t} = \text{sample MedianAD } \hat{\xi}_{n,t}$	$-\frac{\sigma(1-p-2 \max(0, 3/4 - \max(1/4, p)))}{4\phi(\Phi^{-1}(p)) \phi(\Phi^{-1}(3/4)) \Phi^{-1}(p)}$	$-\frac{\sigma\Phi^{-1}(3/4)}{2}$
	$\text{Cor} \left( \log \left  \frac{q_{n,t+1y}(p)}{q_{n,t}(p)} \right , \hat{D}_{i,n,t} \right)$	$\text{Cor} \left( \log \left  \frac{q_{n,\hat{\sigma},t+1y}(p)}{q_{n,\hat{\sigma},t}(p)} \right , \hat{D}_{i,n,t} \right)$
$\hat{D}_{1,n,t} = \text{sample s.d. } \hat{\sigma}_{n,t}$	$-\frac{1}{\sqrt{2}} \frac{\phi(\Phi^{-1}(p)) \Phi^{-1}(p) }{\sqrt{2p(1-p)}}$	$-\frac{1}{\sqrt{2}} \approx -0.71$
$\hat{D}_{2,n,t} = \text{sample MAD } \hat{\theta}_{n,t}$	$-\frac{1}{\sqrt{2}} \frac{ \phi(\Phi^{-1}(p)) - (1-p)\sqrt{\frac{2}{\pi}} }{\sqrt{p(1-p)}\sqrt{1-\frac{2}{\pi}}}$	$-\frac{1}{\sqrt{2}} \sqrt{\frac{1}{\pi-2}} \approx -0.66$
$\hat{D}_{3,n,t} = \text{sample MedianAD } \hat{\xi}_{n,t}$	$-\frac{1}{\sqrt{2}} \frac{ 1-p-2 \max(0, 3/4 - \max(1/4, p)) }{\sqrt{p(1-p)}}$	$-\frac{1}{\sqrt{2}} \frac{4\Phi^{-1}(3/4)\phi(\Phi^{-1}(3/4))}{\sqrt{2}} \approx -0.43$

In Table 5, we observe that in the case of using the sample standard deviation, the covariance is the

same when using the sample quantile or the location-scale quantile with known mean  $\mu = 0$  (first row). Further, we see that all different covariances with the location-scale quantile are constant (independent of  $p$ ) and thus can be written as a multiple of each other (second column).

In the fourth row of Table 5, we can see that, in contrast to the covariances, the correlation with  $q_{n,\hat{\sigma}}$  is not equal to the one with  $q_n$  in the case of the sample standard deviation: The former is constant, whereas the one with the sample quantile is a function of  $p$ . This remark holds for all dispersion measures, i.e. the correlation with the location-scale quantile are constant (and negative) while it is a function of  $p$  for the sample quantile.

To further understand those quantities, we plot them in Figure 5.

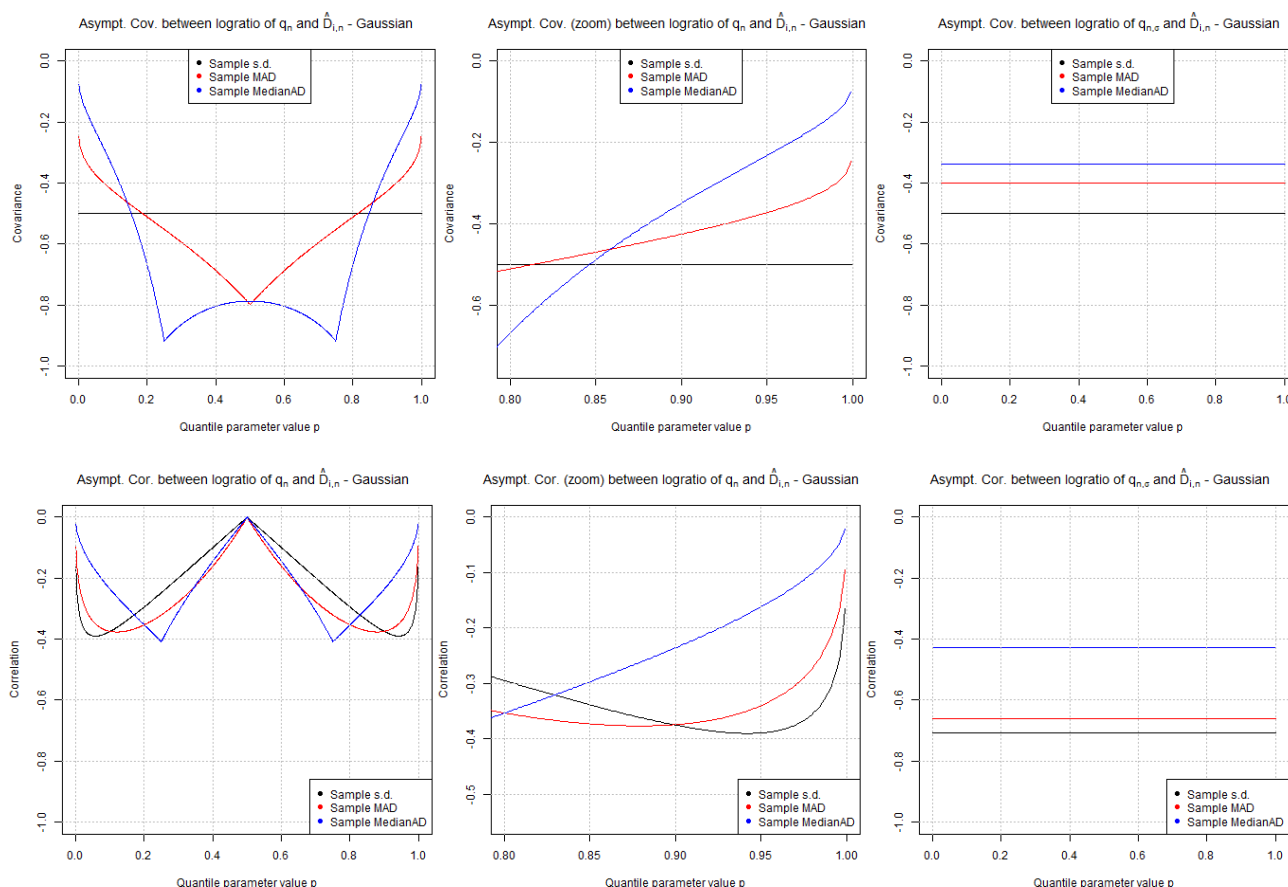


Figure 5: Comparison of the covariances (first row) and correlations (2nd row) between log-ratios of quantile estimators and each of three measure of dispersion estimators in the same plot, in the case of a Gaussian distribution. Left: Using the sample quantile, Middle: Zoomed into the right tail ( $p \geq 0.8$ ) using the sample quantile, Right: Using the location-scale quantile with known mean  $\mu = 0$

In Figure 5, we see in the left column the two plots of covariances and correlations when using the sample quantile and in the right column the two plots, when using the location-scale quantile - in the middle column we zoomed into the tail values ( $p \geq 0.8$ ) in the case with the sample quantile, as those are especially of interest in applications to quantitative risk management. When looking at the covariance with the sample quantile (first row, left plot), we see for all three measures of dispersion the symmetry around the  $p = 0.5$ -axis, and that all values are negative. The covariance is constant (i.e. independent of  $p$ ) for the sample standard deviation with a value of  $-0.5$ , bi-modal for the sample MedianAD (minimum

of  $-0.92$  at  $p = 0.75$ ) and uni-modal for the sample MAD (having its minimum of  $-0.8$  at  $p = 0.5$ ). For intermediate values of  $p$ , say  $0.2 < p < 0.8$ , the covariance is the biggest in absolute terms for the sample MedianAD. In this interval, the difference between MedianAD and MAD/standard deviation is the highest around the minimum of the MedianAD. For all values in this interval, the difference between standard deviation and MedianAD is large. An exception to the MedianAD having the biggest value in absolute terms is for values near  $p = 0.5$  where it is larger for the sample MAD - but in this neighbourhood MAD and MedianAD are very close (and it is the smallest for the sample standard deviation). If we focus on tail-values, i.e. high values of  $p$ , the behaviour changes (by symmetry this applies also for small values of  $p$ ). This can be seen well in the middle plot. Therein the covariance is the highest (in absolute terms) for the sample standard deviation: It is higher than the sample MAD for  $p > 0.81$  and for  $p > 0.85$  in the case of the sample MedianAD. From  $p > 0.86$  on, the covariance is the lowest for the sample MedianAD. For values of  $p$  around  $(0.81, 0.86)$  we can say that the covariance is similar for all three measures of dispersion, while the difference between them increases for increasing  $p > 0.86$ . Looking at the covariance using the location-scale quantile (right plot), we can make the same comments as those of Table 5. The values are constant for all three measures of dispersion and hence independent of  $p$ . The values for the sample standard deviation are the same with both quantile estimators. In absolute terms, the covariance with the sample standard deviation is the highest, and the one with the sample MedianAD, the lowest in the right plot. The values of MAD and MedianAD are more similar than with the standard deviation. Comparing both quantile estimators, we can say for the sample MAD and sample MedianAD that, for  $p > 0.93$  and  $p > 0.90$  respectively, the covariance is lower (in absolute terms) with the sample quantile, while, for smaller values, this behaviour reverses. Also, we see that the covariance values for those two measures of dispersion can be, depending on  $p$ , quite different with respect to the chosen quantile estimator.

When turning to the correlations (2nd row of Figure 5), we see that, in the case of the correlation with the sample quantile (left plot), all three measures of dispersion have the same tendencies. They are symmetric around the axis  $p = 0.5$ , have low values for  $p$  around  $0.5$  and  $p \rightarrow 0$  and  $1$  while being bi-modal and having two minima (one between  $(0, 0.5)$  and one between  $(0.5, 1)$  by symmetry). Further, all values are negative, while for  $p = 0.5$ , the correlation is undefined and set by continuity equal to  $0$ . Also, their range of values is very similar. Thus, the maximum in absolute terms is only a little bit bigger for the sample MedianAD ( $-0.41$  at  $p = 0.75$ ) than for the sample standard deviation ( $-0.39$  at  $p = 0.94$ ) and is the smallest for the sample MAD ( $-0.38$  at  $p = 0.89$ ). Looking at intermediate values of  $p$ , say  $0.2 < p < 0.8$ , we see that the values for all three measures of dispersion are similar around  $p = 0.5$ , while, in absolute terms, from this point to the respective maxima, the correlation increases for all measures of dispersion. The maximum in absolute terms of the sample MedianAD correlation falls in this interval and is achieved for  $p = 0.75$ . Around that point, the difference between the three measures of dispersion is the biggest, the correlation with the sample standard deviation being the smallest. If we focus again on tail values, as in the middle plot, we see how the behaviour changes. Already from  $p = 0.8$ , the MAD is bigger than the MedianAD, and, from  $p > 0.83$  on, the sample standard deviations correlation is bigger than the sample MedianAD. The MAD reaches its maximum value of  $-0.38$  at  $p = 0.88$ . Consequently, from  $p > 0.90$ , the standard deviation is bigger than the sample MAD (note that those thresholds are different than in the case of the covariance). The maximum in absolute terms for the sample standard deviation correlation,  $-0.39$ , is near  $p = 0.94$ . From this point on, it decreases again but still being bigger than the two other measures of dispersion. We can also see that, around this last maximum, the difference between the three measures of dispersion is the highest (pronounced with the MedianAD, rather small between standard deviation and MAD). They are the most similar around their intersection points, i.e. in the range of  $(0.8, 0.83)$  and, additionally, in the case of standard deviation and MAD, also around  $0.9$  where the two latter intersect. Getting back from the tails to the general behaviour and looking now at the correlations with the location-scale quantile (right plot), we see that each

of them is bigger than the ones with the sample quantile. Comparing the three (constant) correlations using the location-scale quantile, we note it is the biggest in absolute terms with the sample standard deviation (-0.71), intermediate for the MAD (-0.66) and the smallest with the sample MedianAD (-0.43). In contrast to the covariance, the correlations for the MAD and standard deviation are now similar, while quite different for the MedianAD.

#### 4.2.2 Student Distribution

Again, we start by presenting and comparing the different covariances and correlations. Before we do so in Table 6, we want to refer to the equations needed for the calculation. For the covariances we use equations (42) and (61) for the sample standard deviation, (49), and (67) for the sample MAD, and (55), and (75) for the sample MedianAD. For the asymptotic correlations correspondingly, equations (43) and (62) in the case of the sample standard deviation, (50) and (68) with the sample MAD, and (56), (77) for the case with the sample MedianAD.

Table 6: All different cases considered for the covariance (three first rows), then the correlation (three last rows), between the log-ratio of quantile estimators and three measure of dispersion estimators  $\hat{D}_{i,n,t}$  ( $i = 1, 2, 3$ ) each, for a Student distribution with  $\nu$  degrees of freedom (restrictions on  $\nu$  depending on the choice of estimator as before)

	$\text{Cov} \left( \log \left  \frac{q_{n,t+1y}(p)}{q_{n,t}(p)} \right , \hat{D}_{i,n,t} \right)$	$\text{Cov} \left( \log \left  \frac{q_{n,\hat{\sigma},t+1y}(p)}{q_{n,\hat{\sigma},t}(p)} \right , \hat{D}_{i,n,t} \right)$
$\hat{D}_{1,n,t} = \text{sample s.d. } \hat{\sigma}_{n,t}$	$-\frac{\sigma}{2} \left( 1 + \frac{q_{\hat{Y}}^2(p)}{\nu} \right)$	$-\frac{\sigma}{2} \frac{\nu-1}{\nu-4}$
$\hat{D}_{2,n,t} = \text{sample MAD } \hat{\theta}_{n,t}$	$-\frac{\sigma \left( \frac{\sqrt{\nu(\nu-2)}}{\nu-1} f_{\hat{Y}}(q_{\hat{Y}}(p)) \left( 1 + \frac{q_{\hat{Y}}^2(p)}{\nu} \right) - (1-p) \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)} \right)}{f_{\hat{Y}}(q_{\hat{Y}}(p)) q_{\hat{Y}}(p)}$	$-\frac{\sigma}{2} \frac{\nu-1}{\nu-3} (\nu-2) \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)} \sqrt{\frac{1}{\nu\pi}}$
$\hat{D}_{3,n,t} = \text{sample MedianAD } \hat{\xi}_{n,t}$	$-\sigma \frac{1-p-2 \max(0, 3/4 - \max(1/4, p))}{4 f_{\hat{Y}}(q_{\hat{Y}}(p)) f_{\hat{Y}}(q_{\hat{Y}}(3/4)) \sqrt{\frac{\nu}{\nu-2}} q_{\hat{Y}}(p)}$	$-\frac{\sigma}{2} \frac{q_{\hat{Y}}(3/4) \left( 1 + \frac{q_{\hat{Y}}^2(3/4)}{\nu} \right)}{\sqrt{\frac{\nu}{\nu-2}}}$
	$\text{Cor} \left( \log \left  \frac{q_{n,t+1y}(p)}{q_{n,t}(p)} \right , \hat{D}_{i,n,t} \right)$	$\text{Cor} \left( \log \left  \frac{q_{n,\hat{\sigma},t+1y}(p)}{q_{n,\hat{\sigma},t}(p)} \right , \hat{D}_{i,n,t} \right)$
$\hat{D}_{1,n,t} = \text{sample s.d. } \hat{\sigma}_{n,t}$	$-\frac{f_{\nu}(q_{\hat{Y}}(p))  q_{\hat{Y}}(p)  \left( 1 + \frac{q_{\hat{Y}}^2(p)}{\nu} \right)}{2 \sqrt{\frac{\nu-1}{\nu-4}} p(1-p)}$	$-\frac{1}{\sqrt{2}} \approx -0.71$
$\hat{D}_{2,n,t} = \text{sample MAD } \hat{\theta}_{n,t}$	$-\frac{\left  \frac{\sqrt{\nu(\nu-2)}}{\nu-1} f_{\hat{Y}}(q_{\hat{Y}}(p)) \left( 1 + \frac{q_{\hat{Y}}^2(p)}{\nu} \right) - (1-p) \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)} \right }{\sqrt{2} \sqrt{p(1-p)} \sqrt{1 - \frac{\nu-2}{\pi} \frac{\Gamma^2(\frac{\nu-1}{2})}{\Gamma^2(\nu/2)}}}$	$-\frac{\frac{\sqrt{\nu-1}}{\nu-3} \sqrt{\frac{\nu-2}{\nu}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)} \sqrt{\frac{\nu}{\pi}}}{2 \sqrt{1 - \frac{\Gamma^2(\frac{\nu-1}{2})}{\Gamma^2(\nu/2)} \frac{\nu-2}{\pi} \sqrt{\frac{1}{\nu-4}}}} \approx -0.54$
$\hat{D}_{3,n,t} = \text{sample MedianAD } \hat{\xi}_{n,t}$	$-\frac{ 1-p-2 \max(0, 3/4 - \max(1/4, p)) }{\sqrt{2} \sqrt{p(1-p)}}$	$-\frac{4 f_{\hat{Y}}(q_{\hat{Y}}(3/4)) \left( 1 + \frac{q_{\hat{Y}}^2(3/4)}{\nu} \right) q_{\hat{Y}}(3/4)}{2 \sqrt{\frac{\nu-1}{\nu-4}}} \approx -0.23$

Again, the values are summarised in Table 6. Apart from the correlation of the sample standard deviation in the case of the location-scale quantile, which equals  $-\frac{1}{\sqrt{2}} \approx -0.71$ , the expressions look more complex than in the case with the Gaussian distribution. In contrast to the Gaussian case, we can observe in Table 6 that, when using the sample standard deviation, the covariance is not the same when using the sample quantile or the location-scale quantile with known mean  $\mu = 0$  (first column). But again, we see that all different covariances with the location-scale quantile are constant (i.e. independent of  $p$ ), hence also multiple of each other (last column), although this multiple depends on the degrees of freedom  $\nu$ . To further understand the quantities, we plot them. We start by considering the case  $\nu = 5$  in Figure 6 since we need  $\nu \geq 4$  for  $(M_2)$  to hold (as we also consider the location-scale quantile  $q_{n,\hat{\sigma}}$ ). As the



behaviour changes with  $\nu$ , in a second step, we look in Figure 7 at the covariances and correlations as a function of  $\nu$  by comparing the cases  $\nu = 3, 4, 5, 10, 40$  with the Gaussian limiting case.

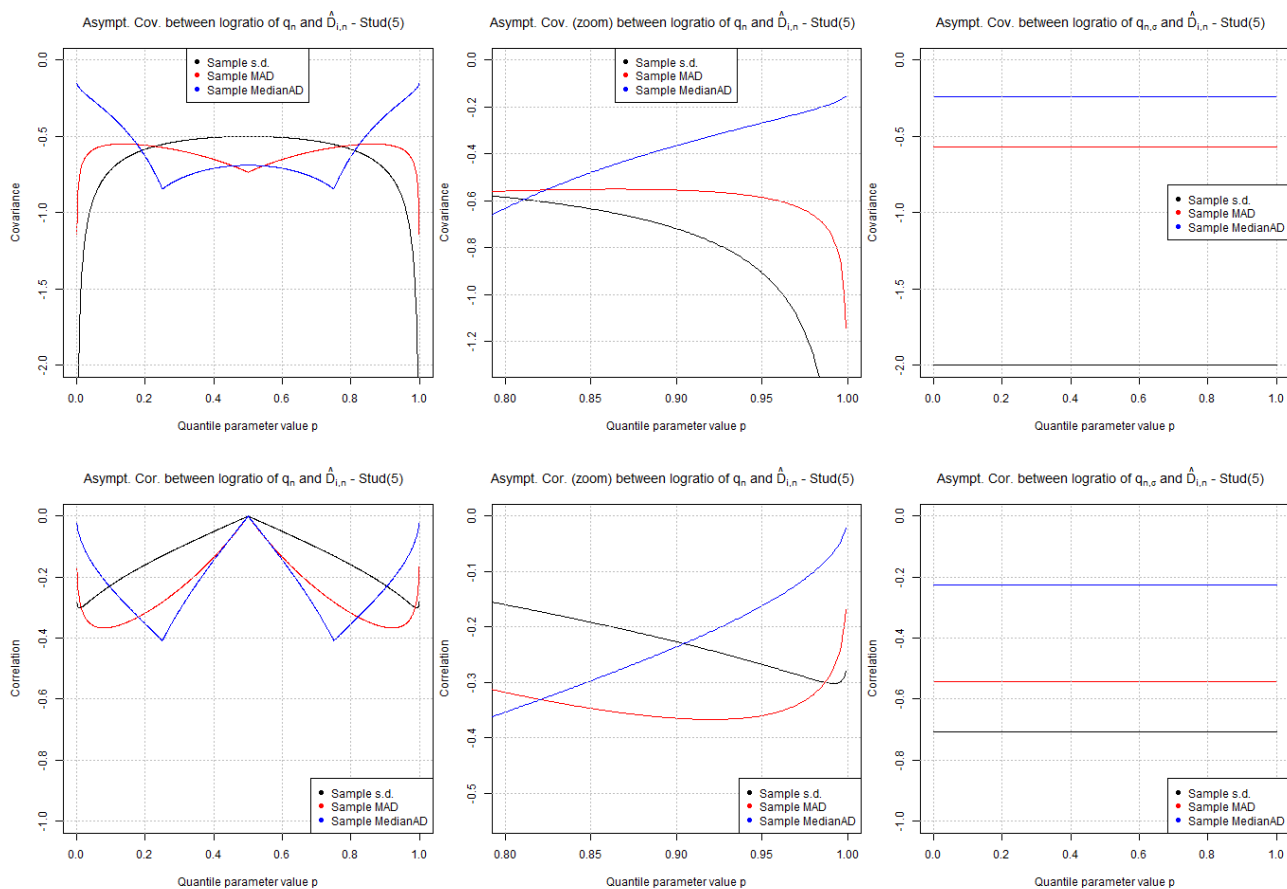


Figure 6: Comparison of the covariances (first row) and correlations (second row) between log-ratios of quantile estimators and each of three measure of dispersion estimators in the same plot, in the case of a Student distribution with  $\nu = 5$  degrees of freedom. Left: Using the sample quantile, Middle: Zoomed into the right tail ( $p \geq 0.8$ ) using the sample quantile, Right: Using the location-scale quantile with known mean  $\mu = 0$

Let us start with the first row of Figure 6. Therein we see in the left two plots the covariances when using the sample quantile and in the right plot when using the location-scale quantile. When looking at the covariance with the sample quantile (left plot), we see that the behaviour is as with the Gaussian distribution: In all cases, the value is negative and symmetric around the  $p = 0.5$ -axis. The sample MedianAD is still bi-modal and for intermediate values of  $p$ , i.e.  $0.2 < p < 0.8$ , the sample MAD has its maximum, in absolute terms, of 0.73 at  $p = 0.5$ . Apart from this maximum, the sample MedianAD has the highest (up to -0.84) and the sample standard deviation the lowest (minimum of -0.5) covariance values in this interval in absolute terms. The three measures of dispersion are similar in values in a very small neighbourhood around their intersection points (between 0.77 and 0.81) and the MedianAD with the MAD also around  $p = 0.5$ . Around its two maxima and especially in the limit  $p \rightarrow 0$  and 1 the MedianAD is very different from the standard deviation and the MAD, whereas the latter two are rather similar (with their biggest difference being around  $p = 0.5$ ). If we turn to the middle plot, we can observe the behaviour for tail values of  $p$ . We see that the MedianAD and the standard deviation intersect at  $p = 0.81$  with a value of -0.59 and the MedianAD with MAD at 0.82 (value of -0.55). As noticed before, this is the region the three measures of dispersion are the most similar. While we see that for  $p < 0.81$ , the MedianAD is the biggest in absolute terms, from  $p > 0.81$  on, the standard deviation

is bigger than the MedianAD (for  $p > 0.8$ , the standard deviation is all the time bigger than the sample MAD anyway). Further, from  $p > 0.82$  on, the covariance with the sample MAD is larger in absolute terms than the sample MedianAD.

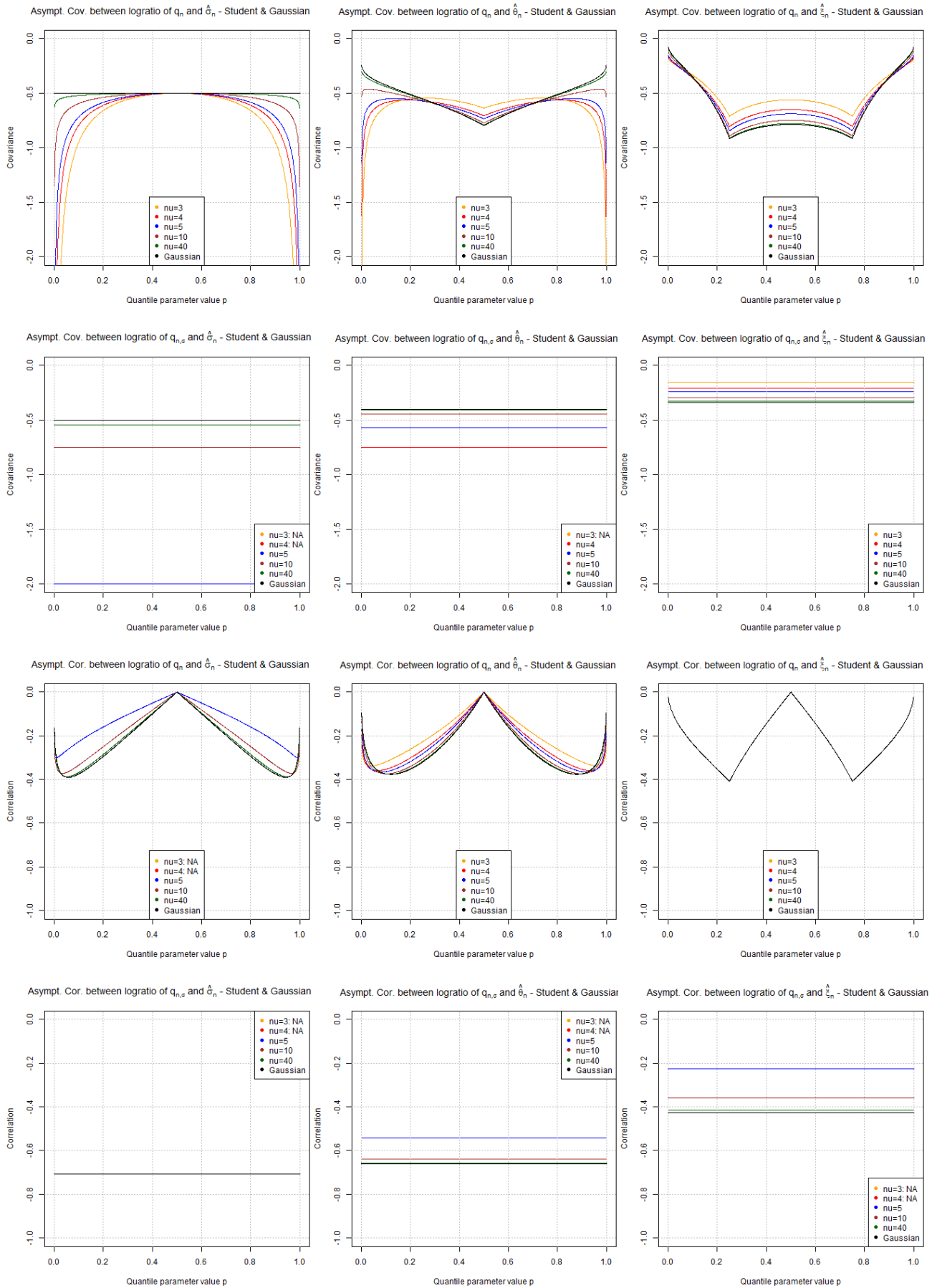
A difference to the Gaussian distribution is, as already pointed out, that the covariance is not constant with the sample standard deviation. Also, the range of values for the sample MAD, and especially for the sample standard deviation, is bigger than in the Gaussian case (the covariances seem to explode near the boundary). We also see now in the limit  $p \rightarrow 0$  and 1 the MedianAD increases, whereas MAD and standard deviation decrease, i.e. they have opposite trends.

We consider now the covariance when using the location-scale quantile (right plot). The same comments as those already made for Table 6 hold. The values are constant, i.e. independent of  $p$ . In absolute terms, the covariance with the sample standard deviation is the highest (-2), and the one with the sample MedianAD (-0.24) the lowest, while -0.57 for the MAD, i.e. same ordering as in the Gaussian case, but here the difference between the MAD and the standard deviation is much more pronounced.

When comparing the covariances with respect to the quantile estimators, we can say that in the case of the sample MedianAD for  $p$  near the boundary of  $(0, 1)$ ,  $p > 0.97$ , the covariance is lower (in absolute terms) with the sample quantile, while this behaviour reverses for all the other values of  $p$  more in the interior of  $(0, 1)$ . For the sample standard deviation, the covariance is always higher in absolute terms for the location-scale quantile (except for  $p$  limiting to the boundary, i.e.  $p > 0.99$ ). For the sample MAD, it is higher for the sample quantile, except between  $0.76 < p < 0.94$ .

Let us turn to the correlations in the second row of Figure 6. We see that, in the case of the correlation with the sample quantile (left plot), all three measures of dispersion have the same tendencies, they are negative and symmetric around the axis  $p = 0.5$ , with their smallest values in absolute terms being around  $p = 0.5$ . They also all have two minima and comparable range of values. In short, the behaviour is very similar to the Gaussian case (recall Figure 5). The minimum is a little bit bigger for the sample MedianAD (-0.41 at  $p = 0.75$ ) than for the sample MAD (-0.37 at 0.92) and is the smallest for the sample standard deviation (-0.30 at 0.99). Then, we focus again on tail values, as in the middle plot. We see that, for  $p < 0.82$ , the MedianAD is the largest, from  $p = 0.82$  on, the MAD correlation is then the largest and, only from  $p > 0.99$ , the standard deviation (note that those thresholds are different than in the case of the covariance and also a bit different to the Gaussian case). The maximum in absolute terms for the sample standard deviation correlation is near  $p = 0.99$  with a value of  $-0.30$ , for the MAD it is around  $p = 0.92$  with a value of  $-0.37$ . All three measures of dispersion do not have very similar values in the tail. Looking at the right plot we can say that, measured in absolute terms, the correlations with the location-scale quantile are bigger than the ones with the sample quantile (for all  $p$ ) for the sample standard deviation (-0.71) and the sample MAD (-0.54), but not for the sample MedianAD where the covariance with  $q_{n,\hat{\sigma}}$  is very low (-0.23) and the correlation is bigger in absolute terms except in the intervals  $(0.10, 0.35)$  and  $(0.65, 0.90)$ . Comparing the three (constant) correlations using the location-scale quantile, it is the biggest with the sample standard deviation and the smallest with the sample MedianAD. Again, as in the Gaussian case, the difference between MAD and standard deviation is smaller, whereas it is big to the MedianAD (differently than for the covariance).

As mentioned, we also want to study the convergence of the Student covariance and correlation to the Gaussian case with respect to the degrees of freedom  $\nu$ . Thus, we look in Figure 7 at the covariance and correlation for each measure separately but showing the cases  $\nu = 3, 4, 5, 10, 40$  and  $\infty$  (Gaussian case) in the same plot. As before, we first look at the covariance. In the first row we show the covariances when using the sample quantile. There are three plots, one for each measure of dispersion. In the case with the sample standard deviation (left plot), we see that the convergence for intermediate values of  $p$  is quick and slows down significantly the further the values are near the boundary  $p \rightarrow 0$  and 1.



49  
 Figure 7: Asymptotic covariances (first two rows) and correlations (last two rows) between the sample quantile or the location-scale quantile respectively and three different measure of dispersion estimators each (from left to right: sample standard deviation, sample MAD, sample MedianAD) in the case of a Student distribution -with  $\nu = 3, 4, 5, 10$  and  $40$  degrees of freedom as well as the Gaussian distribution as comparison

For the MAD (middle plot), the convergence slows down in direction of the boundaries, but interestingly it is quicker near the maxima of the covariance in the case  $\nu = 3$  (i.e.  $p=0.26$ ,  $p=0.74$ ), while slower in between. Further, we see a change in the shape of the covariance: For low degrees of freedom ( $\nu = 3, 4, 5$ ), the maximum value in absolute terms is at the boundaries, whereas, for  $\nu = 10, 40$  and the Gaussian case, this maximum is near  $p = 0.5$ . Also, for  $\nu = 3, 4, 5, 10$ , the covariance has inflection points near the boundary; this is not the case for  $\nu = 40$  and the Gaussian distribution. Contrary to the MAD, the convergence for the MedianAD is quick on the boundary but slower in the interval  $(0.25, 0.75)$ . Comparing the three plots, the convergence is slower at the boundaries for the sample standard deviation and MAD, whereas quicker for the sample MedianAD. For interior values, say  $0.2 < p < 0.8$ , we see the opposite behaviour, the convergence is quick for standard deviation and MAD, but slow for the MedianAD.

In the second row of Figure 7, we see the convergence for the case of using the location-scale quantile. Note that for some covariances, we cannot compute the quantities for  $\nu = 3, 4$  (because of the necessary moment conditions). While the values are always constant for all  $p$ , one can point out that the convergence with the sample standard deviation is the slowest and with the sample MedianAD the fastest. Also, for the sample standard deviation and sample MAD the values for the Student distribution are more negative than for the Gaussian distribution. This is the contrary for the sample MedianAD. Thus, one can say that the convergence to the Gaussian case is ‘from the left’ (coming from more negative values), while it is the reverse for the sample MedianAD.

Finally, as for the covariance, we want to look at the convergence behaviour of the correlation with an underlying Student distribution to the Gaussian case. This is done as a function of the degrees of freedom  $\nu$  in the last two rows of Figure 7 for each of the three different measures of dispersion separately. First, we look at the case with the sample quantile. For the sample standard deviation (left plot), we see that the convergence, for values near the boundaries and  $p = 0.5$ , is quicker as for the other intermediate values. This is something we observe also with the sample MAD (middle plot). Also, the behaviour of the convergence with  $p$  is similar in both cases. These two findings are contrary to what we observed for the standard deviation and MAD for the covariance. Still, comparing it to the case with the sample MAD, the convergence of the correlation for the standard deviation is slower - we observed this ordering in the speed of convergence already for the covariances. Further, the convergence with the sample MAD is smoother than with the sample standard deviation. E.g. the shape and values from  $\nu = 5$  to  $\nu = 10$  change more than with the sample MAD. Comparing the three plots, we see that the convergence is the quickest for the sample MedianAD (right plot). The values for the Student and Gaussian cases are equal (thus one sees only one function in the plot). This was due to the fact that the correlation does not depend on the underlying distribution for symmetric location-scale distributions!

In the last row of Figure 7, we see the convergence for the case of using the location-scale quantile. Again, this is in all three cases only well-defined for  $\nu \geq 5$ . We see that the correlation with the sample standard deviation (left plot) is constant and equal to the Gaussian case for all the well-defined Student cases. For the sample MAD (middle plot), we see that the convergence is quicker than for the sample MedianAD (right plot). Also, the convergence to the Gaussian case is, for both, sample MAD and sample MedianAD ‘from the right’ (i.e. coming from less negative values) - for the covariance we had observed different behaviours for MAD and MedianAD respectively.

### 4.2.3 Impact of Using Longer Samples

We want to understand how this effect of negative correlation for the log-ratio of sample quantiles with the measures of dispersion (sample MAD, sample standard deviation, sample MedianAD) depends on the chosen sample size. To do so, we analyse what happens to the correlation (and also covariance) when

using longer sample sizes than one year of observations (*i.e.*  $n = 252$ ). [43] looked at this empirically for some cases, and in Proposition 10 we already looked at the scaling behaviour in general, the difference here being that we analyse a specific setting. For simplicity, we restrict ourselves to multiples of the base sample size  $n = 252$  and denote them by  $un$  for an integer-valued  $u$ .

Again, instead of looking separately at all the different cases when either using historical estimation of the quantile or the location scale model, and one of the three measures of dispersion, we consider a unified approach.

The only condition we require for this analysis with longer sample sizes, is that neither the measure of dispersion estimator  $\hat{D}_{i,n,t}$  at time  $t$ , nor the quantile prediction  $\hat{q}_{n,t}(p)$  computed at time  $t$ , overlaps (with respect to the sample they are computed on) with the other quantile estimator  $\hat{q}_{n,t+wy}$  (usually computed at time  $t + 1y$ , here in more generality at time  $t$  plus  $w$  years), hence we choose  $\hat{q}_{un,t+wy}$  in a way such that this is fulfilled by construction. It means that, in most generality, we consider for positive integers  $w, u, v$  the following covariance and correlation - in analogy to equation (104):

$$\text{Cov} \left( \log \left| \frac{\hat{q}_{un,t+wy}(p)}{\hat{q}_{un,t}(p)} \right|, \hat{D}_{i,vn,t} \right) = - \text{Cov} \left( \log |\hat{q}_{un,t}(p)|, \hat{D}_{i,vn,t} \right) \quad (105)$$

$$\text{Cor} \left( \log \left| \frac{\hat{q}_{un,t+wy}(p)}{\hat{q}_{un,t}(p)} \right|, \hat{D}_{i,vn,t} \right) = \frac{- \text{Cov} \left( \log |\hat{q}_{un,t}(p)|, \hat{D}_{i,vn,t} \right)}{\sqrt{\text{Var}(\log |\hat{q}_{un,t+wy}(p)|) + \text{Var}(\log |\hat{q}_{un,t}(p)|)} \sqrt{\text{Var}(\hat{D}_{i,vn,t})} \quad (106)$$

As in Subsection 4.2, note that equations (105), (106) are for finite sample sizes while the theoretical results hold asymptotically.

From Proposition 10, it follows in this case (given the respective conditions on the underlying distribution):

$$\lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} \log \left| \frac{\hat{q}_{un,t+wy}(p)}{\hat{q}_{un,t}(p)} \right|, \sqrt{n} \hat{D}_{i,vn,t} \right) = \frac{1}{\max(u, v)} \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} \log \left| \frac{\hat{q}_{n,t+1y}(p)}{\hat{q}_{n,t}(p)} \right|, \sqrt{n} \hat{D}_{i,n,t} \right). \quad (107)$$

This means that the choice of  $w$  does not play a role at all (as the quantile estimator at time  $t + wy$  is by construction disjoint from the other quantities) and the exact values of  $u, v$  itself neither, but only their maximum! Accordingly, one can show for the correlation that

$$\lim_{n \rightarrow \infty} \text{Cor} \left( \log \left| \frac{\hat{q}_{un,t+wy}(p)}{\hat{q}_{un,t}(p)} \right|, \hat{D}_{i,vn,t} \right) = \frac{-1}{\sqrt{1 + \frac{u \max(u, v)}{w \min(u, v)}}} \lim_{n \rightarrow \infty} \text{Cor}(\log |\hat{q}_{n,t}(p)|, \hat{D}_{i,n,t}). \quad (108)$$

**Exemplary Discussion on Sample Size -** Let us discuss what we can deduce from the two equations, (107) and (108), focusing on the correlation (108). As we are interested in the change of the value of the correlation for different sample sizes (and not the value itself), for illustration, let us simply consider the case of the location-scale quantile model with known mean  $\mu = 0$  and the sample standard deviation as measure of dispersion, as in this case (108) simplifies to:

$$\lim_{n \rightarrow \infty} \text{Cor} \left( \log \left| \frac{q_{un,\hat{\sigma},t+wy}(p)}{q_{un,\hat{\sigma},t}(p)} \right|, \hat{\sigma}_{vn,t} \right) = -1 / \sqrt{1 + \frac{u}{w} \times \frac{\max(u, v)}{\min(u, v)}}$$

since the correlation between the logarithm of the quantile estimator and the standard deviation is equal to 1, as one can deduce from (60).

Recall that in the case of using equal sample size  $n$ , *i.e.*  $w = u = v = 1$ , we have a correlation value of  $-1/\sqrt{2}$ . Thus we compare this latter value (named base value) to the values obtained with other choices for  $w, u, v$ , for samples of size  $n_w, n_u, n_v$  respectively. We give an overview of the different cases in Table 7.

Table 7: Correlation values (using the location-scale quantile model with known mean  $\mu = 0$ ) for different sample sizes  $n_w, n_u, n_v$  depending on the choice of  $w, u, v$  in comparison to the standard choice of  $w = u = v = 1$ .

	$w = u = v$		$w = 1$		
	$w = u = v = 1$	$w = u = v > 1$	$u = v > 1$	$1 \leq u < v$	$u > v \geq 1$
$\lim_{n \rightarrow \infty} \text{Cor} \left( \log \left  \frac{\hat{q}_{wn, \hat{\sigma}, t+wy}(p)}{\hat{q}_{un, \hat{\sigma}, t}(p)} \right , \hat{\sigma}_{vn, t} \right)$	$-\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{2}}$	$-\frac{1}{\sqrt{1+u}}$	$-\frac{1}{\sqrt{1+v}}$	$-\frac{1}{\sqrt{1+u^2/v}}$

We observe that, when all sample sizes are equal, taking the size larger than 1 (second column) does not have any impact on the value of the correlation, which remains  $-1/\sqrt{2}$ . This is something one can e.g. observe also from the plots in [43]. If we set  $w = 1$  which is usual in regulation, see [11], we consider three different choices. If  $u = v > 1$ , the correlation depends only on the value of  $u$  (third column) and will be smaller (in absolute values) than the base value. A similar behaviour holds if  $1 \leq u < v$ , although not a realistic case in view of applications (as one wants  $\hat{\sigma}_{vn, t}$  to be a marker of  $q_{un, \hat{\sigma}, t}$  - thus they should be computed over the same sample). Moreover the correlation only depends on  $v$  in such a case (fourth column). In the last case, where  $u > v \geq 1$ , the correlation depends on the factor  $u^2/v > k$ , so is the smallest (in absolute values) of all observed cases (but depends on both,  $u$  and  $v$ , see last column). Note that this last case also incorporates  $w = v = 1$ , which is a realistic setting, as developed e.g. in [9]. We see that, in such a case, the correlation depends only on  $u$ , and will be  $-1/\sqrt{1+u^2}$ .

#### 4.2.4 Comparison with Empirical Work

As final part, we want to assess the finite sample performance of the theoretical results: While we already presented a more general finite sample performance analysis in Subsection 2.4, here we want to evaluate the adequateness of a framework which exactly is the same setup as in [9]. Thus, we simulate an iid sample with mean  $\mu = 0$  of size  $N = 7562$ . Then, using a sample size of  $n = 252$  (corresponding to one year of data), we compute with a monthly rolling-window the time series of quantile estimates  $\hat{q}_{n, t}(p)$  for quantiles of order  $p = 0.95$  and  $p = 0.99$  each, and a monthly time series of measure of dispersion estimates  $\hat{D}_{i, n, t}$ .  $\hat{q}_{n, t}$  is either the sample quantile time series  $\hat{q}_{n, t}$  or the location-scale quantile time series with known  $\mu = 0$  time series  $q_{n, \hat{\sigma}, t}$ ,  $\hat{D}_{i, n, t}$  can be either the sample standard deviation, the sample MAD, or the sample MedianAD. This gives a time series of 349 values each for the quantile estimators and measure of dispersion estimators. Recalling the correlation of interest,  $\text{Cor} \left( \log \left| \frac{\hat{q}_{n, t+1y}(p)}{\hat{q}_{n, t}(p)} \right|, \hat{D}_{i, n, t} \right)$  in (104), we see that for every time point  $t$ , we have to build the ratio of quantile estimates  $\frac{\hat{q}_{n, t+1y}(p)}{\hat{q}_{n, t}(p)}$ . As we are using a monthly rolling-window, we cannot build this ratio for the last twelve points in the time series (as we do not have further points in the future  $\hat{q}_{n, t+1y}(p)$ ). Thus, the time series of log-ratios consists of 337 observations and correspondingly we reduce the time series of dispersion measure estimates to 337. With those  $m = 337$  observations, we then estimate the correlation of interest. This procedure is repeated 100'000-fold.

We consider iid samples coming from three different distributions: From a Gaussian distribution and

from two Student distributions with 3 and 5 degrees of freedom, respectively. Note, that the specification of  $\mu, \sigma^2$  is not needed as the correlation results for location-scale distributions are independent of its parameters. Then, we report in Table 8 the averages of the 100'000-fold repetition with, in brackets, the corresponding empirical 95% confidence interval values. Further, we provide the theoretical value below the estimated value. Also, we provide confidence intervals for the sample Pearson linear correlation coefficient (using the classical variance-stabilizing Fisher transform of the correlation coefficient for a bivariate normal distribution to compute the confidence intervals -see the original paper [18] or e.g. a standard encyclopedia entry [37]). Note, that those confidence interval values have to be considered with care. Firstly, recall that the bivariate normality of our quantile estimator and measure of dispersion estimator hold asymptotically. Hence, it is not clear if for the sample size  $n = 252$  considered we can assume bivariate normality (this could be tested). Secondly, the procedure in [18] assumes an independent sample for the correlation estimation while our sample of  $m = 337$  pairs is clearly dependent (because we computed it with a monthly rolling window giving us overlapping samples). This would reduce the sample size on which to compute the confidence intervals to an effective sample size  $m_{eff} < m$ . Still, we provide those theoretical confidence intervals as approximate guidance.

We recall that when working with the sample standard deviation, the existence of the fourth moment is a condition we require, also when working with the location-scale quantile. Thus, as they do not exist for a Student distribution with 3 degrees of freedom, we simply write 'NA' as theoretical value instead.

Table 8: Average values from a 100'000-fold repetition. Comparing the correlation for our three measures of dispersion and two different quantile estimators. Average empirical values are written first (with empirical 95% confidence interval in brackets), the corresponding theoretical values, mean and 95% confidence interval, being in the row below. We consider two thresholds  $p = 0.95, 0.99$ , overall sample size  $N = 7562$  and the time series of estimators computed each on a sample size of  $n = 252$ .

Quantile estimator $\hat{q}_n$ :	Gaussian distr.		Student distr. (5df)		Student distr. (3df)	
	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$
$p = 0.95$						
Cor( $\log \frac{\hat{q}_{t+1}}{\hat{q}_t} , \hat{\sigma}_t$ )	-40 (-59,-18)	-72 (-82,-61)	-31 (-54,-6)	-72 (-81,-61)	-19 (-45,+8)	-71 (-81,-60)
	-39 (-48,-30)	-71 (-76,-65)	-27 (-37,-17)	-71 (-76,-65)	NA	NA
Cor( $\log \frac{\hat{q}_{t+1}}{\hat{q}_t} , \hat{\theta}_t$ )	-34 (-55,-12)	-67 (-79,-55)	-37 (-57,-14)	-62 (-75,-46)	-35 (-56,-11)	-60 (-74,-43)
	-34 (-43,-24)	-66 (-72,-60)	-36 (-45,-26)	-54 (-61,-46)	-34 (-43,-24)	NA
Cor( $\log \frac{\hat{q}_{t+1}}{\hat{q}_t} , \hat{\xi}_t$ )	-16 (-40,+8)	-43 (-62,-23)	-17 (-40,+8)	-27 (-50,-2)	-16 (-40,+8)	-15 (-40,+11)
	-16 (-26,-5)	-43 (-51,-34)	-16 (-26,-5)	-23 (-33,-13)	-16 (-26,-5)	NA
$p = 0.99$						
Cor( $\log \frac{\hat{q}_{t+1}}{\hat{q}_t} , \hat{\sigma}_t$ )	-32 (-53,-9)	-72 (-82,-61)	-36 (-57,-11)	-72 (-81,-61)	-29 (-54,0)	-71 (-81,-60)
	-31 (-40,-21)	-71 (-76,-65)	-30 (-39,-20)	-71 (-76,-65)	NA	NA
Cor( $\log \frac{\hat{q}_{t+1}}{\hat{q}_t} , \hat{\theta}_t$ )	-23 (-45,+1)	-67 (-79,-55)	-29 (-51,-6)	-62 (-75,-46)	-33 (-54,-9)	-60 (-74,-43)
	-22 (-32,-12)	-66 (-72,-60)	-28 (-38,-18)	-54 (-61,-46)	-31 (-40,-21)	NA
Cor( $\log \frac{\hat{q}_{t+1}}{\hat{q}_t} , \hat{\xi}_t$ )	-8 (-32,+17)	-43 (-62,-23)	-7 (-32,+18)	-27 (-50,-2)	-7 (-32,+18)	-15 (-40,+12)
	-7 (-18,+4)	-43 (-51,-34)	-7 (-18,+4)	-23 (-33,-13)	-7 (-18,+4)	NA

From Table 8 we see that the empirical averages and the theoretical values seem to agree well in general.

Before going into detail, remember that in Table 2 in Subsection 2.4 we had a general extensive simulation study on the effect of sample size. We can set the theoretical asymptotic covariance of the former into relation to the present setting: E.g. when considering as dispersion measure the sample standard deviation, we have

$$\lim_{n \rightarrow \infty} \text{Cor} \left( \log \left| \frac{\hat{q}_{t+1}}{\hat{q}_t} \right|, \hat{\sigma}_t \right) = -\frac{1}{\sqrt{2}} \lim_{n \rightarrow \infty} \text{Cor}(\hat{q}_t, \hat{\sigma}_t)$$

Meaning, all theoretical values in this Table 8 multiplied by  $-\sqrt{2}$  give the theoretical values of Table 2. Further, while the sample correlation in Table 2 was computed on disjoint samples, the setting here uses a rolling-window which results in overlapping samples. Comparing the values in both tables, one can see that this affects (to a varying and sometimes negligible extent) the quality of approximation of the theoretical values.

Let us now examine the results in Table 8 closer.

We start by comparing again the behaviour between the two quantile estimators. It might seem obvious by now, after having presented all the theoretical results, but in the table we can see that the correlation of the same measure of dispersion with the different quantile estimators (comparing always the two columns for each distribution) has different values (and the correlation is higher in absolute terms with  $q_{n,\hat{\sigma}}$  than with  $q_n$  for such tail values - see also Figures 5, 6). Also, the confidence intervals are larger (or at least not smaller) with  $q_n$  in comparison with  $q_{n,\hat{\sigma}}$ , as expected.

Comparing the cases of  $p = 0.95$  and  $p = 0.99$ , we can see that  $q_n$  has usually higher correlation in absolute terms with  $p = 0.95$  (except for the Student distribution with 5 degrees of freedom). Recall that  $q_{n,\hat{\sigma}}$  does not depend on  $p$  (as we already saw in Tables 5, 6), thus one has the same values for  $p = 0.95, 0.99$ . Further, the size of the empirical confidence intervals are similar for both values of  $p$  for all measures of dispersion.

Comparing the behaviour for the three different distributions considered, we first recall that in the case of a Student distribution with 3 degrees of freedom, we do not have all theoretical values because of the moment conditions. Also, we know that the correlation of the MedianAD with the sample quantile is independent of the distribution (follows from Proposition 17), which gives the same values for the three different distributions (still, the values depend on  $p$ ). Similarly, the correlation between the sample standard deviation and the location-scale quantile is independent of the distribution (and even of  $p$ !). In general, we can say that the empirical values behave as analysed in the theoretical formulae in Subsection 4.2: E.g. for the correlation between the sample quantile with the sample standard deviation, it is clearly bigger in absolute terms for the Gaussian case than the Student distribution with 5 degrees of freedom, and reverse (although similar values) for the sample quantile with the MAD (in that case also the value for the Student distribution with 3 degrees of freedom is similar). The size of the confidence intervals changes depending on the distribution: For the sample quantile, it increases with heavier tails when considering the correlation with the sample standard deviation, while staying the same with the MAD or MedianAD. For the location-scale quantile, it is the other way round. It increases with the heaviness of the tail with the MAD and MedianAD, whereas staying constant with the sample standard deviation.

We conclude these empirical observations comparing the behaviour of the three measures of dispersion (the sample standard deviation  $\hat{\sigma}_t$ , the sample MAD  $\hat{\theta}_t$  and the sample MedianAD  $\hat{\xi}_t$ ). As already seen in Subsection 4.2, for  $p = 0.95, 0.99$ , the correlation in absolute terms will be in the following order (for all cases in which its theoretical counterpart is defined): The highest with  $\hat{\sigma}_t$ , then with  $\hat{\theta}_t$  and the lowest with  $\hat{\xi}_t$ . The only exception to this, is the case of the Student distribution with 5 degrees of freedom for  $p = 0.95$ , recall Figure 6 (plots in second row). Also, the increasing size of the confidence intervals



(for defined quantities) is in the same ordering, i.e. the one for  $\hat{\sigma}_t$  being the smallest and with  $\hat{\xi}_t$  the biggest (the only exception being the sample quantile in the case of a Student distribution with 5 degrees of freedom).

## 5 Concluding Remarks: Implications for Use in Practice

To recapitulate, we give a brief overview over the results presented: In this paper we showed the joint asymptotic normality of functionals of two quantile estimators and three measure of dispersion estimators each - for underlying iid models. Apart from providing explicit results and analyses in the cases of the Gaussian and Student distributions, we also compared the difference in speed of convergence when using either the sample quantile or the parametric location-scale quantile as quantile estimator. Further, we applied the theoretical results in to risk measures as VaR, ES and expectile as well as to related empirical work in the literature - [9], [44], [43]. Also we verified a good finite sample performance through simulations. Keeping the focus on the practical applicability of the results, we conclude by summarising and comparing the different theoretical properties and restrictions of the estimators and how this impacts their use in applications.

While we specified in each theorem which moment and smoothness conditions the underlying parent random variable has to fulfil (although we did not not assume minimal conditions, for simplicity), we offer in Table 9 a complete overview which we analyse now.

Let us start with discussing the choice of the quantile estimator  $q_n$  or  $q_{n,\hat{\sigma}}$ . We see that asymptotics with  $q_n$  do not give a restriction on the moments of its distribution (they will only be imposed by the corresponding measure of dispersion estimator) and apart from the density needing to exist at  $q_X(p)$  (condition (H1)) we have some (not very restrictive) further smoothness and differentiability conditions, (H3). In contrast to that, the location-scale quantile can only be used when assuming to have a location-scale distribution with finite fourth moment and, as for any parametric estimator, assume a specific parametric family of (location-scale) distributions - which is a restrictive condition. In return, we have seen in the finite sample analysis in Section 2 that the latter restriction results in narrower confidence intervals - and a higher degree of correlation, as seen in the examples of Section 4, as expected (an exception to this are intermediate values, around  $p = 0.25, 0.75$  for the MedianAD). Also, it is useful to understand the joint asymptotics with the location-scale quantile as it acts as a benchmark to measure the efficiency of any other quantile estimator (recall Subsection 3.3).

With respect to the choice of the measure of dispersion estimator, we can observe similar trade-offs. When using the sample standard deviation, we need a finite fourth moment of the distribution. But in exchange, we usually have the best empirical confidence intervals and the strongest correlation when considering tail values for  $p$ . A needed fourth moment discourages from using the sample standard deviation, although it is the most popular measure of dispersion in finance. We have the options of using the two other measures of dispersion that require at most a finite second moment: The sample MAD requires finite variance and additionally continuity at  $F_X(\mu)$  for its asymptotics (condition (Q1) in the table). If we want to use it in joint asymptotics with a quantile estimator, we additionally need  $F_X$  to be Hölder continuous at  $\mu$  (which, for a continuously differentiable distribution function, is the case). Going back to empirical observations, we can say that it has empirically the second best confidence intervals and usually the second strongest correlation when considering tail values.

In the case of the sample MedianAD, we do not even need a finite mean (these are the least restrictive moment conditions), but rather some continuity and differentiability conditions (see (Q2) in the table)! For its joint asymptotics with a quantile estimator, we further need differentiability of the density at

certain points  $(\nu, \nu \pm \xi)$ . But with this, comes the disadvantage of having the largest uncertainty in the empirical confidence intervals and, in the examples considered, the weakest correlation when considering tail values of  $p$ . Worthwhile recalling is also the fact that its correlation with the sample quantile for symmetric location-scale distributions will be independent of the distribution (but still a function of  $p$ ), meaning that one cannot infer on a possible underlying distribution based on its correlation value.

Overall it seems that using the asymptotics with the sample MAD (and the sample quantile as quantile estimator) seem to be a good compromise between empirical performance and theoretical restrictions.

Table 9: Overview of minimal conditions the distribution have to fulfil for the different quantile (first part) and measure of dispersion (second part) estimators. For each, we split the conditions into three categories: First, the conditions needed for the asymptotics of each estimator. Second, for its representation as iid-sum (usually Bahadur representation) and thirdly for its use in computation of asymptotic covariances and correlations (i.e. joint asymptotics with a corresponding quantile or measure of dispersion estimator)

Quantile Estimator	Asymptotic Normality	Bahadur Representation	Joint asymptotics (with a measure of dispersion estimator)
$q_n(p)$	(H1): $0 < f_X(q_X(p)) < \infty$	(H1)	(H1), (H3): $\begin{cases} F_X \text{ twice differentiable in neighbourhood of } q_X(p), \\ F_X'' \text{ bounded in that neighbourhood,} \\ F_X(q_X(p)) = p \end{cases}$
$q_{n,\hat{\mu},\hat{\sigma}}(p)$	(H2): $\begin{cases} \mathbb{E}[X^4] < \infty, \\ (X - \mu)^2 \text{ not constant} \end{cases}$	—	(H2)
$q_{n,\hat{\sigma}}(p)$	(H2)	—	(H2)
Measure of Dispersion Estimator	Asymptotic Normality	Bahadur Representation	Joint asymptotics (with a quantile estimator)
$\hat{\sigma}_n^2$	(H2)	—	(H2)
$\hat{\theta}_n$	(Q1): $\begin{cases} \mathbb{E}[X^2] < \infty, \\ F_X \text{ continuous at } \mu \end{cases}$	(Q1)	$\begin{cases} \text{(Q1),} \\ \text{(Q3): } F_X \text{ Hölder-continuous at } \mu \end{cases}$
$\hat{\xi}_n$	(Q2): $\begin{cases} F_X \text{ continuous in} \\ \text{neighbourhood of } \nu \pm \xi, \\ F_X \text{ differentiable at } \nu, \nu \pm \xi \\ f_X(\nu) > 0, f_{ X-\nu }(\xi) > 0, \end{cases}$	(Q2)	$\begin{cases} \text{(Q2),} \\ \text{(Q4): } F_X \text{ twice differentiable at } \nu, \nu \pm \xi \end{cases}$

## References

- [1] ACERBI, C., AND TASCHE, D. On the coherence of expected shortfall. *Journal of Banking & Finance* 26, 7 (2002), 1487–1503.
- [2] ARISMENDI, J. Multivariate truncated moments. *Journal of Multivariate Analysis* 117 (2013), 41–75.
- [3] ARTZNER, P., DELBAEN, F., EBER, J.-M., AND HEATH, D. Coherent measures of risks. *Mathematical Finance* 9 (1999), 203–228.
- [4] BABU, G., AND RAO, C. Expansions for statistics involving the mean absolute deviations. *Annals of the institute of statistical mathematics* 44, 2 (1992), 387–403.
- [5] BAHADUR, R. A note on quantiles in large samples. *The Annals of Mathematical Statistics* 37, 3 (1966), 577–580.
- [6] BERA, A., GALVAO, A., WANG, L., AND XIAO, Z. A new characterization of the normal distribution and test for normality. *Econometric Theory* 32, 5 (2016), 1216–1252.
- [7] BIRKES, D., AND DODGE, Y. *Alternative methods of regression*, vol. 190. John Wiley & Sons, 2011.
- [8] BOS, C., AND JANUS, P. A quantile-based realized measure of variation: New tests for outlying observations in financial data. *Tinbergen Institute Discussion Paper 13-155/III* (2013).
- [9] BRÄUTIGAM, M., DACOROGNA, M., AND KRATZ, M. Predicting risk with risk measures: an empirical study. *ESSEC Working Paper 1803* (2018).
- [10] CHEN, J. On exactitude in financial regulation: Value-at-risk, expected shortfall, and expectiles. *Risks* 6, 2 (2018), 1–29.
- [11] COMMITTEE, B., ET AL. Basel III: A global regulatory framework for more resilient banks and banking systems. *Basel Committee on Banking Supervision, Basel* (2010).
- [12] COUPERIER, O., AND LEYMARIE, J. Backtesting expected shortfall via multi-quantile regression. *HAL working paper*; available on <https://halshs.archives-ouvertes.fr/halshs-01909375> (2018).
- [13] DACOROGNA, M., GENÇAY, R., MÜLLER, U., PICTET, O., AND OLSEN, R. *An introduction to high-frequency finance*. Academic Press, New York, 2001.
- [14] DENNEBERG, D. Premium calculation: why standard deviation should be replaced by absolute deviation. *ASTIN Bulletin: The Journal of the IAA* 20, 2 (1990), 181–190.
- [15] EMMER, S., KRATZ, M., AND TASCHE, D. What is the best risk measure in practice? a comparison of standard risk measures. *Journal of Risk* 18, 2 (2015), 31–60.
- [16] FAMA, E. The behavior of stock-market prices. *The journal of Business* 38, 1 (1965), 34–105.
- [17] FERGUSON, T. Asymptotic joint distribution of sample mean and a sample quantile. *unpublished: <http://www.math.ucla.edu/~tom/papers/unpublished/meanmed.pdf>* (1999).
- [18] FISHER, R. On the probable error of a coefficient of correlation deduced from a small sample. *Metron* 1 (1921), 3–32.
- [19] GONZÁLEZ, C., AND COLLAZOS, J. Moments of truncated skew-t distribution and graph theory applied to the shortest path problem. *British Journal of Mathematics & Computer Science* 13, 3 (2016), 1–16.
- [20] GORARD, S. Revisiting a 90-year-old debate: the advantages of the mean deviation. *British Journal of Educational Studies* 53, 4 (2005), 417–430.
- [21] HO, H., LIN, T.-I., CHEN, H.-Y., AND WANG, W.-L. Some results on the truncated multivariate t distribution. *Journal of Statistical Planning and Inference* 142, 1 (2012), 25–40.
- [22] JAMALIZADEH, A., KHOSRAVI, M., AND BALAKRISHNAN, N. Recurrence relations for distributions of a skew-t and a linear combination of order statistics from a bivariate-t. *Computational Statistics & Data Analysis* 53, 4 (2009), 847–852.

- [23] JAMALIZADEH, A., POURMOUSA, R., AND BALAKRISHNAN, N. Truncated and limited skew-normal and skew-t distributions: properties and an illustration. *Communications in Statistics-Theory and Methods* 38, 16-17 (2009), 2653–2668.
- [24] KIM, H.-J. Moments of truncated student-t distribution. *Journal of the Korean Statistical Society* 37, 1 (2008), 81–87.
- [25] KOENKER, R., AND BASSETT JR, G. Regression quantiles. *Econometrica* 46 (1978), 33–50.
- [26] KONNO, H., AND YAMAZAKI, H. Mean-absolute deviation portfolio optimization model and its applications to tokyo stock market. *Management science* 37, 5 (1991), 519–531.
- [27] KUAN, C.-M., YEH, J.-H., AND HSU, Y.-C. Assessing value at risk with care, the conditional autoregressive expectile models. *Journal of Econometrics* 150, 2 (2009), 261–270.
- [28] LEYS, C., LEY, C., KLEIN, O., BERNARD, P., AND LICATA, L. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology* 49, 4 (2013), 764–766.
- [29] LIN, P.-E., WU, K.-T., AND AHMAD, I. Asymptotic joint distribution of sample quantiles and sample mean with applications. *Communications in Statistics-Theory and Methods* 9, 1 (1980), 51–60.
- [30] MARKOWITZ, H. Portfolio selection. *The journal of finance* 7, 1 (1952), 77–91.
- [31] MAZUMDER, S., AND SERFLING, R. Bahadur representations for the median absolute deviation and its modifications. *Statistics & Probability Letters* 79, 16 (2009), 1774–1783.
- [32] MORGAN, J., AND REUTERS. Riskmetrics - technical document. Available on: <https://www.mscom.com/documents/10199/5915b101-4206-4ba0-ae2-3449d5c7e95a> (1996).
- [33] NEWEY, W., AND POWELL, J. Asymmetric least squares estimation and testing. *Econometrica: Journal of the Econometric Society* (1987), 819–847.
- [34] PHAM-GIA, T., AND HUNG, T. The mean and median absolute deviations. *Mathematical and Computer Modelling* 34, 7-8 (2001), 921–936.
- [35] POPOVICIU, T. Sur les équations algébriques ayant toutes leurs racines réelles. *Mathematica* 9 (1935), 129–145.
- [36] PYKE, R. Spacings. *Journal of the Royal Statistical Society. Series B (Methodological)* (1965), 395–449.
- [37] RODRIGUEZ, R. Correlation,. In *Encyclopedia of statistical sciences, 2nd edition*, S. Kotz, N. Balakrishnan, C. Read, B. Vidakovic, and N. Johnson, Eds. Wiley, New York, 1982, pp. 1375–1385.
- [38] ROUSSEEUW, P. Least median of squares regression. *Journal of the American statistical association* 79, 388 (1984), 871–880.
- [39] SEGERS, J. On the asymptotic distribution of the mean absolute deviation about the mean. *arXiv preprint arXiv:1406.4151* (2014).
- [40] SERFLING, R., AND MAZUMDER, S. Exponential probability inequality and convergence results for the median absolute deviation and its modifications. *Statistics & Probability Letters* 79, 16 (2009), 1767–1773.
- [41] SIMAAN, Y. Estimation risk in portfolio selection: the mean variance model versus the mean absolute deviation model. *Management science* 43, 10 (1997), 1437–1446.
- [42] YAO, Q., AND TONG, H. Asymmetric least squares regression estimation: a nonparametric approach. *Journal of Nonparametric Statistics* 6, 2-3 (1996), 273–292.
- [43] ZUMBACH, G. Correlations of the realized volatilities with the centered volatility increment. [http://www.finanscopics.com/figuresPage.php?figCode=corr\\_vol\\_r\\_VsDV0](http://www.finanscopics.com/figuresPage.php?figCode=corr_vol_r_VsDV0), 2012. [Online; accessed 1-October-2018].
- [44] ZUMBACH, G. *Discrete Time Series, Processes, and Applications in Finance*. Springer Science & Business Media, 2012.

## APPENDIX

The appendix has four sections. The first three correspond each to one of the three main sections of the paper and contain its proofs respectively. Of the biggest mathematical interest are the proofs of the main theorems (asymptotic distributions of quantile estimators and measure of dispersion estimators) in Appendix A. As the results in Section 3 are only special cases of the results in Section 2, the proofs in Appendix B are simply its computations. Still, it is worthwhile pointing out that the computations of the examples throughout this Appendix section can be of interest: To understand the examples better, but also as the quantities involved (truncated/absolute moments of Gaussian and Student distribution) have been sometimes wrongly computed in the literature. Then, Appendix C contains the explicit computations of asymptotic correlations between the measure of dispersion estimators and the expected shortfall estimator which were not presented in Subsection 4.1 because of their length. Finally, Appendix D presents extensions of the main asymptotic theorems, as well as some additional tables from the simulation study in Section 2.4.

### Appendix A Proofs of Section 2

In Appendix A.1 we cover the proofs of the asymptotics of the sample quantile with the three measure of dispersion estimators: We present the proofs of Theorem 1 (asymptotic distribution of sample quantile with either sample variance or sample MAD) and Proposition 3 (asymptotic distribution of sample quantile with the sample MedianAD). Both results will be proved using the Bahadur representation (Appendix A.1.1). But in the case of Theorem 1, we also offer an alternative proof via Taylor expansion (Appendix A.1.2). Then, Appendix A.2 contains the corresponding proofs of the asymptotics when using the location-scale quantile estimator. In Appendix A.4 we do not present a proof of Proposition 8, as this is only an application of the Delta method, but the proof of the scaling law (Proposition 10).

#### A.1 Proofs of Subsection 2.1

##### A.1.1 Bahadur's Method

This approach is used to prove both, Theorem 1 and Proposition 3.

**Proof of Theorem 1.** It consists of two parts. In the first part, we assume the mean  $\mu$  to be known. Using the Bahadur representation, then the CLT, we show the asymptotic joint normality of the sample quantile and the sample measure of dispersion with known mean,  $\tilde{m}(X, n, r)$  (recall Table 1), for any integer  $r > 0$ . Then, in a second part, we use this result to show the asymptotic joint normality of the sample quantile with either the sample variance or the sample MAD in the case of an unknown mean.

Part 1 - Known Mean -

- *Bahadur representation*

The Bahadur representation for sample quantiles from an iid sample is given in [5]:

$$q_n(p) = q_X(p) + \frac{1 - F_n(q_X(p)) - (1 - p)}{f_X(q_X(p))} + R_{n,p}, \quad (109)$$

where  $R_{n,p}$  is a 'rest' of order  $O(n^{-3/4} \log n)$  almost surely (which means that  $|R_{n,p}| \leq Mn^{-3/4} \log(n)$  a.s. for all  $n \geq n_0$  for some constants  $M, n_0 < \infty$ ). With this Bahadur representation we are able to use the bivariate central limit theorem for the sample quantile  $q_n(p)$  and the sample measure of dispersion with known mean  $\mu$ ,  $\tilde{m}(X, n, r)$ .

- *Central Limit Theorem*

Under condition  $(M_r)$ ,  $r > 0$ , we get

$$n^{-1/2} \sum_{i=1}^n \left( \left( \frac{\mathbf{1}_{(X_i > q_X(p))}}{|X_i - \mu|^r} \right) - \left( \frac{1 - p}{m(X, r)} \right) \right) = n^{1/2} \left( \left( \frac{1 - F_n(q_X(p))}{\frac{1}{n} \sum |X_i - \mu|^r} \right) - \left( \frac{1 - p}{m(X, r)} \right) \right) \xrightarrow{d} \mathcal{N}(0, \hat{\Sigma}^{(r)}), \quad (110)$$

where  $\hat{\Sigma}^{(r)} = \begin{pmatrix} \text{Var}(\mathbf{1}_{(X_i > q_X(p))}) & \text{Cov}(\mathbf{1}_{(X_i > q_X(p))}, |X - \mu|^r) \\ \text{Cov}(\mathbf{1}_{(X_i > q_X(p))}, |X - \mu|^r) & \text{Var}(|X - \mu|^r) \end{pmatrix}$ .

Then, we need to pre-multiply (i.e. from the left side) equation (110) by  $\begin{bmatrix} 1/(f_X(q_X(p))) & 0 \\ 0 & 1 \end{bmatrix}$  to use the Bahadur representation (109) of the sample quantile. One gets (as in [8], just with a different notation),

$$n^{1/2} \left( \frac{1 - F_n(q_X(p)) - (1-p)}{f_X(q_X(p))} \right) = n^{1/2} \left( \frac{q_n(p) - q_X(p) - R_{n,p}}{\tilde{m}(X, n, r) - m(X, r)} \right) \xrightarrow{d} \mathcal{N}(0, \tilde{\Sigma}^{(r)}) \quad (111)$$

where now

$$\tilde{\Sigma}^{(r)} = \begin{pmatrix} \frac{\text{Var}(\mathbf{1}_{(X_i > q_X(p))})}{f_X^2(q_X(p))} & \frac{\text{Cov}(\mathbf{1}_{(X_i > q_X(p))}, |X - \mu|^r)}{f_X(q_X(p))} \\ \frac{\text{Cov}(\mathbf{1}_{(X_i > q_X(p))}, |X - \mu|^r)}{f_X(q_X(p))} & \text{Var}(|X - \mu|^r) \end{pmatrix}. \quad (112)$$

Because of the fact that  $R_{n,p}$  has rate  $O(n^{-3/4} \log(n))$  a.s., we can ignore it in an asymptotic analysis (as it follows from Popovicius inequality for variances, [35], that  $\text{Var}(\sqrt{n}R_{n,p}) \leq M \frac{\log(n)}{n^{1/2}}$  for some  $M < \infty$ ).

Now let us compute the asymptotic covariance matrix  $\tilde{\Sigma}^{(r)}$ . As we assumed  $(C_1)$ , we have  $F_X(q_X(p)) = p$  and hence

$$\mathbb{E}[\mathbf{1}_{(X_i > q_X(p))}] = 1 - p, \quad \text{Var}(\mathbf{1}_{(X_i > q_X(p))}) = p(1 - p),$$

hence, we obtain,  $\tau_k(\eta(X), p)$  being defined in(3),

$$\text{Cov}(\mathbf{1}_{(X_i > q_X(p))}, |X_i - \mu|^r) = (1 - p) \mathbb{E}[|X_i - \mu|^r | X_i > q_X(p)] - (1 - p) \mathbb{E}[|X_i - \mu|^r] = \tau_r(|X_i - \mu|, p). \quad (113)$$

Thus, we arrive at the asymptotic joint distribution of the sample quantile and the sample measure of dispersion with known  $\mu$  for any integer  $r > 0$ .

## Part 2 - Unknown Mean -

We analyse what happens with respect to the joint asymptotic distribution if we consider  $\hat{m}(X, n, r)$  instead of  $\tilde{m}(X, n, r)$ . For this, we restrict the framework to  $r = 1, 2$ , and treat the two cases separately.

- Case  $r = 2$ . Note that we can write, using the sample variance  $\tilde{\sigma}_n^2$  with known mean  $(\mu)$ ,

$$\hat{\sigma}_n^2 = \frac{n}{n-1} \left( \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 - (\bar{X}_n - \mu)^2 \right) = \frac{n}{n-1} (\tilde{\sigma}_n^2 - (\bar{X}_n - \mu)^2). \quad (114)$$

Since we know that  $(\bar{X}_n - \mu) \xrightarrow{P} 0$  and  $\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$ , we deduce by Slutsky's theorem

$$\sqrt{n}(\bar{X}_n - \mu)^2 \xrightarrow{d, P} 0. \quad (115)$$

Thus, in our asymptotic analysis, we do not need any more to consider this term. Now, as we have

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\hat{\sigma}_n^2, \sqrt{n}q_n(p)) = \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\tilde{\sigma}_n^2, \sqrt{n}q_n(p)) + \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}(\bar{X}_n - \mu)^2, \sqrt{n}q_n(p)),$$

it follows, by Cauchy-Schwarz inequality,

$$\lim_{n \rightarrow \infty} |\text{Cov}(\sqrt{n}(\bar{X}_n - \mu)^2, \sqrt{n}q_n(p))| \leq \lim_{n \rightarrow \infty} \sqrt{\text{Var}(\sqrt{n}(\bar{X}_n - \mu)^2)} \sqrt{\text{Var}(\sqrt{n}q_n(p))} = 0 \times \sqrt{\frac{p(1-p)}{f_X^2(q_X(p))}} = 0.$$

We conclude that we can consider without loss of generality  $\tilde{\sigma}_n^2$  instead of  $\hat{\sigma}_n^2$  for our asymptotic analysis.

- Case  $r = 1$ . In contrast to the case  $r = 2$ , the asymptotics of  $\hat{\theta}_n = \hat{m}(X, n, 1)$  and  $\tilde{\theta}_n := \tilde{m}(X, n, 1)$  are, in general, not the same, as we are going to see.

From [4] - as also mentioned in [39] - we can see that for a distribution  $F_X$  with finite second moment and

whose cdf is Hölder-continuous at  $\mu$ , such that  $|F_X(x) - F_X(\mu)| \leq c|x - \mu|^\beta \forall x \in \mathbb{R}, c > 1, 0 < \beta \leq 1$ , defining

$$R_n := 2 \int_0^1 [F_n(\mu + (\bar{X}_n - \mu)y) - F_n(\mu) - F_X(\mu + (\bar{X}_n - \mu)y) + F_X(\mu)] dy, \quad (116)$$

it holds

$$\sqrt{n}(\hat{\theta}_n - \theta) = \sqrt{n}(\tilde{\theta}_n - \theta) + \sqrt{n}(2F_n(\mu) - 1)(\bar{X}_n - \mu) + \sqrt{n}L_n, \quad (117)$$

$$\text{with } L_n := (\bar{X}_n - \mu)R_n + \sqrt{n}2(\bar{X}_n - \mu) \int_0^1 [F_X(\mu + (\bar{X}_n - \mu)y) - F_X(\mu)] dy$$

and for some  $k > 0$ , again from [4],

$$\mathbb{P} \left( |\bar{X}_n - \mu| \leq n^{-1/2} \log n, |R_n| > k(\log n)^{(1+\beta)/2} n^{-(2+\beta)/4} \right) = O(n^{-2}). \quad (118)$$

Further, [4] show that asymptotically

$$\sqrt{n}(\tilde{\theta}_n - \theta) + (2F_X(\mu) - 1)\sqrt{n}(\bar{X}_n - \mu) + \sqrt{n}L_n \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \sigma_\theta^2) \quad (119)$$

with  $\sigma_\theta^2 = \text{Var}(|X - \mu| + (2F_X(\mu) - 1)X)$ . Those two relations (118) and (119) will be useful later on in the computation. Note that we can write (117) as

$$\sqrt{n}(\hat{\theta}_n - \theta) = \sqrt{n}(\tilde{\theta}_n - \theta) + (2F_X(\mu) - 1)\sqrt{n}(\bar{X}_n - \mu) + (2F_n(\mu) - 2F_X(\mu))\sqrt{n}(\bar{X}_n - \mu) + \sqrt{n}L_n. \quad (120)$$

Since  $F_n(\mu) - F_X(\mu) \xrightarrow[n \rightarrow \infty]{P} 0$  and  $\sqrt{n}(\bar{X}_n - \mu) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \sigma^2)$ , hence by Slutsky Theorem

$$(2F_n(\mu) - 2F_X(\mu))\sqrt{n}(\bar{X}_n - \mu) \xrightarrow[n \rightarrow \infty]{d, P} 0 \quad (121)$$

such that this term does not change the asymptotics of  $\theta_n$ . But clearly, this impacts the following asymptotic joint distribution:

$$\sqrt{n} \begin{pmatrix} q_n(p) - q_X(p) \\ \hat{\theta}_n - \theta \end{pmatrix} = \sqrt{n} \begin{pmatrix} q_n(p) - q_X(p) \\ \tilde{\theta}_n - \theta + (2F_X(\mu) - 1)(\bar{X}_n - \mu) + 2(F_n(\mu) - F_X(\mu))(\bar{X}_n - \mu) + L_n \end{pmatrix} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Sigma^{(r)}) \quad (122)$$

where we already know  $\Sigma_{11}^{(r)}$  from (111)/(112) and  $\Sigma_{22}^{(r)}$  from (119), thus having

$$\Sigma^{(r)} = \begin{pmatrix} \frac{p(1-p)}{f_X^2(q_X(p))} & \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\hat{q}_n(p), \sqrt{n}\hat{\theta}_n) \\ \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\hat{q}_n(p), \sqrt{n}\hat{\theta}_n) & \text{Var}(|X - \mu| + (2F_X(\mu) - 1)X) \end{pmatrix}.$$

We conclude with the computation of the asymptotic covariance: Using (120), then (111)/(112)/(113) for the first part, and the expression of  $\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\bar{X}_n, \sqrt{n}q_n(p))$  given in [17] for the second part of the covariance respectively, we obtain

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}(\hat{\theta}_n), \sqrt{n}q_n(p)) &= \frac{\tau_1(|X - \mu|, p)}{f_X(q_X(p))} + (2F_X(\mu) - 1) \frac{\tau_1(p)}{f_X(q_X(p))} \\ &+ \lim_{n \rightarrow \infty} \text{Cov}(2(F_n(\mu) - F_X(\mu))\sqrt{n}(\bar{X}_n - \mu), \sqrt{n}q_n(p)) + \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}L_n, \sqrt{n}q_n(p)). \end{aligned}$$

The last two covariance terms on the right hand side of this equality can be shown to be asymptotically negligible, using Cauchy-Schwarz. We have, using (121),

$$\lim_{n \rightarrow \infty} |\text{Cov}(2(F_n(\mu) - F_X(\mu))\sqrt{n}(\bar{X}_n - \mu), \sqrt{n}q_n(p))| \leq \lim_{n \rightarrow \infty} [\text{Var}(2(F_n(\mu) - F_X(\mu))\sqrt{n}(\bar{X}_n - \mu)) \text{Var}(\sqrt{n}q_n(p))] = 0.$$

The covariance with  $L_n$  requires more work:

$$\begin{aligned} \lim_{n \rightarrow \infty} |\text{Cov}(\sqrt{n}L_n, \sqrt{n}q_n(p))| &= \\ &\leq \lim_{n \rightarrow \infty} \left( \text{Var}(\sqrt{n}(\bar{X}_n - \mu)R_n) + \text{Var}\left(\sqrt{n}2(\bar{X}_n - \mu) \int_0^1 [F_X(\mu + (\bar{X}_n - \mu)y) - F_X(\mu)] dy\right) \right) \text{Var}(\sqrt{n}q_n(p)), \end{aligned}$$

where we want to show that

$$\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n}(\bar{X}_n - \mu)R_n) = 0 \quad (123)$$

$$\lim_{n \rightarrow \infty} \text{Var}\left(\sqrt{n}2(\bar{X}_n - \mu) \int_0^1 [F_X(\mu + (\bar{X}_n - \mu)y) - F_X(\mu)]dy\right) = 0 \quad (124)$$

Let us first consider (123). Clearly,  $\mathbb{E}[\sqrt{n}(\bar{X}_n - \mu)] = 0$ ,  $\text{Var}(\sqrt{n}(\bar{X}_n - \mu)) = \sigma^2$ . If we go back to the definition of  $R_n$ , (116), by the boundedness of the integral and integrand itself, we can interchange the order of integration and then use the fact that  $\mathbb{E}[F_n(x)] = F(x)$  for any  $x$

$$\mathbb{E}[R_n] = 2 \int_0^1 \mathbb{E}[F_n(\mu + (\bar{X}_n - \mu)y) - F_n(\mu) - F_X(\mu + (\bar{X}_n - \mu)y) + F_X(\mu)]dy = 0. \quad (125)$$

By analogous arguments, we get

$$\begin{aligned} \mathbb{E}[R_n^2] &= \mathbb{E}\left[\left(2 \int_0^1 [F_n(\mu + (\bar{X}_n - \mu)y) - F_n(\mu) - F_X(\mu + (\bar{X}_n - \mu)y) + F_X(\mu)]dy\right)^2\right] \\ &= 4 \int_0^1 \int_0^1 (\text{Var}(F_n(\mu)) - \text{Cov}(F_n(\mu), F_n(\mu + (\bar{X}_n - \mu)y)) - \text{Cov}(F_n(\mu), F_n(\mu + (\bar{X}_n - \mu)x)) \\ &\quad + \text{Cov}(F_n(\mu + (\bar{X}_n - \mu)x), F_n(\mu + (\bar{X}_n - \mu)y))) dx dy. \end{aligned}$$

As we know that the variance of the empirical distribution function is of order  $n^{-1}$ , using again the Cauchy-Schwarz inequality, gives  $\lim_{n \rightarrow \infty} \mathbb{E}[R_n^2] = 0$  Again by Cauchy-Schwarz,

$$\lim_{n \rightarrow \infty} |\mathbb{E}[\sqrt{n}(\bar{X}_n - \mu)R_n]| \leq \lim_{n \rightarrow \infty} \sqrt{\mathbb{E}[R_n^2]} \sqrt{\mathbb{E}[(\sqrt{n}(\bar{X}_n - \mu))^2]} = 0,$$

thus,  $\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n}(\bar{X}_n - \mu)R_n) = \lim_{n \rightarrow \infty} \mathbb{E}[(\sqrt{n}(\bar{X}_n - \mu)R_n)^2]$ . To evaluate this limit, we use the boundedness of  $R_n$ , i.e.

$$\begin{aligned} |R_n| &\leq 2 \int_0^1 |F_n(\mu + (\bar{X}_n - \mu)y) - F_n(\mu) - F(\mu + (\bar{X}_n - \mu)y) + F(\mu)| dy \\ &\leq 2 \int_0^1 (|F_n(\mu + (\bar{X}_n - \mu)y)| + |F_n(\mu)| + |F(\mu + (\bar{X}_n - \mu)y)| + |F(\mu)|) dy \leq 8. \end{aligned}$$

For notational simplicity we write  $X = \sqrt{n}(\bar{X}_n - \mu)$ ,  $Y = R_n$  and denote by  $f$  the joint density of  $(X, Y)$ :

$$\begin{aligned} \mathbb{E}[X^2 Y^2] &= \int_{\mathbb{R}^2} x^2 y^2 f(x, y) d(x, y) = \int_{-\log(n)}^{+\log(n)} \int_{y \notin [-k \frac{\log(n)^{(1+\beta)/2}}{n^{(2+\beta)/4}}, k \frac{\log(n)^{(1+\beta)/2}}{n^{(2+\beta)/4}}]} x^2 y^2 f(x, y) d(x, y) \\ &\quad + \int_{-\log(n)}^{+\log(n)} \int_{y \in [-k \frac{\log(n)^{(1+\beta)/2}}{n^{(2+\beta)/4}}, k \frac{\log(n)^{(1+\beta)/2}}{n^{(2+\beta)/4}}]} x^2 y^2 f(x, y) d(x, y) + \int_{x \notin [-\log(n), \log(n)]} \int_{\mathbb{R}} x^2 y^2 f(x, y) d(x, y). \end{aligned} \quad (126)$$

Using the bounds of the integral with respect to  $x$  and the boundedness of  $Y = R_n$ , we can write

$$\begin{aligned} &\int_{-\log(n)}^{+\log(n)} \int_{y \notin [-k \frac{\log(n)^{(1+\beta)/2}}{n^{(2+\beta)/4}}, k \frac{\log(n)^{(1+\beta)/2}}{n^{(2+\beta)/4}}]} x^2 y^2 f(x, y) d(x, y) \\ &\leq 64(2 \log(n)) \log(n)^2 \int_{-\log(n)}^{+\log(n)} \int_{y \notin [-k \frac{\log(n)^{(1+\beta)/2}}{n^{(2+\beta)/4}}, k \frac{\log(n)^{(1+\beta)/2}}{n^{(2+\beta)/4}}]} f(x, y) d(x, y) \\ &= (128 \log(n))^3 \mathbb{P}\left(|\bar{X}_n - \mu| \leq n^{-1/2} \log n, |R_n| > k(\log n)^{(1+\beta)/2} n^{-(2+\beta)/4}\right) = (128 \log(n))^3 O(n^{-2}) \rightarrow 0 \end{aligned}$$

where the last equality followed from (118). For the other two double-integrals in (126), we notice for each of them (also, because by definition  $0 < \beta \leq 1$ ), that, for  $n \rightarrow \infty$ , one of the integral limits collapses, thus the integral itself too. Hence, (123) holds.



Let us look at the second quantity, (124). We proceed in the same way, showing by Cauchy-Schwarz that

$\lim_{n \rightarrow \infty} |\mathbb{E}[\sqrt{n}2(\bar{X}_n - \mu) \int_0^1 [F_X(\mu + (\bar{X}_n - \mu)y) - F_X(\mu)]dy]| = 0$  such that the variance is the expectation of the square of the given quantity.

Hence, we first look at the integral part and use the Hölder continuity (with a constant  $c > 0$ ):

$$\begin{aligned} \int_0^1 [F_X(\mu + (\bar{X}_n - \mu)y) - F_X(\mu)]dy &\leq \int_0^1 |F_X(\mu + (\bar{X}_n - \mu)y) - F_X(\mu)|dy \leq \int_0^1 c|\mu + (\bar{X}_n - \mu)y - \mu|dy \\ &= \frac{c}{2}|\bar{X}_n - \mu|, \end{aligned} \quad (127)$$

from which we deduce that

$$\mathbb{E} \left[ \left( \int_0^1 [F_X(\mu + (\bar{X}_n - \mu)y) - F_X(\mu)]dy \right)^2 \right] \leq \mathbb{E}[c^2(\bar{X}_n - \mu)^2/4] = \frac{c^2}{4n} \text{Var}(X_1) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Thus,  $\lim_{n \rightarrow \infty} \left| \mathbb{E} \left[ \sqrt{n}2(\bar{X}_n - \mu) \int_0^1 [F_X(\mu + (\bar{X}_n - \mu)y) - F_X(\mu)]dy \right] \right| = 0$  follows by Cauchy-Schwarz, as desired. We are left with showing that the variance, which is expressed as the expectation of the square of the given quantity, vanishes:

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \left( \sqrt{n}2(\bar{X}_n - \mu) \int_0^1 [F_X(\mu + (\bar{X}_n - \mu)y) - F_X(\mu)]dy \right)^2 \right] = 0.$$

Using that  $|\bar{X}_n - \mu| \leq \frac{\log n}{n}$  a.s. for large  $n$  (see Lemma A.4, [4]), we have for  $n \geq n_0$ :

$$\mathbb{E} \left[ \left( \sqrt{n}2(\bar{X}_n - \mu) \right)^2 \left( \int_0^1 [F_X(\mu + (\bar{X}_n - \mu)y) - F_X(\mu)]dy \right)^2 \right] \leq \mathbb{E} \left[ \left( \sqrt{n}2(\bar{X}_n - \mu) \right)^2 \frac{c^2}{4} \left( \frac{\log n}{n} \right)^2 \right]$$

and hence, in the limit

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \left( \sqrt{n}2(\bar{X}_n - \mu) \right)^2 \left( \int_0^1 [F_X(\mu + (\bar{X}_n - \mu)y) - F_X(\mu)]dy \right)^2 \right] \leq \lim_{n \rightarrow \infty} \left( \frac{c \log n}{n} \right)^2 \text{Var}(\sqrt{n}(\bar{X}_n - \mu)) = 0.$$

Putting both cases together, we arrive at

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_n(p), \sqrt{n}\hat{m}(X, n, r)) = \frac{\tau_r(|X - \mu|, p) + (2 - r)(2F_X(\mu) - 1)\tau_1(p)}{f_X(q_X(p))}, \quad (128)$$

$$\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{m}(X, n, r)) = \frac{\tau_r(|X - \mu|, p) + (2 - r)(2F_X(\mu) - 1)\tau_1(p)}{\sqrt{p(1 - p)}\sqrt{\text{Var}(|X - \mu|^r + (2 - r)(2F_X(\mu) - 1)X)}}. \quad (129)$$

Getting the expressions involving the functionals  $h_1, h_2$  is then only application of the Delta method.  $\square$

**Proof of Proposition 3.** Besides the Bahadur representation of the sample quantile, (109), we also use a Bahadur representation for the sample MedianAD, shown in equation (7) of [31]:

$$\hat{\xi}_n - \xi = \frac{1/2 - (F_n(\nu + \xi) - F_n(\nu - \xi))}{f_X(\nu + \xi) + f_X(\nu - \xi)} - \frac{f_X(\nu + \xi) - f_X(\nu - \xi)}{f_X(\nu + \xi) + f_X(\nu - \xi)} \frac{1/2 - F_n(\nu)}{f_X(\nu)} + \Delta_n \quad (130)$$

where  $\Delta_n = O(n^{-3/4} \log(n)^{3/4})$  almost surely. Clearly, (130) can be rewritten in terms of an iid sum as

$$\hat{\xi}_n - \xi = \frac{\frac{1}{n} \sum_{i=1}^n (\alpha \mathbf{I}_{(x \leq \nu)} - f_X(\nu) \mathbf{I}_{(\nu - \xi < x \leq \nu + \xi)}) - \frac{1}{2}(\alpha - f_X(\nu))}{\beta f_X(\nu)} + \Delta_n, \quad (131)$$

where, for notational simplification,  $\alpha := f_X(\nu + \xi) - f_X(\nu - \xi)$  and  $\beta := f_X(\nu + \xi) + f_X(\nu - \xi)$  respectively.

Using equations (109), (131), and the fact that, by definition of  $\nu$  and  $\xi$ ,  $\mathbb{P}(X \leq \nu) = F_X(\nu) = 1/2$  and  $\mathbb{P}(\nu - \xi < X \leq \nu + \xi) = F_{|X-\nu|}(\xi) = 1/2$ , we apply the bivariate CLT and obtain:

$$\begin{aligned} & n^{-1/2} \sum_{i=1}^n \left( \left( \begin{array}{c} \mathbf{1}_{(X_i > q_X(p))} \\ \alpha \mathbf{1}_{(X_i \leq \nu)} - f_X(\nu) \mathbf{1}_{(\nu - \xi < X_i \leq \nu + \xi)} \end{array} \right) - \left( \begin{array}{c} 1-p \\ 1/2(\alpha - f_X(\nu)) \end{array} \right) \right) \\ &= n^{1/2} \left( \left( \begin{array}{c} 1 - F_n(q_X(p)) \\ \frac{1}{n} \sum (\alpha \mathbf{1}_{(X \leq \nu)} - f_X(\nu) \mathbf{1}_{(\nu - \xi < X \leq \nu + \xi)}) \end{array} \right) - \left( \begin{array}{c} 1-p \\ 1/2(\alpha - f_X(\nu)) \end{array} \right) \right) \xrightarrow{d} \mathcal{N}(0, \tilde{\Gamma}), \end{aligned} \quad (132)$$

where  $\tilde{\Gamma} = \begin{pmatrix} p(1-p) & cov_{ind:q_n, \hat{\xi}_n} \\ cov_{ind:q_n, \hat{\xi}_n} & \text{Var}(\alpha \mathbf{1}_{(X_i \leq \nu)} - f_X(\nu) \mathbf{1}_{(\nu - \xi < X_i \leq \nu + \xi)}) \end{pmatrix}$

with  $cov_{ind:q_n, \hat{\xi}_n} := \alpha \max(0, p - 1/2) - f_X(\nu) (\max(0, F_X(\nu + \xi) - \max(F_X(\nu - \xi), p)) - (1-p)/2)$ , as we are going to prove below.

Then, we need to pre-multiply (i.e. from the left side) equation (132) by  $\begin{bmatrix} 1/(f_X(q_X(p))) & 0 \\ 0 & 1/(\beta f_X(\nu)) \end{bmatrix}$  to use the Bahadur representation of the sample quantile and of the sample MedianAD (recall (109), (131)). We obtain:

$$n^{1/2} \left( \begin{array}{c} \frac{1 - F_n(q_X(p)) - (1-p)}{f_X(q_X(p))} \\ \frac{\frac{1}{n} \sum (\alpha \mathbf{1}_{(X \leq \nu)} - f_X(\nu) \mathbf{1}_{(\nu - \xi < X \leq \nu + \xi)}) - 1/2(\alpha - f_X(\nu))}{\beta f_X(\nu)} \end{array} \right) = n^{1/2} \left( \begin{array}{c} q_n(p) - q_X(p) - R_{n,p} \\ \hat{\xi}_n - \xi - \Delta_n \end{array} \right) \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Gamma) \quad (133)$$

where, ignoring  $R_{n,p}$  and  $\Delta_n$  since they are of order  $O(n^{-3/4} \log(n))$  a.s. (same argumentation as for  $R_{n,p}$  in the Proof of Theorem 1, part 1),

$$\Gamma = \begin{pmatrix} \frac{p(1-p)}{f_X^2(q_X(p))} & \frac{cov_{ind:q_n, \hat{\xi}_n}}{\beta f_X(\nu) f_X(q_X(p))} \\ \frac{cov_{ind:q_n, \hat{\xi}_n}}{\beta f_X(\nu) f_X(q_X(p))} & \frac{\text{Var}(\alpha \mathbf{1}_{(X_i \leq \nu)} - f_X(\nu) \mathbf{1}_{(\nu - \xi < X_i \leq \nu + \xi)})}{\beta^2 f_X^2(\nu)} \end{pmatrix}.$$

We are left with computing the covariance  $cov_{ind:q_n, \hat{\xi}_n}$  and the following variance:

$$\begin{aligned} & \text{Var}(\alpha \mathbf{1}_{(X_i \leq \nu)} - f_X(\nu) \mathbf{1}_{(\nu - \xi < X_i \leq \nu + \xi)}) \\ &= \alpha^2 \text{Var}(\mathbf{1}_{(X_i \leq \nu)}) + f_X^2(\nu) \text{Var}(\mathbf{1}_{(\nu - \xi < X_i \leq \nu + \xi)}) + 2\alpha f_X(\nu) \text{Cov}(\mathbf{1}_{(X_i \leq \nu)}, -\mathbf{1}_{(\nu - \xi < X_i \leq \nu + \xi)}) \\ &= \frac{1}{4} \left( \alpha^2 + f_X^2(\nu) - 8\alpha f_X(\nu) (\mathbb{E}[\mathbf{1}_{(\nu - \xi < X_i \leq \nu + \xi)}] - \frac{1}{4}) \right) \\ &= \frac{1}{4} \left( \alpha^2 + f_X^2(\nu) - 4\alpha f_X(\nu) \left( \frac{1}{2} - 2F_X(\nu - \xi) \right) \right) = \frac{1}{4} (f_X^2(\nu) + \gamma) \end{aligned}$$

where  $\gamma := \alpha^2 - 4\alpha f_X(\nu)(1 - F_X(\nu - \xi) - F_X(\nu + \xi))$ . Let us turn to the computation of  $cov_{ind:q_n, \hat{\xi}_n}$ :

$$\begin{aligned} & \text{Cov}(\mathbf{1}_{(X_i > q_X(p))}, \alpha \mathbf{1}_{(X_i \leq \nu)} - f_X(\nu) \mathbf{1}_{(\nu - \xi < X_i \leq \nu + \xi)}) = \\ & \alpha \mathbb{E}[\mathbf{1}_{(X_i > q_X(p))} \mathbf{1}_{(X_i \leq \nu)}] - f_X(\nu) \mathbb{E}[\mathbf{1}_{(X_i > q_X(p))} \mathbf{1}_{(\nu - \xi < X_i \leq \nu + \xi)}] - (1-p) \frac{1}{2} (\alpha - f_X(\nu)). \end{aligned} \quad (134)$$

Let us consider one after the other the two expectations in (134). Note that we can write (using the definition of  $\nu$ )

$$\mathbf{1}_{(X_i > q_X(p))} \mathbf{1}_{(X_i \leq \nu)} = \begin{cases} 0 & \text{if } \nu \leq q_X(p) \quad (\Leftrightarrow p \geq 1/2) \\ \mathbf{1}_{(q_X(p) < X_i \leq \nu)} & \text{if } \nu > q_X(p) \quad (\Leftrightarrow p < 1/2) \end{cases},$$

from which we deduce  $\mathbb{E}[\mathbf{1}_{(X_i > q_X(p))} \mathbf{1}_{(X_i \leq \nu)}] = \max(1/2 - p, 0)$ . Analogously,

$$\mathbf{1}_{(X_i > q_X(p))} \mathbf{1}_{(\nu - \xi < X_i \leq \nu + \xi)} = \begin{cases} 0 & \text{if } q_X(p) > \nu + \xi \quad (\Leftrightarrow p > F_X(\nu + \xi)), \\ \mathbf{1}_{(q_X(p) < X_i \leq \nu + \xi)} & \text{if } \nu - \xi \leq q_X(p) \leq \nu + \xi \quad (\Leftrightarrow F_X(\nu - \xi) \leq p \leq F_X(\nu + \xi)), \\ \mathbf{1}_{(\nu - \xi < X_i \leq \nu + \xi)} & \text{if } q_X(p) < \nu - \xi \quad (\Leftrightarrow p < F_X(\nu - \xi)). \end{cases}$$

Thus we have  $\mathbb{E}[\mathbf{1}_{(X_i > q_X(p))} \mathbf{1}_{(\nu - \xi < X_i \leq \nu + \xi)}] = \max(0, F_X(\nu + \xi) - \max(F_X(\nu - \xi), p))$ .

Combining these two expressions in (134) provides:

$$\begin{aligned} cov_{ind:q_n, \hat{\xi}_n} &= \alpha \max(1/2 - p, 0) - f_X(\nu) \max(0, F_X(\nu + \xi) - \max(F_X(\nu - \xi), p)) - (1-p) \frac{1}{2} (\alpha - f_X(\nu)) \\ &= \alpha \max(-p/2, (p-1)/2) - f_X(\nu) (\max(0, F_X(\nu + \xi) - \max(F_X(\nu - \xi), p)) - (1-p)/2). \end{aligned}$$

This concludes the computations. Nevertheless, to be explicit, let us write out the overall asymptotic covariance and correlation:

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_n(p), \sqrt{n}\hat{\xi}_n) &= \frac{\text{cov}_{\text{ind}:q_n, \hat{\xi}_n}}{\beta f_X(\nu) f_X(q_X(p))} \\ &= \frac{-\max(0, F_X(\nu + \xi) - \max(F_X(\nu - \xi), p)) + \frac{1-p}{2} + \frac{f_X(\nu + \xi) - f_X(\nu - \xi)}{f_X(\nu)} \max(p - \frac{1}{2}, 0)}{(f_X(\nu + \xi) + f_X(\nu - \xi)) f_X(q_X(p))} \end{aligned}$$

which is exactly the covariance in (10) for the case  $h_1(x) = h_2(x) = x$  (the case with general functionals  $h_1, h_2$  follows directly by the application of the Delta method), whereas the correlation is as in (11):

$$\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\xi}_n) = \frac{-\max(0, F_X(\nu + \xi) - \max(F_X(\nu - \xi), p)) + \frac{1-p}{2} + \frac{f_X(\nu + \xi) - f_X(\nu - \xi)}{f_X(\nu)} \max(-\frac{p}{2}, \frac{p-1}{2})}{\sqrt{\frac{p(1-p)}{4}} \sqrt{1 + \frac{\gamma}{f_X^2(\nu)}}}.$$

As expected, the above computed asymptotic variance of the sample MedianAD, i.e.

$$\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n}\hat{\xi}_n) = \frac{1 + \gamma/f_X^2(\nu)}{4(f_X(\nu + \xi) + f_X(\nu - \xi))^2}, \quad (135)$$

exactly equals the variance of the sample MedianAD as in equation (11) of [40] (while in [31] they seem to have some typos in their definition of their quantity  $\gamma$  such that one does not get the same result).  $\square$

## A.1.2 Taylor's Method

As mentioned, we offer an additional proof for Theorem 1 which is based on a Taylor expansion and extends the ideas of [17]. The proof consists of two parts. In the first and main part, we show the Taylor expansion and asymptotic normality in the case of estimating the measures of dispersion with known mean  $\mu$  for any integer valued  $r$  (in analogy to the first part in the proof of Theorem 1). The second part consists of extending the previous result to the case where we estimate the measures of dispersion in the case of an unknown mean  $\mu$  for  $r = 1, 2$  and is identical to Part 2 in the proof of Theorem 1. Hence we focus here only on the case where  $\mu$  is known.

We start showing the asymptotic normality in the case of estimating the measure of dispersion by  $\tilde{m}(X, n, r) = \frac{1}{n} \sum_{i=1}^n |X_i - \mu|^r$ . This is done in three steps. The first step is to use a representation such that our quantities of interest, the sample quantile and the measure of dispersion estimator, are functions of the uniform order statistics. Then, we use the Taylor expansion to show the asymptotic normality of each of the estimators. Finally, in a third step we compute the covariance (and then the correlation) between the measure of dispersion estimator and the sample quantile.

### Step 1: Functions of the uniform order statistics

Recall that for a standard exponentially distributed iid sample of size  $n + 1$ ,  $Z_1, \dots, Z_{n+1}$ , we have

$$U_j := \frac{\sum_{i=1}^j Z_i}{\sum_{k=1}^{n+1} Z_k}, j = 1, \dots, n$$

and  $(U_1, \dots, U_n)$  has the same distribution as the order statistics from a sample of size  $n$  from a standard uniform distribution (see e.g. [36]). This allows us to express the sample quantile  $q_n(p)$  and the sample measure of dispersion  $\tilde{m}(X, n, r)$  as follows:

$$q_n(p) = X_{(\lceil np \rceil)} = q_X(U_{\lceil np \rceil}) \quad (136)$$

$$\tilde{m}(X, n, r) = \frac{1}{n} \sum_{i=1}^n |q_X(U_i) - \mu|^r \quad (137)$$

### Step 2: Taylor expansions

Using this, we can proceed with the Taylor expansion. We start with  $\tilde{m}(X, n, r)$  and expand for each  $q_X(U_i)$  in

$$(137) \text{ around } i/(n+1), i = 1, \dots, n: \tilde{m}(X, n, r) = \frac{1}{n} \sum_{i=1}^n \left( \left| q_X \left( \frac{i}{n+1} \right) - \mu \right|^r + r \left| q_X \left( \frac{i}{n+1} \right) - \mu \right|^{r-1} q'_X \left( \frac{i}{n+1} \right) \operatorname{sgn} \left( q_X \left( \frac{i}{n+1} \right) - \mu \right) \left( U_i - \frac{i}{n+1} \right) + O(n^{-2}) \right)$$

The terms of order  $n^{-2}$  are negligible in the asymptotic analysis (i.e. will vanish asymptotically).

Then, in analogy to  $\mu_n$  in [17], we define  $\mu_n(X, r) := \frac{1}{n} \sum_{i=1}^n \left| q_X \left( \frac{i}{n+1} \right) - \mu \right|^r$ . We can interpret it as the right

$$\text{Riemann sum: } \mu_n(X, r) = \frac{n+1}{n} \times \frac{1}{n+1} \sum_{i=1}^n \left| q_X \left( \frac{i}{n+1} \right) - \mu \right|^r \xrightarrow{n \rightarrow \infty} \int_0^1 |q_X(t) - \mu|^r dt.$$

Using the transformation  $t = F_X(x)$  (and  $dt = dF_X(x)$ ), we obtain:

$$\int_0^1 |q_X(t) - \mu|^r dt = \int_{-\infty}^{+\infty} |q_X(F_X(x)) - \mu|^r dF_X(x) = \int_{-\infty}^{+\infty} |x - \mu|^r dF_X(x) = m(X, r),$$

from which we conclude that  $\lim_{n \rightarrow \infty} \mu_n(X, r) = m(X, r)$ .

Also, by the order of the error term of the right Riemann sum approximation,  $O(n^{-1})$ , we know that  $\lim_{n \rightarrow \infty} \sqrt{n} (\mu_n(X, r) - m(X, r)) = 0$ . Hence,  $m(X, r)$  can be replaced by  $\mu_n(X, r)$ , even in asymptotics when multiplied by  $\sqrt{n}$ , and we can write (with the notation  $a_n \underset{n \rightarrow \infty}{\sim} b_n$  whenever  $\lim_{n \rightarrow \infty} a_n/b_n = 1$ ):

$$\sqrt{n}(\tilde{m}(X, n, r) - \mu_n(X, r)) \underset{n \rightarrow \infty}{\sim} \sqrt{n} \left( \frac{1}{n} \sum_{i=1}^n r \left| q_X \left( \frac{i}{n+1} \right) - \mu \right|^{r-1} q'_X \left( \frac{i}{n+1} \right) \operatorname{sgn} \left( q_X \left( \frac{i}{n+1} \right) - \mu \right) \left( U_i - \frac{i}{n+1} \right) \right).$$

We can then conclude the following convergence in distribution, by using the asymptotics calculated in [17] (see eq. (12),(14) and (16) therein),

$$\sqrt{n}(\tilde{m}(X, n, r) - \mu_n(X, r)) \xrightarrow{d} \int_0^1 r |q_X(t) - \mu|^{r-1} q'_X(t) \operatorname{sgn}(q_X(t) - \mu) B(t) dt, \quad (138)$$

where  $B(t) := W(t) - tW(1)$  is the Brownian bridge,  $W$  denoting the standard Wiener process. Hence the asymptotic normality of the measure of dispersion.

For the sample quantile, we can copy the result from [17] one to one. We expand the sample quantile  $q_n(p) = q_X(U_{\lceil np \rceil})$  around  $p$ :  $q_n(p) = q_X(p) + q'_X(p)(U_{\lceil np \rceil} - p) + O(n^{-2})$ . As before, those higher order terms are asymptotically negligible.

Thus, we get, following equations (11),(13) and (15) of [17],

$$\sqrt{n}(q_n - q_X(p)) \underset{n \rightarrow \infty}{\sim} q'_X(p) \left( \frac{\sum_{j=1}^{\lceil np \rceil} Z_j}{\sum_{k=1}^{n+1} Z_k} - p \right) \xrightarrow[n \rightarrow \infty]{d} q'_X(p) B(p) \quad (139)$$

Hence the asymptotic normality of the sample quantile.

We can now conclude the normal joint distribution by using the Cramer-Wold device (the increments of the Brownian motion being independent and normally distributed).

**Step 3: Asymptotic Covariance and Correlation.** We have, using the first two moments of the Brownian bridge,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \operatorname{Cov} \left( \sqrt{n}(q_n(p) - q_X(p)), \sqrt{n}(\tilde{m}(X, n, r) - \mu_n(X, r)) \right) = \\ & \operatorname{Cov} \left( q'_X(p) B(p), \int_0^1 r |q_X(t) - \mu|^{r-1} q'_X(t) \operatorname{sgn}(q_X(t) - \mu) B(t) dt \right) \\ & = q'_X(p) \int_0^1 r |q_X(t) - \mu|^{r-1} q'_X(t) \operatorname{sgn}(q_X(t) - \mu) \mathbb{E}[B(p)B(t)] dt \\ & = q'_X(p) \int_0^1 r |q_X(t) - \mu|^{r-1} q'_X(t) \operatorname{sgn}(q_X(t) - \mu) q'_X(t) (\min(p, t) - pt) dt. \end{aligned}$$

Hence, we are left with computing the integral:

$$\begin{aligned}
& \int_0^1 r|q_X(t) - \mu|^{r-1} q'_X(t) \operatorname{sgn}(q_X(t) - \mu) q'_X(t) (\min(p, t) - pt) dt \\
&= \int_0^p r|q_X(t) - \mu|^{r-1} q'_X(t) \operatorname{sgn}(q_X(t) - \mu) q'_X(t) t(1-p) dt + \int_p^1 r|q_X(t) - \mu|^{r-1} q'_X(t) \operatorname{sgn}(q_X(t) - \mu) q'_X(t) p(1-t) dt \\
&= (1-p) \left( (|q_X(t) - \mu|^r)_0^p - \int_0^p |q_X(t) - \mu|^r dt \right) + p \left( (|q_X(t) - \mu|^r(1-t))_p^1 + \int_p^1 |q_X(t) - \mu|^r dt \right) \\
&= p \int_p^1 |q_X(t) - \mu|^r dt - (1-p) \int_0^p |q_X(t) - \mu|^r dt = p \int_{q_X(p)}^\infty |x - \mu|^r dF_X(x) - (1-p) \int_{-\infty}^{q_X(p)} |x - \mu|^r dF_X(x)
\end{aligned}$$

using partial integration for each integral (with  $u' = r|q_X(t) - \mu|^{r-1} q'_X(t) \operatorname{sgn}(q_X(t) - \mu) q'_X(t)$ , i.e.  $u = |q_X(t) - \mu|^r$  and  $v$  being  $t$  or  $1-t$  respectively) for the second equality, and  $t = F_X(x)$  ( $dt = dF_X(x)$ ) in the last one. Thus, we have overall, recalling the definition of  $\tau_r$  in (4),

$$\lim_{n \rightarrow \infty} \operatorname{Cov}(\sqrt{n}(q_n(p) - q_X(p)), \sqrt{n}(\tilde{m}(X, n, r) - m(X, r))) = q'_X(p) \tau_r(|X - \mu|, p) = \frac{1}{f_X(q_X(p))} \tau_r(|X - \mu|, p),$$

from which we can deduce the asymptotic correlation, namely

$$\lim_{n \rightarrow \infty} \operatorname{Cor}(q_n(p), \tilde{m}(X, n, r)) = \frac{\tau_r(|X - \mu|, p)}{\sqrt{p(1-p)} \sqrt{\operatorname{Var}(|X - \mu|^r)}}.$$

## A.2 Proofs of Subsection 2.2

In the following we present the analogous proofs to Appendix A.1 but with the location-scale quantile estimator. The main task is to compute the asymptotic covariances of the respective joint asymptotic distributions.

**Proof of Proposition 4.** Let us recall two relations we can deduce about  $\hat{m}(X, n, r)$  from the proof of Theorem 1:

$$\hat{m}(X, n, 2) := \hat{\sigma}_n^2 = \tilde{\sigma}_n^2 + o_P(1) =: \tilde{m}(X, n, 2) + o_P(1) \text{ (obtained from equations (114), and (115))}$$

$$\hat{m}(X, n, 1) := \hat{\theta}_n = \tilde{\theta}_n + (2F_X(\mu) - \mu)(\bar{X}_n - \mu) + o_P(1) =: \tilde{m}(X, n, 1) + (2F_X(\mu) - \mu)(\bar{X}_n - \mu) + o_P(1)$$

(from equation (120) and what was shown in the following),

which can be rewritten, for any  $r = 1, 2$ , as:

$$\hat{m}(X, n, r) = \tilde{m}(X, n, r) + (2-r)(2F_X(\mu) - 1)(\bar{X}_n - \mu) + o_P(1);$$

whereas the location-scale quantile, as:

$$q_{n, \hat{\mu}, \hat{\sigma}}(p) = \bar{X}_n + q_Y(p) \tilde{\sigma}_n + o_P(1).$$

Since we have iid sums (and finite fourth moment by assumption), we can apply the bivariate CLT:

$$\begin{aligned}
n^{1/2} \left( \begin{pmatrix} q_{n, \hat{\mu}, \hat{\sigma}}(p) \\ \hat{m}(X, n, r) \end{pmatrix} - \begin{pmatrix} q_X(p) \\ m(X, r) \end{pmatrix} \right) &= n^{1/2} \left( \begin{pmatrix} \bar{X}_n + q_Y(p) \tilde{\sigma}_n + o_P(1) \\ \tilde{m}(X, n, r) + (2F_X(\mu) - 1)(\bar{X}_n - \mu) + o_P(1) \end{pmatrix} - \begin{pmatrix} q_X(p) \\ m(X, r) \end{pmatrix} \right) \\
&\xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Lambda^{(r)})
\end{aligned} \tag{140}$$

where the covariance matrix  $\Lambda^{(r)} = (\Lambda_{ij}^{(r)})$ ,  $i, j = 1, 2$  has to be determined. The component  $\Lambda_{22}^{(r)}$  is already known from equation (129) in the Proof of Theorem 1:

$$\Lambda_{22}^{(r)} = \lim_{n \rightarrow \infty} \operatorname{Var}(\hat{m}(X, n, r)) = \sigma^{2r} \operatorname{Var}(|Y|^r + (2-r)(2F_Y(0) - 1)Y).$$

Let us compute the other components directly. We have

$$\begin{aligned}\Lambda_{11}^{(r)} &= \lim_{n \rightarrow \infty} \text{Var}(q_{n,\hat{\mu},\hat{\sigma}}(p)) = \lim_{n \rightarrow \infty} \text{Var}(\bar{X}_n + q_Y(p)\hat{\sigma}_n) = \lim_{n \rightarrow \infty} (\text{Var}(\bar{X}_n) + q_Y^2(p) \text{Var}(\hat{\sigma}_n) + 2q_Y(p) \text{Cov}(\bar{X}_n, \hat{\sigma}_n)) \\ &= \sigma^2 + q_Y^2(p) \frac{\mu_4 - \sigma^4}{(2\sigma)^2} + 2q_Y(p) \frac{\mu_3}{2\sigma} = \sigma^2 \left( 1 + q_Y^2(p) \frac{\mathbb{E}[Y^4] - 1}{4} + q_Y(p) \mathbb{E}[Y^3] \right)\end{aligned}\quad (141)$$

where we used the Delta-method to derive the variance and covariance in the case of  $\hat{\sigma}_n$ , from  $\text{Var}(\hat{\sigma}_n^2) = \mu_4 - \sigma^4$  and  $\text{Cov}(\bar{X}_n, \hat{\sigma}_n^2) = \mu_3$ ,

$$\begin{aligned}\text{and } \Lambda_{12}^{(r)} &= \Lambda_{21}^{(r)} = \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_{n,\hat{\mu},\hat{\sigma}}(p), \sqrt{n}\hat{m}(X, n, r)) \\ &= \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}(\bar{X}_n + q_Y(p)\hat{\sigma}_n), \sqrt{n}(\tilde{m}(X, n, r) + (2-r)(2F_X(\mu) - 1)(\bar{X}_n - \mu))) \\ &= \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}(\bar{X}_n), \sqrt{n}(\tilde{m}(X, n, r) + (2-r)(2F_X(\mu) - 1)(\bar{X}_n - \mu))) \\ &\quad + q_Y(p) \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\hat{\sigma}_n, \sqrt{n}(\tilde{m}(X, n, r) + (2-r)(2F_X(\mu) - 1)(\bar{X}_n - \mu))).\end{aligned}\quad (142)$$

We do this in two steps, considering separately the asymptotic covariance with the sample mean (first covariance term in (142)) and that with the sample standard deviation (second covariance term of (142)).

In both steps we proceed with the same techniques as in the proof of Theorem 1, when using the bivariate central limit theorem. This means, we compute the covariances by looking at the  $i$ -th element of the iid sums:

$$\text{Cov}(X_i, |X_i - \mu|^r + (2-r)(2F_X(\mu) - 1)(X_i - \mu)) \text{ and } \text{Cov}((X_i - \mu)^2, |X_i - \mu|^r + (2-r)(2F_X(\mu) - 1)(X_i - \mu)).$$

*Step 1: Covariance with the sample mean*

$$\begin{aligned}\text{Cov}(X_i, |X_i - \mu|^r + (2-r)(2F_X(\mu) - 1)(X_i - \mu)) &= \mathbb{E}[X_i |X_i - \mu|^r] - \mu \mathbb{E}[|X_i - \mu|^r] + (2-r)(2F_X(\mu) - 1) \text{Var}(X_i) \\ &= \mathbb{E}[(X_i - \mu) |X_i - \mu|^r] + (2-r)(2F_Y(0) - 1) \sigma^2 \\ &= \mathbb{E}[|X_i - \mu|^{r+1} \mathbf{1}_{(X > \mu)}] - \mathbb{E}[|X_i - \mu|^{r+1} \mathbf{1}_{(X < \mu)}] + (2-r)(2F_Y(0) - 1) \sigma^2 \\ &= \sigma^{r+1} \mathbb{E}[|Y_i|^{r+1} \mathbf{1}_{(Y_i > 0)}] - \sigma^{r+1} \mathbb{E}[|Y_i|^{r+1} \mathbf{1}_{(Y_i < 0)}] + (2-r)(2F_Y(0) - 1) \sigma^2 \\ &= \sigma^{r+1} \mathbb{E}[Y_i^{r+1}] - \sigma^{r+1} \mathbb{E}[Y_i^{r+1} \mathbf{1}_{(Y_i < 0)}(1 + (-1)^{r+1})] + (2-r)(2F_Y(0) - 1) \sigma^{r+1} \\ &= \sigma^{r+1} \mathbb{E}[Y_i^{r+1}] - \sigma^{r+1} 2(2-r) \mathbb{E}[Y_i^{r+1} \mathbf{1}_{(Y_i < 0)}] + (2-r)(2F_Y(0) - 1) \sigma^{r+1} \\ &= \sigma^{r+1} (\mathbb{E}[Y_i^{r+1}] + (2-r)(2F_Y(0) - 1 - 2 \mathbb{E}[Y_i^{r+1} \mathbf{1}_{(Y_i < 0)}]))\end{aligned}$$

where, for the transformation in the last three lines, we need to recall that, since we only consider  $r = 1, 2$ , we can write  $(2-r)\sigma^2 = (2-r)\sigma^{r+1}$  and  $(1 + (-1)^{r+1}) = 2(2-r)$ .

*Step 2: Covariance with the sample standard deviation*

Considering the covariance with the sample variance, we can write

$$\begin{aligned}\text{Cov}((X_i - \mu)^2, |X_i - \mu|^r + (2-r)(2F_X(\mu) - 1)(X_i - \mu)) &= \mathbb{E}[|X_i - \mu|^{r+2}] - \sigma^2 \mathbb{E}[|X_i - \mu|^r] + (2-r)(2F_X(\mu) - 1) \mathbb{E}[(X_i - \mu)^3] \\ &= \sigma^{r+2} (\mathbb{E}[|Y_i|^{r+2}] - \mathbb{E}[|Y_i|^r] + (2-r)(2F_Y(0) - 1) \mathbb{E}[Y_i^3])\end{aligned}$$

where we used in the last line the fact that  $(2-r)\sigma^3 = (2-r)\sigma^{r+2}$  for  $r = 1, 2$ .

Thus,  $\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\hat{\sigma}_n^2, \sqrt{n}\hat{m}(X, n, r)) = \sigma^{r+2} (\mathbb{E}[|Y|^{r+2}] - \mathbb{E}[|Y|^r] + (2-r)(2F_Y(0) - 1) \mathbb{E}[Y^3])$ , from which we obtain the covariance with the sample standard deviation, by applying the Delta-method:

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\hat{\sigma}_n, \sqrt{n}\hat{m}(X, n, r)) = \frac{\sigma^{r+1}}{2} (\mathbb{E}[|Y|^{r+2}] - \mathbb{E}[|Y|^r] + (2-r)(2F_Y(0) - 1) \mathbb{E}[Y^3]). \quad (143)$$

Putting together the results of both steps in equation (142) gives  $\Lambda_{12}^{(r)}$ . The asymptotic correlation follows by dividing by the asymptotic variances  $\Lambda_{11}^{(r)}$  and  $\Lambda_{22}^{(r)}$  computed before. The asymptotics involving general functionals  $h_1, h_2$  follows by applying the Delta method. To end, we point out the symmetry of the asymptotic correlation around  $p$  with a value of 0 for  $p = 0.5$  for a symmetric location-scale distribution. For this it is simply enough to consider the numerator of the asymptotic correlation, as the denominator is symmetric around  $p = 0.5$  as in such a case  $\mathbb{E}[Y^3] = 0$ , see (18):

$$\begin{aligned} & \mathbb{E}[Y^{r+1}] + (2-r)(2F_Y(0) - 1 - 2\mathbb{E}[|Y|^{r+1}\mathbf{1}_{(Y<0)}]) + \frac{q_Y(p)}{2} (\mathbb{E}[|Y|^{r+2}] - \mathbb{E}[|Y|^r] + (2-r)(2F_Y(0) - 1) \mathbb{E}[Y^3]) \\ &= \mathbb{E}[Y^{r+1}] + (2-r)(-2\mathbb{E}[|Y|^{r+1}\mathbf{1}_{(Y<0)}]) + \frac{q_Y(p)}{2} (\mathbb{E}[|Y|^{r+2}] - \mathbb{E}[|Y|^r]) \\ &= \begin{cases} \frac{q_Y(p)}{2} (\mathbb{E}[|Y|^3] - \mathbb{E}[|Y|]) & \text{for } r = 1, \\ \frac{q_Y(p)}{2} \text{Var}(Y) & \text{for } r = 2. \end{cases} \end{aligned}$$

Hence, we have shown the symmetry around  $p = 0.5$  and that the correlation equals to zero for  $p = 0.5$ .  $\square$

**Proof of Corollary 5.** Note that this proof is one-to-one as the Proof of Proposition 4, only that we have  $q_{n,\hat{\sigma}}(p)$  instead of  $q_{n,\hat{\mu},\hat{\sigma}}(p)$ . Thus, there is only little change. For the covariance note that, using (143) in the second equality,

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_{n,\hat{\sigma}}(p), \sqrt{n}\hat{m}(X, n, r)) &= q_Y(p) \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\tilde{\sigma}_n, \sqrt{n}\hat{m}(X, n, r)) \\ &= \frac{\sigma^{r+1}}{2} (\mathbb{E}[|Y|^{r+2}] - \mathbb{E}[|Y|^r] + (2-r)(2F_Y(0) - 1) \mathbb{E}[Y^3]), \end{aligned}$$

and for the variances,  $\lim_{n \rightarrow \infty} \text{Var}(q_{n,\hat{\sigma}}(p)) = q_Y^2(p) \lim_{n \rightarrow \infty} \text{Var}(\tilde{\sigma}_n) = \sigma^2 q_Y^2(p) \frac{\mathbb{E}[Y^4] - 1}{4}$ , and  $\lim_{n \rightarrow \infty} \text{Var}(\hat{m}(X, n, r))$  is given in equation (129).  $\square$

**Proof of Proposition 6.** To use the bivariate CLT in this case, recall the Bahadur representation (131) for the sample MedianAD, and the asymptotic equivalence of  $\hat{\sigma}_n^2$  and  $\tilde{\sigma}_n^2$  (see (114) and (115)).

Thus, as we have iid sums (and finite fourth moment by assumption), we can apply the bivariate CLT and obtain:

$$\begin{aligned} n^{1/2} \left( \begin{pmatrix} q_{n,\hat{\mu},\hat{\sigma}}(p) \\ \hat{\xi}_n \end{pmatrix} - \begin{pmatrix} q_X(p) \\ \xi \end{pmatrix} \right) &= n^{1/2} \left( \begin{pmatrix} \bar{X}_n + q_Y(p)\tilde{\sigma} + o_P(1) \\ \frac{\frac{1}{n} \sum_{i=1}^n (\alpha \mathbf{1}_{(x \leq \nu)} - f_X(\nu) \mathbf{1}_{(\nu - \xi < x \leq \nu + \xi)}) - \frac{1}{2}(\alpha - f_X(\nu))}{\beta f_X(\nu)} + \Delta_n \end{pmatrix} - \begin{pmatrix} q_X(p) \\ \xi \end{pmatrix} \right) \\ &\xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Pi), \end{aligned}$$

with  $\alpha = f_X(\nu + \xi) - f_X(\nu - \xi)$ ,  $\beta = f_X(\nu + \xi) + f_X(\nu - \xi)$  and  $\Pi = (\Pi_{ij})$ ,  $i, j = 1, 2$ , the covariance matrix that we compute in the following.

Since  $\Pi_{11}$  and  $\Pi_{22}$  have been already evaluated in (141) and (135), respectively, in the proof of Proposition 3, we are left with computing the asymptotic covariance  $\Pi_{12} = \Pi_{21} = \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_{n,\hat{\mu},\hat{\sigma}}, \sqrt{n}\hat{\xi}_n)$ :

$$\Pi_{12} = \text{Cov} \left( \sqrt{n}(\bar{X}_n + q_Y(p)\tilde{\sigma}_n), \sqrt{n} \frac{\frac{1}{n} \sum_{i=1}^n (\alpha \mathbf{1}_{(x \leq \nu)} - f_X(\nu) \mathbf{1}_{(\nu - \xi < x \leq \nu + \xi)}) - \frac{1}{2}(\alpha - f_X(\nu))}{\beta f_X(\nu)} \right). \quad (144)$$

We do this in two steps, looking separately at  $\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\bar{X}_n, \sqrt{n}\hat{\xi}_n)$  and  $\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\tilde{\sigma}_n^2, \sqrt{n}\hat{\xi}_n)$ . For the latter, we then use the Delta-method to obtain  $\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\tilde{\sigma}_n, \sqrt{n}\hat{\xi}_n)$  instead.

Since we have iid sums, we are left with computing the covariance of the  $i$ -th element of their two sums each.

*Step 1: Covariance with the sample mean.* Recall that  $\mathbb{P}(X \leq \nu) = 1/2$  and  $\mathbb{P}(|X - \nu| \leq \xi) = 1/2$ . Then,

$$\begin{aligned} \text{Cov} \left( X_i, \frac{\alpha \mathbb{1}_{(x \leq \nu)} - f_X(\nu) \mathbb{1}_{(\nu - \xi < X_i \leq \nu + \xi)}}{\beta f_X(\nu)} \right) &= \frac{1}{\beta f_X(\nu)} \left( \alpha \mathbb{E}[X_i \mathbb{1}_{(X_i \leq \nu)}] - f_X(\nu) \mathbb{E}[X_i \mathbb{1}_{(\nu - \xi < X_i \leq \nu + \xi)}] - \frac{\mu}{2} (\alpha - f_X(\nu)) \right) \\ &= \frac{\sigma}{\beta f_X(\nu)} \left( \alpha \mathbb{E}[Y_i \mathbb{1}_{(Y_i \leq \frac{\nu - \mu}{\sigma})}] - f_X(\nu) \mathbb{E}[Y_i \mathbb{1}_{(\frac{\nu - \xi - \mu}{\sigma} < Y_i \leq \frac{\nu + \xi - \mu}{\sigma})}] \right), \end{aligned}$$

using  $X_i = \mu + \sigma Y_i$  for the second equality. Therefrom we deduce,

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \bar{X}_n, \sqrt{n} \hat{\xi}_n) = \frac{\sigma}{\beta} \left( \frac{\alpha}{f_X(\nu)} \mathbb{E}[Y \mathbb{1}_{(Y \leq \frac{\nu - \mu}{\sigma})}] - \mathbb{E}[Y \mathbb{1}_{(\frac{\nu - \xi - \mu}{\sigma} < Y \leq \frac{\nu + \xi - \mu}{\sigma})}] \right). \quad (145)$$

*Step 2: Covariance with the sample variance.* Analogously we can proceed for the covariance with the sample variance:

$$\begin{aligned} &\text{Cov} \left( (X_i - \mu)^2, \frac{\alpha \mathbb{1}_{(X_i \leq \nu)} - f_X(\nu) \mathbb{1}_{(\nu - \xi < X_i \leq \nu + \xi)}}{\beta f_X(\nu)} \right) \\ &= \frac{1}{\beta f_X(\nu)} \left( \alpha \mathbb{E}[(X_i - \mu)^2 \mathbb{1}_{(X_i \leq \nu)}] - f_X(\nu) \mathbb{E}[(X_i - \mu)^2 \mathbb{1}_{(\nu - \xi < X_i \leq \nu + \xi)}] - \frac{\sigma^2}{2} (\alpha - f_X(\nu)) \right) \\ &= \frac{\sigma^2}{\beta f_X(\nu)} \left( \alpha \mathbb{E}[Y_i^2 \mathbb{1}_{(Y_i \leq \frac{\nu - \mu}{\sigma})}] - f_X(\nu) \mathbb{E}[Y_i^2 \mathbb{1}_{(\frac{\nu - \xi - \mu}{\sigma} < Y_i \leq \frac{\nu + \xi - \mu}{\sigma})}] - \frac{1}{2} (\alpha - f_X(\nu)) \right) \end{aligned}$$

Hence,

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \hat{\sigma}_n^2, \sqrt{n} \hat{\xi}_n) = \frac{\sigma^2}{\beta} \left( \frac{\alpha}{f_X(\nu)} \mathbb{E}[Y^2 \mathbb{1}_{(Y \leq \frac{\nu - \mu}{\sigma})}] - \mathbb{E}[Y^2 \mathbb{1}_{(\frac{\nu - \xi - \mu}{\sigma} < Y \leq \frac{\nu + \xi - \mu}{\sigma})}] - \frac{1}{2} \left( \frac{\alpha}{f_X(\nu)} - 1 \right) \right)$$

and by the Delta-method

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \hat{\sigma}_n, \sqrt{n} \hat{\xi}_n) = \frac{\sigma}{2\beta} \left( \frac{\alpha}{f_X(\nu)} \mathbb{E}[Y^2 \mathbb{1}_{(Y \leq \frac{\nu - \mu}{\sigma})}] - \mathbb{E}[Y^2 \mathbb{1}_{(\frac{\nu - \xi - \mu}{\sigma} < Y \leq \frac{\nu + \xi - \mu}{\sigma})}] - \frac{1}{2} \left( \frac{\alpha}{f_X(\nu)} - 1 \right) \right). \quad (146)$$

Combining (144), (145) and (146) provides the desired asymptotic covariance:

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_{n, \hat{\mu}, \hat{\sigma}}, \sqrt{n} \hat{\xi}_n) &= \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} (\bar{X}_n + q_Y(p) \hat{\sigma}_n), \sqrt{n} \hat{\xi}_n) \\ &= \frac{\sigma}{2\beta} \left( \frac{\alpha}{f_X(\nu)} \mathbb{E}[2Y \mathbb{1}_{(Y \leq \frac{\nu - \mu}{\sigma})}] - \mathbb{E}[2Y \mathbb{1}_{(\frac{\nu - \xi - \mu}{\sigma} < Y \leq \frac{\nu + \xi - \mu}{\sigma})}] \right. \\ &\quad \left. + q_Y(p) \left( \frac{\alpha}{f_X(\nu)} \mathbb{E}[Y^2 \mathbb{1}_{(Y \leq \frac{\nu - \mu}{\sigma})}] - \mathbb{E}[Y^2 \mathbb{1}_{(\frac{\nu - \xi - \mu}{\sigma} < Y \leq \frac{\nu + \xi - \mu}{\sigma})}] - \frac{1}{2} \left( \frac{\alpha}{f_X(\nu)} - 1 \right) \right) \right) \end{aligned}$$

from which (26) follows, by plugging in the explicit expressions for  $\beta, \alpha, f_X(\nu)$  in terms of  $f_Y$ , e.g.  $f_X(\nu) = \frac{1}{\sigma} f_Y(\frac{\nu - \mu}{\sigma})$ . The asymptotics involving general functionals  $h_1, h_2$  follows by applying the Delta method.  $\square$

**Proof of Corollary 7.** This proof is one-to-one as the Proof of Proposition 6, replacing  $q_{n, \hat{\mu}, \hat{\sigma}}(p)$  by  $q_{n, \hat{\sigma}}(p)$ , thus, there is only little change. To get the covariance, (31), we use that  $\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_{n, \hat{\sigma}}, \sqrt{n} \hat{\xi}_n) = q_Y(p) \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} \hat{\sigma}_n, \sqrt{n} \hat{\xi}_n)$ , where the latter covariance has been computed in (146). Also, we use the expression in terms of  $Y$  for  $\beta, \alpha, f_X(\nu)$ . Then, the correlation follows directly as the variance  $\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n} \hat{\xi}_n)$  is given in (135) and  $\lim_{n \rightarrow \infty} \text{Var}(q_{n, \hat{\sigma}}) = q_Y^2(p) (\mathbb{E}[Y^4] - 1) \sigma^2 / 4$ .  $\square$

### A.3 Proofs of Subsection 2.3

**Proof (Proposition 8)** This proof applies simply the Delta method.  $\square$



## A.4 Proof of Subsection 2.4

The only proof in this subsection is Proposition 10. To better structure the proof, we formulate the first result of Proposition 10, equation (33), as a lemma and prove it separately. Using this lemma, we then prove Proposition 10.

**Lemma 28** *Let  $k, l$  be positive integers and consider an iid sample with parent rv  $X$  (with mean  $\mu$  and variance  $\sigma^2$ , if defined). Given the respective smoothness and moment conditions, the asymptotic covariance between functionals of a quantile estimator  $\hat{q}_{ln}$  with sample size  $ln$  (be it  $h_1(q_{ln}), h_1(q_{ln, \hat{\mu}, \hat{\sigma}})$  or  $h_1(q_{ln, \hat{\sigma}})$ ) and the functional of the measure of dispersion estimator with sample size  $kn$ ,  $h_2(\hat{D}_{i, kn})$  for  $i \in \{1, 2, 3\}$ , simply is*

$$\lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} h_1(\hat{q}_{ln}), \sqrt{n} h_2(\hat{D}_{i, kn}) \right) = \frac{1}{\max(k, l)} \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} h_1(\hat{q}_n), \sqrt{n} h_2(\hat{D}_{i, n}) \right).$$

Note that by ‘the respective smoothness and moment conditions’ we refer to the corresponding cases in the results of Subsection 3.1 and B.2.

**Proof of Lemma 28.** We want to prove the scaling law

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(\hat{q}_{ln}), \sqrt{n} h_2(\hat{D}_{i, kn})) = \frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} h_1(\hat{q}_n), \sqrt{n} h_2(\hat{D}_{i, n}))}{\max(k, l)}$$

in the general case considering all three dispersion measure estimators ( $i = 1, 2, 3$ ) and the three possible quantile estimators. This gives us nine different cases to consider. But all the proofs in those different cases share a common approach that we present first (before considering each case).

*General procedure* - Consider two sequences of random variables  $A_n = \frac{1}{n} \sum_{i=1}^n a(X_i) + a_{r, n}$  and  $B_n = \frac{1}{n} \sum_{i=1}^n b(X_i) + b_{r, n}$ , which are functionals of  $X_i$  and consist of a sum of two given parts: One linear part, an iid sum of functions  $a$  and  $b$ , respectively, of  $X_i$ , denoted by  $a(X_i)$  or  $b(X_i)$ , and a second part called the ‘rest’ denoted by  $a_{r, n}$  and  $b_{r, n}$  respectively.

Let us compute the asymptotic covariance  $\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} A_{ln}, \sqrt{n} B_{kn})$  assuming  $k > l$ , the reverse case being shown analogously. We proceed in two steps. The first step consists of splitting the longer sample as  $B_{kn} = B_{ln} + \text{‘rest’}$  to have a covariance of equal sample size that we already know how to handle, as  $\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} A_{ln}, \sqrt{n} B_{ln}) = \frac{1}{l} \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} A_n, \sqrt{n} B_n)$ . The second step consists of showing why the ‘rest’, when splitting  $B_{kn}$ , is negligible in the calculation of the covariance. Assuming this second step, more precisely that

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} A_{ln}, \sqrt{n} b_{r, kn}) = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} A_{ln}, \sqrt{n} b_{r, ln}) = 0, \quad (147)$$

and noticing that  $\lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} A_{ln}, \frac{\sqrt{n}}{kn} \sum_{i=ln+1}^{kn} b(X_i) \right) = 0$  because this is the covariance of iid random variables over disjoint samples, we can write

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} A_{ln}, \sqrt{n} B_{kn}) &= \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} A_{ln}, \sqrt{n} \left( \frac{1}{kn} \sum_{i=1}^{kn} b(X_i) + b_{r, kn} \right) \right) = \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} A_{ln}, \sqrt{n} \frac{1}{kn} \sum_{i=1}^{kn} b(X_i) \right) \\ &= \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} A_{ln}, \sqrt{n} \times \frac{l}{k} \left( \frac{1}{ln} \sum_{i=1}^{ln} b(X_i) + b_{r, ln} \right) \right) + \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} A_{ln}, \sqrt{n} \left( \frac{1}{kn} \sum_{i=ln+1}^{kn} b(X_i) - \frac{l}{k} b_{r, ln} \right) \right) \\ &= \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} A_{ln}, \sqrt{n} \times \frac{l}{k} \times B_{ln} \right) = \frac{\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} A_n, \sqrt{n} B_n)}{\max(k, l)}. \end{aligned}$$

By Cauchy-Schwarz, equations (147) will equal to zero if we have that  $\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n} b_{r, n}) = 0$ .

Thus, to show the scaling law for the quantile and measure of dispersion estimators, we will prove the following: (i) We can express the quantile and measure of dispersion estimators in the form of  $A_n, B_n$  respectively; (ii)  $\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n} b_{r, n}) = 0$  (covers the case  $k > l$ ); (iii)  $\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n} a_{r, n}) = 0$  (covers the case  $k < l$ ). Let us finish the proof commenting on why our six different estimators fulfil those properties. Meaning that for each estimator we will show that the representation (i) is possible and that (ii) holds (which is equivalent to (iii)).

- *Sample Variance*  $\hat{\sigma}_n^2$ . Recall from (114) that  $\hat{\sigma}_n^2 = \frac{n}{n-1}\tilde{\sigma}_n^2 + \frac{n}{n-1}(\bar{X}_n - \mu)^2$ . Hence, we have  $\frac{n-1}{n}A_n = \tilde{\sigma}_n^2$  with  $a_{r,n} = \frac{n}{n-1}(\bar{X}_n - \mu)^2$  (thus, (i) holds) and we know, by (115),  $\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n}(\bar{X}_n - \mu)^2) = 0$  (i.e. (ii) holds).

- *Sample MAD*  $\hat{\theta}_n$ . The calculations were done in Case  $r = 1$  in Part 2 of the Proof of Theorem 1: From (120) we know that  $\hat{\theta}_n = \tilde{\theta}_n + (2F_X(\mu) - 1)(\bar{X}_n - \mu) + (2F_n(\mu) - 2F_X(\mu))(\bar{X}_n - \mu) + L_n$ , where recall that  $\tilde{\theta}_n$  is an iid sum. Hence, (i) is fulfilled. After, (120) we showed that  $\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n}(2F_n(\mu) - 2F_X(\mu))(\bar{X}_n - \mu)) = 0$ . Further, following the proof of the claims (123), (124), we can conclude that  $\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n}\frac{L_n}{\sqrt{n}}) = 0$ . The two facts together meaning that (ii) holds.

- *Sample MedianAD*  $\hat{\xi}_n$ . The Bahadur for the sample MedianAD gives us directly (i) (see (131)):

$$\hat{\xi}_n - \hat{\xi} = \frac{\frac{1}{n} \sum_{i=1}^n (\alpha \mathbf{I}_{(x \leq \nu)} - f_X(\nu) \mathbf{I}_{(\nu - \xi < x \leq \nu + \xi)}) - \frac{1}{2} (\alpha - f_X(\nu))}{\beta f_X(\nu)} + \Delta_n$$

As by this representation it also holds that  $\Delta_n \stackrel{a.s.}{=} O(n^{-3/4}(\log(n))^{3/4})$ , it follows from Popovicius inequality for variances, [35], that there exists  $M < \infty$  such that

$$\text{Var}(\sqrt{n}\Delta_n) \leq M \frac{(\log n)^{3/2}}{n^{1/2}}, \text{ which converges to zero as } n \rightarrow \infty \text{ giving us (ii) as desired.}$$

- *Sample quantile*  $q_n$ . As for the Sample MedianAD, by the Bahadur representation we can show that (i) and (ii) are fulfilled: We have from (109)

$$q_n(p) = q_X(p) + \frac{1 - F_{n,X}(q_X(p)) - (1-p)}{f_X(q_X(p))} + R_{n,p} \quad \text{with } R_{n,p} \stackrel{a.s.}{=} O(n^{-3/4}(\log(n))^{3/4})$$

Thus, with Popoviciu's inequality for variances again, we get for some  $M < \infty$

$$\text{Var}(\sqrt{n}R_{n,p}) \leq M \frac{(\log n)^{3/2}}{n^{1/2}} \xrightarrow{n \rightarrow \infty} 0.$$

- *Location scale quantile (known mean)*  $q_{n,\hat{\sigma}}$ . This case can be seen as a functional of the sample variance:  $q_{n,\hat{\sigma}} = q_Y(p)\sqrt{\hat{\sigma}_n^2}$ . Thus, we simply apply the Delta method to the result from the case with  $\hat{\sigma}_n^2$  and we are done (no need to verify (i) and (ii)).
- *Location scale quantile (unknown mean)*  $q_{n,\hat{\mu},\hat{\sigma}}$ . Recall, that  $q_{n,\hat{\mu},\hat{\sigma}} = \bar{X}_n + \hat{\sigma}_n q_{n,\hat{\sigma}}$  and  $\bar{X}_n$  is already an iid sum. Thus, in comparison with the case of  $q_{n,\hat{\sigma}}$  nothing changes.

For general functions  $h_1, h_2$ , we simply need to use the first order of the Taylor expansion (and the shown argumentation then holds for this linear approximation) -as higher orders are asymptotically negligible.  $\square$

**Proof Proposition 10.** Using the result from Lemma 28, it is straightforward to show the relation for the correlation: We note that, by the asymptotic normality results for all the different quantile estimators and measure of dispersion estimators, we obtain for a fixed integer  $k > 0$

$$\lim_{n \rightarrow \infty} \text{Var}(q_{kn}(p)) = \lim_{n \rightarrow \infty} \frac{\text{Var}(q_n(p))}{k}; \quad \lim_{n \rightarrow \infty} \text{Var}(q_{kn,\hat{\mu},\hat{\sigma}}(p)) = \lim_{n \rightarrow \infty} \frac{\text{Var}(q_{n,\hat{\mu},\hat{\sigma}}(p))}{k}; \quad \lim_{n \rightarrow \infty} \text{Var}(q_{kn,\hat{\sigma}}(p)) = \lim_{n \rightarrow \infty} \frac{\text{Var}(q_{n,\hat{\sigma}}(p))}{k};$$

$$\lim_{n \rightarrow \infty} \text{Var}(\hat{\sigma}_{kn}) = \lim_{n \rightarrow \infty} \frac{\text{Var}(\hat{\sigma}_n)}{k}; \quad \lim_{n \rightarrow \infty} \text{Var}(\hat{\theta}_{kn}) = \lim_{n \rightarrow \infty} \frac{\text{Var}(\hat{\theta}_n)}{k}; \quad \lim_{n \rightarrow \infty} \text{Var}(\hat{\xi}_{kn}) = \lim_{n \rightarrow \infty} \frac{\text{Var}(\hat{\xi}_n)}{k}.$$

Thus, directly

$$\lim_{n \rightarrow \infty} \text{Cor} \left( \sqrt{nh_1}(\hat{q}_{kn}), \sqrt{nh_2}(\hat{D}_{i,ln}) \right) = \lim_{n \rightarrow \infty} \frac{\text{Cov} \left( \sqrt{nh_1}(\hat{q}_{kn}), \sqrt{nh_2}(\hat{D}_{i,ln}) \right)}{\sqrt{\text{Var}(h_1(\hat{q}_{kn}))} \sqrt{\text{Var}(h_2(\hat{D}_{i,ln,t}))}} = \lim_{n \rightarrow \infty} \frac{\frac{\text{Cov}(\sqrt{nh_1}(\hat{q}_n), \sqrt{nh_2}(\hat{D}_{i,n}))}{\max(k,l)}}{\sqrt{\frac{\text{Var}(h_1(\hat{q}_n))}{k}} \sqrt{\frac{\text{Var}(h_2(\hat{D}_{i,n,t}))}{l}}}$$

$$= \sqrt{\frac{kl}{\max^2(k,l)}} \times \lim_{n \rightarrow \infty} \text{Cor} \left( \sqrt{nh_1}(\hat{q}_n), \sqrt{nh_2}(\hat{D}_{i,n}) \right) = \sqrt{\frac{\min(k,l)}{\max(k,l)}} \times \lim_{n \rightarrow \infty} \text{Cor} \left( \sqrt{nh_1}(\hat{q}_n), \sqrt{nh_2}(\hat{D}_{i,n}) \right)$$

where the second equality uses Lemma 28 and the aforementioned scaling of the asymptotic variances.  $\square$

## Appendix B Proofs of Section 3

The proofs in this Appendix B will be an application of Theorem 1 and Proposition 3 for Section 3.1 and Corollaries 5, 7 for Section 3.2 to more specific cases.

### B.1 Proofs of Subsection 3.1

**Proof of Proposition 11.** We simply need to show how equations (6), (7) and (8) transform for location-scale random variables, i.e. when  $X = \mu + \sigma Y$ . Let us start with the covariance in equation (7), and observe, that by definition, we have (for any integer  $r > 0$ )

$$\begin{aligned}\tau_r(|X - \mu|, p) &= \tau_r(|\sigma Y|, p) = (1 - p) (\mathbb{E}[|\sigma Y|^r | X > q_X(p)] - \mathbb{E}[|\sigma Y|^r]) \\ &= (1 - p) \sigma^r (\mathbb{E}[|Y|^r | X > q_X(p)] - \mathbb{E}[|Y|^r]) = \sigma^r \tau_r(|Y|, p)\end{aligned}$$

and besides, for  $r = 1$ , we can write

$$\tau_1(p) = \tau_1(X, p) = (1 - p) (\mathbb{E}[\mu + \sigma Y | X > q_X(p)] - \mathbb{E}[\mu + \sigma Y]) = (1 - p) \sigma (\mathbb{E}[Y | X > q_X(p)] - \mathbb{E}[Y]) = \sigma \tau_1(Y, p) = \sigma^r \tau_1(Y, p).$$

By the basic properties of a location-scale random variable, we have that  $F_X(\mu) = F_Y(0)$  and  $f_X(q_X(p)) = \frac{1}{\sigma} f_Y(q_Y(p))$ . Hence, we get the desired covariance:

$$\begin{aligned}\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_n(p), \sqrt{n} \hat{m}(X, n, r)) &= \frac{\tau_r(|X - \mu|, p) + (2 - r)(2F_X(\mu) - 1)\tau_1(p)}{f_X(q_X(p))} \\ &= \frac{\sigma^{r+1} \left( \tau_r(|Y|, p) + (2 - r)(2F_Y(0) - 1)\tau_1(Y, p) \right)}{f_Y(q_Y(p))}.\end{aligned}$$

The correlation follows directly as we can write, for  $r = 1, 2$ ,

$$\begin{aligned}\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n} q_n(p)) &= \frac{p(1 - p)}{f_X^2(q_X(p))} = \sigma^2 \frac{p(1 - p)}{f_Y^2(q_Y(p))} \quad \text{and} \\ \lim_{n \rightarrow \infty} \text{Var}(\sqrt{n} \hat{m}(X, n, r)) &= \text{Var}(|\sigma Y|^r + (2 - r)(2F_Y(0) - 1)(\mu + \sigma Y)) = \text{Var}(|\sigma Y|^r + (2 - r)(2F_Y(0) - 1)\sigma Y)\end{aligned}$$

As we are considering  $r = 1, 2$  we can write in the second part of the variance  $\sigma = \sigma^r$  such that we get

$$\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n} \hat{m}(X, n, r)) = \sigma^{2r} \text{Var}(|Y|^r + (2 - r)(2F_Y(0) - 1)Y). \quad (148)$$

Thus, we deduce

$$\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{m}(X, n, r)) = \frac{\tau_r(|Y|, p) + (2 - r)(2F_Y(0) - 1)\tau_1(Y, p)}{\sqrt{p(1 - p)} \sqrt{\text{Var}(|Y|^r + (2 - r)(2F_Y(0) - 1)Y)}}. \quad (149)$$

As a last step we verify the following two claims about the asymptotic correlation between the sample quantile and the measure of dispersion estimator in the case of a underlying symmetric location-scale distribution:

- it has its minimum correlation value (in absolute terms) of 0 for  $p = 0.5$
- it is point-symmetric with respect to  $p = 0.5$

If we have a symmetric location scale distribution, the asymptotic correlation, (149), simplifies as follows:

$$\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{m}(X, n, r)) = \frac{\tau_r(|Y|, p)}{\sqrt{p(1 - p)} \sqrt{\text{Var}(|Y|^r)}}.$$

First let us show the value of zero correlation for  $p = 0.5$ . We have

$$\begin{aligned}\tau_r(|Y|, p) = 0 &\Leftrightarrow p(1 - p)(\mathbb{E}[|Y|^r | X > q_X(p)] - \mathbb{E}[|Y|^r | X \leq q_X(p)]) = 0 \\ &\Leftrightarrow \mathbb{E}[|Y|^r | Y > q_Y(p)] = \mathbb{E}[|Y|^r | Y \leq q_Y(p)] \Leftrightarrow p = 0.5.\end{aligned}$$

where the last equivalence follows by the symmetry of the underlying distribution.

To show the point symmetry of the asymptotic correlation, we need to show that

$$\lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{m}(X, n, r)) = - \lim_{n \rightarrow \infty} \text{Cor}(q_n(1-p), \hat{m}(X, n, r)).$$

This follows by straightforward manipulation as, for a symmetric distribution, it holds  $\mathbb{E}[|Y|^r | X > q_X(p)] = \mathbb{E}[|Y|^r | X \leq q_X(1-p)]$ . Hence,

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{m}(X, n, r)) &= \frac{\tau_r(|Y|, p)}{\sqrt{p(1-p)}\sqrt{\text{Var}(|Y|^r)}} \\ &= \frac{p(1-p) (\mathbb{E}[|Y|^r | X \leq q_X(1-p)] - \mathbb{E}[|Y|^r | X > q_X(1-p)])}{\sqrt{p(1-p)}\sqrt{\text{Var}(|Y|^r)}} \\ &= \frac{-\tau_r(|Y|, 1-p)}{\sqrt{p(1-p)}\sqrt{\text{Var}(|Y|^r)}} = - \lim_{n \rightarrow \infty} \text{Cor}(q_n(1-p), \hat{m}(X, n, r)). \quad \square \end{aligned}$$

**Proof of Corollary 12.** It is a direct consequence of Proposition 11, by plugging in  $r = 2$ . □

We think it is worth it to provide insight into the calculation of the two examples:

**Proof of Example 13.** We simply need to compute the quantities in (37), (38) for the two distributions explicitly in the case of the dependence between the sample quantile and the sample variance. The related cases with the logarithm follows directly from this by the Delta method.

- (i) Looking at the Gaussian distribution, we can compute, using partial integration,

$$\mathbb{E}[Y^2 | Y \leq q_Y(p)] = \frac{1}{p} \int_{-\infty}^{q_Y(p)} y^2 \phi(y) dy = \frac{1}{p} (-\phi(\Phi^{-1}(p))\Phi^{-1}(p) + p) \quad (150)$$

and, as  $f_Y(q_Y(p)) = \phi(\Phi^{-1}(p))$  in this case, equation (37) becomes

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\hat{\sigma}_n^2, \sqrt{n}q_n(p)) = \sigma^3 \Phi^{-1}(p),$$

and accordingly, as  $\mathbb{E}[Y^4] = 3$  in the case of a standard Gaussian distribution, equation (38) becomes

$$\lim_{n \rightarrow \infty} \text{Cor}(\hat{\sigma}_n^2, q_n(p)) = \frac{\phi(\Phi^{-1}(p))\Phi^{-1}(p)}{\sqrt{2p(1-p)}}.$$

- (ii) Note that the standard Student t-distribution has variance  $\nu/(\nu-2)$ . Thus, while we would like to work with the random variable  $Y$  that is standardised to have variance 1, we will stick to the standard Student random variable  $\tilde{Y}$ , which is more favourable from an implementation point of view. Thus, when considering equations (37),(38), we need to be careful about this. Analogously to the Gaussian distribution, we look at

$$\mathbb{E}[Y^2 | Y \leq q_Y(p)] = \frac{1}{p} \int_{-\infty}^{q_Y(p)} y^2 f_Y(y) dy = \frac{1}{p} \frac{\nu-2}{\nu} \int_{-\infty}^{q_{\tilde{Y}}(p)} z^2 f_{\tilde{Y}}(z) dz, \quad (151)$$

which can be seen as the second moment of a truncated Student distribution. Results for this are readily available. Unfortunately, the provided results are sometimes wrong (see e.g. [24] for the univariate case, [2] for the multivariate case or [23] when using a truncated skew-Student distribution; the approach via hypergeometric series in [19], in the case of truncated skew-t-distributions remains unchecked because it is difficult to reduce it to the non-skewed case). So we turn to a correct derivation, found in [21], where we have that, for real-valued truncation points  $-\infty < a < b < \infty, \nu > 2$

$$\frac{1}{F_{\tilde{Y}}(b) - F_{\tilde{Y}}(a)} \int_a^b z^2 f_{\tilde{Y}}(x) dx = \frac{\nu(\nu-1)}{\nu-2} \times \frac{F_{\tilde{Y}, \nu-2}(b\sqrt{\frac{\nu-2}{\nu}}) - F_{\tilde{Y}, \nu-2}(a\sqrt{\frac{\nu-2}{\nu}})}{F_{\tilde{Y}}(b) - F_{\tilde{Y}}(a)} - \nu \quad (152)$$

where  $F_{\tilde{Y}}$  denotes the standard Student cdf with  $\nu$  degrees of freedom and  $F_{\tilde{Y}, k}$  the standard Student cdf with  $k$  degrees of freedom. Note that the condition  $-\infty < a < b < \infty$  in [21] is actually not necessary for their proof, i.e. we can have  $-\infty \leq a < b \leq \infty$ .

The above expression can be modified (which yields a simplification for our calculations) using a recurrence

relation for the cumulative distribution function of the Student distribution (as presented e.g. in [22] in the more general case of a skew-t distribution), here for  $\nu > 2$ ,  $a \in \mathbb{R}$ :

$$F_{\tilde{Y}, \nu-2} \left( a \sqrt{\frac{\nu-2}{\nu}} \right) = F_{\tilde{Y}}(a) - \frac{\Gamma(\frac{\nu-1}{2}) \nu^{\frac{\nu-2}{2}}}{2\sqrt{\pi}\Gamma(\frac{\nu}{2})} a(\nu+a^2)^{-\frac{\nu-1}{2}},$$

which for our purposes, is simplified in a form where we get the density involved:

$$F_{\tilde{Y}, \nu-2} \left( a \sqrt{\frac{\nu-2}{\nu}} \right) = F_{\tilde{Y}}(a) - \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{a^2}{\nu}\right)^{-\frac{\nu+1}{2}} \frac{(1 + \frac{a^2}{\nu})a}{\nu-1} = F_{\tilde{Y}}(a) - \frac{1}{\nu-1} f_{\tilde{Y}}(a) \left(1 + \frac{a^2}{\nu}\right) a. \quad (153)$$

Combining this latter equation and (152), taking  $a = -\infty$ ,  $b = q_{\tilde{Y}}(p)$ , we obtain:

$$\int_{-\infty}^{q_{\tilde{Y}}(p)} y^2 f_{\tilde{Y}}(y) dy = \frac{\nu(\nu-1)}{\nu-2} \left( F_{\tilde{Y}}(q_{\tilde{Y}}(p)) - f_{\nu}(q_{\tilde{Y}}(p)) \frac{(1 + \frac{q_{\tilde{Y}}^2(p)})q_{\tilde{Y}}(p)}{\nu+1} \right) - \nu p = \frac{\nu}{\nu-2} \left( p - f_{\tilde{Y}}(q_{\tilde{Y}}(p))q_{\tilde{Y}}(p)(1 + q_{\tilde{Y}}^2(p)/\nu) \right), \quad (154)$$

or, equivalently,

$$\int_{q_{\tilde{Y}}(p)}^{\infty} y^2 f_{\tilde{Y}}(y) dy = \frac{\nu}{\nu-2} - \frac{\nu}{\nu-2} \left( p - f_{\tilde{Y}}(q_{\tilde{Y}}(p))q_{\tilde{Y}}(p)(1 + q_{\tilde{Y}}^2(p)/\nu) \right). \quad (155)$$

Hence, recalling (151), we can conclude that

$$\mathbb{E}[Y^2 | Y \leq q_Y(p)] = 1 - \frac{1}{p} f_{\tilde{Y}}(F_{\tilde{Y}}(p))q_{\tilde{Y}}(p)(1 + q_{\tilde{Y}}^2(p)/\nu), \quad (156)$$

and, as  $f_Y(q_Y(p)) = c f_{\tilde{Y}}(q_{\tilde{Y}}(p))$  in this case, we get from equation (37) that

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}\hat{\sigma}_n^2, \sqrt{n}q_n(p)) = \frac{\sigma^3 q_{\tilde{Y}}(p)(1 + q_{\tilde{Y}}^2(p)/\nu)}{\sqrt{\frac{\nu}{\nu-2}}}.$$

For the correlation, one needs to recall that in the case of the Student distribution

$$\mu_4 = \mathbb{E}[(\sigma Y)^4] = \sigma^4 \left( \frac{\nu-2}{\nu} \right)^2 \mathbb{E}[\tilde{Y}^4] = \sigma^4 \left( \frac{\nu-2}{\nu} \right)^2 \frac{3\nu^2}{(\nu-2)(\nu-4)} = 3\sigma^4 \frac{\nu-2}{\nu-4}, \quad (157)$$

$$\text{i.e. } \mathbb{E}[Y^4] = 3\frac{\nu-2}{\nu-4}, \text{ which gives us, by (38), } \lim_{n \rightarrow \infty} \text{Cor}(\hat{\sigma}_n^2, q_n(p)) = \frac{f_{\tilde{Y}}(q_{\tilde{Y}}(p))q_{\tilde{Y}}(p)(1 + \frac{q_{\tilde{Y}}^2(p)}{\nu})}{\sqrt{2\frac{\nu-1}{\nu-4}p(1-p)}}. \quad \square$$

**Proof of Corollary 14.** It follows from Proposition 11 by plugging in  $r = 1$ . □

**Proof of Example 15.** As both distributions are symmetric around the mean, and are location-scale distributions, the asymptotic covariance and correlation already simplify from (44) and (45) to:

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}(q_n(p), \sqrt{n}(\hat{\theta}_n))) &= \frac{p\sigma^2}{f_Y(q_Y(p))} \left( \frac{\theta}{\sigma} - \mathbb{E}[|Y| | Y \leq q_Y(p)] \right) = \frac{(1-p)\sigma^2}{f_Y(q_Y(p))} (\mathbb{E}[|Y| | Y > q_Y(p)] - \theta/\sigma), \\ \lim_{n \rightarrow \infty} \text{Cor}(q_n(p), \hat{\theta}_n) &= \frac{p(\frac{\theta}{\sigma} - \mathbb{E}[|Y| | Y \leq q_Y(p)])}{\sqrt{p(1-p)}\sqrt{1-\theta^2/\sigma^2}} = \frac{(1-p)(\mathbb{E}[|Y| | Y > q_Y(p)] - \theta/\sigma)}{\sqrt{p(1-p)}\sqrt{1-\theta^2/\sigma^2}}. \end{aligned}$$

The main task is to find an explicit expression for either  $\mathbb{E}[|Y| | Y \leq q_Y(p)]$  or  $\mathbb{E}[|Y| | Y > q_Y(p)]$ . The easiest way to compute this is to restrict oneself to the case of  $p < 0.5$  in the former or  $p \geq 0.5$  in the latter. As we know that the correlation (and hence also the covariance) is point-symmetric, we can then deduce the case of  $p \geq 0.5$  or  $p < 0.5$  respectively. As in applications we are usually interested in high values  $p$  of the quantile, we show the expressions for  $p > 0.5$ :

As by construction,  $Y \geq 0$  a.s. for  $p \geq 0.5$ , we have  $\mathbb{E}[|Y| | Y > q_Y(p)] = \mathbb{E}[Y | Y > q_Y(p)] = \frac{\mathbb{E}[Y \mathbf{1}_{(Y > q_Y(p))}]}{1-p}$ , which simply is the truncated first moment of the corresponding distribution.

(i) In the case of the Gaussian distribution,  $\theta = \mathbb{E}|X - \mu| = \sigma \sqrt{\frac{2}{\pi}}$ . Hence  $\mathbb{E}|Y| = \theta/\sigma = \sqrt{\frac{2}{\pi}}$  and thus

$$\text{Var}(|Y|) = 1 - \frac{\theta^2}{\sigma^2} = 1 - \frac{2}{\pi} \quad (158)$$

For  $p \geq 0.5$ , we have by direct calculation that

$$\mathbb{E}[|Y| \mathbf{1}_{(Y > q_Y(p))}] = \mathbb{E}[Y \mathbf{1}_{(Y > q_Y(p))}] = \phi(\Phi^{-1}(p)). \quad (159)$$

Hence, we have

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_n, \sqrt{n}\hat{\theta}_n) = \frac{\sigma^2 \left( \phi(\Phi^{-1}(p)) - (1-p) \sqrt{\frac{2}{\pi}} \right)}{\phi(\Phi^{-1}(p))} \quad \text{and} \quad \lim_{n \rightarrow \infty} \text{Cor}(q_n, \hat{\theta}_n) = \frac{\phi(\Phi^{-1}(p)) - (1-p) \sqrt{\frac{2}{\pi}}}{\sqrt{p(1-p)} \sqrt{1 - \frac{2}{\pi}}}.$$

(ii) For the case of a Student distribution, we need to pay attention as  $\theta = \mathbb{E}|X - \mu| = \sigma \mathbb{E}|Y| = \frac{\sigma}{\sqrt{\frac{\nu}{\nu-2}}} \mathbb{E}|\tilde{Y}|$ .

Recall that  $E|\tilde{Y}| = \sqrt{\frac{\nu}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)}$ , hence we have  $\mathbb{E}|Y| = \frac{\theta}{\sigma} = \sqrt{\frac{\nu}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)} \sqrt{\frac{\nu-2}{\nu}}$  and thus

$$\text{Var}(|Y|) = 1 - \frac{\theta^2}{\sigma^2} = 1 - \frac{\nu-2}{\pi} \frac{\Gamma^2(\frac{\nu-1}{2})}{\Gamma^2(\frac{\nu}{2})} \quad (160)$$

Finally, using the formula for the truncated moments from [21], we get

$$\begin{aligned} \mathbb{E}[\tilde{Y} \mathbf{1}_{(\tilde{Y} > q_{\tilde{Y}}(p))}] &= \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\nu/2) \sqrt{\nu\pi}} \frac{\nu}{\nu-1} \left( 1 + \frac{q_{\tilde{Y}}^2(p)}{\nu} \right)^{-(\nu-1)/2} = \frac{\nu}{\nu-1} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\nu/2) \sqrt{\nu\pi}} \left( 1 + \frac{q_{\tilde{Y}}^2(p)}{\nu} \right) \left( 1 + \frac{q_{\tilde{Y}}^2(p)}{\nu} \right)^{-(\nu+1)/2} \\ &= \frac{\nu}{\nu-1} f_{\tilde{Y}}(q_{\tilde{Y}}(p)) (1 + q_{\tilde{Y}}^2(p)/\nu). \end{aligned} \quad (161)$$

Hence, as  $\mathbb{E}[Y \mathbf{1}_{(Y > F_Y^{-1}(p))}] = \frac{1}{\sqrt{\frac{\nu}{\nu-2}}} \mathbb{E}[\tilde{Y} \mathbf{1}_{(\tilde{Y} > q_{\tilde{Y}}(p))}]$ , we conclude

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_n, \sqrt{n}\hat{\theta}_n) &= \frac{\sigma^2 \left( \frac{\sqrt{\nu(\nu-2)}}{\nu-1} f_{\tilde{Y}}(q_{\tilde{Y}}(p)) (1 + q_{\tilde{Y}}^2(p)/\nu) - (1-p) \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)} \right)}{f_{\tilde{Y}}(q_{\tilde{Y}}(p)) \sqrt{\nu/(\nu-2)}} \\ \text{and} \quad \lim_{n \rightarrow \infty} \text{Cor}(q_n, \hat{\theta}_n) &= \frac{\frac{\sqrt{\nu(\nu-2)}}{\nu-1} f_{\tilde{Y}}(q_{\tilde{Y}}(p)) (1 + q_{\tilde{Y}}^2(p)/\nu) - (1-p) \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)}}{\sqrt{p(1-p)} \sqrt{1 - \frac{\nu-2}{\pi} \frac{\Gamma^2(\frac{\nu-1}{2})}{\Gamma^2(\nu/2)}}}. \quad \square \end{aligned}$$

**Proof of Proposition 16.** We need to show how covariance and correlation transform from equations (9), (10) and (11). For the covariance it suffices to notice the following - which comes from the definition of  $X$  and  $Y$ , i.e. from  $X = \mu + \sigma Y$ :

$$F_X(\nu \pm \xi) = P(\mu + \sigma Y \leq \nu \pm \xi) = F_Y\left(\frac{\nu \pm \xi - \mu}{\sigma}\right); \quad f_X(q_X(p)) = \frac{1}{\sigma} f_Y(q_Y(p)); \quad f_X(\nu \pm \xi) = \frac{1}{\sigma} f_Y\left(\frac{\nu \pm \xi - \mu}{\sigma}\right).$$

For the correlation we notice, additionally to the above, that  $\gamma = \sigma^2 \tilde{\gamma}$  (with  $\gamma$  defined in Proposition 3, and  $\tilde{\gamma}$  in Proposition 16) and hence  $\frac{\gamma}{f_X^2(\nu)} = \frac{\tilde{\gamma}}{f_Y^2(\frac{\nu-\mu}{\sigma})}$ .  $\square$

**Proof of Proposition 17.** We will show how to transform the general location-scale equations (51), (52) for covariance and correlation when assuming symmetry. For the MedianAD, we have

$$1/2 = \mathbb{P}(|X - \nu| \leq \xi) = \mathbb{P}(|X - \mu| \leq \xi) = \mathbb{P}(|\sigma Y| \leq \xi) = F_Y(\xi/\sigma) - F_Y(-\xi/\sigma).$$

By symmetry, we can write that  $F_Y(-x) = 1 - F_Y(x)$  and  $f_Y(x) = f_Y(-x)$ , for any  $x \in \mathbb{R}$ . So  $3/4 = F_Y(\xi/\sigma) \Leftrightarrow \xi = q_Y(3/4)\sigma$ , and thus, as  $\nu = \mu$  and  $\xi = q_Y(3/4)\sigma$ , we have

$$F_Y\left(\frac{\nu + \xi - \mu}{\sigma}\right) = F_Y(q_Y(3/4)) = 3/4, \quad F_Y\left(\frac{\nu - \xi - \mu}{\sigma}\right) = F_Y(-q_Y(3/4)) = 1 - F_Y(q_Y(3/4)) = 1/4,$$

$$\text{and } f_Y\left(\frac{\nu \pm \xi - \mu}{\sigma}\right) = f_Y(\pm q_Y(3/4)).$$

Therefore the covariance (51) simplifies as follows:

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}\left(\sqrt{n} h_1(q_n(p)), \sqrt{n} h_2(\hat{\xi}_n)\right) &= \sigma^2 \frac{(-\max(0, 3/4 - \max(1/4, p)) + \frac{1-p}{2})}{f_Y(q_Y(p))2f_Y(q_Y(3/4))} h'_1(q_X(p)) h'_2(\xi) \\ &= \sigma^2 \frac{1-p-2\max(0, 3/4 - \max(1/4, p))}{4f_Y(q_Y(p))f_Y(q_Y(3/4))} h'_1(q_X(p)) h'_2(\xi) \end{aligned}$$

As for a symmetric distribution  $\gamma = \tilde{\gamma} = 0$ , the correlation follows directly with the above from equation (52).  $\square$

**Proof of Example 18.** As the Gaussian and Student distribution are both symmetric, the correlation does not depend on the underlying distribution and is equal to the one in equation (54). From equation (53) we straightforwardly compute the covariances. In the Student case, just recall that  $Y = \frac{\tilde{Y}}{\sqrt{\nu/(\nu-2)}}$  hence  $f_Y(q_Y(p)) = \sqrt{\frac{\nu}{\nu-2}} f_{\tilde{Y}}(q_{\tilde{Y}}(p))$ .  $\square$

## B.2 Proofs of Subsection 3.2

**Proof of Proposition 19.** From Proposition 4 by plugging in  $r = 2$  and noticing that  $\mu_i = \sigma^i \mathbb{E}[Y^i]$  and  $\mathbb{E}[Y^2] = \mathbb{E}[|Y|^2] = 1$ , the covariance follows directly from (17). For the correlation, we have from (18),

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cor}(q_{n, \hat{\mu}, \hat{\sigma}}(p), \hat{\sigma}_n^2) &= \frac{\mathbb{E}[Y^3] + \frac{q_Y(p)}{2}(\mathbb{E}[Y^4] - 1)}{\sqrt{1 + q_Y^2(p) \frac{\mathbb{E}[Y^4] - 1}{4} + q_Y(p) \mathbb{E}[Y^3] \sqrt{\mathbb{E}[Y^4] - 1}}} \\ &= \frac{\left(\frac{2\mathbb{E}[Y^3]}{(\mathbb{E}[Y^4] - 1)q_Y(p)} + 1\right) \frac{q_Y(p) \sqrt{(\mathbb{E}[Y^4] - 1)}}{2}}{|q_Y(p)| \sqrt{\frac{\mathbb{E}[Y^4] - 1}{4}} \sqrt{\frac{4}{q_Y^2(p)(\mathbb{E}[Y^4] - 1)} + 1} + \frac{4\mathbb{E}[Y^3]}{q_Y(p)(\mathbb{E}[Y^4] - 1)}} \\ &= \frac{\text{sgn}(p - F_Y(0))}{\sqrt{1 + \frac{4(1 + \mathbb{E}[Y^3]q_Y(p))}{(\mathbb{E}[Y^4] - 1)q_Y^2(p)}}} \left(1 + \frac{2\mathbb{E}[Y^3]}{(\mathbb{E}[Y^4] - 1)q_Y(p)}\right). \quad \square \end{aligned}$$

**Proof of Examples 20.**

(i) The Gaussian case is direct from Proposition 19.

(ii) For the Student distribution, we just need to recall that for  $\tilde{Y} \sim t(0, 1, \nu)$ ,  $\mathbb{E}[\tilde{Y}^4] = \frac{3\nu^2}{(\nu-2)(\nu-4)}$ , from

$$\text{which we deduce } \mu_4 = \mathbb{E}[(X - \mu)^4] = \mathbb{E}[\sigma^4 Y^4] = \mathbb{E}\left[\left(\sigma \frac{\tilde{Y}}{\sqrt{\nu/(\nu-2)}}\right)^4\right] = 3\sigma^4 \frac{\nu-2}{\nu-4},$$

$$\text{and, accordingly, } \mu_4 - \sigma^4 = \sigma^4 \left(3 \frac{\nu-2}{\nu-4} - 1\right) = 2\sigma^4 \frac{\nu-1}{\nu-4}. \quad \square$$

**Proof of Proposition 21.** Direct by plugging in  $r = 1$  in Proposition 4.  $\square$

**Proof of Example 22.**

(i) For the Gaussian distribution, we need to recall its absolute moments, namely  $\mathbb{E}[|X|^p] = \sigma^p \frac{2^{p/2}}{\sqrt{\pi}} \Gamma\left(\frac{p+1}{2}\right)$ ; in particular  $\mathbb{E}[|Y|^3] = 2\sqrt{2/\pi}$  and  $\mathbb{E}|Y| = \theta/\sigma = \sqrt{2/\pi}$ . Hence, plugging into (63), (64) respectively

$$\lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n} q_{n, \hat{\sigma}}(p), \sqrt{n} \hat{\theta}_n) = \sigma^2 \frac{q_Y(p)}{2} (\mathbb{E}[|Y|^3] - \theta/\sigma) = \sigma^2 \frac{\Phi^{-1}(p)}{\sqrt{2\pi}} \quad \text{and}$$

$$\lim_{n \rightarrow \infty} \text{Cor}(q_{n,\hat{\sigma}}(p), \hat{\theta}_n) = \frac{\mathbb{E}[|Y|^3] - \frac{\theta}{\sigma}}{\frac{\mu_4 - \sigma^4}{\sigma^4} \text{Var}(|Y|)} \text{sgn}(p - F_Y(0)) = \frac{\sqrt{2/\pi}}{\sqrt{2(1 - \frac{2}{\pi})}} \text{sgn}(p - 1/2) = \frac{1}{\sqrt{\pi - 2}} \text{sgn}(p - 1/2).$$

- (ii) For the Student case, we first need to compute  $\mathbb{E}[|\tilde{Y}|^3] (= \left(\frac{\nu}{\nu-2}\right)^{3/2} \mathbb{E}[|Y|^3])$ . Notice that for any symmetric distribution around its mean (which equals 0 in this case), we have

$$\mathbb{E}[|\tilde{Y}|^3] = \int_{-\infty}^{\infty} |y|^3 f_{\tilde{Y}}(y) dy = 2 \int_0^{\infty} y^3 f_{\tilde{Y}}(y) dy$$

Truncated moments of a standard Student distribution being computed in Theorem 1 of [21], we apply their result for a lower truncation point equal to 0 and obtain, keeping their notation  $\kappa = \frac{2\Gamma(\frac{\nu+1}{2})}{\Gamma(\nu/2)\sqrt{\nu\pi}}$ ,

$$\mathbb{E}[|\tilde{Y}|^3] = 2 \int_0^{\infty} y^3 f_{\tilde{Y}}(y) dy = \kappa \nu^2 \left( \frac{-1}{\nu-1} + \frac{1}{\nu-3} \right) = \frac{2\kappa\nu^2}{(\nu-1)(\nu-3)}.$$

Thus,

$$\mathbb{E}[|Y|^3] = \left(\frac{\nu-2}{\nu}\right)^{3/2} \mathbb{E}[|\tilde{Y}|^3] = \left(\frac{\nu-2}{\nu}\right)^{3/2} \frac{2\kappa\nu^2}{(\nu-1)(\nu-3)} = \frac{2\kappa(\nu-2)^{3/2}\sqrt{\nu}}{(\nu-1)(\nu-3)} = \frac{(\nu-2)^{3/2}}{\nu-3} \frac{2\Gamma(\frac{\nu-1}{2})}{\sqrt{\pi}\Gamma(\nu/2)}$$

and, as for a Student distribution (see e.g. [34]) we have for the absolute first moment  $\mathbb{E}|\tilde{Y}| = \sqrt{\frac{\nu}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)}$ ,

we obtain for  $Y = \sqrt{\frac{\nu-2}{\nu}} \tilde{Y}$  that  $\frac{\theta}{\sigma} = \mathbb{E}[|Y|] = \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)}$ . Hence, for the covariance, again plugging into (63),

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_{n,\hat{\sigma}}(p), \sqrt{n}\hat{\theta}_n) &= \sigma^2 \frac{q_Y(p)}{2} (\mathbb{E}[|Y|^3] - \theta/\sigma) = \sigma^2 \frac{q_Y(p)}{2} \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)} \left( 2\frac{\nu-2}{\nu-3} - 1 \right) \\ &= \sigma^2 \frac{q_Y(p)}{2} \frac{\nu-1}{\nu-3} \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)} = q_{\tilde{Y}}(p) \frac{\sigma^2}{2\sqrt{\nu\pi}} \frac{(\nu-1)(\nu-2)}{\nu-3} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\nu/2)}. \end{aligned}$$

where for the last equality we used the relation between  $q_Y$  and  $q_{\tilde{Y}}$  in the case of a Student distribution.

Recalling that  $\mathbb{E}[\tilde{Y}^4] = \frac{3\nu^2}{(\nu-2)(\nu-4)}$ , we get for the fourth moment

$$\mu_4 = \mathbb{E}[(X - \mu)^4] = \mathbb{E}[\sigma^4 Y^4] = \sigma^4 \left(\frac{\nu-2}{\nu}\right)^2 \mathbb{E}[\tilde{Y}^4] = 3\sigma^4 \frac{\nu-2}{\nu-4}, \text{ accordingly, } \mu_4 - \sigma^4 = 2\sigma^4 \frac{\nu-1}{\nu-4},$$

hence we obtain for the correlation, from (64),

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cor}(q_{n,\hat{\sigma}}(p), \hat{\theta}_n) &= \text{sgn}(p - F_Y(0)) \times \frac{\mathbb{E}[|Y|^3] - \theta/\sigma}{\sqrt{\frac{\mu_4 - \sigma^4}{\sigma^4} \text{Var}(|Y|)}} = \text{sgn}(p - 1/2) \times \frac{\mathbb{E}[|Y|^3] - \theta/\sigma}{\sqrt{\frac{\mu_4 - \sigma^4}{\sigma^4} (1 - (\theta/\sigma)^2)}} \\ &= \text{sgn}(p - 1/2) \times \frac{\sqrt{(\nu-1)(\nu-2)}}{(\nu-3) \sqrt{\frac{\pi\Gamma^2(\nu/2)}{\Gamma^2(\frac{\nu-1}{2})} - (\nu-2) \sqrt{\frac{2}{\nu-4}}}}. \quad \square \end{aligned}$$

**Proof of Proposition 23.** This follows straightforward from Proposition 6 and Corollary 7. Starting from the covariances (26) and (31) respectively, we simply note that, for a symmetric distribution,  $\nu = \mu$  and  $\xi = q_Y(3/4)\sigma$  and  $\alpha = f_Y(\frac{\nu+\xi-\mu}{\sigma}) - f_Y(\frac{\nu-\xi-\mu}{\sigma}) = f_Y(q_Y(3/4)) - f_Y(-q_Y(3/4)) = 0$ ,  $\beta = 2f_Y(q_Y(3/4))$  and  $\mathbb{E}[Y \mathbf{1}_{(-a < Y \leq a)}] = 0$ ,  $\mathbb{E}[Y^2 \mathbf{1}_{(-a < Y \leq a)}] = 2\mathbb{E}[Y \mathbf{1}_{(0 < Y \leq a)}]$  for any  $a \in \mathbb{R}$ . This gives us (69) and (71) respectively.

Also, since for a symmetric distribution  $\mathbb{E}[Y^3] = 0$ , the correlations in (70) and (72) follow from (27), (32) respectively. For completeness, we can state that for a symmetric distribution  $\lim_{n \rightarrow \infty} \text{Var}(\sqrt{n}\hat{\xi}_n) = \frac{1}{4(f_X(\nu + \xi) + f_X(\nu - \xi))^2}$ ,

and, as  $\mu_3 = 0$ ,  $\lim_{n \rightarrow \infty} \text{Var}(q_{n,\hat{\mu},\hat{\sigma}}(p)) = \sigma^2 \left( 1 + q_Y^2(p) \frac{\mathbb{E}[Y^4] - 1}{4} \right)$ , and  $\lim_{n \rightarrow \infty} \text{Var}(q_{n,\hat{\sigma}}(p)) = \sigma^2 q_Y^2(p) \frac{\mathbb{E}[Y^4] - 1}{4}$ .  $\square$



**Proof of Example 24.**

(i) **Gaussian distribution**

By partial integration, for  $a > 0$ , we can write  $2 \int_0^{+a} x^2 e^{-x^2/2} dx = -2a\phi(a) - 1 + 2\Phi(a)$ , which value is  $-2\Phi^{-1}(3/4)\phi(\Phi^{-1}(3/4)) + 1/2$  for  $a = \Phi^{-1}(3/4)$ . Thus, we obtain, from equations (71) and (72),

$$\lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} q_{n,\hat{\sigma}}(p), \sqrt{n} \hat{\xi}_n \right) = \frac{\sigma^2 \Phi^{-1}(3/4) \Phi^{-1}(p)}{2} \quad \text{and} \quad \lim_{n \rightarrow \infty} \text{Cor}(q_{n,\hat{\sigma}}(p), \hat{\xi}_n) = \text{sgn}(p-1/2) \times 2\sqrt{2} \Phi^{-1}(3/4) \phi(\Phi^{-1}(3/4)).$$

(ii) **Student-t distribution**

Here we need once again the formula for the truncated second moment. A correct derivation can be found in [21] where we have that for  $b > 0$ ,

$$\begin{aligned} \frac{1}{F_{\tilde{Y}}(b) - F_{\tilde{Y}}(-b)} \int_{-b}^b z^2 f_{\tilde{Y}}(x) dx &= \frac{\nu(\nu-1)}{\nu-2} \frac{F_{\tilde{Y},\nu-2} \left( b\sqrt{\frac{\nu-2}{\nu}} \right) - F_{\tilde{Y},\nu-2} \left( -b\sqrt{\frac{\nu-2}{\nu}} \right)}{F_{\tilde{Y}}(b) - F_{\tilde{Y}}(-b)} - \nu \\ \Leftrightarrow \int_{-b}^b z^2 f_{\tilde{Y}}(x) dx &= \nu \left( \frac{\nu-1}{\nu-2} \left( 2F_{\tilde{Y},\nu-2} \left( b\sqrt{\frac{\nu-2}{\nu}} \right) - 1 \right) - (2F_{\tilde{Y}}(b) - 1) \right). \end{aligned}$$

As shown in the proof of Example 13 in (153), it follows from [23] that we can write

$$F_{\tilde{Y},\nu-2} \left( b\sqrt{\frac{\nu-2}{\nu}} \right) = F_{\tilde{Y}}(b) - f_{\tilde{Y}}(b) \frac{b(1+b^2/\nu)}{\nu-1},$$

hence,  $\int_{-b}^b z^2 f_{\tilde{Y}}(x) dx = \frac{\nu}{\nu-2} (2F_{\tilde{Y}}(b) - 2f_{\tilde{Y}}(b) b(1+b^2/\nu) - 1)$ .

Noticing that  $\mathbb{E}[Y^2 | |Y| \leq q_Y(3/4)] = \frac{\nu-2}{\nu} \mathbb{E}[\tilde{Y}^2 | |\tilde{Y}| \leq q_{\tilde{Y}}(3/4)]$ , we obtain

$$\mathbb{E}[Y^2 | |Y| \leq q_Y(3/4)] = 1/2 - 2 f_{\tilde{Y}}(q_{\tilde{Y}}(3/4)) q_{\tilde{Y}}(3/4) (1 + q_{\tilde{Y}}^2(3/4)/\nu)$$

and thus, in total, following equations (71) and (72),

$$\begin{aligned} \lim_{n \rightarrow \infty} \text{Cov} \left( \sqrt{n} q_{n,\hat{\sigma}}(p), \sqrt{n} \hat{\xi}_n \right) &= \sigma^2 \frac{q_{\tilde{Y}}(p) q_{\tilde{Y}}(3/4) (1 + q_{\tilde{Y}}^2(3/4)/\nu)}{2 \frac{\nu}{\nu-2}} \\ \text{and} \quad \lim_{n \rightarrow \infty} \text{Cor}(q_{n,\hat{\sigma}}, \hat{\xi}_n) &= \text{sgn}(p-1/2) \times \frac{4 f_{\tilde{Y}}(q_{\tilde{Y}}(3/4)) (1 + q_{\tilde{Y}}^2(3/4)/\nu) q_{\tilde{Y}}(3/4)}{\sqrt{2 \frac{\nu-1}{\nu-4}}}. \quad \square \end{aligned}$$

## Appendix C Proofs of Section 4

**Proof (Example 27)** To show how we obtain the expressions in Example 27, we only need to focus on the quantities with the sample ES.

Indeed, for the correlations including the sample VaR, i.e. with the sample variance ((84), (85)), the sample MAD ((90), (91)) or the sample MedianAD ((96), (97)) there is nothing to do as they are simply the asymptotic correlation of the sample quantile with the sample variance (see (41), (43)), the sample MAD (see (47), (50)) or the sample MedianAD, respectively (see (56)).

The same remarks hold for the expectile estimator, as it is the sample quantile at level  $\kappa^{-1}(p)$  with  $\kappa(\alpha)$  being defined in (81), which simplifies for location-scale distributions, as follows:

$$\kappa(\alpha) = \frac{\alpha q_Y(\alpha) - \int_{-\infty}^{q_Y(\alpha)} y dF_Y(y)}{-2 \int_{-\infty}^{q_Y(\alpha)} y dF_Y(y) - (1 - 2\alpha) q_Y(\alpha)}.$$

This gives us, in the case of the Gaussian distribution (recall the first truncated moment (159)),

$$\kappa_{norm}(p) = \frac{p\Phi^{-1}(p) + \phi(\Phi^{-1}(p))}{2\phi(\Phi^{-1}(p)) - (1-2p)\Phi^{-1}(p)}.$$

For the Student distribution (assumed to be with mean 0, and recalling the first truncated moment computed in (161)), we obtain

$$\kappa_{stud}(p) = \frac{pq_{\bar{Y}}(p) + \frac{\nu}{\nu-1}f_{\bar{Y}}(q_{\bar{Y}}(p))(1+q_{\bar{Y}}^2(p)/\nu)}{2\frac{\nu}{\nu-1}f_{\bar{Y}}(q_{\bar{Y}}(p))(1+q_{\bar{Y}}^2(p)/\nu) - (1-2p)q_{\bar{Y}}(p)}.$$

**ES with Sample Variance or sample MAD** - First, realise that for location-scale distributions we can rewrite the asymptotic correlation of equation (82) independently of their mean  $\mu$  and variance  $\sigma^2$ , analogously to Proposition 11.

$$\lim_{n \rightarrow \infty} \text{Cor}(\widetilde{\text{ES}}_n(p), \hat{m}(X, n, r)) = \frac{\int_p^1 \frac{\tau_r(|Y|, u) + (2-r)(2F_Y(0)-1)\tau_1(Y, u)}{f_Y(q_Y(u))} du}{\sqrt{2(\int_p^1 \int_v^1 \frac{v(1-u)}{f_Y(q_Y(v))f_Y(q_Y(u))} dudv) \sqrt{\text{Var}(|Y|^r + (2-r)(2F_Y(0)-1)Y)}}}.$$

(162)

In a next step, we look separately at the sample variance ( $r = 2$ ) and sample MAD ( $r = 1$ ), each for the Gaussian and Student distribution separately.

▷ *ES with Sample Variance - Gaussian distribution.* Recall from equation (40) the asymptotic covariance between sample quantile and sample variance in the case of the Gaussian distribution (in this case of a standard normal distribution with  $\sigma^2 = 1$ ). Thus, the numerator of (162) simplifies as  $\int_p^1 \Phi^{-1}(u) du$ , which we already solved by change of variables before (truncated first moment, (159))

$$\int_p^1 \Phi^{-1}(u) du = \int_{\Phi^{-1}(p)}^{\infty} y\phi(y) dy = \phi(\Phi^{-1}(p))$$

More work is needed for the computation of the double-integral in the denominator of (162). We first consider the inner integral,  $\int_v^1 \frac{1-u}{\phi(\Phi^{-1}(u))} du$ . Again, change of variable and partial integration give the following:

$$\int_v^1 \frac{1-u}{\phi(\Phi^{-1}(u))} du = \int_{\Phi^{-1}(v)}^{+\infty} (1-\Phi(y)) dy = y(1-\Phi(y))|_{\Phi^{-1}(v)}^{\infty} + \int_{\Phi^{-1}(v)}^{\infty} y\phi(y) dy = -(1-v)\Phi^{-1}(v) + \phi(\Phi^{-1}(v)).$$

(163)

Thus, plugging this in, the double integral transforms to

$$\int_p^1 \int_v^1 \frac{v(1-u)}{f_Y(q_Y(v))f_Y(q_Y(u))} dudv = - \int_p^1 v \frac{\Phi^{-1}(v)}{\phi(\Phi^{-1}(v))} dv + \int_p^1 v^2 \frac{\Phi^{-1}(v)}{\phi(\Phi^{-1}(v))} + \frac{1}{2} v^2 |_p^1.$$

(164)

We compute the two integrals of (164) one after the other, again using change of variables and partial integration (with  $\phi'(y) = -y\phi(y)$ ), obtaining:

$$\begin{aligned} \int_p^1 v \frac{\Phi^{-1}(v)}{\phi(\Phi^{-1}(v))} dv &= \int_{\Phi^{-1}(p)}^{\infty} y \Phi(y) dy = \frac{1}{2} y^2 \Phi(y) |_{\Phi^{-1}(p)}^{\infty} - \int_{\Phi^{-1}(p)}^{\infty} \frac{y^2}{2} \phi(y) dy \\ &= \frac{1}{2} y^2 \Phi(y) |_{\Phi^{-1}(p)}^{\infty} - \frac{1}{2} (y(-\phi(y))) |_{\Phi^{-1}(p)}^{\infty} + \frac{1}{2} \int_{\Phi^{-1}(p)}^{+\infty} (-\phi(y)) dy \\ &= \frac{1}{2} y^2 \Phi(y) |_{\Phi^{-1}(p)}^{\infty} - \frac{1}{2} \Phi^{-1}(p) \phi(\Phi^{-1}(p)) - \frac{1}{2} (1-p) \end{aligned}$$

and

$$\begin{aligned} \int_p^1 v^2 \frac{\Phi^{-1}(v)}{\phi(\Phi^{-1}(v))} dv &= \int_{\Phi^{-1}(p)}^{\infty} y \Phi^2(y) dy = \frac{y^2}{2} \Phi^2(y) |_{\Phi^{-1}(p)}^{\infty} - \int_{\Phi^{-1}(p)}^{\infty} y^2 \phi(y) \Phi(y) dy \\ &= \frac{1}{2} y^2 \Phi^2(y) |_{\Phi^{-1}(p)}^{\infty} + y \Phi(y) \phi(y) |_{\Phi^{-1}(p)}^{\infty} - \int_{\Phi^{-1}(p)}^{\infty} \phi(y) (\Phi(y) + y\phi(y)) dy \\ &= \frac{1}{2} y^2 \Phi^2(y) |_{\Phi^{-1}(p)}^{\infty} + y \phi(y) \Phi(y) |_{\Phi^{-1}(p)}^{\infty} - \frac{1}{2} \Phi^2(y) |_{\Phi^{-1}(p)}^{\infty} + \frac{1}{2} \phi^2(y) |_{\Phi^{-1}(p)}^{\infty}. \end{aligned}$$

Hence, putting this together gives

$$(164) = \frac{1}{2}\Phi^{-1}(p)^2(p-p^2) + \Phi^{-1}(p)\phi(\Phi^{-1}(p))\left(\frac{1}{2}-p\right) + \frac{1}{2}(1-p) - \frac{1}{2}\phi^2(\Phi^{-1}(p)) \quad (165)$$

and overall, for (162) in the case of correlation with the sample variance (as  $\mathbb{E}[Y^4] = 3$  for the Gaussian distribution, i.e.  $\text{Var}(Y^2) = 2$ ),

$$\lim_{n \rightarrow \infty} \text{Cor}(\widetilde{\text{ES}}_n(p), \sigma_n^2) = \frac{\phi(\Phi^{-1}(p))}{\sqrt{(\Phi^{-1}(p))^2(p-p^2) - \phi^2(\Phi^{-1}(p)) + \Phi^{-1}(p)\phi(\Phi^{-1}(p))(1-2p) + 1-p}\sqrt{2}}.$$

▷ *ES with Sample Variance - Student distribution.* The case of the Student distribution works analogously, but needs in some details more care. Recall from equation (42) the asymptotic covariance between sample quantile and sample variance in the case of the Student distribution with  $\nu > 3$ , thus the numerator of (162) equals to:

$$\int_p^1 \sqrt{\frac{\nu-2}{\nu}} q_{\tilde{Y}}(u) (1 + q_{\tilde{Y}}^2(u)/\nu) du = \sqrt{\frac{\nu-2}{\nu}} \left( \int_{q_{\tilde{Y}}(p)}^{\infty} y f_{\tilde{Y}}(y) dy + \frac{1}{\nu} \int_{q_{\tilde{Y}}(p)}^{\infty} y^3 f_{\tilde{Y}}(y) dy \right) \quad (166)$$

where the equality follows by change of variables. This first truncated moment was computed already in equation (161). For the third truncated moment we, again, use and simplify the formula provided in [21], which gives us:

$$E[\tilde{Y}^3 \mathbf{1}_{(\tilde{Y} > q_{\tilde{Y}}(p))}] = \int_{q_{\tilde{Y}}(p)}^{\infty} y^3 f_{\tilde{Y}}(y) dy = \frac{\nu^2}{(\nu-1)(\nu-3)} f_{\tilde{Y}}(q_{\tilde{Y}}(p)) \left( 2 + q_{\tilde{Y}}^2(p) \frac{\nu-1}{\nu} \right) \left( 1 + \frac{q_{\tilde{Y}}^2(p)}{\nu} \right),$$

giving us for (166) (and hence the numerator of (162)):

$$(166) = \sqrt{\frac{\nu-2}{\nu}} \int_p^1 q_{\tilde{Y}}(u) (1 + q_{\tilde{Y}}^2(u)/\nu) du = \frac{\sqrt{\nu(\nu-2)}}{\nu-3} (1 + q_{\tilde{Y}}^2(p)/\nu)^2 f_{\tilde{Y}}(q_{\tilde{Y}}(p)).$$

We now turn to the double-integral in the denominator of (162), expressed as  $\frac{\nu-2}{\nu} \int_p^1 \int_v^1 \frac{v(1-u)}{f_{\tilde{Y}}(q_{\tilde{Y}}(v)) f_{\tilde{Y}}(q_{\tilde{Y}}(u))} dudv$ .

The inner integral follows one-to-one from the Gaussian case above (as we already know the truncated first moment, see (161)):

$$\int_v^1 \frac{1-u}{f_{\tilde{Y}}(q_{\tilde{Y}}(u))} du = \int_{q_{\tilde{Y}}(v)}^{+\infty} (1 - F_{\tilde{Y}}(y)) dy = -(1-v)q_{\tilde{Y}}(v) + \frac{\nu}{\nu-1} f_{\tilde{Y}}(q_{\tilde{Y}}(v)) (1 + q_{\tilde{Y}}^2(v)/\nu). \quad (167)$$

Plugging this in the double integral, transforms to

$$\int_p^1 \int_v^1 \frac{v(1-u)}{f_{\tilde{Y}}(q_{\tilde{Y}}(v)) f_{\tilde{Y}}(q_{\tilde{Y}}(u))} dudv = \int_p^1 v(v-1) \frac{q_{\tilde{Y}}(v)}{f_{\tilde{Y}}(q_{\tilde{Y}}(v))} dv + \frac{\nu}{\nu-1} \int_p^1 v (1 + q_{\tilde{Y}}^2(v)/\nu) dv. \quad (168)$$

But, using change of variables, then partial integration and then the knowledge of the truncated second moment (155), we obtain:

$$\begin{aligned} \int_p^1 (v^2 - v) \frac{q_{\tilde{Y}}(v)}{f_{\tilde{Y}}(q_{\tilde{Y}}(v))} dv &= \int_{q_{\tilde{Y}}(p)}^{\infty} (F_{\tilde{Y}}^2(y) - F_{\tilde{Y}}(y)) y dy \\ &= \frac{1}{2} y^2 F_{\tilde{Y}}(y) (F_{\tilde{Y}}(y) - 1) \Big|_{q_{\tilde{Y}}(p)}^{\infty} - \int_{q_{\tilde{Y}}(p)}^{\infty} y^2 f_{\tilde{Y}}(y) F_{\tilde{Y}}(y) dy + \frac{1}{2} \int_{q_{\tilde{Y}}(p)}^{\infty} y^2 f_{\tilde{Y}}(y) dy \\ &= \frac{1}{2} \left( \frac{\nu}{\nu-2} + p \left( q_{\tilde{Y}}^2(p) - \frac{\nu}{\nu-2} \right) - q_{\tilde{Y}}^2(p) p^2 + \frac{\nu}{\nu-2} f_{\tilde{Y}}(q_{\tilde{Y}}(p)) q_{\tilde{Y}}(p) (1 + q_{\tilde{Y}}^2(p)/\nu) \right) - \int_{q_{\tilde{Y}}(p)}^{\infty} y^2 f_{\tilde{Y}}(y) F_{\tilde{Y}}(y) dy \end{aligned}$$

$$\text{and } \int_p^1 v q_{\tilde{Y}}^2(v) dv = \int_{q_{\tilde{Y}}(p)}^{\infty} F_{\tilde{Y}}(y) y^2 f_{\tilde{Y}}(y) dy,$$

hence, (168) becomes:

$$\frac{1}{2} \left( \frac{\nu(2\nu-3)}{(\nu-2)(\nu-1)} + p \left( q_{\bar{Y}}^2(p) - \frac{\nu}{\nu-2} \right) - p^2 \left( q_{\bar{Y}}^2(p) + \frac{\nu}{\nu-1} \right) + \frac{\nu}{\nu-2} f_{\bar{Y}}(q_{\bar{Y}}(p)) q_{\bar{Y}}(p) \left( 1 + \frac{q_{\bar{Y}}^2(p)}{\nu} \right) \right) - \frac{\nu-2}{\nu-1} \int_{q_{\bar{Y}}(p)}^{\infty} y^2 f_{\bar{Y}}(y) F_{\bar{Y}}(y) dy. \quad (169)$$

So, we are left with the integral of (169). Again, using partial integration with the fact that the anti-derivative

$\int y f_{\bar{Y}}(y) dy = \frac{-\nu}{\nu-1} f_{\bar{Y}}(y)(1+y^2/\nu)$ , we obtain

$$\begin{aligned} \int_{q_{\bar{Y}}(p)}^{\infty} y^2 f_{\bar{Y}}(y) F_{\bar{Y}}(y) dy &= \frac{-\nu}{\nu-1} f_{\bar{Y}}(y)(1+y^2/\nu) y F_{\bar{Y}}(y) \Big|_{q_{\bar{Y}}(p)}^{\infty} + \frac{\nu}{\nu-1} \int_{q_{\bar{Y}}(p)}^{\infty} f_{\bar{Y}}(y)(1+y^2/\nu) (F_{\bar{Y}}(y) + y f_{\bar{Y}}(y)) dy \\ &= \frac{\nu}{\nu-1} \left\{ p q_{\bar{Y}}(p) \left( 1 + q_{\bar{Y}}^2(p)/\nu \right) f_{\bar{Y}}(q_{\bar{Y}}(p)) + \int_{q_{\bar{Y}}(p)}^{\infty} (1+y^2/\nu) F_{\bar{Y}}(y) f_{\bar{Y}}(y) dy + \int_{q_{\bar{Y}}(p)}^{\infty} y(1+y^2/\nu) f_{\bar{Y}}^2(y) dy \right\}. \end{aligned} \quad (170)$$

Since we have  $\left( 1 + \frac{y^2}{\nu} \right) f_{\bar{Y}}(y) = \frac{\nu-1}{\sqrt{\nu(\nu-2)}} f_{\bar{Y},\nu-2} \left( y \sqrt{\frac{\nu-2}{\nu}} \right)$  (using the notation  $f_{\bar{Y},k}$  introduced after (152)), we can write, proceeding again with partial integration,

$$\begin{aligned} \int_{q_{\bar{Y}}(p)}^{\infty} (1+y^2/\nu) f_{\bar{Y}}(y) F_{\bar{Y}}(y) dy &= \frac{\nu-1}{\sqrt{\nu(\nu-2)}} \int_{q_{\bar{Y}}(p)}^{\infty} f_{\bar{Y},\nu-2} \left( y \sqrt{\frac{\nu-2}{\nu}} \right) F_{\bar{Y}}(y) dy \\ &= \frac{\nu-1}{\nu-2} \left\{ F_{\bar{Y},\nu-2} \left( y \sqrt{\frac{\nu-2}{\nu}} \right) F_{\bar{Y}}(y) \Big|_{q_{\bar{Y}}(p)}^{\infty} - \int_{q_{\bar{Y}}(p)}^{\infty} f_{\bar{Y}}(y) F_{\bar{Y},\nu-2} \left( y \sqrt{\frac{\nu-2}{\nu}} \right) dy \right\}. \end{aligned} \quad (171)$$

Now we use the recurrence relation (153) for the distribution functions, to evaluate the integral in (171), namely

$$\begin{aligned} \int_{q_{\bar{Y}}(p)}^{\infty} f_{\bar{Y}}(y) F_{\bar{Y},\nu-2} \left( y \sqrt{\frac{\nu-2}{\nu}} \right) dy &= \int_{q_{\bar{Y}}(p)}^{\infty} f_{\bar{Y}}(y) \left( F_{\bar{Y}}(y) - \frac{1}{\nu-1} y(1+y^2/\nu) f_{\bar{Y}}(y) \right) dy \\ &= \frac{1}{2} F_{\bar{Y}}^2(y) \Big|_{q_{\bar{Y}}(p)}^{\infty} - \frac{1}{\nu-1} \int_{q_{\bar{Y}}(p)}^{\infty} y(1+y^2/\nu) f_{\bar{Y}}^2(y) dy. \end{aligned} \quad (172)$$

But, by partial integration and with the expression of the antiderivative  $\int y f_{\bar{Y}}(y)$  from above, we can write the integral of (172)

$$\int_{q_{\bar{Y}}(p)}^{\infty} y(1+y^2/\nu) f_{\bar{Y}}^2(y) dy = -\frac{\nu}{\nu-1} (1+y^2/\nu)^2 f_{\bar{Y}}^2(y) \Big|_{q_{\bar{Y}}(p)}^{\infty} - \int_{q_{\bar{Y}}(p)}^{\infty} y(1+y^2/\nu) f_{\bar{Y}}^2(y) dy,$$

which is equivalent to

$$\int_{q_{\bar{Y}}(p)}^{\infty} y(1+y^2/\nu) f_{\bar{Y}}^2(y) dy = \frac{1}{2} \frac{\nu}{\nu-1} (1+q_{\bar{Y}}^2(p)/\nu)^2 f_{\bar{Y}}^2(q_{\bar{Y}}(p)). \quad (173)$$

Hence we obtain for (172)

$$\int_{q_{\bar{Y}}(p)}^{\infty} f_{\bar{Y}}(y) F_{\bar{Y},\nu-2} \left( y \sqrt{\frac{\nu-2}{\nu}} \right) dy = \frac{1}{2} (1-p^2) - \frac{1}{2} \frac{\nu}{(\nu-1)^2} (1+q_{\bar{Y}}^2(p)/\nu)^2 f_{\bar{Y}}^2(q_{\bar{Y}}(p)),$$

so that (171) becomes (using again the recurrence relation (153))

$$\begin{aligned} \int_{q_{\bar{Y}}(p)}^{\infty} (1+y^2/\nu) f_{\bar{Y}}(y) F_{\bar{Y}}(y) dy &= \frac{1}{2} \frac{\nu-1}{\nu-2} \left( 1+p^2 - 2p F_{\bar{Y},\nu-2} \left( q_{\bar{Y}}(p) \sqrt{\frac{\nu-2}{\nu}} \right) + \frac{\nu}{(\nu-1)^2} (1+q_{\bar{Y}}^2(p)/\nu)^2 f_{\bar{Y}}^2(q_{\bar{Y}}(p)) \right) \\ &= \frac{1}{2} \frac{\nu-1}{\nu-2} \left( 1-p^2 + \frac{2}{\nu-1} p q_{\bar{Y}}(p) (1+q_{\bar{Y}}^2(p)/\nu) f_{\bar{Y}}(q_{\bar{Y}}(p)) + \frac{\nu}{(\nu-1)^2} (1+q_{\bar{Y}}^2(p)/\nu)^2 f_{\bar{Y}}^2(q_{\bar{Y}}(p)) \right) \end{aligned} \quad (174)$$

Combining (174) with (173) in (170) provides:

$$\begin{aligned} \int_{q_{\bar{Y}}(p)}^{\infty} y^2 f_{\bar{Y}}(y) F_{\bar{Y}}(y) dy &= \frac{\nu}{\nu-1} p q_{\bar{Y}}(p) (1 + q_{\bar{Y}}^2(p)/\nu) f_{\bar{Y}}(q_{\bar{Y}}(p)) + \frac{1}{2} \frac{\nu}{\nu-2} \left( 1 - p^2 + \frac{2p}{\nu-1} q_{\bar{Y}}(p) (1 + q_{\bar{Y}}^2(p)/\nu) f_{\bar{Y}}(q_{\bar{Y}}(p)) \right) \\ &\quad + \frac{1}{2} \frac{\nu^2}{(\nu-1)(\nu-2)} (1 + q_{\bar{Y}}^2(p)/\nu)^2 f_{\bar{Y}}^2(q_{\bar{Y}}(p)). \end{aligned} \quad (175)$$

Now we are ready to look at the final expression for the double integral as shown in (168), using the results in equations (169) and (175).

$$\begin{aligned} &\int_p^1 \int_v^1 \frac{v(1-u)}{f_{\bar{Y}}(q_{\bar{Y}}(v)) f_{\bar{Y}}(q_{\bar{Y}}(u))} dudv \\ &= \frac{1}{2} \frac{\nu}{\nu-2} \left\{ 1 + q_{\bar{Y}}(p) (1 + q_{\bar{Y}}^2(p)/\nu) f_{\bar{Y}}(q_{\bar{Y}}(p)) - \frac{\nu(\nu-2)}{(\nu-1)^2} (1 + q_{\bar{Y}}^2(p)/\nu)^2 f_{\bar{Y}}^2(q_{\bar{Y}}(p)) \right. \\ &\quad \left. + p \left( \frac{\nu-2}{\nu} q_{\bar{Y}}^2(p) - 1 - \frac{2(\nu-2)}{\nu-1} q_{\bar{Y}}(p) (1 + q_{\bar{Y}}^2(p)/\nu) f_{\bar{Y}}(q_{\bar{Y}}(p)) \right) - \frac{\nu-2}{\nu} p^2 q_{\bar{Y}}^2(p) \right\} \end{aligned} \quad (176)$$

Recalling,  $\mathbb{E}[Y^4] = 3 \frac{\nu-2}{\nu-4}$ , see (157) for the Student distribution (and hence  $\text{Var}(Y^2) = 2 \frac{\nu-1}{\nu-4}$ ), we get overall  $\lim_{n \rightarrow \infty} \text{Cor}(\widetilde{\text{ES}}_n(p), \hat{\sigma}_n^2) =$

$$\frac{\frac{\sqrt{\nu(\nu-2)}}{\nu-3} f_{\bar{Y}}(q_{\bar{Y}}(p))(1 + q_{\bar{Y}}^2(p)/\nu)^2}{\sqrt{1 + q_{\bar{Y}}(p) (1 + q_{\bar{Y}}^2(p)/\nu) f_{\bar{Y}}(q_{\bar{Y}}(p)) - \frac{\nu(\nu-2)}{(\nu-1)^2} (1 + q_{\bar{Y}}^2(p)/\nu)^2 f_{\bar{Y}}^2(q_{\bar{Y}}(p)) + p \left( \frac{\nu-2}{\nu} q_{\bar{Y}}^2(p) - 1 - \frac{2(\nu-2)}{\nu-1} q_{\bar{Y}}(p) (1 + q_{\bar{Y}}^2(p)/\nu) f_{\bar{Y}}(q_{\bar{Y}}(p)) \right) - \frac{\nu-2}{\nu} p^2 q_{\bar{Y}}^2(p)} \sqrt{2 \frac{\nu-1}{\nu-4}}}$$

and we get back the expression for the Gaussian distribution for  $\nu \rightarrow \infty$ .

▷ *ES with Sample MAD - Gaussian distribution.* Recall that we computed asymptotic correlation and covariance between the sample quantile and the sample MAD only for  $p \geq 0.5$ , as the case  $p < 0.5$  can be deduced using the point-symmetry around  $p = 0.5$ . For ES, this argument converts to a symmetry around the  $p = 0.5$ -axis when integrating over the asymptotic covariance, i.e. for any  $p \in (0, 1)$  we can write

$$\int_p^1 \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_n(u), \sqrt{n}\hat{\theta}_n) du = \int_{1-p}^1 \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_n(u), \sqrt{n}\hat{\theta}_n) du.$$

Clearly, this symmetry does not hold for the double-integral in the denominator of (162). Thus, the asymptotic correlation with the sample ES, in contrast to the one with the sample VaR, is not symmetric around  $p = 0.5$ . Nevertheless, for the ease of presentation, we will only consider the case  $p \geq 0.5$ . The only quantity we need to compute of (162) is the integral over the covariance, which comes back, via (46), to evaluate  $\int_p^1 \left( 1 - \frac{1-u}{\phi(\Phi^{-1}(u))} \sqrt{\frac{2}{\pi}} \right) du$ .

Using (163), we obtain

$$\int_p^1 \left( 1 - \frac{1-u}{\phi(\Phi^{-1}(u))} \sqrt{2/\pi} \right) du = (1-p) \left( 1 + \Phi^{-1}(p) \sqrt{2/\pi} \right) - \phi(\Phi^{-1}(p)) \sqrt{2/\pi}.$$

Recalling the asymptotic variance of the sample MAD, (158), and the solution of the double integral, (165), in the Gaussian case, this gives overall:

$$\lim_{n \rightarrow \infty} \text{Cor}(\widetilde{\text{ES}}_n(p), \hat{\theta}_n) = \frac{(1-p) \left( 1 + \Phi^{-1}(p) \sqrt{2/\pi} \right) - \phi(\Phi^{-1}(p)) \sqrt{2/\pi}}{\sqrt{\Phi^{-1}(p)^2(p-p^2) - \phi(\Phi^{-1}(p))^2 + \Phi^{-1}(p)\phi(\Phi^{-1}(p))(1-2p) + 1 - p} \sqrt{1-2/\pi}}$$

▷ *ES with Sample MAD - Student distribution.* The remarks made in the Gaussian case hold also for the Student distribution. We proceed analogously. Again, we first compute the integral in the numerator of the asymptotic covariance given in (48) (with  $\sigma^2 = 1$ ):

$$\int_p^1 \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_n(u), \sqrt{n}\hat{\theta}_n(u)) du = (\nu-2) \int_p^1 \left( \frac{1}{\nu-1} \left( 1 + \frac{q_{\bar{Y}}^2(u)}{\nu} \right) - \frac{\Gamma(\frac{\nu-1}{2})}{\sqrt{\pi\nu}\Gamma(\frac{\nu}{2})} \frac{1-u}{f_{\bar{Y}}(q_{\bar{Y}}(u))} \right) du \quad (177)$$

Since by change of variables  $\int_p^1 q_{\tilde{Y}}^2(u) du = \int_{q_{\tilde{Y}}(p)}^\infty y^2 f_{\tilde{Y}}(y) dy$ , using the knowledge of the second truncated moment, (155), for the first integral of (177) and (167) for the second integral, provides

$$\begin{aligned} & \int_p^1 \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_n(u), \sqrt{n}\hat{\theta}_n(u)) du \\ &= \frac{1}{\nu-1} \left\{ (\nu-2)(1-p) + 1-p + f_{\tilde{Y}}(q_{\tilde{Y}}(p))q_{\tilde{Y}}(p) \left(1 + q_{\tilde{Y}}^2(p)/\nu\right) \right\} \\ &+ \frac{(\nu-2)\Gamma(\frac{\nu-1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left( (1-p)q_{\tilde{Y}}(p) - \frac{\nu}{\nu-1} f_{\tilde{Y}}(q_{\tilde{Y}}(p))(1 + q_{\tilde{Y}}^2(p)/\nu) \right) \\ &= 1 + q_{\tilde{Y}}(p) \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})} \frac{\nu-2}{\sqrt{\pi\nu}} + \frac{1}{\nu-1} \left[ \left(1 + \frac{q_{\tilde{Y}}^2(p)}{\nu}\right) f_{\tilde{Y}}(q_{\tilde{Y}}(p)) \left( q_{\tilde{Y}}(p) - \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})} \frac{(\nu-2)\sqrt{\nu}}{\sqrt{\pi}} \right) \right] - p \left( 1 + q_{\tilde{Y}}(p) \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})} \frac{\nu-2}{\sqrt{\pi\nu}} \right) \end{aligned}$$

which gives us back the Gaussian case for  $\nu \rightarrow \infty$ , recalling an asymptotic property of the Gamma function

$$\lim_{n \rightarrow \infty} \frac{\Gamma(n+\alpha)}{\Gamma(n)n^\alpha} = 1 \text{ that we need to use here with } n = \nu\alpha \text{ and } \alpha = 1/2. \text{ Recalling the asymptotic variance of the sample MAD in the Student case, (160), and the solution of the double integral, (176), this gives overall}$$

$$\lim_{n \rightarrow \infty} \text{Cor}(\widehat{\text{ES}}_n(p), \hat{\theta}_n) = \frac{1}{\sqrt{1 - \frac{\nu-2}{\pi} \frac{\Gamma((\nu-1)/2)^2}{\Gamma(\nu/2)^2}}} \times$$

$$\frac{1 + q_{\tilde{Y}}(p) \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})} \frac{\nu-2}{\sqrt{\pi\nu}} + \frac{1}{\nu-1} \left[ \left(1 + \frac{q_{\tilde{Y}}^2(p)}{\nu}\right) f_{\tilde{Y}}(q_{\tilde{Y}}(p)) \left( q_{\tilde{Y}}(p) - \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})} \frac{(\nu-2)\sqrt{\nu}}{\sqrt{\pi}} \right) \right] - p \left( 1 + q_{\tilde{Y}}(p) \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})} \frac{\nu-2}{\sqrt{\pi\nu}} \right)}{\sqrt{1 + q_{\tilde{Y}}(p) \left(1 + q_{\tilde{Y}}^2(p)/\nu\right) f_{\tilde{Y}}(q_{\tilde{Y}}(p)) - \frac{\nu(\nu-2)}{(\nu-1)^2} \left(1 + q_{\tilde{Y}}^2(p)/\nu\right)^2 f_{\tilde{Y}}^2(q_{\tilde{Y}}(p)) + p \left(\frac{\nu-2}{\nu} q_{\tilde{Y}}^2(p) - 1 - \frac{2(\nu-2)}{\nu-1} q_{\tilde{Y}}(p) \left(1 + q_{\tilde{Y}}^2(p)/\nu\right) f_{\tilde{Y}}(q_{\tilde{Y}}(p))\right) - \frac{\nu-2}{\nu} p^2 q_{\tilde{Y}}^2(p)}}$$

and we get back the expression for the Gaussian distribution for  $\nu \rightarrow \infty$ .

**ES with Sample MedianAD** - For the asymptotic correlation of the sample ES with the sample MedianAD, we only need to compute the integral in the numerator, as the rest is known already.

Further, recall that for symmetric location-scale distributions we can rewrite the asymptotic correlation of equation (83) independently of their mean  $\mu$  and variance  $\sigma^2$ , and they simplify a lot analogously to Proposition 17.

$$\lim_{n \rightarrow \infty} \text{Cor}(\widehat{\text{ES}}_n(p), \hat{\xi}_n) = \frac{\int_p^1 \frac{(1-u-2 \max(0, 3/4 - \max(1/4, u))) du}{f_Y(q_Y(u))}}{\sqrt{2 \int_p^1 \int_v^1 \frac{v(1-u)}{f_Y(q_Y(v))f_Y(q_Y(u))} dudv}}. \quad (178)$$

Also, as discussed in the case of the ES with the sample MAD, the asymptotic covariance with the ES and the sample MedianAD is symmetric around the  $p = 0.5$ -axis, i.e., for any  $p \in (0, 1)$ , we can write

$$\int_p^1 \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_n(u), \sqrt{n}\hat{\xi}_n) du = \int_{1-p}^1 \lim_{n \rightarrow \infty} \text{Cov}(\sqrt{n}q_n(u), \sqrt{n}\hat{\xi}_n) du.$$

But again, the asymptotic correlation with the sample ES is not symmetric around  $p = 0.5$  because of the double-integral in the denominator. Thus, as previously, for the ease of presentation, we will only consider the case  $p \geq 0.5$ . For this, we split the covariance integral in the numerator as follows.

Introducing  $g(u) := 1 - u - 2 \max(3/4 - \max(1/4, u), 0)$ , we can write, for any  $p \geq 0.5$ ,

$$\int_p^1 \frac{g(u)}{f_Y(q_Y(u))} du = \begin{cases} \int_p^{3/4} \frac{u-1/2}{f_Y(q_Y(u))} du + \int_{3/4}^1 \frac{1-u}{f_Y(q_Y(u))} du & \text{if } 3/4 > p \geq 1/2, \\ \int_p^1 \frac{1-u}{f_Y(q_Y(u))} du & \text{if } p \geq 3/4. \end{cases} \quad (179)$$

In the following we need to compute both integrals of (179). For this, we plug-in the explicit first moments in the case of the Gaussian and Student distribution.

▷ *ES with Sample MedianAD - Gaussian distribution.* We have, via change of variables for the first integral and usin (163) for the second,

$$\begin{aligned} \int_p^{3/4} \frac{u-1/2}{\phi_Y(\Phi^{-1}(u))} du &= y\Phi(y)|_{\Phi^{-1}(p)}^{\Phi^{-1}(3/4)} - \frac{1}{2} y|_{\Phi^{-1}(p)}^{\Phi^{-1}(3/4)} - \int_{\Phi^{-1}(p)}^{\Phi^{-1}(3/4)} y\phi(y) dy \\ &= \frac{1}{4} \Phi^{-1}(3/4) - (p-1/2)\Phi^{-1}(p) + \phi(\Phi^{-1}(3/4)) - \phi(\Phi^{-1}(p)) \end{aligned}$$

$$\text{and } \int_p^1 \frac{1-u}{\phi_Y(\Phi^{-1}(u))} du = -(1-p)\Phi^{-1}(p) + \phi(\Phi^{-1}(p)),$$

giving, for (179),

$$\int_p^1 \frac{g(u)}{f_Y(q_Y(u))} du = \begin{cases} -\Phi^{-1}(p)(p-1/2) + 2\phi(\Phi^{-1}(3/4)) - \phi(\Phi^{-1}(p)) & \text{if } 3/4 > p \geq 1/2, \\ -\Phi^{-1}(p)(1-p) + \phi(\Phi^{-1}(p)) & \text{if } p \geq 3/4. \end{cases}$$

As the denominator of (178) was computed in (165), we are done: For  $1/2 \leq p < 3/4$  we have,

$$\lim_{n \rightarrow \infty} \text{Cor}(\widetilde{\text{ES}}_n(p), \hat{\xi}_n) = \frac{-\Phi^{-1}(p)(p-1/2) + 2\phi(\Phi^{-1}(3/4)) - \phi(\Phi^{-1}(p))}{\sqrt{\frac{1}{2}\Phi^{-1}(p)^2(p-p^2) + \Phi^{-1}(p)\phi(\Phi^{-1}(p))(\frac{1}{2}-p) + \frac{1}{2}(1-p) - \frac{1}{2}\phi^2(\Phi^{-1}(p))}},$$

and for  $p \geq 3/4$

$$\lim_{n \rightarrow \infty} \text{Cor}(\widetilde{\text{ES}}_n(p), \hat{\xi}_n) = \frac{-\Phi^{-1}(p)(1-p) + \phi(\Phi^{-1}(p))}{\sqrt{\frac{1}{2}\Phi^{-1}(p)^2(p-p^2) + \Phi^{-1}(p)\phi(\Phi^{-1}(p))(\frac{1}{2}-p) + \frac{1}{2}(1-p) - \frac{1}{2}\phi^2(\Phi^{-1}(p))}}.$$

▷ *ES with Sample MedianAD - Student distribution.* Analogously, for the Student distribution, we can write, using the expression of the anti-derivative  $\int y f_{\bar{Y}}(y) dy = \frac{-\nu}{\nu-1} f_{\bar{Y}}(y)(1+y^2/\nu)$  for the first integral,

$$\begin{aligned} \int_p^{3/4} \frac{u-1/2}{f_Y(q_Y(u))} du &= \sqrt{\frac{\nu-2}{\nu}} \int_p^{3/4} \frac{u-1/2}{f_{\bar{Y}}(q_{\bar{Y}}(u))} du = \sqrt{\frac{\nu-2}{\nu}} \left( y F_{\bar{Y}}(y)|_{q_{\bar{Y}}(p)}^{q_{\bar{Y}}(3/4)} - \frac{1}{2} y|_{q_{\bar{Y}}(p)}^{q_{\bar{Y}}(3/4)} - \int_{q_{\bar{Y}}(p)}^{q_{\bar{Y}}(3/4)} y f_{\bar{Y}}(y) dy \right) \\ &= \sqrt{\frac{\nu-2}{\nu}} \left( \frac{1}{4} q_{\bar{Y}}(3/4) - (p-1/2) q_{\bar{Y}}(p) + \frac{\nu}{\nu-1} \left[ \left( 1 + \frac{q_{\bar{Y}}^2(3/4)}{\nu} \right) f_{\bar{Y}}(q_{\bar{Y}}(3/4)) - \left( 1 + \frac{q_{\bar{Y}}^2(p)}{\nu} \right) f_{\bar{Y}}(q_{\bar{Y}}(p)) \right] \right) \end{aligned}$$

and (167) for the second integral:

$$\text{and } \int_p^1 \frac{1-u}{f_Y(q_Y(u))} du = \sqrt{\frac{\nu-2}{\nu}} \int_p^1 \frac{1-u}{f_{\bar{Y}}(q_{\bar{Y}}(u))} du = \sqrt{\frac{\nu-2}{\nu}} \left( -(1-p) q_{\bar{Y}}(p) + \frac{\nu}{\nu-1} \left( 1 + \frac{q_{\bar{Y}}^2(p)}{\nu} \right) f_{\bar{Y}}(q_{\bar{Y}}(p)) \right),$$

giving, for (179),

$$\int_p^1 \frac{g(u)}{f_Y(q_Y(u))} du = \sqrt{\frac{\nu-2}{\nu}} \begin{cases} \left( -(p-1/2) q_{\bar{Y}}(p) + \frac{2\nu}{\nu-1} \left( 1 + \frac{q_{\bar{Y}}^2(3/4)}{\nu} \right) f_{\bar{Y}}(q_{\bar{Y}}(3/4)) - \frac{\nu}{\nu-1} \left( 1 + \frac{q_{\bar{Y}}^2(p)}{\nu} \right) f_{\bar{Y}}(q_{\bar{Y}}(p)) \right) & \text{if } 3/4 > p \geq 1/2, \\ \left( -(1-p) q_{\bar{Y}}(p) + \frac{\nu}{\nu-1} \left( 1 + \frac{q_{\bar{Y}}^2(p)}{\nu} \right) f_{\bar{Y}}(q_{\bar{Y}}(p)) \right), & \text{if } p \geq 3/4. \end{cases}$$

As the denominator of (178) was computed in (176), we are done: For  $1/2 \leq p < 3/4$ , we have,  $\lim_{n \rightarrow \infty} \text{Cor}(\widetilde{\text{ES}}_n(p), \hat{\xi}_n) =$

$$\frac{\sqrt{\frac{\nu-2}{\nu}} \left( -(p-1/2) q_{\bar{Y}}(p) + \frac{2\nu}{\nu-1} \left( 1 + \frac{q_{\bar{Y}}^2(3/4)}{\nu} \right) f_{\bar{Y}}(q_{\bar{Y}}(3/4)) - \frac{\nu}{\nu-1} \left( 1 + \frac{q_{\bar{Y}}^2(p)}{\nu} \right) f_{\bar{Y}}(q_{\bar{Y}}(p)) \right)}{\sqrt{\left\{ 1 + q_{\bar{Y}}(p) \left( 1 + \frac{q_{\bar{Y}}^2(p)}{\nu} \right) f_{\bar{Y}}(q_{\bar{Y}}(p)) - \frac{\nu(\nu-2)}{(\nu-1)^2} \left( 1 + \frac{q_{\bar{Y}}^2(p)}{\nu} \right)^2 f_{\bar{Y}}^2(q_{\bar{Y}}(p)) + p \left( \frac{\nu-2}{\nu} q_{\bar{Y}}^2(p) - 1 - \frac{2(\nu-2)}{\nu-1} q_{\bar{Y}}(p) \left( 1 + \frac{q_{\bar{Y}}^2(p)}{\nu} \right) f_{\bar{Y}}(q_{\bar{Y}}(p)) \right) - \frac{\nu-2}{\nu} p^2 q_{\bar{Y}}^2(p) \right\}}},$$

and for  $p \geq 3/4$ ,  $\lim_{n \rightarrow \infty} \text{Cor}(\widetilde{\text{ES}}_n(p), \hat{\xi}_n) =$

$$\frac{\sqrt{\frac{\nu-2}{\nu}} \left( -(1-p) q_{\bar{Y}}(p) + \frac{\nu}{\nu-1} \left( 1 + \frac{q_{\bar{Y}}^2(p)}{\nu} \right) f_{\bar{Y}}(q_{\bar{Y}}(p)) \right)}{\sqrt{\left\{ 1 + q_{\bar{Y}}(p) \left( 1 + \frac{q_{\bar{Y}}^2(p)}{\nu} \right) f_{\bar{Y}}(q_{\bar{Y}}(p)) - \frac{\nu(\nu-2)}{(\nu-1)^2} \left( 1 + \frac{q_{\bar{Y}}^2(p)}{\nu} \right)^2 f_{\bar{Y}}^2(q_{\bar{Y}}(p)) + p \left( \frac{\nu-2}{\nu} q_{\bar{Y}}^2(p) - 1 - \frac{2(\nu-2)}{\nu-1} q_{\bar{Y}}(p) \left( 1 + \frac{q_{\bar{Y}}^2(p)}{\nu} \right) f_{\bar{Y}}(q_{\bar{Y}}(p)) \right) - \frac{\nu-2}{\nu} p^2 q_{\bar{Y}}^2(p) \right\}}}$$

□

## Appendix D Miscellanea

### D.1 Extensions of Theorem 1

As mentioned in Subsection 2.1, there are different direct extensions of Theorem 1 (note that we could also provide in the same way extensions of Proposition 3, 4, 6). First, one can consider a more general function  $h(x, y)$ . Second, as in [8], we can look at the joint distribution of a vector of sample quantiles, instead of only at one sample quantile. For this, denote by  $\mathbf{q}_X(\mathbf{p})$  the  $m$ -vector of quantiles evaluated at  $p_i, i = 1, \dots, m$ , where  $0 < p_1 < \dots < p_m < 1$ , and by  $\mathbf{q}_n(\mathbf{p})$  the corresponding  $m$ -vector of sample quantiles  $q_n(p_i), i = 1, \dots, m$ . And, one can also combine the two ideas and look at more general functions applied on a vector of sample quantiles:

**Theorem 29** Consider an iid sample with parent rv  $X$  having mean  $\mu$ , variance  $\sigma^2$ . Further, consider a function

$$h : \mathbb{R}^{m+1} \mapsto \mathbb{R}^{m+1}, \text{ i.e. } h(x_1, \dots, x_m, y) = \begin{pmatrix} h_1(x_1, \dots, x_m, y) \\ \dots \\ h_{m+1}(x_1, \dots, x_m, y) \end{pmatrix} \text{ with continuous real-valued components}$$

$h_i(x_1, \dots, x_m, y), i = 1, \dots, m$ , and existing partial derivatives denoted by  $\partial_i h_j, i, j \in \{1, \dots, m+1\}$ . Assume conditions  $(C_1), (C_3)$  in each neighbourhood of  $q_X(p_i), i = 1, \dots, m$ , and  $(M_r)$  for  $r = 1, 2$  respectively as well as  $(C_2)$  at  $\mu$  for  $r = 1$ . Then, the joint behaviour of the functional  $h$  of the sample quantile vector  $\mathbf{q}_n(\mathbf{p})$  and of the measure of dispersion  $\hat{m}(X, n, r)$  (defined in Table 1) is asymptotically normal:

$$\sqrt{n} h \begin{pmatrix} \mathbf{q}_n(\mathbf{p}) \\ \hat{m}(X, n, r) \end{pmatrix} - h \begin{pmatrix} \mathbf{q}_X(\mathbf{p}) \\ m(X, r) \end{pmatrix} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, J(h(\mathbf{v})) \Sigma^{(m,r)} J(h(\mathbf{v}))'), \quad (180)$$

where the asymptotic covariance matrix  $\Sigma^{(m,r)}$  of dimension  $(m+1) \times (m+1)$  can be written as

$$\Sigma^{(m,r)} = \begin{bmatrix} \Sigma^{(m)} & \mathbf{s}(\mathbf{m}, r) \\ \mathbf{s}(\mathbf{m}, r)' & \text{Var} \left( (|X - \mu|^r + (2-r)(2F_X(\mu) - 1)X) \right) \end{bmatrix} \quad (181)$$

with  $\Sigma_{ij}^{(m)} = \Sigma_{ji}^{(m)} = \frac{p_i(1-p_j)}{f_X(q_X(p_i))f_X(q_X(p_j))}$  for  $i, j \in \{1, \dots, m\}$  and the  $i$ -th element of  $\mathbf{s}(\mathbf{m}, r)$  being  $\frac{\tau_r(|X - \mu|, p_i) + (2-r)(2F_X(\mu) - 1)\tau_1(p_i)}{f_X(q_X(p_i))}$ ,  $i = 1, \dots, m$  where  $\tau_r$  is defined in (3),  $J(h(\mathbf{v}))$  is the Jacobian matrix of  $h(\mathbf{v})$ , and we denote with a ' the transpose of a vector.

As a corollary of the theorem, we can state how the result explicitly looks like if we go back to the one-dimensional sample quantile case (with a general function  $h(x, y)$ ).

**Corollary 30** Consider an iid sample with parent rv  $X$  having mean  $\mu$ , variance  $\sigma^2$  and a function  $h(x, y) = \begin{pmatrix} h_1(x, y) \\ h_2(x, y) \end{pmatrix}$  with continuous real-valued components  $h_1(x, y), h_2(x, y)$  and existing partial derivatives denoted by  $\partial_i h_j, i, j \in \{1, 2\}$ . Assume conditions  $(C_1), (C_3)$  in a neighbourhood of  $q_X(p)$ ,  $(M_r)$  for  $r = 1, 2$  respectively, as well as  $(C_2)$  at  $\mu$  for  $r = 1$ . Then, the joint behaviour of the functional  $h$  of the sample quantile  $q_n(p)$  (for  $p \in (0, 1)$ ) and of the measure of dispersion  $\hat{m}(X, n, r)$  (defined in Table 1) is asymptotically normal:

$$\sqrt{n} h \begin{pmatrix} q_n(p) \\ \hat{m}(X, n, r) \end{pmatrix} - h \begin{pmatrix} \hat{q}_X(p) \\ m(X, r) \end{pmatrix} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, \Sigma^{(h,r)}),$$

where the asymptotic covariance matrix  $\Sigma^{(h,r)} = (\Sigma_{ij}^{(h,r)}, 1 \leq i, j \leq 2)$  satisfies, denoting by abuse of notation,  $h_1 = h_1(q_X(p), m(X, r))$ , and  $h_2 = h_2(q_X(p), m(X, r))$ ,

$$\begin{aligned} \Sigma_{11}^{(h,r)} &= \text{Var}(q_n(p)) (\partial_1 h_1)^2 + 2\partial_1 h_1 \partial_2 h_1 \text{Cov}(q_n(p), \hat{m}(X, n, r)) + \text{Var}(\hat{m}(X, n, r)) (\partial_2 h_1)^2; \\ \Sigma_{22}^{(h,r)} &= \text{Var}(\hat{m}(X, n, r)) (\partial_2 h_2)^2 + 2\partial_1 h_2 \partial_2 h_2 \text{Cov}(q_n(p), \hat{m}(X, n, r)) + \text{Var}(q_n(p)) (\partial_1 h_2)^2; \end{aligned}$$

$$\Sigma_{12}^{(h,r)} = \Sigma_{21}^{(h,r)} = \text{Cov}(q_n(p), \hat{m}(X, n, r)) (\partial_1 h_1 \partial_2 h_2 + \partial_2 h_1 \partial_1 h_2) + \text{Var}(q_n(p)) \partial_1 h_1 \partial_1 h_2 + \text{Var}(\hat{m}(X, n, r)) \partial_2 h_1 \partial_2 h_2,$$



with

$$\begin{aligned}\text{Var}(q_n(p)) &= \frac{p(1-p)}{f_X^2(q_X(p))}, & \text{Var}(\hat{m}(X, n, r)) &= \text{Var}(|X - \mu|^r + (2-r)(2F_X(\mu) - 1)X), \\ \text{Cov}(q_n(p), \hat{m}(X, n, r)) &= \frac{\tau_r(|X - \mu|, p) + (2-r)(2F_X(\mu) - 1)\tau_1(p)}{f_X(q_X(p))}.\end{aligned}$$

The asymptotic correlation between the functional  $h$  of the measure of dispersion and the sample quantile can be deduced from the above expressions. In the special case of having  $\partial_2 h_1 = \partial_1 h_2 = 0$ , it is identical -up to its sign (and apart of the case that  $\text{sgn}(\partial_1 h_1 \partial_2 h_2) = 0$ ) - whatever the choice of  $h$  (under that restriction):

$$\lim_{n \rightarrow \infty} \text{Cor}(h(q_n(p), \hat{m}(X, n, r))) = \frac{\tau_r(|X - \mu|, p) + (2-r)(2F(\mu) - 1)\tau_1(p)}{\sqrt{\text{Var}(|X - \mu|^r)p(1-p)}} \times \text{sgn}(\partial_1 h_1 \partial_2 h_2).$$

## D.2 Tables of Finite Sample Size Simulation Study

For the Appendix to be self-contained, we present the procedure already explained in Subsection 2.4 again:

To assess the finite sample performance, we conduct a simulation study in the following way: We simulate an iid sample with mean  $\mu = 0$  of varying size  $N$ . The overall sample size is determined by the fact that we use different sample sizes  $n$  for the estimation of either the quantile or the dispersion measure, with  $n = 126, 252, 504, 1008$  (being multiples or fractions of one year of data, i.e. 252 data points), and different lengths of time-series  $l$  which we use to compute the sample correlation,  $l = 10, 25, 50, 100, 250, 500$ . In each of the cases, the overall sample size needed is  $N = nl$ . We compute the time series of quantile estimates  $\hat{q}_{n,t}(p)$  on disjoint samples for quantiles of order  $p = 0.95$  and  $p = 0.99$  each, and accordingly the time series of measure of dispersion estimates  $\hat{D}_{i,n,t}$  too. Thus,  $\hat{q}_{n,t}$  is either the sample quantile time series  $q_{n,t}$  or the location-scale quantile time series with known  $\mu = 0$   $q_{n,\hat{\sigma},t}$ ,  $\hat{D}_{i,n,t}$  can be either the sample standard deviation, the sample MAD, or the sample MedianAD. Recalling the correlation of interest,  $\text{Cor}(\hat{q}_{n,t}(p), \hat{D}_{i,n,t})$ , we then estimate the linear Pearson correlation using these two time series of  $l$  estimates. This procedure is repeated 1'000-fold in each case. Then, we report in Tables 10, 11, 12 the averages of the 1'000-fold repetition with, in brackets, the corresponding empirical 95% confidence interval values. Further, we provide the theoretical asymptotic value in the last column. Also, we provide confidence intervals for the sample Pearson linear correlation coefficient (using the classical variance-stabilizing Fisher transform of the correlation coefficient for a bivariate normal distribution to compute the confidence intervals -see the original paper [18] or e.g. a standard encyclopedia entry [37]). Note, that those confidence interval values have to be considered with care. Recall that the bivariate normality of our quantile estimator and measure of dispersion estimator holds asymptotically. Hence, it is not clear if for the sample sizes  $n$  considered we can assume bivariate normality (this could be tested). Still, we provide those theoretical confidence intervals as approximate guidance.

We consider iid samples coming from three different distributions: A Gaussian distribution and two Student distributions with 3 and 5 degrees of freedom, respectively. Recall, the explicit expressions of the asymptotic correlation used to calculate the theoretical values in the tables are presented in Section 3. Also, note that the specification of  $\mu, \sigma^2$  is not needed as the correlation results for location-scale distributions are independent of its parameter (again, see Section 3).

We also recall that when working with the sample standard deviation, the existence of the fourth moment is a condition we require, as well as when working with the location-scale quantile. Thus, as they do not exist for a Student distribution with 3 degrees of freedom, we simply write 'NA' as theoretical value instead.

Table 10: Average values from a 1'000-fold repetition. Comparing the correlation of the measure of dispersion estimator with two different quantile estimators each, as a function of both, the sample size on which the quantile is estimated  $n$  and the length of the time-series  $l$  used to estimate the correlation. Underlying samples are simulated from a Gaussian distribution. Average empirical values are written first (with empirical 95% confidence interval in brackets). The corresponding theoretical values, mean and 95% confidence interval, are provided in the last column. We consider two thresholds  $p = 0.95, 0.99$ . The upper part of the table considers the sample variance as measure of dispersion estimator, the middle part the sample MAD and the lower part the sample MedianAD.

$\widehat{\text{Cor}}(\hat{\sigma}_n^2, \hat{q}_n)$	n=126		n=252		n=504		n=1008		theoretical values	
	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$
$p = 0.95$										
10	52 (-9,88)	100 (100,100)	52 (-11,89)	100 (100,100)	54 (-7,89)	100 (100,100)	53 (-5,89)	100 (100,100)	55 (-12,88)	100(100,100)
25	54 (22,79)	100 (100,100)	54 (23,79)	100 (100,100)	55 (24,79)	100 (100,100)	54 (22,78)	100 (100,100)	55 (20,78)	100(100,100)
50	55 (33,71)	100 (100,100)	55 (34,73)	100 (100,100)	55 (34,73)	100 (100,100)	55 (34,71)	100 (100,100)	55 (32,72)	100(100,100)
100	56 (41,68)	100 (100,100)	55 (41,67)	100 (100,100)	55 (41,68)	100 (100,100)	55 (40,67)	100 (100,100)	55 (40,67)	100(100,100)
250	56 (46,64)	100 (100,100)	55 (46,63)	100 (100,100)	55 (46,63)	100 (100,100)	55 (46,63)	100 (100,100)	55 (46,63)	100(100,100)
500	56 (49,61)	100 (100,100)	55 (49,61)	100 (100,100)	55 (49,61)	100 (100,100)	55 (49,61)	100 (100,100)	55 (49,61)	100(100,100)
$p = 0.99$										
10	44 (-20,87)	100 (100,100)	42 (-22,83)	100 (100,100)	44 (-18,85)	100 (100,100)	42 (-20,86)	100 (100,100)	44 (-26,84)	100 (100,100)
25	47 (10,75)	100 (100,100)	43 (5,72)	100 (100,100)	44 (9,71)	100 (100,100)	44 (6,71)	100 (100,100)	44 (6,71)	100 (100,100)
50	47 (21,67)	100 (100,100)	44 (19,64)	100 (100,100)	45 (20,65)	100 (100,100)	44 (21,64)	100 (100,100)	44 (18,64)	100 (100,100)
100	47 (31,61)	100 (100,100)	44 (28,60)	100 (100,100)	45 (28,59)	100 (100,100)	44 (27,59)	100 (100,100)	44 (27,59)	100 (100,100)
250	47 (37,57)	100 (100,100)	45 (35,54)	100 (100,100)	45 (35,55)	100 (100,100)	44 (35,54)	100 (100,100)	44 (33,54)	100 (100,100)
500	47 (40,54)	100 (100,100)	45 (38,52)	100 (100,100)	45 (38,52)	100 (100,100)	45 (37,52)	100 (100,100)	44 (37,51)	100 (100,100)
$\widehat{\text{Cor}}(\hat{\theta}_n, \hat{q}_n)$	n=126		n=252		n=504		n=1008		theoretical values	
	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$
$p = 0.95$										
10	45 (-21,87)	93 (78,99)	44 (-16,84)	93 (77,99)	47 (-18,86)	93 (76,99)	47 (-17,86)	93 (81,98)	48(-21,85)	94(75,99)
25	47 (13,75)	93 (86,97)	47 (16,73)	94 (87,97)	48 (11,75)	93 (86,97)	48 (13,74)	93 (86,97)	48(11,74)	94(86,97)
50	48 (26,66)	94 (89,96)	48 (26,69)	94 (89,96)	48 (25,69)	94 (90,96)	48 (26,66)	94 (89,96)	48(23,67)	94(89,96)
100	49 (32,62)	94 (91,96)	48 (33,62)	94 (91,96)	48 (33,62)	94 (91,96)	48 (31,62)	94 (91,96)	48(31,62)	94(91,96)
250	49 (39,58)	94 (92,95)	48 (39,58)	94 (92,95)	48 (39,58)	94 (92,95)	48 (39,57)	94 (92,95)	48(38,57)	94(92,95)
500	49 (42,55)	94 (93,95)	48 (41,55)	94 (92,95)	48 (41,55)	94 (92,95)	48 (42,55)	94 (92,95)	48(41,45)	94(92,95)
$p = 0.99$										
10	32 (-37,81)	93 (78,99)	29 (-39,78)	93 (77,99)	32 (-30,79)	93 (76,99)	30 (-36,80)	93 (81,98)	31(-40,79)	94(75,99)
25	34 (-3,66)	93 (86,97)	30 (-11,64)	94 (87,97)	32 (-8,65)	93 (86,97)	31 (-7,63)	93 (86,97)	31(-10,63)	94(86,97)
50	34 (8,57)	94 (89,96)	31 (5,55)	94 (89,96)	32 (5,56)	94 (90,96)	31 (3,54)	94 (89,96)	31(4,54)	94(89,96)
100	34 (17,51)	94 (91,96)	32 (14,50)	94 (91,96)	32 (14,49)	94 (91,96)	31 (13,48)	94 (91,96)	31(12,48)	94(91,96)
250	34 (22,45)	94 (92,95)	32 (21,42)	94 (92,95)	32 (22,44)	94 (92,95)	32 (21,42)	94 (92,95)	31(19,42)	94(92,95)
500	34 (26,41)	94 (93,95)	32 (24,40)	94 (92,95)	32 (25,40)	94 (92,95)	32 (24,39)	94 (92,95)	31(23,39)	94(92,95)
$\widehat{\text{Cor}}(\hat{\xi}_n, \hat{q}_n)$	n=126		n=252		n=504		n=1008		theoretical values	
	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$
$p = 0.95$										
10	21 (-48,75)	59 (5,90)	20 (-42,73)	58 (3,90)	22 (-46,73)	58 (-4,91)	24 (-42,78)	60 (2,90)	23(-47,75)	61(-4,89)
25	22 (-18,60)	60 (29,81)	23 (-14,57)	60 (31,81)	23 (-18,60)	60 (28,82)	22 (-18,57)	60 (27,81)	23(-18,57)	61(28,81)
50	23 (-4,48)	61 (42,76)	23 (-3,48)	60 (40,76)	23 (-4,49)	61 (41,76)	23 (-4,45)	60 (39,76)	23(-5,48)	61(39,76)
100	24 (6,41)	61 (48,72)	23 (4,41)	61 (47,72)	23 (5,43)	61 (47,72)	23 (4,41)	60 (47,72)	23(3,41)	61(47,72)
250	24 (12,35)	61 (53,68)	23 (11,35)	61 (53,68)	23 (11,35)	61 (52,67)	23 (11,34)	61 (52,68)	23(11,34)	61(52,68)
500	24 (15,32)	61 (55,66)	23 (15,32)	61 (55,66)	23 (15,32)	61 (55,66)	23 (15,32)	61 (55,66)	23(14,31)	61(55,66)
$p = 0.99$										
10	11 (-56,72)	59 (5,90)	8 (-57,69)	58 (3,90)	11 (-51,68)	58 (-4,91)	10 (-55,68)	60 (2,90)	10(-56,69)	61(-4,89)
25	12 (-30,49)	60 (29,81)	10 (-30,47)	60 (31,81)	11 (-29,49)	60 (28,82)	10 (-29,46)	60 (27,81)	10(-31,48)	61(28,81)
50	12 (-16,38)	61 (42,76)	10 (-18,37)	60 (40,76)	11 (-16,37)	61 (41,76)	10 (-19,37)	60 (39,76)	10(-18,37)	61(39,76)
100	12 (-8,30)	61 (48,72)	10 (-8,29)	61 (47,72)	11 (-9,29)	61 (47,72)	10 (-10,29)	60 (47,72)	10(-10,29)	61(47,72)
250	12 (-2,23)	61 (53,68)	10 (-2,22)	61 (53,68)	11 (-2,24)	61 (52,67)	10 (-2,22)	61 (52,68)	10(-2,22)	61(52,68)
500	12 (3,19)	61 (55,66)	10 (2,19)	61 (55,66)	11 (2,19)	61 (55,66)	11 (2,19)	61 (55,66)	10(1,19)	61(55,66)

Table 11: Average values from a 1'000-fold repetition. Comparing the correlation of the measure of dispersion estimator with two different quantile estimators each, as a function of both, the sample size on which the quantile is estimated  $n$  and the length of the time-series  $l$  used to estimate the correlation. Underlying samples are simulated from a Student distribution with 5 degrees of freedom. Average empirical values are written first (with empirical 95% confidence interval in brackets). The corresponding theoretical values, mean and 95% confidence interval, are provided in the last column. We consider two thresholds  $p = 0.95, 0.99$ . The upper part of the table considers the sample variance as measure of dispersion estimator, the middle part the sample MAD and the lower part the sample MedianAD.

$\widehat{\text{Cor}}(\hat{\sigma}_n^2, \hat{q}_n)$	n=126		n=252		n=504		n=1008		theoretical values	
	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$
$p = 0.95$										
10	43 (-20,87)	100 (100,100)	41 (-24,85)	100 (100,100)	41 (-23,83)	100 (100,100)	40 (-25,83)	100 (100,100)	38(-33,81)	100 (100,100)
25	42 (0,74)	100 (99,100)	42 (4,72)	100 (100,100)	41 (3,72)	100 (100,100)	40 (3,70)	100 (100,100)	38(-2,67)	100 (100,100)
50	42 (14,67)	100 (99,100)	42 (16,64)	100 (99,100)	41 (13,65)	100 (100,100)	40 (12,63)	100 (100,100)	38(11,59)	100 (100,100)
100	42 (21,59)	100 (98,100)	42 (22,59)	100 (99,100)	40 (20,57)	100 (99,100)	40 (20,56)	100 (100,100)	38(20,53)	100 (100,100)
250	42 (24,54)	99 (97,100)	41 (25,52)	100 (98,100)	40 (26,51)	100 (99,100)	40 (26,50)	100 (100,100)	38(27,48)	100 (100,100)
500	41 (26,51)	99 (96,100)	41 (28,49)	100 (98,100)	40 (28,49)	100 (99,100)	39 (29,48)	100 (99,100)	38(30,45)	100 (100,100)
$p = 0.99$										
10	52 (-14,89)	100 (100,100)	49 (-17,89)	100 (100,100)	47 (-19,87)	100 (100,100)	46 (-21,86)	100 (100,100)	43(-28,83)	100 (100,100)
25	51 (11,81)	100 (99,100)	49 (9,79)	100 (100,100)	47 (8,75)	100 (100,100)	46 (9,75)	100 (100,100)	43(4,70)	100 (100,100)
50	51 (19,75)	100 (99,100)	49 (19,71)	100 (99,100)	47 (19,68)	100 (100,100)	46 (20,67)	100 (100,100)	43(17,63)	100 (100,100)
100	51 (26,69)	100 (98,100)	48 (27,66)	100 (99,100)	47 (25,63)	100 (99,100)	45 (26,61)	100 (100,100)	43(25,57)	100 (100,100)
250	50 (31,64)	99 (97,100)	48 (32,60)	100 (98,100)	46 (31,57)	100 (99,100)	45 (31,56)	100 (100,100)	43(32,52)	100 (100,100)
500	50 (34,60)	99 (96,100)	47 (34,56)	100 (98,100)	46 (34,54)	100 (99,100)	45 (35,53)	100 (99,100)	43(35,50)	100 (100,100)

$\widehat{\text{Cor}}(\hat{\theta}_n, \hat{q}_n)$	n=126		n=252		n=504		n=1008		theoretical values	
	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$
$p = 0.95$										
10	48 (-11,86)	86 (58,98)	48 (-16,87)	84 (49,97)	49 (-12,85)	83 (47,97)	49 (-11,86)	82 (46,97)	51(-18,86)	77(27,94)
25	49 (13,75)	87 (70,95)	50 (16,76)	85 (67,95)	50 (16,77)	83 (62,94)	50 (17,75)	82 (60,93)	51(14,75)	77(53,89)
50	50 (27,71)	86 (73,93)	50 (27,69)	84 (70,92)	50 (27,70)	83 (67,92)	51 (27,69)	82 (64,91)	51(27,69)	77(62,86)
100	51 (35,65)	86 (76,91)	51 (35,65)	84 (72,90)	51 (35,64)	82 (69,90)	51 (36,64)	81 (69,89)	51(35,64)	77(67,84)
250	51 (42,60)	85 (77,90)	51 (42,59)	84 (74,89)	51 (41,59)	82 (73,88)	51 (41,59)	81 (70,87)	51(41,60)	77(71,81)
500	51 (45,58)	85 (79,89)	51 (44,57)	83 (76,87)	51 (44,57)	82 (74,86)	51 (44,57)	81 (73,85)	51(44,57)	77(73,80)
$p = 0.99$										
10	41 (-27,84)	86 (58,98)	40 (-32,84)	84 (49,97)	41 (-19,84)	83 (47,97)	39 (-28,84)	82 (46,97)	40(-30,82)	77(27,94)
25	42 (5,72)	87 (70,95)	41 (-2,73)	85 (67,95)	41 (3,70)	83 (62,94)	41 (5,69)	82 (60,93)	40(1,69)	77(53,89)
50	42 (16,65)	86 (73,93)	41 (13,64)	84 (70,92)	41 (17,62)	83 (67,92)	41 (15,62)	82 (64,91)	40(14,61)	77(62,86)
100	43 (24,59)	86 (76,91)	40 (23,56)	84 (72,90)	41 (25,57)	82 (69,90)	40 (23,55)	81 (69,89)	40(22,55)	77(67,84)
250	42 (31,52)	85 (77,90)	40 (30,51)	84 (74,89)	41 (30,52)	82 (73,88)	41 (30,50)	81 (70,87)	40(29,50)	77(71,81)
500	42 (34,49)	85 (79,89)	40 (33,47)	83 (76,87)	41 (34,48)	82 (74,86)	41 (34,47)	81 (73,85)	40(33,47)	77(73,80)

$\widehat{\text{Cor}}(\hat{\xi}_n, \hat{q}_n)$	n=126		n=252		n=504		n=1008		theoretical values	
	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$
$p = 0.95$										
10	23 (-46,77)	38 (-33,83)	21 (-45,76)	36 (-30,81)	23 (-47,77)	34 (-36,82)	22 (-49,77)	33 (-38,82)	23(-47,75)	32(-39,79)
25	23 (-19,57)	38 (-3,71)	22 (-19,56)	36 (-3,69)	22 (-20,57)	35 (-4,67)	23 (-17,56)	34 (-5,66)	23(-18,57)	32(-9,63)
50	23 (-6,50)	38 (8,62)	23 (-6,47)	36 (7,59)	23 (-6,47)	35 (7,58)	23 (-6,48)	34 (7,58)	23(-5,48)	32(4,55)
100	23 (2,41)	38 (17,54)	23 (4,40)	36 (15,52)	23 (3,41)	35 (16,52)	23 (4,40)	34 (14,51)	23(3,41)	32(13,49)
250	23 (12,35)	37 (23,48)	23 (12,34)	36 (23,47)	23 (11,34)	35 (23,46)	23 (12,35)	34 (22,45)	23(11,34)	32(20,43)
500	24 (15,32)	37 (27,45)	23 (15,31)	36 (27,44)	23 (14,31)	35 (26,43)	23 (15,31)	34 (24,42)	23(14,31)	32(24,40)
$p = 0.99$										
10	11 (-55,71)	38 (-33,83)	11 (-56,69)	36 (-30,81)	12 (-49,69)	34 (-36,82)	10 (-54,68)	33 (-38,82)	10(-56,69)	32(-39,79)
25	12 (-32,52)	38 (-3,71)	11 (-33,49)	36 (-3,69)	10 (-29,48)	35 (-4,67)	10 (-29,49)	34 (-5,66)	10(-31,48)	32(-9,63)
50	11 (-19,40)	38 (8,62)	10 (-19,38)	36 (7,59)	10 (-17,37)	35 (7,58)	10 (-16,38)	34 (7,58)	10(-18,37)	32(4,55)
100	11 (-10,31)	38 (17,54)	10 (-10,28)	36 (15,52)	11 (-8,29)	35 (16,52)	10 (-9,29)	34 (14,51)	10(-10,29)	32(13,49)
250	11 (-2,23)	37 (23,48)	10 (-2,22)	36 (23,47)	10 (-2,22)	35 (23,46)	10 (-2,23)	34 (22,45)	10(-2,22)	32(20,43)
500	11 (2,20)	37 (27,45)	10 (2,18)	36 (27,44)	10 (2,19)	35 (26,43)	10 (2,19)	34 (24,42)	10(1,19)	32(24,40)

Table 12: Average values from a 1'000-fold repetition. Comparing the correlation of the measure of dispersion estimator with two different quantile estimators each, as a function of both, the sample size on which the quantile is estimated  $n$  and the length of the time-series  $l$  used to estimate the correlation. Underlying samples are simulated from a Student distribution with 3 degrees of freedom. Average empirical values are written first (with empirical 95% confidence interval in brackets). The corresponding theoretical values, mean and 95% confidence interval, are provided in the last column (if inexistent, we write 'NA'). We consider two thresholds  $p = 0.95, 0.99$ . The upper part of the table considers the sample variance as measure of dispersion estimator, the middle part the sample MAD and the lower part the sample MedianAD.

$\widehat{\text{Cor}}(\hat{\sigma}_n^2, \hat{q}_n)$	n=126		n=252		n=504		n=1008		theoretical values	
	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$
$p = 0.95$										
10	31 (-32,82)	100 (99,100)	27 (-40,80)	100 (99,100)	22 (-43,77)	100 (100,100)	21 (-45,74)	100 (100,100)	NA	NA
25	28 (-15,68)	99 (98,100)	24 (-18,64)	99 (99,100)	21 (-19,56)	100 (99,100)	19 (-24,56)	100 (99,100)	NA	NA
50	25 (-8,55)	99 (97,100)	22 (-9,52)	99 (98,100)	19 (-9,47)	99 (98,100)	17 (-14,44)	100 (99,100)	NA	NA
100	23 (-2,47)	98 (96,100)	20 (-4,42)	99 (97,100)	17 (-5,38)	99 (97,100)	15 (-7,37)	99 (98,100)	NA	NA
250	20 (0,37)	97 (94,99)	17 (-1,34)	97 (94,99)	15 (-2,31)	98 (96,100)	13 (-4,28)	98 (96,100)	NA	NA
500	18 (2,31)	95 (91,99)	16 (1,29)	96 (92,99)	13 (0,26)	97 (93,99)	12 (-1,23)	98 (94,99)	NA	NA
$p = 0.99$										
10	45 (-27,93)	100 (99,100)	42 (-32,88)	100 (99,100)	35 (-33,84)	100 (100,100)	31 (-35,82)	100 (100,100)	NA	NA
25	43 (-2,82)	99 (98,100)	38 (-7,77)	99 (99,100)	31 (-14,70)	100 (99,100)	28 (-16,67)	100 (99,100)	NA	NA
50	40 (-1,76)	99 (97,100)	35 (0,70)	99 (98,100)	29 (-5,59)	99 (98,100)	26 (-7,57)	100 (99,100)	NA	NA
100	37 (4,69)	98 (96,100)	32 (2,59)	99 (97,100)	27 (0,50)	99 (97,100)	24 (-1,47)	99 (98,100)	NA	NA
250	33 (6,60)	97 (94,99)	28 (4,50)	97 (94,99)	23 (2,43)	98 (96,100)	20 (1,38)	98 (96,100)	NA	NA
500	30 (6,52)	95 (91,99)	26 (5,44)	96 (92,99)	21 (1,37)	97 (93,99)	18 (1,33)	98 (94,99)	NA	NA

$\widehat{\text{Cor}}(\hat{\theta}_n, \hat{q}_n)$	n=126		n=252		n=504		n=1008		theoretical values	
	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$
$p = 0.95$										
10	47 (-17,87)	84 (53,97)	45 (-19,87)	81 (46,97)	45 (-20,86)	77 (35,96)	45 (-26,87)	73 (29,95)	48(-22,85)	NA
25	48 (10,76)	84 (66,95)	47 (10,74)	80 (57,92)	46 (8,73)	76 (53,91)	47 (13,75)	71 (44,89)	48(10,73)	NA
50	48 (21,68)	83 (70,92)	47 (23,67)	78 (64,89)	47 (20,68)	74 (56,87)	47 (23,67)	70 (49,84)	48(23,67)	NA
100	48 (29,62)	82 (73,90)	48 (30,62)	77 (66,87)	47 (29,62)	73 (60,83)	47 (31,61)	68 (53,80)	48(31,61)	NA
250	48 (33,57)	81 (76,87)	47 (35,57)	76 (69,83)	47 (37,57)	71 (63,79)	47 (36,57)	66 (56,76)	48(37,57)	NA
500	47 (37,55)	81 (77,86)	48 (38,55)	76 (70,82)	47 (39,55)	70 (64,77)	47 (40,54)	65 (57,72)	48(41,54)	NA
$p = 0.99$										
10	45 (-19,87)	84 (53,97)	43 (-19,85)	81 (46,97)	43 (-19,86)	77 (35,96)	44 (-23,85)	73 (29,95)	44(-26,84)	NA
25	46 (11,76)	84 (66,95)	44 (5,74)	80 (57,92)	44 (6,73)	76 (53,91)	45 (9,71)	71 (44,89)	44(6,71)	NA
50	46 (20,70)	83 (70,92)	45 (20,67)	78 (64,89)	45 (19,65)	74 (56,87)	45 (20,65)	70 (49,84)	44(19,64)	NA
100	46 (27,64)	82 (73,90)	45 (27,61)	77 (66,87)	45 (28,59)	73 (60,83)	44 (27,59)	68 (53,80)	44(27,59)	NA
250	46 (34,58)	81 (76,87)	45 (32,55)	76 (69,83)	45 (34,55)	71 (63,79)	45 (35,54)	66 (56,76)	44(34,54)	NA
500	46 (36,55)	81 (77,86)	45 (36,53)	76 (70,82)	45 (37,52)	70 (64,77)	45 (37,52)	65 (57,72)	44(37,51)	NA

$\widehat{\text{Cor}}(\hat{\xi}_n, \hat{q}_n)$	n=126		n=252		n=504		n=1008		theoretical values	
	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$	$q_n$	$q_{n,\hat{\sigma}}$
$p = 0.95$										
10	22 (-48,76)	22 (-49,78)	20 (-51,75)	18 (-52,74)	21 (-46,76)	16 (-53,73)	21 (-42,74)	13 (-53,74)	23(-47,75)	NA
25	23 (-15,58)	21 (-20,59)	22 (-21,59)	18 (-27,55)	21 (-20,58)	15 (-25,52)	22 (-20,57)	13 (-29,52)	23(-18,57)	NA
50	23 (-4,49)	21 (-9,47)	22 (-7,48)	17 (-14,45)	22 (-7,47)	15 (-13,41)	23 (-7,49)	12 (-20,39)	23(-5,48)	NA
100	23 (4,41)	20 (-2,40)	22 (2,40)	16 (-7,36)	23 (2,40)	14 (-8,34)	23 (2,41)	11 (-10,32)	23(3,41)	NA
250	23 (11,35)	18 (4,31)	23 (10,34)	15 (1,28)	23 (10,34)	13 (-2,27)	23 (11,34)	11 (-3,25)	23(11,34)	NA
500	23 (14,31)	18 (7,28)	23 (14,31)	15 (4,25)	23 (14,31)	13 (1,23)	23 (15,31)	11 (0,20)	23(14,31)	NA
$p = 0.99$										
10	11 (-53,71)	22 (-49,78)	7 (-58,71)	18 (-52,74)	9 (-58,74)	16 (-53,73)	9 (-59,70)	13 (-53,74)	10(-56,69)	NA
25	10 (-29,51)	21 (-20,59)	9 (-31,48)	18 (-27,55)	10 (-32,47)	15 (-25,52)	10 (-29,48)	13 (-29,52)	10(-31,48)	NA
50	10 (-19,38)	21 (-9,47)	9 (-20,35)	17 (-14,45)	10 (-20,39)	15 (-13,41)	10 (-19,38)	12 (-20,39)	10(-18,37)	NA
100	10 (-11,30)	20 (-2,40)	9 (-11,28)	16 (-7,36)	10 (-10,31)	14 (-8,34)	10 (-9,29)	11 (-10,32)	10(-10,29)	NA
250	10 (-2,22)	18 (4,31)	9 (-3,21)	15 (1,28)	10 (-3,22)	13 (-2,27)	10 (-2,22)	11 (-3,25)	10(-2,22)	NA
500	10 (1,19)	18 (7,28)	9 (0,18)	15 (4,25)	10 (1,19)	13 (1,23)	10 (2,19)	11 (0,20)	10(1,19)	NA

**ESSEC Business School**

3 avenue Bernard-Hirsch  
CS 50105 Cergy  
95021 Cergy-Pontoise Cedex  
France  
Tel. +33 (0)1 34 43 30 00  
[www.essec.edu](http://www.essec.edu)

**ESSEC Executive Education**

CNIT BP 230  
92053 Paris-La Défense  
France  
Tel. +33 (0)1 46 92 49 00  
[www.executive-education.essec.edu](http://www.executive-education.essec.edu)

**ESSEC Asia-Pacific**

5 Nepal Park  
Singapore 139408  
Tel. +65 6884 9780  
[www.essec.edu/asia](http://www.essec.edu/asia)

ESSEC | CPE Registration number 200511927D  
Period of registration: 30 June 2017 - 29 June 2023  
Committee of Private Education (CPE) is part of SkillsFuture Singapore (SSG)

**ESSEC Africa**

Plage des Nations - Golf City  
Route de Kénitra - Sidi Bouknadel (Rabat-Salé)  
Morocco  
Tel. +212 (0)5 37 82 40 00  
[www.essec.edu](http://www.essec.edu)

**CONTACT**

**Centre de Recherche**  
Tel. +33 (0)1 34 43 30 91  
[research.center@essec.fr](mailto:research.center@essec.fr)