VaR and CVaR optimization with Large deviation results (working)

Ryan Dai
Quantitative Finance, National Tsing Hua University

October 3, 2025

Abstract

This paper aims to provide a point of view of VaR and CVaR optimization problem with large deviation results. Fist, we provide the dual problem of VaR and CVaR optimization under normal return assumption using basic Calculus and Statistics knowledge. Second, we consider an order statistics which converges to VaR (α-quantile) and give some large deviation results. Third, we use similar way to describe CVaR via L-statistcs.

1 VaR and CVaR optimization via basic Calculus and Statistics

Consider a single period $t \in \{0, 1\}$ with stock prices $\mathbf{S}(t) = [S_1(t)\ S_2(t)\ \dots\ S_n(t)]^\top$ and portfolio value $P_t = \mathbf{N}^\top \mathbf{S}(t)$ where position $\mathbf{N} = [N_1, \dots, N_n]^\top$. Define:

- Portfolio weights $\mathbf{w} = [w_1, \dots, w_n]^\top$ where $w_i = \frac{N_i S_i(0)}{P_0}$.
- Simple returns $\mathbf{r} = [r_1, \dots, r_n]^\top$ where $r_i = \frac{S_i(1) - S_i(0)}{S_i(0)}$.
- Portfolio simple return

$$r_p = \frac{P_1 - P_0}{P_0} = \frac{\mathbf{N}^\top(\mathbf{S}(1) - \mathbf{S}(0))}{P_0} = \sum_{i=1}^{n} \frac{N_i[S_i(1) - S_i(0)]}{P_0} \frac{S_i(0)}{S_i(0)} = \sum_{i=1}^{n} w_i r_i = \mathbf{w}^\top \mathbf{r}.$$

Assume that $r_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ for all i, then we have $\mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ where $\boldsymbol{\mu} = [\mu_1, \dots, \mu_n]^\top$, $\Sigma = \mathbb{E}[(\mathbf{r} - \boldsymbol{\mu})(\mathbf{r} - \boldsymbol{\mu})^\top]$, and

$$r_p = \mathbf{w}^\top \mathbf{r} \sim \mathcal{N}(\mu_p, \sigma_p^2), \quad \mu_p = \mathbf{w}^\top \boldsymbol{\mu},\ \sigma_p^2 = \mathbf{w}^\top \Sigma \mathbf{w}.$$

1.1 VaR optimization

Define $\mathrm{VaR}_\alpha = \inf\{l : \mathbb{P}(P_0 - P_1 \geq l) \leq 1 - \alpha\}$ with $\mathbb{P}(P_0 - P_1 \geq \mathrm{VaR}_\alpha) = 1 - \alpha$. Then

$$\begin{aligned}
\mathbb{P}\left(\frac{P_1 - P_0}{P_0} \leq \frac{-\mathrm{VaR}_\alpha}{P_0}\right) &= \mathbb{P}\left(r_p \leq \frac{-\mathrm{VaR}_\alpha}{P_0}\right) \\
&= \mathbb{P}\left(\frac{r_p - \mu_p}{\sigma_p} \leq \frac{\frac{-\mathrm{VaR}_\alpha}{P_0} - \mu_p}{\sigma_p}\right) \\
&= \Phi\left(\frac{\frac{-\mathrm{VaR}_\alpha}{P_0} - \mu_p}{\sigma_p}\right) = 1 - \alpha,
\end{aligned}$$

where $\Phi(\cdot)$ is the standard normal CDF. Hence

$$\frac{\frac{-\mathrm{VaR}_\alpha}{P_0} - \mu_p}{\sigma_p} = \Phi^{-1}(1 - \alpha) \quad \Rightarrow \quad \mathrm{VaR}_\alpha = P_0[-\mu_p - \Phi^{-1}(1 - \alpha)\sigma_p].$$

For $\alpha > 0.5$, $\Phi^{-1}(1-\alpha) < 0$, so

$$\begin{aligned}
\arg\min_{\mathbf{w}} \mathrm{VaR}_\alpha &= \arg\min_{\mathbf{w}}[-\mu_p - \Phi^{-1}(1-\alpha)\sigma_p] \\
&= \arg\max_{\mathbf{w}}[\mu_p + \Phi^{-1}(1-\alpha)\sigma_p] \\
&= \arg\min_{\mathbf{w}}\left[\frac{\mu_p}{\Phi^{-1}(1-\alpha)} + \sigma_p\right].
\end{aligned}$$

Since $\lim_{\alpha\uparrow 1} \frac{\mu_p}{\Phi^{-1}(1-\alpha)} = 0$, we get $\arg\min_{\mathbf{w}} \mathrm{VaR}_\alpha = \arg\min_{\mathbf{w}} \sigma_p$ as $\alpha \uparrow 1$.

1.2 CVaR optimization

Define $\mathrm{CVaR}_\alpha = \mathbb{E}[P_0 - P_1 \mid P_0 - P_1 \geq \mathrm{VaR}_\alpha]$. From above, $\mathrm{VaR}_\alpha = P_0[-\mu_p - \Phi^{-1}(1-\alpha)\sigma_p]$. Therefore

$$\begin{aligned}
\mathrm{CVaR}_\alpha &= \mathbb{E}[P_0 - P_1 \mid P_0 - P_1 \geq P_0(-\mu_p - \Phi^{-1}(1-\alpha)\sigma_p)] \\
&= -P_0\,\mathbb{E}\left[\frac{P_1 - P_0}{P_0} \,\middle|\, \frac{P_1 - P_0}{P_0} \leq \mu_p + \Phi^{-1}(1-\alpha)\sigma_p\right] \\
&= -P_0\,\mathbb{E}[r_p \mid r_p \leq \mu_p + \Phi^{-1}(1-\alpha)\sigma_p] \\
&= -P_0\,\mathbb{E}\left[r_p \,\middle|\, \frac{r_p - \mu_p}{\sigma_p} \leq \Phi^{-1}(1-\alpha)\right] \\
&= -P_0\sigma_p\,\mathbb{E}\left[\frac{r_p - \mu_p}{\sigma_p} \,\middle|\, \frac{r_p - \mu_p}{\sigma_p} \leq \Phi^{-1}(1-\alpha)\right] - P_0\mu_p \\
&= -P_0\sigma_p\,\frac{1}{1-\alpha}\int_{-\infty}^{\Phi^{-1}(1-\alpha)} x\,\phi(x)\,dx - P_0\mu_p,
\end{aligned}$$

where $\phi(\cdot)$ is the standard normal PDF. Noting $\frac{d}{dx}\phi(x) = -x\phi(x)$,

$$\begin{aligned}
\Rightarrow\ \mathrm{CVaR}_\alpha &= -\frac{P_0\sigma_p}{1-\alpha}\left[-\phi(x)\right]_{-\infty}^{\Phi^{-1}(1-\alpha)} - P_0\mu_p \\
&= P_0\left[\frac{\phi(\Phi^{-1}(1-\alpha))}{1-\alpha}\,\sigma_p - \mu_p\right].
\end{aligned}$$

Thus

$$\begin{aligned}
\arg\min_{\mathbf{w}} \mathrm{CVaR}_\alpha &= \arg\min_{\mathbf{w}}\left[\frac{\phi(\Phi^{-1}(1-\alpha))}{1-\alpha}\,\sigma_p - \mu_p\right] \\
&= \arg\min_{\mathbf{w}}\left[\sigma_p - \frac{1-\alpha}{\phi(\Phi^{-1}(1-\alpha))}\,\mu_p\right].
\end{aligned}$$

By L'Hôpital's rule,

$$\lim_{\alpha\uparrow 1}\frac{1-\alpha}{\phi(\Phi^{-1}(1-\alpha))} = \lim_{\alpha\uparrow 1}\frac{-1}{\Phi^{-1}(1-\alpha)} = 0,$$

hence $\arg\min_{\mathbf{w}} \mathrm{CVaR}_\alpha = \arg\min_{\mathbf{w}} \sigma_p$ as $\alpha \uparrow 1$.

1.3 Conclusion

Under the one-period normal-return assumption,

$$\arg\min_{\mathbf{w}} \mathrm{VaR}_\alpha = \arg\min_{\mathbf{w}} \mathrm{CVaR}_\alpha = \arg\min_{\mathbf{w}} \sigma_p \quad \text{as } \alpha \uparrow 1.$$

In this section, Var and CVaR are regarded as quantile and conditional expectation, respectively, derived form a distribution(Normal distribution). In the next section, we will use specific random variables to describe them and derive some large deviation results.

2 VaR Optimization

By previous definition, we know that VaR_α is a quantile derived from the distribution of portfolio loss. In this section, we consider an order statistics which converges to the VaR_α, or α-quantile, and optimize its relative probability to construct a resilient portfolio optimization problem.

2.1 Large devaition principle for Order statistics

In [1] and [2], we have the following theoretical results.

Condition 2.1. *(i)* $\{X_n : n \geq 1\}$ *is a sequence of i.i.d. random variables with CDF* F *continuous and strictly increasing on* (a,b)*, where* $-\infty \leq a < b \leq \infty$.

(ii) $\{k_n : n \geq 1\}$ *is a sequence satisfying* $k_n \in \{1,\ldots,n\}$ *and* $\lim_{n\to\infty} \frac{k_n}{n} = \alpha \in (0,1)$*. For example, taking* $k_n = \lfloor 0.99n \rfloor$ *yields* $\alpha = 0.99$*. Here* k_n *is the order and* α *is the quantile, and they align as* $n \to \infty$.

Proposition 2.1. *Under Condition 2.1, let* $X_{1:n} \leq \cdots \leq X_{n:n}$ *be order statistics, and consider a sequence* $\{X_{k_n:n} : n \geq 1\}$*. Then* $\{X_{k_n:n} : n \geq 1\}$ *satisfies an LDP with rate function*

$$I_{\alpha,F}(x) = H(\alpha \mid F(x)), \quad x \in (a,b),$$

where $H(p \mid q) = p \log \frac{p}{q} + (1-p)\log\frac{1-p}{1-q}$ *for* $p, q \in (0,1)$*. That is, for* $x > F^{-1}(\alpha)$*, we have*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X_{k_n:n} \geq x) = -I_{\alpha,F}(x).$$

More intuitively,

$$\mathbb{P}(X_{k_n:n} \geq x) \approx e^{-n\cdot I_{\alpha,F}(x)} \text{ for } n \text{ large.}$$

2.2 Using large deviations for VaR optimization

In section 1, we define $\mathrm{VaR}_\alpha = \inf\{l : \mathbb{P}(P_0 - P_1 \geq l) \leq 1-\alpha\}$. Here, we let X be a random variable representing the portfolio loss with CDF F and assume Condition 2.1. Then we can define

$$\mathrm{VaR}_\alpha = \inf\{l : \mathbb{P}(X \geq l) \leq 1-\alpha\} = \inf\{l : F(l) \geq \alpha\} = F^{-1}(\alpha).$$

By Proposition 2.1, for $x > F^{-1}(\alpha) = \mathrm{VaR}_\alpha$, we have

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(X_{k_n:n} \geq x) = -I_{\alpha,F}(x) = -H(\alpha \mid F(x)).$$

Rather than minimizing VaR_α, here we minimize the probability where the empirical VaR (order statistics) exceeds some large threshold, that is minimize $\mathbb{P}(X_{k_n:n} \geq x)$ and we get, for n large,

$$\begin{aligned}\arg\min_{\mathbf{w}} \mathbb{P}(X_{k_n:n} \geq x) &= \arg\min_{\mathbf{w}} \frac{1}{n} \log \mathbb{P}(X_{k_n:n} \geq x) \\ &= \arg\max_{\mathbf{w}} H(\alpha \mid F(x)).\end{aligned}$$

Note that $F(x)$ depends on $\mathbf{w}$. Since $x > F^{-1}(\alpha)$ implies $F(x) > \alpha$ and note that

$$\frac{\partial}{\partial q} H(\alpha \mid q) = \frac{q-\alpha}{q(1-q)},$$

so for $q > \alpha$ the mapping

$$q \longmapsto H(\alpha \mid q)$$

is strictly increasing. Thus

$$\arg\min_{\mathbf{w}} \mathbb{P}(X_{k_n:n} \geq x) = \arg\max_{\mathbf{w}} H(\alpha \mid F(x)) = \arg\max_{\mathbf{w}} F(x) \text{ for n large.}$$

Example 2.1. *(Normal distribution) If* $X \sim \mathcal{N}(-\mu_p, \sigma_p^2)$ *with* $\mu_p = \mathbf{w}^\top \boldsymbol{\mu}$ *and* $\sigma_p^2 = \mathbf{w}^\top \Sigma \mathbf{w}$*, then*

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}\left(\frac{X+\mu_p}{\sigma_p} \leq \frac{x+\mu_p}{\sigma_p}\right) = \Phi\left(\frac{x+\mu_p}{\sigma_p}\right),$$

and

$$\arg\min_{\mathbf{w}} \mathbb{P}(X_{k_n:n} \geq x) = \arg\max_{\mathbf{w}} F(x) = \arg\max_{\mathbf{w}} \Phi\left(\frac{x+\mu_p}{\sigma_p}\right) = \arg\max_{\mathbf{w}} \frac{x+\mu_p}{\sigma_p}.$$

For α close to 1, we have $x > 0$ and

$$\arg\min_{\mathbf{w}} \mathbb{P}(X_{k_n:n} \geq x) = \arg\max_{\mathbf{w}} \frac{x + \mu_p}{\sigma_p} = \arg\max_{\mathbf{w}} \frac{1}{\sigma_p} + \frac{\mu_p}{x\sigma_p}.$$

Since $\lim_{\alpha\uparrow 1} VaR_\alpha = \infty$ and $\lim_{x\uparrow\infty} \frac{\mu_p}{x\sigma_p} = 0$, we get

$$\arg\min_{\mathbf{w}} \mathbb{P}(X_{k_n:n} \geq x) \approx \arg\min_{\mathbf{w}} \sigma_p$$

for α very close to 1 .

3 CVaR Optimization

Similar to the previous section, we know CVaR is a conditional expectation and we will use L-statistics to represent it. Assume that the portfolio loss is a random variable X with pdf f and cdf F. We can define

$$\mathrm{VaR}_\alpha = \inf\{L : P(X \geq L) \leq 1-\alpha\} \quad \text{for } \alpha \in (0,1)$$

and

$$\mathrm{CVaR}_\alpha = E[X \mid X > \mathrm{VaR}_\alpha]$$

By definition of conditional expectation, for discrete random variable X,

$$\begin{aligned} E[X \mid X > \mathrm{VaR}_\alpha] &= \sum_x x \cdot P(X = x \mid X > \mathrm{VaR}_\alpha) \\ &= \sum_x x \cdot \frac{P(\{X = x\} \cap \{X > \mathrm{VaR}_\alpha\})}{P(X > \mathrm{VaR}_\alpha)} \end{aligned}$$

Note that, by definition of VaR_α, we have $P(X > \mathrm{VaR}_\alpha) = 1-\alpha$ and $\mathrm{VaR}_\alpha = F^{-1}(\alpha)$. Then

$$E[X \mid X > \mathrm{VaR}_\alpha] = \frac{1}{1-\alpha} \sum_x x \cdot P(\{X = x\} \cap \{X \geq \mathrm{VaR}_\alpha\})$$

For continuous case, we can write

$$E[X \mid X \geq \mathrm{VaR}_\alpha] = \frac{1}{1-\alpha} \int_{\mathbb{R}} x \cdot f(x) \cdot I\{x > \mathrm{VaR}_\alpha\}\, dx$$

where I is indicator function. Moreover, let $u = F(x)$.

$$\begin{aligned} E[X \mid X \geq \mathrm{VaR}_\alpha] &= \frac{1}{1-\alpha} \int_{\mathrm{VaR}_\alpha}^{\infty} x \cdot f(x)\, dx \\ &= \frac{1}{1-\alpha} \int_{F(\alpha)}^{\infty} x \cdot f(x)\, dx \\ &= \frac{1}{1-\alpha} \int_{\alpha}^{1} F^{-1}(u)\, du \\ &= \int_0^1 \frac{I\{u > \alpha\}}{1-\alpha} F^{-1}(u)\, du \end{aligned}$$

In [3], the author gives the large deviation results for L-statistics in the form

$$\sum_{i=1}^{n} c_{i,n} \cdot X_{i:n}$$

where $c_{i,n}$ corresponds to $\frac{I\{u>\alpha\}}{1-\alpha}$ and $X_{i:n}$ corresponds to $F^{-1}(u)$.

Proposition 3.1 (Non-measure-theoretic statement)**.** *Let X be a real-valued r.v. with cdf F and pdf f. Denote by $\mathrm{CVaR}_\alpha(F)$ the α-CVaR of X for $\alpha \in (0,1)$. For $c > \mathrm{CVaR}_\alpha(F)$ define the constraint set*

$$\Omega_c = \Big\{ G : \ G \textit{ is a cdf with pdf } g \textit{ and } \mathrm{CVaR}_\alpha(G) \geq c \Big\}.$$

Let $X_{1:n} \le \cdots \le X_{n:n}$ be the order statistics of an i.i.d. sample from F. Consider the empirical CVaR *at level α,*

$$\widehat{\mathrm{CVaR}}_\alpha^{(n)} \;=\; \frac{1}{n-\lfloor n\alpha \rfloor} \sum_{i=1}^{n} \mathbf{1}\{i > \lfloor n\alpha \rfloor\}\; X_{i:n},$$

namely the average of the top $(1-\alpha)$-fraction of order statistics.

Under suitable regularity conditions,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\Big(\widehat{\mathrm{CVaR}}_\alpha^{(n)} \ge c\Big) \;=\; -\inf_{G\in\Omega_c} K(G\,\|\,F)\,,$$

where the Kullback–Leibler divergence is

$$K(G\,\|\,F) \;=\; \int_{\mathbb{R}} g(x)\, \log \frac{g(x)}{f(x)}\, \mathrm{d}x.$$

Remark. The statistic $\frac{1}{n-\lfloor n\alpha\rfloor}\sum_{i=1}^{n} \mathbf{1}\{i > \lfloor n\alpha\rfloor\}\; X_{i:n}$ is the sample mean of the upper $(1-\alpha)$ proportion of order statistics, hence it is a natural empirical version of $\mathrm{CVaR}_\alpha(F)$. The theoretical proof of this proposition is too technical and hard to understand. Thus we conduct some experiments to verify this convergence.

Example 3.1 (Gaussian case)**.** *Assume $F = \mathcal{N}(0,1)$ and restrict G to the Gaussian family $G = \mathcal{N}(\mu,\sigma^2)$, determined by (μ,σ). The following optimization problems are equivalent:*

$$\text{(i)} \quad \min_{G}\; K(G\,\|\,F) \quad \text{s.t.} \quad \mathrm{CVaR}_\alpha(G) \ge c,$$

$$\text{(ii)} \quad \min_{\mu,\sigma>0}\; \Big[\log(1/\sigma) + \tfrac{1}{2}(\mu^2+\sigma^2) - \tfrac{1}{2}\Big] \quad \text{s.t.} \quad \mu + k_\alpha\,\sigma = c,$$

$$k_\alpha \;=\; \frac{\varphi(z_\alpha)}{1-\alpha}, \qquad z_\alpha := \Phi^{-1}(\alpha),$$

$$\text{(iii)} \quad \min_{\sigma>0}\; \Big[\log(1/\sigma) + \tfrac{1}{2}\big(\sigma^2 + c^2 + k_\alpha^2\sigma^2 - 2ck_\alpha\sigma\big) - \tfrac{1}{2}\Big].$$

The objective in (iii) is convex in σ, so the problem admits a unique closed-form optimizer $(\mu^\star,\sigma^\star)$. (see Appendix)

Fix n and repeat M independent experiments. In the m-th replication, draw n i.i.d. samples from $\mathcal{N}(0,1)$, form the order statistics $X_{1:n}^{(m)} \le \cdots \le X_{n:n}^{(m)}$, and compute

$$\widehat{\mathrm{CVaR}}_\alpha^{(m)} \;=\; \frac{1}{n-\lfloor n\alpha \rfloor} \sum_{j=1}^{n} \mathbf{1}\{j > \lfloor n\alpha \rfloor\}\; X_{j:n}^{(m)}.$$

With c chosen (e.g. $c = \mathrm{CVaR}_{0.99}(F)$ in the Gaussian example), estimate the probability by

$$\widehat{p} \;=\; \frac{1}{M} \sum_{m=1}^{M} \mathbf{1}\Big\{\widehat{\mathrm{CVaR}}_\alpha^{(m)} \ge c\Big\}.$$

Finally, in the experiment, we fix the number of replications M (e.g., $M = 8000$) and then increase the sample size n (e.g., $n = 100, 200, \ldots, N_{\max}$). For each n, we estimate

$$\widehat{p}_n \;=\; \frac{1}{M} \sum_{m=1}^{M} \mathbf{1}\Big\{\widehat{\mathrm{CVaR}}_\alpha^{(m)} \ge c\Big\},$$

and record

$$y_n \;=\; \frac{1}{n} \log \widehat{p}_n.$$

According to the proposition, we expect y_n to converge to the theoretical rate:

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\Big(\widehat{\mathrm{CVaR}}_\alpha^{(n)} \ge c\Big) \;=\; -I^\star, \qquad I^\star \;:=\; \inf_{G\in\Omega_c} K(G\,\|\,F)\,.$$

In the Gaussian example, $I^\star$ equals the value of (iii) at the optimizer $\sigma^\star$ (with $\mu^\star = c - k_\alpha\sigma^\star$):

$$I^\star \;=\; \log \frac{1}{\sigma^\star} + \frac{1}{2}\Big(\sigma^{\star 2} + (c - k_\alpha\sigma^\star)^2\Big) - \frac{1}{2}.$$

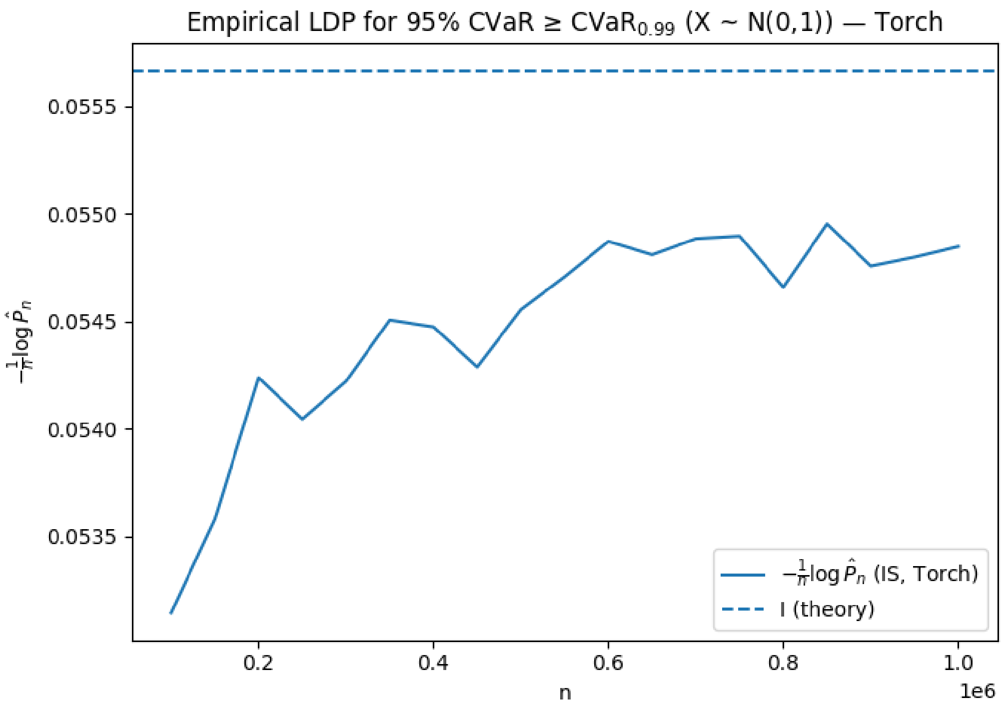

Figure 1: With M fixed, $\frac{1}{n}\log\widehat{p}_n$ (line) versus the theoretical value $-I^{\star}$ (horizontal dashed line) as n increases.

References

[1] Enkelejd Hashorva, Claudio Macci, and Barbara Pacchiarotti (2013). *Large Deviations for Proportions of Observations Which Fall in Random Sets Determined by Order Statistics*. Methodology and Computing in Applied Probability, 15:875–896. doi:10.1007/s11009-012-9290-y.

[2] Valeria Bignozzi, Claudio Macci, and Lea Petrella (2020). *Large deviations for method-of-quantiles estimators of one-dimensional parameters*. Communications in Statistics – Theory and Methods, 49(5):1132–1157. doi:10.1080/03610926.2018.1554134.

[3] P. Groeneboom (1980). *Large Deviations and Asymptotic Efficiencies*. Mathematical Centre Tracts, No. 118. Mathematisch Centrum, Amsterdam. ISBN 90-6196-190-4.