Short Course
Theory and Practice of Risk Measurement

Part 1
Introduction to Risk Measures and Regulatory Capital

Ruodu Wang

Department of Statistics and Actuarial Science
University of Waterloo, Canada

UNIVERSITY OF WATERLOO

Email: wang@uwaterloo.ca

Contents

References

The main reference books are

(i) Föllmer, H. and Schied, A. (2011). *Stochastic Finance: An Introduction in Discrete Time.* Third Edition. Walter de Gruyter.

(ii) Delbaen, F. (2012). *Monetary Utility Functions.* Osaka University Press.

(iii) McNeil, A. J., Frey, R. and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools.* Revised Edition. Princeton University Press.

- You are not required to purchase those books.
- A list of relevant papers will be provided.

Notes

- The depth of the topics will be at the level of recent research advances.
- Preliminary knowledge on (graduate level) probability theory and mathematical statistics is expected.
- Preliminary knowledge on (graduate level) mathematical finance is expected.
- This is a course in mathematics.

Notes

My website:

http://sas.uwaterloo.ca/~wang

(materials will be posted under the teaching tab)

Reference list for this course:

http://sas.uwaterloo.ca/~wang/teaching/S2016/References.pdf

Part 1

- Risk measures and regulatory capital
- Value-at-Risk and Expected Shortfall
- Current debates (2013 - 2016) in regulation
- Basic Extreme Value Theory (EVT) for VaR and ES
- Estimation and modeling issues

Introduction

Key question in mind

A financial institution has a risk (random loss) X in a fixed period. How much capital should this financial institution reserve in order to undertake this risk?

- X can be market risks, credit risks, operational risks, insurance risks, etc.

Risks

In this course, risks are represented by random variables. The realization of a risk is loss/profit.

Some standard notation: fix an atomless probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

- $\mathbb{P}$-a.s. equal random variables are treated as identical
- L^p, $p \in [0, \infty)$ is the set of random variables with finite p-th moment
- L^∞ is the set of bounded random variables

Risks

Let $\mathcal{X} \supset L^\infty$ be a convex cone (closed under addition and $\mathbb{R}^+$-multiplication)

- $\mathcal{X}$ is the set of all risks that we are interested in.

For different types of risks, the time horizon of the measurement procedure might be different: it can be, for instance,

- 1 day (market risk)
- 10 days (liability risk)
- 1 year (credit and operational risks)
- 20 years (life-insurance risk)

In this course we do not specify the type of risks and discuss the theory of risk measurement with generality.

Risk Measures

Risk measures

A risk measure calculates the amount of regulatory capital of a financial institution taking a risk (random loss) X in a fixed period.

A risk measure is a functional $\rho : \mathcal{X} \to (-\infty, \infty]$.

- Typically one requires $\rho(L^\infty) \subset \mathbb{R}$ for obvious reasons.
- In most technical parts of this course, we take $\mathcal{X} = L^\infty$ for convenience.

Some General Mindset

Question

What is a good risk measure to use?

- Regulator's and firm manager's perspectives can be different or even conflicting
  - taxpayers versus shareholders
  - systemic risk in an economy versus risk of a single firm
- How does one know about X?
  - Typically through a model/distribution: simulation, parametric models, expert opinion, ...
  - Information asymmetry, model misspecification, data sparsity, random errors ...

Example: VaR

$p \in (0,1)$, $X \sim F$.

Definition (Value-at-Risk)

$\mathrm{VaR}_p : L^0 \to \mathbb{R}$,

$$\mathrm{VaR}_p(X) = F^{-1}(p) = \inf\{x \in \mathbb{R} : F(x) \geq p\}.$$

In practice, the choice of p is typically close to 1.

- Book on VaR: Jorion (2006)

Proposition

For $X \in L^0$, $\mathrm{VaR}_p(X)$ is increasing[1] *in $p \in (0,1)$.*

[1] In this course, the term "increasing" is in the non-strict sense.

Example: ES

$p \in (0,1)$.

Definition (Expected Shortfall (TVaR, CVaR, CTE, WCE))

$\mathrm{ES}_p : L^0 \to (-\infty, \infty]$,

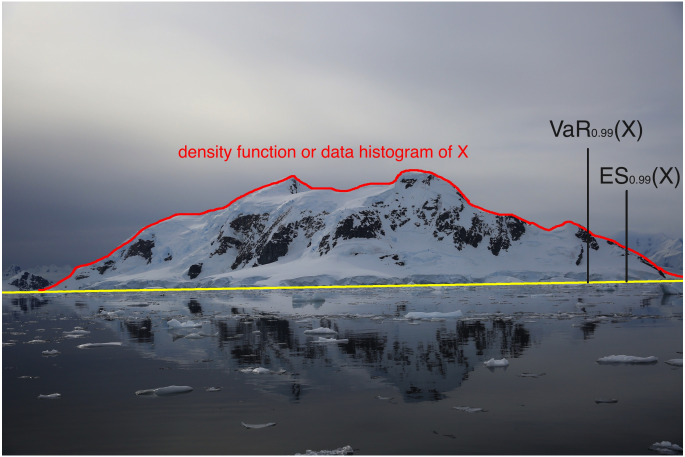
$$\mathrm{ES}_p(X) = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_q(X)\mathrm{d}q \underset{(F \text{ cont.})}{=} \mathbb{E}\left[X | X > \mathrm{VaR}_p(X)\right].$$

In addition, let $\mathrm{VaR}_1(X) = \mathrm{ES}_1(X) = \text{ess-sup}(X)$, and $\mathrm{ES}_0(X) = \mathbb{E}[X]$ (ES_0 is only well-defined on e.g. L^1 or L^0_+).

Proposition

For $X \in L^0$, $\mathrm{ES}_p(X)$ is increasing in $p \in (0,1)$, and $\mathrm{ES}_p(X) \geq \mathrm{VaR}_p(X)$ for $p \in (0,1)$.

Value-at-Risk and Expected Shortfall



Example: Standard Deviation Principle

$b \geq 0$.

Definition (Standard deviation principle)

$\mathrm{SD}_b : L^2 \to \mathbb{R}$,

$$\mathrm{SD}_b(X) = \mathbb{E}[X] + b\sqrt{\mathrm{Var}(X)}.$$

- A small note: for normal risks, one can find p, q, b such that $\mathrm{VaR}_p(X) \approx \mathrm{ES}_q(X) \approx \mathrm{SD}_b(X)$. Example: $p = 0.99$, $q = 0.975$, $b = 2.33$.
- All risk measures can be defined on smaller subset $\mathcal{X}$ than its natural domain, e.g. $\mathcal{X} = L^\infty$.

Functionals: $\mathcal{X} \to (-\infty, \infty]$

Three major perspectives

- Preference of risk: Economic Decision Theory
- Pricing of risk: Insurance and Actuarial Science
- Capital requirement: Mathematical Finance

Preference of Risk

Preference of risk: Economic Decision Theory

- Mathematical theory established since 1940s.
    - Expected utility: von Neumann-Morgenstern (1944)
    - Rank-dependent expected utility: Quiggin (1982, JEBO)
    - Dual utility: Yaari (1987, Econometrica); Schmeidler (1989, Econometrica)
    - Prospect theory: Kahneman-Tversky (1979, Econometrica)
        - Citation: 39000+ (Google, March 2016)
    - Cumulative prospect theory: Tversky-Kahneman (1992, JRU)

Pricing of risk: Insurance and Actuarial Science

- Mathematical theory established since 1970s.
    - Additive principles: Gerber (1974, ASTIN Bulletin)
    - Economic principles: Bühlmann (1980, ASTIN Bulletin)
    - Convex principles: Deprez-Gerber (1985, IME)
    - Choquet principles: Wang-Young-Panjer (1997, IME)

Capital Requirement

Capital requirement: Mathematical Finance

- Mathematical theory established around late 1990s.
  - Coherent measures of risk: Artzner-Delbaen-Eber-Heath (1999, MF)
    - Citation: 6800+ (Google, March 2016)
  - Law-invariant risk measures: Kusuoka (2001, AME)
  - Convex measures of risk: Föllmer-Schied (2002, FS), Frittelli-Rossaza Gianin (2002, JBF)
  - Spectral measures of risk: Acerbi (2002, JBF)
- Mathematically very well developed, and fast expanding in the past $\sim$15 years.
- Value-at-Risk introduced earlier (around 1994): e.g. Duffie-Pan (1997, J. Derivatives).

Caution...

Different perspectives should lead to different principles of desirability.

- Preference of risk: only ordering matters (not precise values), gain and loss matter
- Pricing of risk: precise values matter, gain and loss matter
  - central limit theorem often kicks in (large number effect)
  - typically there is a market
- Capital requirement: precise values matter, only loss matters ($\leftarrow$ our focus)
  - typically there is no market; no large number effect

Of course, very large mathematical overlap ...

Research of Risk Measures

Two major perspectives

- What interesting mathematical/statistical problems arise from this field?
- What risk measures are practical in real life, and what are the practicality issues?

Good research may ideally address both questions, but it often only addresses one of them.

Academic Debates

- ES is generally advocated in academia for desirable properties in the past $\sim$ 15 years
- Some argue: backtesting ES is difficult, whereas backtesting VaR is straightforward
  - Paper before 2012: Gneiting (2011 JASA)
- Some argue: ES is not robust, whereas VaR is
  - Papers before 2012: Cont-Deguest-Scandolo (2010 QF); Kou-Peng-Heyde (2013 MOR)

Review paper: Embrechts et al. (2014, Risks).

Academic Debates

Some more recent papers

- On backtesting of risk measures:
    - Ziegel (2015+ MF)
    - Acerbi-Székely (2014 Risk)
    - Kou-Peng (2015 SSRN)
    - Fissler-Ziegel (2016 AoS)
    - Fissler-Ziegel-Gneiting (2016 Risk)
- On robustness of risk measures:
    - Stahl-Zheng-Kiesel-Rühlicke (2012 SSRN)
    - Krätschmer-Schied-Zähle (2012 JMVA, 2014 FS, 2015 arXiv)
    - Cambou-Filipović (2015+ MF)
    - Embrechts-Wang-Wang (2015 FS)
    - Daníelsson-Zhou (2015 SSRN)

Regulatory Documents

From the **Basel Committee on Banking Supervision**:

R1: Consultative Document, May 2012,
Fundamental review of the trading book

R2: Consultative Document, October 2013,
Fundamental review of the trading book: A revised market risk framework

R3: Standards, January 2016,
Minimum capital requirements for Market Risk

From the **International Association of Insurance Supervisors**:

R4: Consultation Document, December 2014,
Risk-based global insurance capital standard

Questions from Regulation

R1, Page 20, *Choice of risk metric*:

"... However, a number of weaknesses have been identified with VaR, including its inability to capture "tail risk". The Committee therefore believes it is necessary to consider alternative risk metrics that may overcome these weaknesses."

R1, Page 41, *Question 8*:

"What are the likely constraints with moving from VaR to ES, including any challenges in delivering robust backtesting, and how might these be best overcome?"

R3, Page 1. *Executive Summary*:

"... A shift from Value-at-Risk (VaR) to an Expected Shortfall (ES) measure of risk under stress. Use of ES will help to ensure a more prudent capture of "tail risk" and capital adequacy during periods of significant financial market stress."

R4, Page 43. *Question 42*:

"Which risk measure - VaR, Tail-VaR [ES] or another - is most appropriate for ICS [insurance capital standard] capital requirement purposes? Why?"

VaR versus ES

Features/Risk measure	VaR	Tail-VaR
Frequency captured?	Yes	Yes
Severity captured?	No	Yes
Sub-additive?	Not always	Always
Diversification captured?	Issues	Yes
Back-testing?	Straight-forward	Issues
Estimation?	Feasible	Issues with data limitation
Model uncertainty?	Sensitive to aggregation	Sensitive to tail modelling
Robustness I (with respect to "Lévy metric[33]")?	Almost, only minor issues	No
Robustness II (with respect to "Wasserstein metric[34]")?	Yes	Yes

Table from R4, December 2014

VaR versus ES

Centers of discussion:

- backtesting
- estimation
- model uncertainty
- robustness

All refer to uncertainty (or ambiguity)

VaR versus ES

A summary of the current situation (early 2016):

- VaR is globally dominating banking regulation at the moment; in insurance it is popularly used (e.g. Solvency II).
- ES is also widely implemented (e.g. Swiss Solvency Test). In many places ES and VaR coexist (Basel III).
- ES is proposed to replaced VaR in many places of the world.
- The search for alternative risk measures to VaR and ES is on going (mainly academic).

Industry Perspectives

From the **International Association of Insurance Supervisors**:

R5: Document (version June 2015)

Compiled Responses to ICS Consultation 17 Dec 2014 - 16 Feb 2015

In summary

- Responses from insurance organizations and companies in the world.
- 49 responses are public
- 34 commented on Q42: VaR versus ES (TVaR)

Industry Perspectives

- 5 responses are supportive about ES:
  - Canadian Institute of Actuaries, CA
  - Liberty Mutual Insurance Group, US
  - National Association of Insurance Commissioners, US
  - Nematrian Limited, UK
  - Swiss Reinsurance Company, CH
- Some are indecisive; most favour VaR.

The debate will go on for a while

Regulator and firms may have different views

Question from Regulation

We focus on the mathematical and statistical aspects, and try to avoid practicalities and operational issues.

From R1, Page 3:

"The Committee recognises that moving to ES could entail certain operational challenges; nonetheless it believes that these are outweighed by the benefits of replacing VaR with a measure that better captures tail risk."

Value-at-Risk

- Interpretation: only crashes when the worst $100(1-p)\%$ case happens.
- In one-year capital requirement:
  - 0.95 - only crashes in a crisis that happens once in 20 years.
  - 0.99 - only crashes in a crisis that happens once in 100 years. (Survive 2007.)
  - 0.995 - only crashes in a crisis that happens once in 200 years. (Survive 1930.)
- VaR is a measure based on frequency, and "does not capture the tail risk".
- Does not require $\mathbb{E}[X] < \infty$.

Expected Shortfall

- Interpretation: prepare for the worst $100(1-p)\%$ case.
- More conservative than VaR at the same level.
- In one-year capital requirement:
  - 0.95 - prepare for a crisis that happens once in 20 years.
  - 0.99 - prepare for a crisis that happens once in 100 years. (prepare for 1930.)
  - 0.995 - prepare for a crisis that happens once in 200 years. (prepare for something we have not experienced.)
- ES is a measure based on frequency and severity, and "captures the tail risk".
- Requires $\mathbb{E}[X] < \infty$.
- Always keep in mind that the above probabilities are inaccurate

VaR and ES

$VaR_p, p \in (0,1)$ is invariant under increasing transformations: for any strictly increasing function $f : \mathbb{R} \to \mathbb{R}$,
$f(\mathrm{VaR}_p(X)) = \mathrm{VaR}_p(f(X))$.

- This property does not hold generally for ES.
- For instance, one can calculate the VaR of the return of an asset, and then transform it into the VaR of the asset value. This procedure does not work for ES.

$$\text{asset value} : A_T; \qquad \text{return} : \log(A_T/A_0)$$

- If the return follows some heavy-tailed distribution (like t-distribution), the asset value may not have finite expectation. Typically however, one worries about the loss, which is bounded by assuming $A_T \geq 0$.

Extreme Value Theory (EVT)

Definition

An eventually non-negative (that is, $f(x) \geq 0$ for x large enough) measurable function f is said to be regularly varying (RV) with a regularity index $\gamma \in \mathbb{R}$, if

$$\lim_{x\to\infty} \frac{f(tx)}{f(x)} = t^{\gamma}, \quad \text{for all } t > 0.$$

Denote this by $f \in \mathrm{RV}_{\gamma}$.

- An RV function only concerns its behavior close to infinity.
- General reference book on EVT: de Haan and Ferreira (2006).

Proposition

$f \in \mathrm{RV}_\gamma$ if and only if

$$f(x) = x^\gamma L(x),$$

for some slowly varying function L, that is,

$$\lim_{x\to\infty} \frac{L(tx)}{L(x)} = 1, \quad \text{for all } t > 0.$$

For a distribution F, $\overline{F}(\cdot) = 1 - F(\cdot)$ is the survival function.

Lemma (RV Inversion)

Let F be a distribution function. Then for any $\beta > 0$, $\overline{F}(\cdot) \in \mathrm{RV}_{-\beta}$ is equivalent to $F^{-1}(1 - 1/\cdot) \in \mathrm{RV}_{1/\beta}$.

Corollary (VaR Extrapolation*)

Suppose that $X \sim F$ and $\overline{F} \in \mathrm{RV}_{-\beta}$. Then for $t > 0$,

$$\lim_{\epsilon \downarrow 0} \frac{\mathrm{VaR}_{1-t\epsilon}(X)}{\mathrm{VaR}_{1-\epsilon}(X)} = t^{-1/\beta}.$$

*an asterisk always indicates that details (proofs) are planned to be given in the lecture

Example: Pareto distributions, for $\alpha > 0$, $\theta > 0$,

$$F(x) = 1 - \left(\frac{x}{\theta}\right)^{-\alpha}, \quad x \geq \theta.$$

We can easily see that $\overline{F}(\cdot) \in \mathrm{RV}_{-\alpha}$. Moreover,

$$F^{-1}(p) = \theta(1-p)^{-1/\alpha}, \quad p \in (0,1),$$

and hence $F^{-1}(1 - 1/\cdot) \in \mathrm{RV}_{1/\alpha}$.

Note that for $X \sim F$,

- $\mathrm{VaR}_p(X) = F^{-1}(p)$.
- $\mathrm{ES}_p(X) < \infty$ if and only if $\alpha > 1$.

VaR versus ES, Extreme Value Theory

Very common in Quantitative Risk Management (QRM) applications, RV distributions are used to model

- $\alpha \in [0.5, 1]$ for catastrophe insurance,
- $\alpha \in [3, 5]$ for market return data,
- $\alpha > 0.5$ for operational risk.

Typical choices of p are close to 1, so it is natural to study the limiting behavior of VaR_p and ES_p as $p \to 1$.

VaR versus ES, Extreme Value Theory

- For light tailed distributions (such as $X \sim N(\mu, \sigma^2)$),

$$\lim_{p \to 1} \frac{\mathrm{ES}_p(X)}{\mathrm{VaR}_p(X)} = 1.$$

- For heavy tailed distributions:

  Suppose that the function $\overline{F}(x) = \mathbb{P}(X > x)$ is $\mathrm{RV}_{-1/\xi}$, $\xi \in (0, 1)$, then

$$\lim_{p \to 1} \frac{\mathrm{ES}_p(X)}{\mathrm{VaR}_p(X)} = \frac{1}{1 - \xi}.$$

- This remarkable result is known as **Karamata's Theorem**.

Karamata's Theorem

Theorem (Karamata's Theorem)

Suppose that an eventually non-negative and locally bounded function f is $\mathrm{RV}_{-\alpha}$, $\alpha > 1$. Then

$$\lim_{t\to\infty} \frac{tf(t)}{\int_t^\infty f(s)\mathrm{d}s} = \alpha - 1.$$

EVT for VaR/ES Ratio

Theorem (Karamata's Theorem for VaR/ES*)

Suppose that the function $\overline{F}(x) = \mathbb{P}(X > x)$ *is* $\mathrm{RV}_{-1/\xi}$, $\xi \in (0,1)$, *then*

$$\lim_{p \to 1} \frac{\mathrm{VaR}_p(X)}{\mathrm{ES}_p(X)} = 1 - \xi.$$

VaR versus ES, 0.99 vs 0.975

From R4: Page 22, *Moving to expected shortfall*:

"... using an ES model, the Committee believes that moving to a confidence level of 97.5% (relative to the 99th percentile confidence level for the current VaR measure) is appropriate."

$$\boxed{\mathrm{VaR}_{0.99} \text{ vs } \mathrm{ES}_{0.975}}$$

- Example: $X \sim$ Normal(0,1).

$$\mathrm{ES}_{0.975}(X) = 2.3378,$$

$$\mathrm{VaR}_{0.99}(X) = 2.3263.$$

They are quite close for all normal models!

VaR versus ES, 0.99 vs 0.975

From EVT: approximately,

- for heavy-tailed risks, $\mathrm{ES}_{0.975}$ yields a more conservative value than $\mathrm{VaR}_{0.99}$;
- for light-tailed distributions, $\mathrm{ES}_{0.975}$ yields an equivalent regulation principle as $\mathrm{VaR}_{0.99}$;
- for risks that do not have a very heavy tail, it holds $\mathrm{ES}_{0.975}(X) \approx \mathrm{VaR}_{0.99}(X)$.

VaR versus ES, 0.99 vs 0.975

- Via Karamata's Theorem: for $\xi \in [0,1)$ ($\xi = 0$ indicates a light tail),
$$\frac{\mathrm{ES}_{0.975}(X)}{\mathrm{VaR}_{0.975}(X)} \approx \frac{1}{1-\xi},$$
and (via VaR extrapolation)
$$\frac{\mathrm{VaR}_{0.99}(X)}{\mathrm{VaR}_{0.975}(X)} \approx 2.5^{\xi}.$$
Putting the above together,
$$\frac{\mathrm{VaR}_{0.99}(X)}{\mathrm{ES}_{0.975}(X)} \approx 2.5^{\xi}(1-\xi).$$

VaR versus ES, 0.99 vs 0.975

- $\xi \in [0, 1)$,

$$\frac{\mathrm{VaR}_{0.99}(X)}{\mathrm{ES}_{0.975}(X)} \approx 2.5^{\xi}(1-\xi) \leq e^{\xi}(1-\xi) \leq 1.$$

  Approximately, $\mathrm{ES}_{0.975}$ yields a more conservative regulation principle than $\mathrm{VaR}_{0.99}$.

- For a particular X, it is not always $\mathrm{ES}_{0.975}(X) \geq \mathrm{VaR}_{0.99}(X)$.

VaR versus ES, 0.99 vs 0.975

- Light-tailed distributions: as $\xi \to 0$,

$$\frac{\mathrm{VaR}_{0.99}(X)}{\mathrm{ES}_{0.975}(X)} \approx 2.5^{\xi}(1-\xi) \to 1.$$

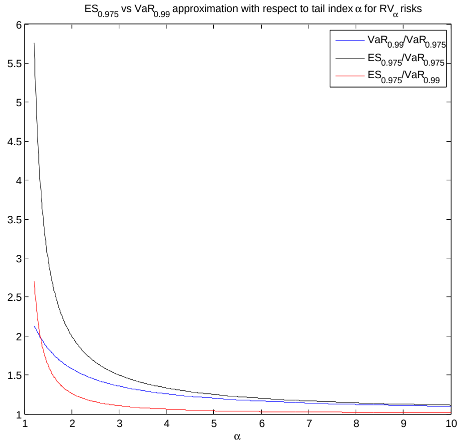
  For light-tailed distributions, $\mathrm{ES}_{0.975}$ yields an (approximately) equivalent regulation principle as $\mathrm{VaR}_{0.99}$.

- It seems that the value

$$c = 2.5 = (1-0.975)/(1-0.99)$$

  is chosen such that c is close to $e \approx 2.72$, so that the approximation $c^{\xi}(1-\xi) \approx 1$ holds most accurate for small ξ; note that $e^{-\xi} \approx 1-\xi$ for small ξ.

VaR versus ES, 0.99 vs 0.975 ($\alpha = 1/\xi$)



Estimation

Suppose that the iid data are $X_1, \ldots, X_n$ from a distribution F. To estimate $\mathrm{VaR}_p(X)$ and $\mathrm{ES}_p(X)$ for $X \sim F$, three basic methods:

- Empirical method
- Parametric (model) method
- EVT (semi-parametric) method

VaR versus ES: Estimation

Empirical method: $[x]$ stands for the integer part of x.

- $\widehat{\mathrm{VaR}}_p(X) = X_{[np]}$: the $[np]$-th largest observation
- $\widehat{\mathrm{ES}}_p(X) = \frac{1}{n-[np]} \sum_{i=[np]}^{n} X_{[i]}$: the largest $n - [np] + 1$ observations
- One may also use $[np] + 1$ instead of $[np]$, or a linear combination of both.
- Problems: p is typically close to 1
  - if n is small and p is close to 1, the estimators $\widehat{\mathrm{VaR}}_p(X)$ and $\widehat{\mathrm{ES}}_p(X)$ may not be viable
  - very large estimation error
  - in real problems in finance, large and iid data set is not easy to find

VaR versus ES: Estimation

Parametric (model) method:

- First fit data to a model, typically parametric
- Then calculate the model VaR and ES
  - This may be through analytical calculation for nice models
    - e.g. for $X \sim \mathrm{N}(\mu, \sigma^2)$,

$$\mathrm{VaR}_p(X) = \mu + \sigma \Phi^{-1}(p),$$

$$\mathrm{ES}_p(X) = \mu + \sigma \frac{\phi(\Phi^{-1}(p))}{1-p},$$

      where Φ and ϕ are the standard normal cdf and pdf, respectively.
  - or Monte Carlo simulation for complicated models
- Problems:
  - model risk
  - very large estimation error
  - data limitation

VaR versus ES: Estimation

EVT (semi-parametric) method:

- First estimate the RV index of the distribution
  - This can be done through for instance the Hill Estimator; see Section 3.2 of de Haan and Ferreira (2006).
- Use the data to obtain credible lower level VaR_q, $q < p$
- Then use VaR extrapolation to obtain VaR_p
- Problems:
  - still model risk: there is no guarantee that the extrapolation is valid
  - depends on the choices of thresholds in the Hill Estimator and in q
  - considerable estimation error
  - data limitation

Summary

Key questions in this course and in reality

- Which risk measure is better for regulation, VaR or ES, and in what situations?
- What other possible risk measure can be used?
- What properties should a good risk measure satisfy?
- Are there any overlooked problems with existing risk measures?
- Are there new methods for estimation, calculation, simulation, model selection or other practical issues?

Summary

From R5, major reasons to favour VaR from the industry

- Implementation of ES is expensive (staff, software, capital)
- ES does not exist for certain heavy-tailed risks
- ES is more costly on distributional information, data and simulation
- ES has trouble with a change of currency (Koch Medina-Munari, 2016 JBF)