

Estimation methods for expected shortfall

1 Introduction

Value at Risk, the most popular measure for financial risk, has been widely used by financial institutes around the world since it was proposed. However, value at risk has several shortcomings. For instance, Artzner *et al.* (1997,1999) have shown that value at risk not only ignores any loss beyond the value at risk level and also can not satisfy one of the axioms of coherence as it is not sub-additive. Furthermore, Yamai and Yoshida (2002) have stated another two disadvantages. One is that rational investors wishing to maximize expected utility may be misled by the information offered by value at risk. The other one is that value at risk is hard to use when investors want to optimize their portfolios. In order to deal with the conceptual problems caused by value at risk, Artzner *et al.* (1999) introduced a new measure of financial risk referred to as the *expected shortfall*. It is defined as follows.

Let $\{X_t, t = 1, 2, \dots, n\}$ denote a stationary financial series with marginal distribution function F and marginal density function f . The Value at Risk (abbreviated as VaR) for a given probability p is defined as

$$\text{VaR}_p(X) = \inf \{u : F(u) \geq p\}.$$

The expected shortfall (abbreviated as ES) for a given probability p is defined as

$$\begin{aligned} \text{ES}_p(X) &= (1/p) \left[E(XI\{X \leq \text{VaR}_p(X)\}) + p\text{VaR}_p(X) \right. \\ &\quad \left. - \text{VaR}_p(X) \Pr(X \leq \text{VaR}_p(X)) \right], \end{aligned} \quad (1)$$

where $I\{\cdot\}$ denotes the indicator function. As we can see, both measures are closely related to each other.

Applications of expected shortfall have been extensive. Some recent applications and application areas include: repowering of existing coastal stations to augment water supplies in Southern California (Sims and Kamal, 1996); risk management of basic social security fund in China (An *et al.*, 2005); futures clearinghouse margin requirements (Cotter and Dowd, 2006); reward-risk stock selection criteria (Rachev *et al.*, 2007); Shanghai stock exchange (Li and Li, 2006, Fan *et al.*, 2008); extreme daily changes in US Dollar London inter-bank offer rates (Krehbiel and Adkins, 2008); exchange rate risk of CNY (Wang and Wu, 2008); financial risk associated with US movie box office earnings (Bi and Giles, 2007, Bi and Giles, 2009); operational risk of Chinese commercial banks (Gao and Li, 2009, Song *et al.*, 2009); cash flow risk measurement for Chinese non-life insurance industry (Teng and Zhang, 2009); risk contribution of different industries in China's stock market (Liu *et al.*, 2008; Yu and Tao, 2008, Fan *et al.*, 2010); operational risk in Taiwanese commercial banks (Lee and Fang, 2010); the exchange rate risk of Chinese Yuan (Wang *et al.*, 2010); extreme dependence between European electricity markets (Lindstrom and Regland, 2012).

The aim of this thesis is to review known methods for estimating (1). The review of methods is divided as follows: general properties (Chapter 2), parametric methods (Chapter 3), nonparametric methods (Chapter 4), semiparametric methods (Chapter 5), and computer software (Chapter 6). A paper based on this material has been submitted to the journal, *Econometric Reviews*.

The review of value of risk presented here is by no means comprehensive. For a fuller account of the theory and applications of value risk, we refer the readers to the following books: Hafner (2004, Chapter 7), Ardia (2008, Chapter 6), Wüthrich *et al.* (2010, Chapter 3), and Ruppert (2011, Chapter 19).

2 General properties

2.1 Basic properties

Let X and Y denote real random variables. Let

$$x^+ = \begin{cases} x, & \text{if } x > 0, \\ 0, & \text{if } x \leq 0 \end{cases}, \quad x^- = (-x)^+.$$

Also let

$$\begin{aligned} x_{(p)} &= \inf \{x \in \mathbb{R} : \Pr(X \leq x) \geq p\}, \\ x^{(p)} &= \inf \{x \in \mathbb{R} : \Pr(X \leq x) > p\}. \end{aligned}$$

Some basic properties of expected shortfall are:

1. if $E[X^-] < \infty$ and X is stochastically greater than Y then $\text{ES}_p(X) > \text{ES}_p(Y)$;
2. if $E[X^-] < \infty$ then $X \geq 0$ then $\text{ES}_p(X) \geq 0$ for all p (Proposition 3.1, Acerbi and Tasche, 2002);
3. if $E[X^-] < \infty$ then $\text{ES}_p(\lambda X) = \lambda \text{ES}_p(X)$ for all $\lambda > 0$ (Proposition 3.1, Acerbi and Tasche, 2002);
4. if $E[X^-] < \infty$ then $\text{ES}_p(X + k) = \lambda \text{ES}_p(X) - k$ for all $-\infty < k < \infty$ (Proposition 3.1, Acerbi and Tasche, 2002);
5. if $E[X^-] < \infty$ then $\text{ES}_p(X)$ is convex, that is $\text{ES}_p(\lambda X + (1-\lambda)Y) \leq \lambda \text{ES}_p(X) + (1-\lambda) \text{ES}_p(Y)$.
6. for $E[X^-] < \infty$, any $\alpha \in (0, 1)$ and any $e > 0$ with $\alpha + e < 1$, $\text{ES}_{\alpha+e}(X) \geq \text{ES}_\alpha(X)$ (Proposition 3.4, Acerbi and Tasche, 2002);
7. if $E[X^-] < \infty$ and $E[Y^-] < \infty$ then $\text{ES}_p(X + Y) \leq \text{ES}_p(X) + \text{ES}_p(Y)$ for any $p \in (0, 1)$ (Proposition A.1, Acerbi and Tasche, 2002);
8. if X is integrable and if $p \in (0, 1)$ then

$$\text{ES}_p(X) = (1/p) [E(XI\{X \leq s\}) + sp - s \Pr(X \leq s)]$$

for $s \in [x_{(p)}, x^{(p)}]$ (Corollary 4.3, Acerbi and Tasche, 2002);

9. if X is integrable and if $p \in (0, 1)$ then

$$\text{ES}_p(X) = (1/p) [E(XI\{X < s\}) + sp - s \Pr(X < s)]$$

for $s \in [x_{(p)}, x^{(p)}]$ (equation (4.12), Acerbi and Tasche, 2002);

- 10. if $E[X^-] < \infty$ then $\text{ES}_p(X) \geq E[X | | \leq x_{(p)}] \geq E[X | | \leq x^{(p)}]$ (Corollary 5.2, Acerbi and Tasche, 2002);
- 11. if $E[X^-] < \infty$ then $\text{ES}_p(X) \geq \inf \{E[X|A] : \Pr(A) > p\} \geq E[X | | \leq x^{(p)}]$ (Corollary 5.2, Acerbi and Tasche, 2002);
- 12. if $E[X^-] < \infty$ then $\text{ES}_p(X) = p \int_0^p x_{(u)} du$ (equation (3.3), Acerbi and Tasche, 2002).

2.2 Upper comonotonicity

Let X_i denote the loss of the i th asset. Let $\mathbf{X} = (X_1, \dots, X_n)$ with joint cdf $F(x_1, \dots, x_n)$. Let $T = X_1 + \dots + X_n$. Suppose all random variables are defined on the probability space $(\Omega, \mathcal{F}, \Pr)$. Then a simple formula for the expected shortfall of T in terms of expected shortfalls of X_i can be established if \mathbf{X} is *upper comonotonic* (Cheung, 2009).

We now define what is meant by upper comonotonicity. A subset $C \subset \mathbb{R}^n$ is said to be comonotonic if $(t_i - s_i)(t_j - s_j) \geq 0$ for all i and j whenever (t_1, \dots, t_n) and (s_1, \dots, s_n) belong to C . The random vector is said to be comonotonic if it has a comonotonic support.

Let \mathcal{N} the collection of all zero probability sets in the probability space. Let $\overline{\mathbb{R}^n} = \mathbb{R}^n \cup (-\infty, \dots, -\infty)$. For a given $(a_1, \dots, a_n) \in \mathbb{R}^n$, let $U(\mathbf{a})$ denote the upper quadrant of $(a_1, \infty) \times \dots \times (a_n, \infty)$ and let $L(\mathbf{a})$ denote the lower quadrant of $(-\infty, a_1] \times \dots \times (-\infty, a_n]$. Let $R(\mathbf{a}) = \mathbb{R}^n \setminus (U(\mathbf{a}) \cup L(\mathbf{a}))$.

Then the random vector \mathbf{X} is said to be upper comonotonic if there exist $\mathbf{a} \in \overline{\mathbb{R}^n}$ and a zero probability set $N(\mathbf{a}) \in \mathcal{N}$ such that

- (a) $\{\mathbf{X}(w) \mid w \in \Omega \setminus N(\mathbf{a})\} \cap U(\mathbf{a})$ is a comonotonic subset of \mathbb{R}^n ;
- (b) $\Pr(\mathbf{X} \in U(\mathbf{a})) > 0$;
- (c) $\{\mathbf{X}(w) \mid w \in \Omega \setminus N(\mathbf{a})\} \cap R(\mathbf{a})$ is an empty set.

If these three conditions are satisfied then the expected shortfall of T can be expressed as

$$\text{ES}_p(T) = \sum_{i=1}^n \text{ES}_p(X_i) \tag{2}$$

for $p \in (F(a_1^*, \dots, a_n^*), 1)$ and $\mathbf{a}^* = (a_1^*, \dots, a_n^*)$ a comonotonic threshold as constructed in Lemma 2 of Cheung (2009).

2.3 Risk decomposition

Suppose a portfolio is made up of n assets. Then portfolio loss, say X , can be written as $X = w_1X_1 + \dots + w_nX_n$, where X_i denotes the loss for asset i and w_i denotes the weight for asset i . Then it can be shown (Fan *et al.*, 2012)

$$\text{ES}_p(X) = \sum_{i=1}^n \frac{\partial \text{ES}_p(X)}{\partial w_i} w_i. \quad (3)$$

This is known as risk decomposition.

2.4 Hürlimann's inequalities

Let X denote a random variable defined over $[A, B]$, $-\infty \leq A < B \leq \infty$ with mean μ , and variance σ^2 . Hürlimann (2002) provides various upper bounds for $\text{ES}_p(X)$: for $p \leq \sigma^2 / \{\sigma^2 + (B - \mu)^2\}$ then

$$\text{ES}_p(X) \leq B;$$

for $\sigma^2 / \{\sigma^2 + (B - \mu)^2\} \leq p \leq (\mu - A)^2 / \{\sigma^2 + (\mu - A)^2\}$ then

$$\text{ES}_p(X) \leq \mu + \sqrt{\frac{1-p}{p}} \sigma;$$

for $p \geq (\mu - A)^2 / \{\sigma^2 + (\mu - A)^2\}$ then

$$\text{ES}_p(X) \leq \mu + (\mu - A) \frac{1-p}{p}.$$

Now suppose X is a random variable defined over $[A, B]$, $-\infty \leq A < B \leq \infty$ with mean μ , variance σ , skewness γ and kurtosis γ_2 . In this case, Hürlimann (2002) provides the following upper bound for $\text{ES}_p(X)$:

$$\text{ES}_p(X) \leq \mu + x_p \sigma,$$

where x_p is the $100(1-p)$ percentile of the standardized Chebyshev-Markov maximal distribution. The latter is defined as the root of

$$p(x_p) = p$$

if $p \leq (1/2)\{1 - \gamma/\sqrt{4 + \gamma^2}\}$ and as the root of

$$p(\psi(x_p)) = 1 - p$$

if $p > (1/2)\{1 - \gamma/\sqrt{4 + \gamma^2}\}$, where

$$\begin{aligned} p(u) &= \frac{\Delta}{q^2(u) + \Delta(1 + u^2)}, \\ \psi(u) &= \frac{1}{2} \left[\frac{A(u) - \sqrt{A^2(u) + 4q(u)B(u)}}{q(u)} \right], \end{aligned}$$

where $\Delta = \gamma_2 - \gamma^2 + 2$, $A(u) = \gamma q(u) + \Delta u$, $B(u) = q(u) + \Delta$ and $q(u) = 1 + \gamma u - u^2$.

Hürlimann (2003) provided further inequalities for expected shortfall based on stop-loss ordering: a random variable X is said to be less than or equal to another random variable Y with respect to stop-loss order if $\int_x^\infty [1 - F_X(t)]dt \leq \int_x^\infty [1 - F_Y(t)]dt$ for all x . Given this ordering, Hürlimann (2003) showed that $\text{ES}_p(X) \leq \text{ES}_p(Y)$ for all p . Similarly, if X_{\min} is less than or equal to X and if X is less than or equal to X_{\max} with respect to stop-loss ordering then Hürlimann (2003) showed that $\text{ES}_p(X_{\min}) \leq \text{ES}_p(X) \leq \text{ES}_p(X_{\max})$ for all p .

3 Parametric methods

3.1 Gaussian distribution

If X_1, X_2, \dots, X_n are observations from a Gaussian distribution with mean μ and variance σ^2 then ES can be estimated by

$$\widehat{\text{ES}}_\alpha = E[X \mid X > s\Phi^{-1}(\alpha)]$$

where s denotes the sample standard deviation

$$s = \sqrt{\frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2}$$

and \bar{X} is the sample mean.

3.2 Johnson family method

An approximation for expected shortfall suggested by Simonato (2011) is based on the Johnson family of distributions due to Johnson (1949).

Let Y denote a standard normal random variable. A Johnson random variable can be expressed as

$$Z = c + dg^{-1}\left(\frac{Y - a}{b}\right),$$

where

$$g^{-1}(u) = \begin{cases} \exp(u), & \text{for the lognormal family,} \\ [\exp(u) - \exp(-u)]/2, & \text{for the unbounded family,} \\ 1/[1 + \exp(-u)], & \text{for the bounded family,} \\ u, & \text{for the normal family.} \end{cases}$$

Here, a , b , c and d are unknown parameters can be determined, for example, by the method of moments, see Hill *et al.* (1976).

With the notation as above, the approximation for expected shortfall is

$$\text{ES}_p = \frac{1}{p} \int_{-\infty}^K \left[c + dg^{-1}\left(\frac{y - a}{b}\right) \right] \phi(y) dy,$$

where

$$K = a + bg \left(\frac{k_J - c}{d} \right),$$

and

$$k_J = c + dg^{-1} \left(\frac{\Phi^{-1}(p) - a}{b} \right),$$

where $\phi(\cdot)$ denotes the standard normal pdf and $\Phi^{-1}(\cdot)$ denotes the standard normal quantile function.

3.3 Azzalini's skewed normal distribution

The major weakness of the normal distribution is its inability to model skewed data. Several skewed extensions of the normal distribution have been proposed in the literature. The most popular and the most widely used of these is the skew-normal distribution due to Azzalini (1985). The pdf of this distribution is given by

$$f_X(x) = \frac{2}{\sigma} \phi \left(\frac{x - \mu}{\sigma} \right) \Phi \left(\lambda \frac{x - \mu}{\sigma} \right) \quad (4)$$

for $x \in \mathbb{R}$, $\lambda \in \mathbb{R}$, $\mu \in \mathbb{R}$ and $\sigma > 0$, where $\phi(x)$ is the standard normal pdf and $\Phi(x)$ is the standard normal cdf. The cdf corresponding to (4) is

$$F_X(x) = \Phi \left(\frac{x - \mu}{\sigma} \right) - 2T \left(\frac{x - \mu}{\sigma}, \lambda \right),$$

where

$$T(h, a) = \frac{1}{2\pi} \int_0^a \frac{\exp \left\{ -h^2 (1 + x^2)/2 \right\}}{1 + x^2} dx$$

is Owen's T function (Owen, 1956). Bernardi (2012) has shown that the expected shortfall of a skew normal random variable X is

$$\text{ES}_p(X) = \mu + \frac{\sigma \sqrt{2}}{p \sqrt{\pi}} \left[\lambda \Phi(z_p) - \sqrt{2\pi} \phi(y_p) \Phi(\lambda y_p) \right],$$

where $\delta = \lambda / \sqrt{1 + \lambda^2}$, $z_p = \sqrt{1 + \lambda^2} y_p$, $y_p = (x_p - \mu) / \sigma$ and x_p satisfies $F_X(x_p) = p$.

3.4 Azzalini's skewed normal mixture distribution

Bernardi (2012) has also considered a mixture of skew normal distribution. Let X be a random variable with the pdf

$$f_X(x) = \frac{2}{\sigma} \sum_{i=1}^L \eta_i \phi \left(\frac{x - \mu_i}{\sigma_i} \right) \Phi \left(\lambda_i \frac{x - \mu_i}{\sigma_i} \right), \quad (5)$$

where the weights η_i are non-negative and sum to one. Bernardi (2012) shows that the expected shortfall of X can be expressed as

$$\text{ES}_p(X) = \sum_{i=1}^L \pi_i \left\{ \mu_i + \frac{\sigma_i \sqrt{2}}{p \sqrt{\pi}} \left[\lambda_i \Phi(z_{p,i}) - \sqrt{2\pi} \phi(y_{p,i}) \Phi(\lambda_i y_{p,i}) \right] \right\},$$

where $\delta_i = \lambda_i / \sqrt{1 + \lambda_i^2}$, $z_{p,i} = \sqrt{1 + \lambda_i^2} y_{p,i}$, $y_{p,i} = (x_{p,i} - \mu_i) / \sigma_i$ and $x_{p,i}$ is the root of

$$\Phi\left(\frac{x - \mu_i}{\sigma_i}\right) - 2T\left(\frac{x - \mu_i}{\sigma_i}, \lambda_i\right) = p.$$

Furthermore,

$$\pi_i = \frac{\eta_i}{p} \left[\Phi\left(\frac{x_p - \mu_i}{\sigma_i}\right) - 2T\left(\frac{x_p - \mu_i}{\sigma_i}, \lambda_i\right) \right],$$

where x_p is the root of $F_X(x_p) = p$.

3.5 Student's t distribution

Let X denote a Student's t random variable with location parameter $-\infty < \mu < \infty$, scale parameter $c > 0$ and degrees of freedom $n > 0$; that is, X has the pdf

$$f_X(x) = \frac{n^{-1/2}}{\sigma B(n/2, 1/2)} \left(1 + \frac{(x - \mu)^2}{n}\right)^{-(n+1)/2}.$$

Let q_p denote the p th quantile of the standard Student's t distribution; that is, q_p is the root $\Pr(X \leq q_p) = p$ when $\mu = 0$ and $\sigma = 1$. Broda and Paoletta (2011, Section 2.2.2) show the expected shortfall for X can be expressed as

$$\text{ES}_p(X) = \frac{1}{p} \text{Ttail}(q_p, n),$$

where

$$\text{Ttail}(c, n) = -\frac{n^{-1/2}}{\sigma B(n/2, 1/2)} \left(1 + \frac{c^2}{n}\right)^{-(n+1)/2} \frac{n + c^2}{n - 1}.$$

3.6 Azzalini's skewed t distribution

Let a random variable X follow Azzalini and Capitanio (2003)'s skewed t distribution given by the pdf

$$f_X(x) = 2\psi_m(x)\Psi_{m+1} \left(\lambda x \sqrt{\frac{m+1}{x^2+m}} \right)$$

for $-\infty < x < \infty$ and $m > 0$, where $\psi_m(\cdot)$ and $\Psi_m(\cdot)$ denote, respectively, the pdf and the cdf of a Student's t random variable with m degrees of freedom. Broda and Paoletta (2011, Section 2.2.2) also show the expected shortfall for X can be expressed as

$$\text{ES}_p(X) = 2 \int_{-\infty}^q x \psi_m(x) \Psi_{m+1} \left(\lambda x \sqrt{\frac{m+1}{x^2+m}} \right) dx$$

for q satisfying $F_X(q) = p$.

3.7 Jones and Faddy's skewed t distribution

Let a random variable X follow Jones and Faddy (2003)'s skewed t distribution given by the pdf

$$f_X(x) = C \left(1 + \frac{x}{y_x}\right)^{a+1/2} \left(1 - \frac{x}{y_x}\right)^{b+1/2},$$

where $y_x = (a + b + x^2)^{1/2}$ and $1/C = B(a, b)(a + b)^{1/2}2^{a+b-1}$. Let q_p denote the p th quantile of X . Broda and Paoletta (2011, Section 2.2.2) also show the expected shortfall for X can be expressed as

$$\text{ES}_p(X) = \frac{\sqrt{a+b}}{B(a, b)} \left\{ B_y(a + 1/2, b - 1/2) - \frac{1}{2} B_y(a - 1/2, b - 1/2) \right\}$$

where $y = 1/2 + c(c^2 + a + b)^{-1/2}/2$, c satisfies $F_X(c) = p$ and

$$B_x(a, b) = \int_0^x t^{a-1} (1-t)^{b-1} dt$$

denotes the incomplete beta function.

3.8 Generalized asymmetric t distribution

Let a random variable X follow Jones and Faddy (2003)'s skewed t distribution given by the pdf

$$f_X(x) = C \begin{cases} \left(1 + \frac{(-x\theta)^d}{\nu}\right)^{-\nu-1/d}, & \text{if } x < 0, \\ \left(1 + \frac{(x/\theta)^d}{\nu}\right)^{-\nu-1/d}, & \text{if } x \geq 0, \end{cases}$$

where $d > 0$, $\nu > 0$, $\theta > 0$ and $1/C = (\theta^{-1} + \theta)d^{-1}\nu^{1/d}B(1/d, \nu)$. Let q_p denote the p th quantile of X . Broda and Paoletta (2011, Section 2.2.2) also show the expected shortfall for X can be expressed as

$$\text{ES}_p(X) = -\nu^{1/d} \frac{1 + \theta^2}{\theta + \theta^3} \frac{B_L(\nu - 1/d, 2/d)}{B_L(\nu, 1/d)},$$

where $L = \nu/\{\nu + (-q_p\theta)^d\}$.

3.9 Noncentral t distribution

A random variable X follows the noncentral t distribution if its pdf is given by

$$f_X(x) = \exp(-\mu^2/2) \frac{\Gamma((k+1)/2) k^{k/2}}{\sqrt{\pi} \Gamma(k/2)} (k + x^2)^{-(k+1)/2} A(x)$$

for $-\infty < x < \infty$, where $k > 0$ denotes the degree of freedom parameter, $-\infty < \mu < \infty$ denotes the non-centrality parameter and

$$A(x) = \sum_{i=0}^{\infty} \frac{(\mu x)^i}{i!} \left(\frac{2}{x^2 + k}\right)^{i/2} \frac{\Gamma((k+i+1)/2)}{\Gamma((k+1)/2)}.$$

Broda and Paoella (2011, Section 2.2.2) also show the expected shortfall for X can be expressed as

$$\text{ES}_p(X) = \exp(-\mu^2/2) \frac{\Gamma((k+1)/2) k^{k/2}}{\sqrt{\pi} \Gamma(k/2)} \int_{-\infty}^q x (k+x^2)^{-(k+1)/2} A(x) dx$$

for q satisfying $F_X(c) = p$. The noncentral t distribution has received much applications in risk management since the paper by Harvey and Siddique (1999).

3.10 Stable distribution

A random variable, say S , is said to have stable distribution with tail index parameter $1 < \alpha \leq 2$ and asymmetry parameter $\beta \in [-1, 1]$ if its characteristic function is given by

$$\ln \phi_S(t) = -|t|^\alpha \left[1 - i\beta \text{sign}(t) \tan \frac{\pi\alpha}{2} \right],$$

where $i = \sqrt{-1}$ is the imaginary unit. We write $S \sim S_{\alpha, \beta}(0, 1)$. Let X denote the location-scale variant $X = \mu + \sigma S$ and let $q_p = F_S^{-1}(p)$ denote the p th quantile of S . Broda and Paoella (2011, Section 2.2.3) also show the expected shortfall for X can be expressed as

$$\text{ES}_p(X) = \frac{1}{p} \text{Stoy}(q_p, \alpha, \beta),$$

where

$$\begin{aligned} \text{Stoy}(c, \alpha, \beta) &= \frac{\alpha}{\alpha - 1} \frac{|c|}{\pi} \int_{-\bar{\theta}_0}^{\pi/2} g(\theta) \exp \left\{ -|c|^{\alpha/(\alpha-1)} v(\theta) \right\} d\theta, \\ g(\theta) &= \frac{\sin \{ \alpha (\bar{\theta}_0 + \theta) - 2\theta \}}{\sin \{ \alpha (\bar{\theta}_0 + \theta) \}} - \frac{\alpha \cos^2 \theta}{\sin^2 \{ \alpha (\bar{\theta}_0 + \theta) \}}, \\ v(\theta) &= \{ \cos (\alpha \bar{\theta}_0) \}^{1/(\alpha-1)} \left[\frac{\cos \theta}{\sin \{ \alpha (\bar{\theta}_0 + \theta) \}} \right]^{\alpha/(\alpha-1)} \frac{\cos \{ \alpha (\bar{\theta}_0 + \theta) - \theta \}}{\cos \theta}, \\ \bar{\theta}_0 &= \frac{1}{\alpha} \arctan \left\{ \bar{\beta} \tan \left(\frac{\pi\alpha}{2} \right) \right\}, \quad \bar{\beta} = \text{sign}(c)\beta. \end{aligned}$$

3.11 Generalized hyperbolic distribution

A random variable X follows the generalized hyperbolic distribution if its has the pdf

$$f_X(x) = \frac{(\eta/\delta)^\lambda}{\sqrt{2\pi} K_\lambda(\delta\eta)} \frac{K_{\lambda-1/2} \left(\alpha \sqrt{\delta^2 + (x-\mu)^2} \right)}{\left\{ \sqrt{\delta^2 + (x-\mu)^2}/\alpha \right\}^{1/2-\lambda}} \exp [\beta(x-\mu)],$$

where $\mu \in \mathbb{R}$ is the location parameter, $\alpha \in \mathbb{R}$ is the shape parameter, $\beta \in \mathbb{R}$ is the asymmetry parameter, $\delta \in \mathbb{R}$ is the scale parameter, $\lambda \in \mathbb{R}$, $\eta = \sqrt{\alpha^2 - \beta^2}$, and $K_\nu(\cdot)$ is the modified Bessel

function of order ν . Broda and Paoletta (2011, Section 2.2.4) also show the expected shortfall for X can be expressed as

$$\text{ES}_p(X) = \frac{(\eta/\delta)^\lambda}{\sqrt{2\pi}K_\lambda(\delta\eta)} \int_{-\infty}^q x \frac{K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)}{\left\{\sqrt{\delta^2 + (x-\mu)^2}/\alpha\right\}^{1/2-\lambda}} \exp[\beta(x-\mu)] dx$$

for q satisfying $F_X(c) = p$.

3.12 Normal mixture distribution

Broda and Paoletta (2011, Section 2.3.2) also derive a formula for expected shortfall for a k -component normal mixture. Let X denote a random variable with the cdf

$$F_X(x) = \sum_{i=1}^k \lambda_i \Phi\left(\frac{x - \mu_i}{\sigma_i}\right),$$

where λ_i are non-negative weights summing to one, $-\infty < \mu_i < \infty$ and $\sigma_i > 0$, where $\Phi(\cdot)$ denotes the standard normal cdf. Let q_p denote the quantile defined by $F_X(q_p) = p$ and let $c_j = (q_p - \mu_j)/\sigma_j$. Broda and Paoletta (2011, Section 2.3.2) show that the expected shortfall of X can be expressed as

$$\text{ES}_p(X) = \sum_{i=1}^k \frac{\lambda_i \Phi(c_i)}{p} \left\{ \mu_i - \sigma_i \frac{\phi(c_i)}{\Phi(c_i)} \right\}.$$

3.13 Stable mixture distribution

Let a random variable X represent a k -component normal mixture of symmetric stable random variables with non-negative weights λ_j , location parameters μ_j , scale parameters σ_j and zero asymmetry parameters. Let q_p denote the p th quantile of X and let $c_j = (q_p - \mu_j)/\sigma_j$. Broda and Paoletta (2011, Section 2.3.3) also show that the expected shortfall of X can be expressed as

$$\text{ES}_p(X) = \frac{1}{p} \sum_{i=1}^k \lambda_i [\sigma_i \text{Stoy}(c_i, \alpha_i) + \mu_i F_S(c_i)],$$

where $S \sim S_{\alpha,0}(0, 1)$.

3.14 Student's t mixture distribution

Let a random variable X represent a k -component normal mixture of Student's t random variables with non-negative weights λ_j , location parameters μ_j , scale parameters σ_j and degrees of freedom ν_j . Let q_p denote the p th quantile of X and let $c_j = (q_p - \mu_j)/\sigma_j$. Broda and Paoletta (2011, Section 2.3.4) also show that the expected shortfall of X can be expressed as

$$\text{ES}_p(X) = \frac{1}{p} \sum_{i=1}^k \lambda_i \left[\sigma_i \text{Ttail}(c_i, \nu_i) + \frac{\mu_i \nu_i^{-1/2}}{B(\nu_i/2, 1/2)} \int_{-\infty}^{c_i} \left(1 + \frac{x^2}{\nu_i}\right)^{-(\nu_i+1)/2} dx \right].$$

3.15 Generalized Pareto distribution

Suppose the financial observations of interest, say X_1, X_2, \dots, X_n , follow the generalized Pareto distribution given by the cdf

$$F(x) = 1 - \left\{ 1 + \xi \frac{x - u}{\sigma} \right\}^{-1/\xi},$$

where either $u < x < \infty$ ($\xi \geq 0$) or $u < x < u - \sigma/\xi$ ($\xi < 0$). In this case, Pattarathammas *et al.* (2008) show that the expected shortfall can be expressed as

$$\text{ES}_p(X) = \frac{1}{1-\xi} \left\{ u + \frac{\sigma}{\xi} \left[\left\{ \frac{n(1-p)}{N_u} \right\}^{-\xi} - 1 \right] \right\} + \frac{\beta - \xi u}{1-\xi}$$

where N_u is the number of observations exceeding u . An estimate of $\text{ES}_p(X)$ can be obtained by replacing the parameters, ξ and σ , by their maximum likelihood estimators.

3.16 Asymmetric exponential power distribution

A random variable X is said to have the asymmetric exponential power distribution if its pdf is given by

$$f(x) = \begin{cases} \frac{\alpha}{\alpha^*} K(p_1) \exp \left[-\frac{1}{p_1} \left| \frac{x}{2\alpha^*} \right|^{p_1} \right], & \text{if } x \leq 0, \\ \frac{1-\alpha}{1-\alpha^*} K(p_2) \exp \left[-\frac{1}{p_2} \left| \frac{x}{2(1-\alpha^*)} \right|^{p_2} \right], & \text{if } x > 0, \end{cases}$$

where $0 < \alpha < 1$ is the skewness parameter, $p_1 > 0$, $p_2 > 0$, $K(v) = 1/[2p^{1/p}\Gamma(1+1/p)]$, $\alpha^* = \alpha K(p_1)/[\alpha K(p_1) + (1-\alpha)K(p_2)]$, and

$$\frac{\alpha}{\alpha^*} K(p_1) = \frac{1-\alpha}{1-\alpha^*} K(p_2) = \alpha K(p_1) + (1-\alpha) K(p_2) = B.$$

For a standard asymmetric exponential power random variable, Zhu and Galbraith (2011) have shown that the expected shortfall of X is given by

$$\begin{aligned} \text{ES}_p(X) &= \frac{2}{F(q)} \left\{ -\alpha \alpha^* E(p_1) \left[1 - G \left(h_1(q), \frac{2}{p_1} \right) \right] \right. \\ &\quad \left. + (1-\alpha)(1-\alpha^*) E(p_2) G \left(h_2(q), \frac{2}{p_2} \right) \right\}, \end{aligned}$$

where $F(\cdot)$ denotes the cdf of X , q is the root of $F(q) = p$, and

$$\begin{aligned} E(p) &= p^{1/p} \Gamma(2/p)/\Gamma(1/p), \\ G(x, a) &= \gamma(a, x)/\Gamma(a), \\ h_1(q) &= \frac{1}{p_1} \left| \frac{\min(q, 0)}{2\alpha^*} \right|^{p_1}, \quad h_2(q) = \frac{1}{p_2} \left| \frac{\max(q, 0)}{2(1-\alpha^*)} \right|^{p_2}, \end{aligned}$$

where $\gamma(a, x) = \int_0^x t^{a-1} \exp(-t) dt$ denotes the incomplete gamma function.

3.17 Generalized asymmetric Student's t distribution

A random variable X is said to have the generalized asymmetric Student's t distribution if its pdf is given by

$$f(x) = \begin{cases} \frac{\alpha}{\alpha^*} K(v_1) \left[1 + \frac{1}{v_1} \left(\frac{x}{2\alpha^*} \right)^2 \right]^{-\frac{v_1+1}{2}}, & \text{if } x \leq 0, \\ \frac{1-\alpha}{1-\alpha^*} K(v_2) \left[1 + \frac{1}{v_2} \left(\frac{x}{2(1-\alpha^*)} \right)^2 \right]^{-\frac{v_2+1}{2}}, & \text{if } x > 0, \end{cases}$$

where $0 < \alpha < 1$ is the skewness parameter, $v_1 > 0$, $v_2 > 0$, $K(v) = \Gamma((v+1)/2)/[\sqrt{\pi v} \Gamma(v/2)]$, $\alpha^* = \alpha K(v_1)/[\alpha K(v_1) + (1-\alpha)K(v_2)]$, and

$$\frac{\alpha}{\alpha^*} K(v_1) = \frac{1-\alpha}{1-\alpha^*} K(v_2) = \alpha K(v_1) + (1-\alpha) K(v_2) = B.$$

For a standard generalized asymmetric Student's t random variable, Zhu and Galbraith (2009) have shown that the expected shortfall of X is given by

$$\begin{aligned} ES_p(X) &= \frac{4B}{F(q)} \left\{ (\alpha^*)^2 \frac{v_1}{v_1-1} \left[1 + \frac{1}{v_1} \left(\frac{\min(q, 0)}{2\alpha^*} \right)^2 \right]^{(1-v_1)/2} \right. \\ &\quad \left. + (1-\alpha^*)^2 \frac{v_2}{v_2-1} \left(\left[1 + \frac{1}{v_2} \left(\frac{\max(q, 0)}{2(1-\alpha^*)} \right)^2 \right]^{(1-v_2)/2} - 1 \right) \right\}, \end{aligned}$$

where $F(\cdot)$ denotes the cdf of X and q is the root of $F(q) = p$.

3.18 Mitnik and Paoella's generalized t distribution

Mitnik and Paoella (2000)'s generalized t distribution has the pdf specified by

$$f(x) = I(x < 0)C \left[1 + \frac{(-x\psi)^d}{\nu} \right]^{-(\nu+1/2)} + I(x \geq 0)C \left[1 + \frac{(x/\psi)^d}{\nu} \right]^{-(\nu+1/2)}$$

where $d > 0$, $\nu > 0$, $\psi > 0$ and $C = [(\psi + 1/\psi)d^{-1}\nu^{1/d}B(1/d, \nu)]^{-1}$. Let X denote a random variable having this distribution. Taylor (2008a) has shown that the expected shortfall of X is given by

$$ES_p(X) = \begin{cases} -\frac{C\nu^{2/d}}{p\psi^2 d} B_L(\nu - 1/d, 2/d), & \text{if } p \leq 1/2, \\ \frac{C\nu^{2/d}\psi}{(1-p)d} [B(2/d, \nu - 1/d) - B_U(2/d, \nu - 1/d)], & \text{if } p > 1/2, \end{cases}$$

where $L = \nu / \{\nu + (-q\psi)^d\}$, $U = (q/\psi)^d / \{\nu + (q/\psi)^d\}$, q is the p th quantile of X and

$$B_x(a, b) = \int_0^x t^{a-1} (1-t)^{b-1} dt$$

is the incomplete beta function.

3.19 Asymmetric Laplace distribution

Lu *et al.* (2010)'s asymmetric Laplace distribution has the pdf specified by

$$f(x) = \frac{b}{\tau} \exp \left[-\frac{b}{\tau} |x - \tau| \left(\frac{1}{c} I(x < \gamma) + \frac{1}{1-c} I(x > \gamma) \right) \right],$$

where $b = \sqrt{c^2 + (1-c)^2}$, γ is a location parameter, τ is a scale parameter, and c is a shape parameter. Let X denote a random variable having this pdf. Chen *et al.* (2012) have shown that the expected shortfall of X is given by

$$\text{ES}_p = \frac{c}{b} \left[\ln \left(\frac{p}{c} \right) - 1 \right]$$

for $0 \leq p < c$.

3.20 Elliptical distribution

Suppose a portfolio loss can be expressed as $X = \delta_1 X_1 + \cdots + \delta_n X_n = \boldsymbol{\delta}^T \mathbf{X}$, where δ_i are non-negative weights summing to one, X_i are assets losses, $\boldsymbol{\delta} = (\delta_1, \dots, \delta_n)$, and $\mathbf{X} = (X_1, \dots, X_n)$. Suppose too that \mathbf{X} has the elliptical distribution given by the pdf $f(\mathbf{x}) = |\boldsymbol{\Sigma}|^{-1} g((\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}))$. In this case, Kamdem (2005) shows that the expected shortfall of X can be expressed as

$$\text{ES}_p(X) = -\boldsymbol{\delta}\boldsymbol{\mu} + |\boldsymbol{\delta}^T \boldsymbol{\Sigma} \boldsymbol{\delta}|^{1/2} \frac{\pi^{(n-1)/2}}{p\Gamma((n+1)/2)} \int_{q^2}^{\infty} (u - q^2)^{(n-1)/2} g(u) du, \quad (6)$$

where q is the root of

$$\frac{\pi^{(n-1)/2}}{\Gamma((n+1)/2)} \int_q^{\infty} \int_z^{\infty} (u - z)^{(n-3)/2} g(u) du dz = p.$$

The multivariate t distribution mean vector $\boldsymbol{\mu}$, covariance matrix $\boldsymbol{\Sigma}$ and degrees of freedom ν is a member of the elliptical family. For this particular case, Kamdem (2005) shows that (6) can be reduced to

$$\text{ES}_p(X) = -\boldsymbol{\delta}\boldsymbol{\mu} + a |\boldsymbol{\delta}^T \boldsymbol{\Sigma} \boldsymbol{\delta}|^{1/2},$$

where a is the root of

$$a = \frac{\Gamma((\nu-1)/2) \nu^{\nu/2}}{p\sqrt{\pi}\Gamma(\nu/2)} (\nu + q^2)^{-(\nu+1)/2},$$

where q is the root of

$$\frac{\nu^{\nu/2}\Gamma((\nu+1)/2)}{\nu\sqrt{\pi}s^{\nu}\Gamma(\nu/2)} {}_2F_1\left(\frac{1+\nu}{2}, \frac{\nu}{2}; 1 + \frac{\nu}{2}; -\frac{\nu}{s^2}\right) = p,$$

where ${}_2F_1(a, b; c; x)$ denotes the Gauss hypergeometric function.

3.21 Multivariate gamma distribution

Suppose a portfolio loss can be expressed as $S = X_1 + \dots + X_n$, where X_i are assets losses. Suppose too that (X_1, \dots, X_n) has Mathai and Moschopoulos (1991)'s multivariate gamma distribution; that is,

$$X_j = \frac{\alpha_0}{\alpha_j} Y_0 + Y_j$$

for $j = 1, 2, \dots, n$, where Y_j , $j = 0, 1, \dots, n$ are independent gamma random variables with shape parameters γ_j and scale parameters α_j . According to Mathai and Moschopoulos (1991, Theorem 2.1), the pdf of S can be expressed as

$$f_S(s) = \sum_{k=0}^{\infty} p_k g(s|\gamma + k, \alpha_{\max}),$$

where $\alpha_{\max} = \max(\alpha_1, \dots, \alpha_n)$, $\gamma = \gamma_1 + \dots + \gamma_n$, $p_k = C\delta_k$, $k = 0, 1, \dots$, where

$$\begin{aligned} C &= \prod_{j=1}^n \left(\frac{\alpha_j}{\alpha_{\max}} \right)^{\gamma_j}, \\ \Delta_{j,i} &= \left(1 - \frac{\alpha_j}{\alpha_{\max}} \right)^i, \quad j = 1, 2, \dots, n, \quad i = 1, 2, \dots, k, \\ \delta_k &= k^{-1} \sum_{i=1}^k \sum_{j=1}^n \gamma_j \Delta_{j,i} \delta_{k-i}, \quad k > 0 \end{aligned}$$

and $\delta_0 = 1$. The cdf of S is a gamma random variable with shape parameter $\gamma + K$ and scale parameter α_{\max} , where K is a discrete random variable with pmf $p_k = C\delta_k$, $k = 0, 1, \dots$.

Furman and Landsman (2005) derive an expression for the expected shortfall of S . Let V an independent convolution of a gamma random variable with shape parameter $\gamma + K + 1$ and scale parameter α_{\max} and another gamma random variable with shape parameter γ_0 and scale parameter α_0/η . Let Z_t , $t = 0, 1, \dots, n$ denote a gamma random variable with unit shape parameter and scale parameter α_t . Let Z_{\max} denote the Z_t for which $\alpha_t = \max(\alpha_0, \alpha_{\max})$. Further, let $E_K(\cdot)$ and $E_V(\cdot)$ denote the expectations with respect to K and V , respectively. Then, according to Furman and Landsman (2005, Theorem 1), the expected shortfall of S can be expressed as

$$\begin{aligned} \text{ES}_p(S) &= \eta \frac{\gamma_0}{\alpha_0} \frac{1 - F_{S+\eta Z_0}(s_p)}{1 - F_S(s_p)} + \frac{\gamma}{\alpha_{\max}} \frac{1 - F_{S+Z_{\max}}(s_p)}{1 - F_S(s_p)} \\ &\quad + \frac{E_K(K E_V(\Gamma^{-1}(\gamma + K + V + 1) \gamma(\gamma + K + V + 1, \alpha_{\max})))}{\alpha_{\max} F_S(s_p)}, \end{aligned}$$

where s_p satisfies $F_S(s_p) = p$ and $\gamma(a, x) = \int_0^x t^{a-1} \exp(-t) dt$ denotes the incomplete gamma function.

3.22 Bayesian approach

Let X_1, X_2, \dots, X_n denote the financial series of interest and let $PL(\cdot)$ denote the profit and loss function associated with the series. Suppose the series is fitted to a model parameterized by θ ,

where θ has a prior pdf $\pi(\theta)$. Then an algorithm for Bayesian computation of expected shortfall due to Hoogerheide and van Dijk (2008) is

1. simulate θ^i , $i = 1, 2, \dots, m$ from the posterior distribution of θ . This can be performed using either Gibbs sampling (Geman and Geman, 1984)) or the Metropolis-Hastings algorithm (Metropolis *et al.*, 1953; Hastings, 1970);
2. simulate future values of the financial series, say $\mathbf{X}^{*i} = \{X_{n+1}^i, X_{n+2}^i, \dots, X_{n+\tau}^i\}$, $i = 1, 2, \dots, m$ given the θ^i and the past observations;
3. compute the profit and loss function values $PL(\mathbf{X}^{*i})$, $i = 1, 2, \dots, m$ and order them in ascending order to yield $PL^{(j)}$, $j = 1, 2, \dots, m$;
4. estimate expected shortfall as

$$\widehat{ES}_p = \frac{1}{n(1-p)} \sum_{j=1}^{n(1-p)} PL^{(j)}.$$

Hoogerheide and van Dijk (2008) also propose a variant of this algorithm incorporating importance sampling.

3.23 Random walks

Suppose that the financial returns $\{X_t\}$ is a random walk; that is, $X_t = X_{t-h} + r_t$, where r_t are independent normal random variables with mean μ and variance σ^2 . Let

$$\widehat{\mu} = \frac{1}{n} \sum_{i=1}^n r_{ih}, \quad \widehat{\sigma} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (r_{ih} - \widehat{\mu})^2}, \quad \widehat{\sigma}^k = \sqrt{k} \widehat{\sigma}, \quad \widehat{\mu}^k = k \widehat{\mu}.$$

Let $r_i^k = \sum_{i=0}^{k-1} r_{t-ih}$ denote the k -period returns. With this notation, Embrechts *et al.* (2005) show that the expected shortfall of k -period returns can be estimated by

$$\widehat{ES}_p = \frac{1}{p} \exp \left(\widehat{\mu}^k + \frac{(\widehat{\sigma}^k)^2}{2} \right) \Phi \left(\Phi^{-1}(p) - \widehat{\sigma}^k \right) - 1,$$

where $\Phi(\cdot)$ denotes the standard normal cdf.

3.24 Autoregressive process

Suppose that the financial returns $\{X_t\}$ is an autoregressive process; that is,

$$X_t = \sum_{i=1}^q a_i X_{t-i} + e_t,$$

or alternatively

$$\bar{X}_t = \sum_{i=1}^q a_i \bar{X}_{t-i} + \bar{e}_t,$$

where a_i are parameters with $|a_i|$ summing to a number less than one, e_i are independent normal random variables with mean $\mu_0 + \mu_1 t$ and constant variance σ^2 , $\bar{X}_t = X_t - \mu t$, $\mu = \mu_1/(1 - a_1 - \dots - a_p)$, and \bar{e}_t are independent normal random variables with zero mean and constant variance σ^2 . Let \hat{q} , \hat{a}_i and $\hat{\sigma}$ denote the maximum likelihood estimates. Let

$$\hat{\mu}^k = kh\hat{\mu} + \hat{m}, \quad \hat{\sigma}^k = \hat{\sigma} \sqrt{\sum_{j=0}^{k-1} \delta_j^2},$$

where

$$\begin{aligned}\hat{\mu} &= (s_{nh} - s_0) / (nh), \\ \hat{m} &= \hat{s}_{t+kh} - \tilde{s}_t, \\ \delta_j &= \sum_{i=1}^j \hat{a}_i I(i=j),\end{aligned}$$

where $\hat{a}_i = 0$ for $i > \hat{q}$ and \hat{s}_{t+kh} satisfy

$$\begin{aligned}\hat{s}_{t+jh} &= \sum_{i=1}^{\hat{q}} \hat{a}_i \hat{s}_{t+(j-i)h}, \quad j = 1, 2, \dots, k, \\ \hat{s}_u &= \bar{s}_u, \quad u \leq t.\end{aligned}$$

With this notation, Embrechts *et al.* (2005) show that the expected shortfall of k -period returns can be estimated by

$$\widehat{\text{ES}}_p = \frac{1}{p} \exp \left(\hat{\mu}^k + \frac{(\hat{\sigma}^k)^2}{2} \right) \Phi \left(\Phi^{-1}(p) - \hat{\sigma}^k \right) - 1,$$

where $\Phi(\cdot)$ denotes the standard normal cdf.

3.25 GARCH (1, 1) process

Suppose that the financial returns $\{X_t\}$ is a GARCH (1, 1) process; that is,

$$X_t = X_{t-h} + r_t,$$

where $r_t = \bar{r}_t + \mu$, $\bar{r}_t = \sigma_t e_t$, $\sigma_t^2 = \alpha_0 + \alpha_1 \bar{r}_{t-h}^2 + \beta_1 \sigma_{t-h}^2$, and e_t are independent standard normal variables. Let $\hat{\mu}$, $\hat{\alpha}_0$, $\hat{\alpha}_1$ and $\hat{\beta}_1$ denote the maximum likelihood estimates. Let

$$\hat{\mu}^k = k\hat{\mu}, \quad \hat{\sigma}^k = \hat{\sigma}(t, t),$$

where $\hat{\sigma}(t, t)$ is specified by

$$\begin{aligned}\hat{\sigma}^2(t^*, t) &= \hat{\alpha}_0 + \hat{\alpha}_1 \left(r_{t^*}^k - \hat{\mu}^k \right)^2 + \hat{\beta}_1 \hat{\sigma}^2(t^* - kh, t), \\ \hat{\sigma}^2(t - nh, t) &= \frac{k}{nk - 1} \sum_{i=0}^{nk-1} (r_{t-ih} - \hat{\mu})^2\end{aligned}$$

for $t^* = t - (n - 1)kh, \dots, t - kh, t$ and n denoting the number of k -period returns. With this notation, Embrechts *et al.* (2005) show that the expected shortfall of k -period returns can be estimated by

$$\widehat{\text{ES}}_p = \frac{1}{p} \int_0^p \exp \left(\hat{\mu}^k + \hat{\sigma}^k x_{\nu^k}^q \right) dq - 1,$$

where x_{ν}^q denotes the q th quantile of a Student's t random variable with ν degrees of freedom and $\hat{\mu}^k$ is calculated using a six-step procedure described in Section 5 of Embrechts *et al.* (2005).

3.26 Quantile regression method

Quantile regression (Koenker and Hallock, 2001; Koenker, 2005) has been used to estimate expected shortfall, see Taylor (2008b). The idea is to regress the financial returns on some known covariates. Let X_t at time t denote the financial return, let \mathbf{z}_t denote a $k \times 1$ vector of covariates at time t , and let $\boldsymbol{\beta}$ denote a $k \times 1$ vector of regression coefficients. Taylor (2008b) provides several quantile based estimators for expected shortfall. Two of them are

$$\widehat{\text{ES}}_p = \frac{1}{pn} \sum_{t=1}^n \left(X_t - \mathbf{z}_t^T \widehat{\boldsymbol{\beta}} \right) \left[p - I(X_t < \mathbf{z}_t^T \widehat{\boldsymbol{\beta}}) \right]$$

and

$$\widehat{\text{ES}}_p = \frac{1}{p \sum_{t=1}^n \lambda^{n-t}} \sum_{t=1}^n \lambda^{n-t} \left(X_t - \mathbf{z}_t^T \widehat{\boldsymbol{\beta}} \right) \left[p - I(X_t < \mathbf{z}_t^T \widehat{\boldsymbol{\beta}}) \right],$$

where $\widehat{\boldsymbol{\beta}}$ is chosen as the $\boldsymbol{\beta}$ that minimizes

$$\sum_{t=1}^n \lambda^{n-t} \left(X_t - \mathbf{z}_t^T \boldsymbol{\beta} \right) \left[p - I(X_t < \mathbf{z}_t^T \boldsymbol{\beta}) \right]$$

for some weighting parameter λ . Another estimator suggested by Taylor (2008b) is

$$\begin{aligned}\widehat{\text{ES}}_p &= \frac{1}{p \sum_{t=1}^n \lambda^{n-t}} \sum_{t=1}^n \lambda^{n-t} \left[p \left(X_t - \mathbf{z}_t^T \widehat{\boldsymbol{\beta}} \right) + \left(\mathbf{z}_t^T \widehat{\boldsymbol{\beta}} - X_t \right) \Phi \left(\frac{\mathbf{z}_t^T \widehat{\boldsymbol{\beta}} - X_t}{h} \right) \right. \\ &\quad \left. + h \phi \left(\frac{\mathbf{z}_t^T \widehat{\boldsymbol{\beta}} - X_t}{h} \right) \right],\end{aligned}$$

where $\hat{\beta}$ is chosen as the β that minimizes

$$\sum_{t=1}^n \lambda^{n-t} \left[p(X_t - \mathbf{z}_t^T \beta) + (\mathbf{z}_t^T \beta - X_t) \Phi\left(\frac{\mathbf{z}_t^T \beta - X_t}{h}\right) + h\phi\left(\frac{\mathbf{z}_t^T \beta - X_t}{h}\right) \right]$$

for some weighting parameter λ and a suitable bandwidth h .

3.27 Location-scale distributions

Suppose X_1, X_2, \dots, X_n is a random sample from the location-scale family with cdf $F_{\mu,\sigma}(x) = F_0((x - \mu)/\sigma)$ and pdf $f_{\mu,\sigma}(x) = (1/\sigma)f_0((x - \mu)/\sigma)$, where $f_0(x) = dF_0(x)/dx$. Then,

$$\text{ES}_p = \mu + e_p \sigma, \quad (7)$$

where $e_p = (1/(1-p)) \int_{z_p}^{\infty} x f_0(x) dx$ and $z_p = F_0^{-1}(p)$. The point estimate for expected shortfall is

$$\widehat{\text{ES}}_p = \widehat{\mu}_n + e_p c_n \widehat{\sigma}_n,$$

where

$$\widehat{\mu}_n = \frac{1}{n} \sum_{i=1}^n X_i,$$

$$\widehat{\sigma}_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \widehat{\mu}_n)^2,$$

and

$$c_n = (E[\widehat{\sigma}/\sigma])^{-1}.$$

Bae and Iscoe (2012) propose various confidence intervals for expected shortfall. Based on $c_n = 1 + O(n^{-1})$ and asymptotic normality, Bae and Iscoe (2012) propose the interval

$$\widehat{\mu}_n + e_p \widehat{\sigma}_n \pm \frac{\widehat{\sigma}_n}{n} z_{(1+\alpha)/2} \sqrt{1 + \frac{e_p^2}{4}(\kappa - 1) + e_p \omega}, \quad (8)$$

where α is the confidence level, κ is the kurtosis of $F_0(x)$, and ω is the skewness of $F_0(x)$. Based on an empirical estimator due to Brazauskas *et al.* (2008) (see Section 4.3), Bae and Iscoe (2012) propose the interval

$$\frac{1}{1-p} \int_p^1 \widehat{F}_{\mu,\sigma}(u) du \pm \frac{1}{\sqrt{n}} z_{(1+\alpha)/2} \sqrt{\frac{\sigma^2}{1-p} [p(e_p - z_p)^2 + \sigma_p^2]},$$

where $\widehat{F}_{\mu,\sigma}(\cdot)$ denotes the empirical cdf and

$$\sigma_p^2 = \frac{1}{1-p} \int_{z_p}^{\infty} z^2 f_0(z) dz - e_p^2.$$

Sometimes the financial series of interest is strictly positive. In this case, if X_1, X_2, \dots, X_n is a random sample from a log location-scale family with distribution function $G_{\mu,\sigma}(x) = \ln F_0((x - \mu)/\sigma)$, then (7) and (8) will generalize to

$$\text{ES}_p = \exp(\mu + h_p(\sigma))$$

and

$$\exp \left[\hat{\mu}_n + h_p(\hat{\sigma}_n) \pm \frac{\hat{\sigma}_n}{\sqrt{n}} z_{(1+\alpha)/2} \sqrt{1 + \frac{\nu_p^2(\hat{\sigma}_n)}{4} (\kappa - 1) + \nu_p(\hat{\sigma}_n) \omega} \right],$$

respectively, as noted by Bae and Iscoe (2012), where

$$h_p(y) = \ln \left[\frac{1}{1-p} \int_{z_p}^{\infty} \exp(ty) f_0(t) dt \right]$$

and

$$\nu_p(\sigma) = \frac{\int_{z_p}^{\infty} t \exp(\sigma t) f_0(t) dt}{\int_{z_p}^{\infty} \exp(\sigma t) f_0(t) dt}.$$

3.28 RiskMetrics model

Let r_t be the financial return at time t and let Ω_t denote the information available up to time t . Then the aggregate return from time t to time $t+h$ is $r_{t+1} + \dots + r_{t+h} = R_{t,h}$ say. The RiskMetrics model (RiskMetrics Group, 1996) supposes $R_{t,h} | \Omega_t$ is normal with mean $h\mu$ and variance $h\sigma_{t+1}^2$, where $\mu = E(r_{t+1} | \Omega_t)$ and $\sigma_{t+1}^2 = \text{Var}(r_{t+1} | \Omega_t)$. So, the corresponding expected shortfall is

$$\widehat{\text{ES}}_p = h\mu - \frac{\sqrt{h}}{p} \sigma_{t+1} \phi(\Phi^{-1}(p)),$$

where $\phi(\cdot)$ denotes the standard normal pdf and $\Phi(\cdot)$ denotes the standard normal cdf.

3.29 QGARCH (1, 1) model

With the notation as in Section 3.28, suppose $R_{t,h} | \Omega_t$ is normal with mean $h\mu$ and variance $\text{Var}(R_{t,h} | \Omega_t)$ unspecified. In this case, one has the QGARCH (1, 1) model (Wong and So, 2010) with the corresponding expected shortfall given by

$$\widehat{\text{ES}}_p = h\mu - \frac{\sqrt{\text{Var}(R_{t,h} | \Omega_t)}}{p} \phi(\Phi^{-1}(p)),$$

where $\phi(\cdot)$ denotes the standard normal pdf and $\Phi(\cdot)$ denotes the standard normal cdf.

3.30 QGARCH (p, q) model

With the notation as in Section 3.28, suppose $R_{t,h}$ follows Theodossiou (1998)'s skew t distribution given by the pdf

$$f(x) = \begin{cases} C \left[1 + \frac{2}{v-2} \left(\frac{x+a}{\theta(1-\tau)} \right)^2 \right]^{-(v+1)/2}, & \text{if } x < -a, \\ C \left[1 + \frac{2}{v-2} \left(\frac{x+a}{\theta(1+\tau)} \right)^2 \right]^{-(v+1)/2}, & \text{if } x \geq -a \end{cases} \quad (9)$$

with $Var(R_{t,h} | \Omega_t)$ unspecified, where

$$\begin{aligned} C &= \frac{\sqrt{B(3/2, v/2-1)}S(\tau)}{B^{3/2}(1/2, v/2)}, \\ \theta &= \frac{\sqrt{2}}{S(\tau)}, \\ a &= \frac{2\tau B(1, (v-1)/2)}{S(\tau)\sqrt{B(1/2, v/2)}\sqrt{B(3/2, v/2-1)}}, \\ S(\tau) &= \sqrt{1 + 3\tau^2 - \frac{4\tau^2 B^2(1, (v-1)/2)}{B(1/2, v/2)B(3/2, v/2-1)}}. \end{aligned}$$

Let $F(\cdot)$ denote the cdf corresponding to (9). In this case, one has the QGARCH (p, q) model (Wong and So, 2010) with the corresponding expected shortfall given by

$$\widehat{ES}_p = h\mu - \sqrt{Var(R_{t,h} | \Omega_t)} \left[\frac{\theta^2(1-\tau)^2 f(F^{-1}(p))}{p} \frac{v-2+2\beta^2(p)}{2(v-1)} + a \right],$$

where $\beta(p) = \{a + F^{-1}(p)\} / [\theta(1-\tau)]$.

3.31 Block minimum method

Let r_t denote the financial return at time t . Let $r_{(1)}$ denote the minimum of the returns over a long enough period of time. According to extreme value theory (see, for example, Leadbetter *et al.* (1987)), the cdf of $r_{(1)}$ can be approximated by

$$F(x) = \exp \left\{ - \left(1 + \xi \frac{x-\mu}{\sigma} \right)^{-1/\xi} \right\} \quad (10)$$

for $1 + \xi(x-\mu)/\sigma > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$ and $\xi \in \mathbb{R}$. The corresponding pdf is

$$f(x) = \frac{1}{\sigma} \left(1 + \xi \frac{x-\mu}{\sigma} \right)^{-1/\xi-1} \exp \left\{ - \left(1 + \xi \frac{x-\mu}{\sigma} \right)^{-1/\xi} \right\} \quad (11)$$

for $1 + \xi(x-\mu)/\sigma > 0$, $\mu \in \mathbb{R}$, $\sigma > 0$ and $\xi \in \mathbb{R}$. The distribution given by (10) and (11) is known as the generalized extreme value distribution. The corresponding expected shortfall can be computed

as

$$\text{ES}_p = \frac{1}{\sigma} \int_{-\infty}^u x \left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-1/\xi-1} \exp \left\{ - \left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-1/\xi} \right\} dx,$$

where

$$u = \mu - \frac{\sigma}{\xi} \left[1 - \{-\ln p\}^{-\xi} \right]. \quad (12)$$

See Ou and Yi (2009).

4 Nonparametric methods

4.1 Historical method

Let $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ denote the order statistics in ascending order corresponding to the original financial returns X_1, X_2, \dots, X_n . The historical method suggests to estimate expected shortfall by

$$\widehat{\text{ES}}_p(X) = \left(\sum_{i=[np]}^n X_{(i)} \right) / (n - [np]),$$

where $[x]$ denotes the largest integer not greater than x .

4.2 Filtered historical method

Let \widehat{e}_i , $i = 1, 2, \dots, n$ denote the residuals after the financial series is fitted to some model like ARMA-GARCH. Then the filtered historical estimator of expected shortfall (Magadia, 2011) is given by

$$\widehat{\text{ES}}_p(X) = \frac{\sum_{\eta_t > q} \eta_t}{\sum_{\eta_t > q} I_{\eta_t > q}},$$

where

$$\eta_t = \widehat{e}_t - \frac{1}{n} \sum_{t=1}^n \widehat{e}_t$$

and $q = \eta_{([pn]+1)}$ is the $([pn]+1)$ th order statistic of $\{\eta_1, \dots, \eta_n\}$.

4.3 Brazauskas *et al.*'s estimator

For the financial series in Section 4.1, let $\widehat{F}(\cdot)$ denote the empirical cdf and $\widehat{F}^{-1}(\cdot)$ its quantile function. Brazauskas *et al.* (2008) suggest the estimator

$$\widehat{\text{ES}}_p(X) = \frac{1}{p} \int_0^p \widehat{F}^{-1}(u) du.$$

4.4 Yamai and Yoshiba's estimator

With the notation as in Section 4.1, Yamai and Yoshiba (2002) suggest the following estimator for expected shortfall

$$\widehat{ES}_p(X) = \frac{1}{n(\alpha - \beta)} \sum_{i=n\beta}^{n\alpha} X_i,$$

where α is assumed to be much greater than β .

4.5 Inui and Kijima's estimator

With the notation as in Section 4.1, Inui and Kijima (2005) suggest the following estimator for expected shortfall

$$\widehat{ES}_p(X) = \begin{cases} -\bar{X}_{k:n}, & \text{if } n(1-p) \text{ is an integer,} \\ -p\bar{X}_{k:n} - (1-p)\bar{X}_{k+1:n}, & \text{if } n(1-p) \text{ is not an integer,} \end{cases}$$

where

$$\bar{X}_{k:n} = \frac{1}{k} (X_{(1)} + \cdots + X_{(k)})$$

for $k = 1, 2, \dots, n$.

4.6 Chen's estimator

With the notation as in Section 4.1, Chen (2008) suggests the following estimator for expected shortfall

$$\widehat{ES}_p(X) = \frac{1}{1 + [np]} \sum_{i=1}^n X_i I(X_i \geq X_{([n(1-p)]+1)}) .$$

4.7 Peracchi and Tanase's estimator

With the notation as in Section 4.1, Peracchi and Tanase (2008) suggest the following estimator for expected shortfall

$$\widehat{ES}_p(X) = \frac{1}{np} \sum_{i=1}^{[np]} X_{(i)} + \left(1 - \frac{[np]}{np}\right) Y_{([np]+1)} .$$

4.8 Jadhav *et al.*'s estimators

Jadhav *et al.* (2009) propose several modifications of the historical estimator for expected shortfall. With the notation as in Section 4.1, one estimator proposed is

$$\widehat{ES}_p(X) = -\frac{\sum_{i=0}^{\lfloor np^{1+a} \rfloor + 1} X_{(i)}}{\lfloor np^{1+a} \rfloor + 2},$$

where $i = \left[(n+1)p'_{(i)} \right]$,

$$p'_{(i)} = p - \frac{ip}{\lfloor np \rfloor + 1}, \quad i = 0, 1, \dots, \lfloor np^{1+a} \rfloor + 1,$$

and a is a constant taking values in $[0, 0.1]$. Another estimator proposed is

$$\widehat{ES}_p(X) = -\frac{\sum_{i=0}^{\lfloor np^{1+a} \rfloor + 1} (1 - h_i X_{(i)} + kh_i X_{i+1})}{\lfloor np^{1+a} \rfloor + 2},$$

where $i = \left[(n+1)p'_{(i)} \right]$, $h_i = (n+1)p'_{(i)} - \left[(n+1)p'_{(i)} \right]$,

$$p'_{(i)} = p - \frac{ip}{\lfloor np \rfloor + 1}, \quad i = 0, 1, \dots, \lfloor np^{1+a} \rfloor + 1,$$

and a is a constant taking values in $[0, 0.1]$.

4.9 Kernel method

Let $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ denote the order statistics in ascending order corresponding to the financial returns X_1, X_2, \dots, X_n . Let $K(\cdot)$ denote a symmetric kernel, h a suitable bandwidth, $K_h(u) = (1/h)K(u/h)$, $A(x) = \int_{-\infty}^x K(u)du$ and $A_h(u) = A(u/h)$. Yu *et al.* (2010) suggest the following formula for kernel estimation of expected shortfall:

$$\widehat{ES}_p(X) = \frac{1}{np} \sum_{i=1}^n X_i A_h(\widehat{q}(p) - X_i),$$

where

$$\widehat{q}(p) = \sum_{i=1}^n \left[\int_{i-1/n}^{i/n} K_h(t-p) dt \right] X_{(i)}.$$

An alternative is to obtain \widehat{q}_p as the solution of

$$\frac{1}{n} \sum_{i=1}^n A_h(x - x_i) = p.$$

Further details on this kernel method can be seen from Scaillet (2004) and Chen (2008).

4.10 Trimmed kernel method

With the notation as in Section 4.9, let $X_t^{(-)} = X_t I(X_t < 0)$ and let $\{k_n\}$ be such that $k_n \rightarrow \infty$ and $k_n/n \rightarrow 0$ as $n \rightarrow \infty$. Hill (2012) suggests several trimmed estimators for expected shortfall. One of them is

$$\widehat{ES}_p(X) = \frac{1}{np} \sum_{i=1}^n X_i I\left(X_{(k_n)}^{(-)} \leq X_i \leq \widehat{q}(p)\right),$$

where $\widehat{q}(p)$ is as defined in Section 4.9. Three other estimators proposed in Hill (2012) are

$$\widehat{ES}_p(X) = \frac{1}{np} \sum_{i=1}^n X_i I\left(X_{(k_n)}^{(-)} \leq X_i \leq \widehat{q}(p)\right) + \frac{1}{p} \left(\frac{\widehat{\kappa}_{k_n}^{(-)}}{\widehat{\kappa}_{k_n}^{(-)} - 1} \frac{k_n}{n} X_{(l_n)}^{(-)} \right),$$

$$\widehat{ES}_p(X) = \frac{1}{np} \sum_{i=1}^n X_i I\left(X_{(k_n)}^{(-)} \leq X_i \leq \widehat{q}(p)\right) + \frac{1}{p} \left(\frac{\widehat{\kappa}_{m_n}^{(-)}}{\widehat{\kappa}_{m_n}^{(-)} - 1} \frac{k_n}{n} X_{(l_n)}^{(-)} \right),$$

and

$$\widehat{ES}_p(X) = \frac{1}{np} \sum_{i=1}^n X_i I\left(X_{(k_n)}^{(-)} \leq X_i \leq \widehat{q}(p)\right) + \frac{1}{p} \left(\frac{\widehat{\kappa}_{m_n(\widehat{\lambda}_n)}^{(-)}}{\widehat{\kappa}_{m_n(\widehat{\lambda}_n)}^{(-)} - 1} \frac{k_n}{n} X_{(l_n)}^{(-)} \right),$$

where

$$\widehat{\kappa}_{k_n}^{(-)} = \left[\frac{1}{k_n} \sum_{i=1}^n \ln \left(\frac{X_{(i)}^{(-)}}{X_{(k_n)}^{(-)}} \right) \right]^{-1}$$

and

$$\begin{aligned} \widehat{\lambda} = \operatorname{argmin}_{\lambda} & \left| \frac{1}{np} \sum_{i=1}^n X_i I\left(X_{(k_n)}^{(-)} \leq X_i \leq \widehat{q}(p)\right) \right. \\ & \left. + \frac{1}{p} \left(\frac{\widehat{\kappa}_{m_n(\widehat{\lambda}_n)}^{(-)}}{\widehat{\kappa}_{m_n(\widehat{\lambda}_n)}^{(-)} - 1} \frac{k_n}{n} X_{(l_n)}^{(-)} \right) - \widehat{q}(p) \right|, \end{aligned}$$

where m_n satisfies $k_n/m_n \rightarrow 0$, $m_n = o(1)$, and l_n satisfies $\Pr(X_t < -l_n) = k_n/n$. The two latter of the three estimators are bias corrected estimators.

4.11 Richardson's method

Fan *et al.* (2010) have suggested an algorithm for estimating expected shortfall based on Richardson's extrapolation method (Richardson, 1911; Richardson and Gaunt, 1927). The algorithm is as follows:

1. generate a samples X_1, X_2, \dots, X_N by a Monte Carlo method;

2. estimate the corresponding expected shortfall by the historical method;
3. repeat steps 1 and 2 say one thousand times and compute the mean

$$m_N = \frac{1}{1000} \sum_{i=1}^{1000} \widehat{\text{ES}}_{N,i},$$

where $\widehat{\text{ES}}_{N,i}$ are the one thousand expected shortfall estimates;

4. set $s_n = m_{N_n}$, $n = 1, 2, \dots, k+1$ for some k and N_1, N_2, \dots, N_{k+1} (for example, $k = 2$, $N_1 = 100$, $N_2 = 200$ and $N_3 = 300$);
5. estimate expected shortfall as

$$\left| \begin{array}{cccc} s_1 & s_2 & \cdots & s_{k+1} \\ 1 & 1/2 & \cdots & 1/(k+1) \\ \vdots & \vdots & \ddots & \vdots \\ 1^k & (1/2)^k & \cdots & (1/(k+1))^k \end{array} \right| \cdot \left| \begin{array}{cccc} 1 & 1 & \cdots & 1 \\ 1 & 1/2 & \cdots & 1/(k+1) \\ \vdots & \vdots & \ddots & \vdots \\ 1^k & (1/2)^k & \cdots & (1/(k+1))^k \end{array} \right|.$$

For example, if $k = 2$, $N_1 = 100$, $N_2 = 200$ and $N_3 = 300$ then one can estimate expected shortfall as

$$\left| \begin{array}{ccc} m_{100} & m_{200} & m_{300} \\ 1 & 1/2 & 1/3 \\ 1 & 1/4 & 1/9 \end{array} \right| \cdot \left| \begin{array}{ccc} 1 & 1 & 1 \\ 1 & 1/2 & 1/3 \\ 1 & 1/4 & 1/9 \end{array} \right|.$$

5 Semiparametric methods

5.1 Heavy tailed processes

Suppose that the financial returns $\{X_t\}$ is a heavy tailed process; that is, $r_t = X_t - X_{t-h}$ satisfies

$$\Pr(r_t < -x) \sim x^{-\alpha} L(x)$$

as $x \rightarrow \infty$, where $\alpha > 0$ and $L(\cdot)$ is a slowly varying function. The Hill estimator of α is given by

$$\widehat{\alpha}_{\ell,n} = \left[\frac{1}{\ell} \sum_{i=1}^{\ell} \ln \left(\frac{r_{(i)}}{r_{(\ell)}} \right) \right]^{-1},$$

where n denotes the number of k -period returns and $r_{(1)} \leq r_{(2)} \leq \dots \leq r_{(n)}$ are the order statistics in ascending order. With this notation, Embrechts *et al.* (2005) show that the expected shortfall of k -period returns can be estimated by

$$\widehat{\text{ES}}_p = \frac{1}{p} \int_0^p \exp \left[\left(\frac{k\ell_{n,p}}{nq} \right)^{1/\widehat{\alpha}_{\ell_{n,p},n}} r_{\ell_{n,p}} \right] dq - 1,$$

where $\ell_{n,p} = [n(p + 0.045 + 0.005h)]$.

5.2 Necir *et al.*'s estimator

Let $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ denote the order statistics in ascending order corresponding to the financial returns X_1, X_2, \dots, X_n . Another semiparametric estimator for expected shortfall suggested by Necir *et al.* (2010) is

$$\widehat{\text{ES}}_p = \frac{1}{p} \int_{k/n}^p \widehat{F}^{-1}(t) dt + \frac{kX_{(n-k)}}{np(1-\widehat{\gamma})},$$

where

$$\widehat{\gamma} = \frac{1}{k} \sum_{i=1}^k \ln \frac{X_{(n-i+1)}}{X_{(n-k)}}$$

is Hill's estimator of tail index and $\widehat{F}^{-1}(\cdot)$ denotes the empirical quantile function.

6 Computer software

Software for computing expected shortfall and related quantities are widely available. Some software available from the R package (R Development Core Team, 2011) are:

- the package **ghyp** due to David Luethi, Wolfgang Breymann. According to the author, this package “provides detailed functionality for working with the univariate and multivariate Generalized Hyperbolic distribution and its special cases (Hyperbolic (hyp), Normal Inverse Gaussian (NIG), Variance Gamma (VG), skewed Student-t and Gaussian distribution). Especially, it contains fitting procedures, an AIC-based model selection routine, and functions for the computation of density, quantile, probability, random variates, expected shortfall and some portfolio optimization and plotting routines as well as the likelihood ratio test. In addition, it contains the Generalized Inverse Gaussian distribution”;
- the package **evir** due to Bernhard Pfaff, Alexander McNeil and Alec Stephenson;
- the package **fAssets** due to the Rmetrics Core Team;
- the package **crp.CSFP** due to Matthias Fischer, Kevin Jakob and Stefan Kolb;
- the package **QRM** due to Bernhard Pfaff, Alexander McNeil and Scott Ulmann.

Some other software available for computing value at risk and related quantities are:

- the package **ALM Optimizer** for asset allocation software due to Bob Korkie from the company RMKorkie & Associates, <http://assetallocationsoftware.org/>. According to the author, this package provides “risk and expected return of Markowitz efficient portfolios but extended to include recent technical advances on the definition of risk, adjustments for input bias, non normal distributions, and enhancements that allow for overlays, risk budgets, and investment horizon adjustments”. Also the package “is a true Portfolio Optimizer with lognormal asset returns and user specified: return or surplus optimization; optimization, risk, and rebalancing horizons; volatility, expected shortfall, and two value at risk (VaR) risk variables tailored to the risk horizon; and user specified portfolio constraints including risk budget constraints”;
- the package **QuantLib** due to StatPro, <http://www.statpro.com/portfolio-analytics-products/risk-management-software/>. According to the authors, this package provides “access to a complete universe of pricing functions for risk assessment covering every asset class from equity, interest rate-linked products to mortgage-backed securities”. The package has key features including “Multiple ex-ante risk measures including Value-at-Risk and CVaR (expected shortfall) at a variety of confidence levels, potential gain, volatility, tracking error and diversification grade. These measures are available in both absolute and relative basis”;
- the package **FinAnalytica’s Cognity** risk management due to FinAnalytica, <http://www.finanalytica.com/daily-risk-statistics/>. According to the authors, this package provide “more accurate fat-tailed VaR estimates that do not suffer from the over-optimism of normal distributions. But Cognity goes beyond VaR and also provides the downside Expected Tail Loss (ETL) measure - the average or expected loss beyond VaR. As compared with volatility and VaR, ETL, also known as Conditional Value at Risk (CVaR) and Expected Shortfall (ES), is a highly informative and intuitive measure of extreme downside losses. By combining ETL with fat-tailed distributions, risk managers have access to the most accurate estimate of downside risk available today”;
- the package **CVaR Expert** due to CVaR Expert Rho - Works Advanced Analytical Systems, <http://www.rhoworks.com/software/detail/cvarxpert.htm>. According to the authors, this package implements “total solution for measuring, analyzing and managing portfolio risk using historical VaR and CVaR methodologies. Traditional Value-at-Risk, Beta VaR, Component VaR, Conditional VaR and backtesting modules are incorporated on the current version, which lets you work with individual assets, portfolios, asset groups and multi currency investments (Enterprise Edition). An integrated optimizer can solve for the minimum CVaR portfolio based on market data and investor preferences, offering the best risk benchmark that can be produced. A module capable of doing Stochastic Simulation allows you to graph the CVaR-Return space for all feasible portfolios”;
- the **Enterprise Risk Management** software (KRM) due to ZSL Inc, <http://www.zsl.com/solutions/banking-finance/enterprise-risk-management-krm>. According to the authors, “Kamakura Risk Manager (KRM) completely integrates credit portfolio management, market risk management, asset and liability management, Basel II and other capital allocation technologies, transfer pricing, and performance measurement. KRM is also directly applicable to operational risk, total risk, and accounting and regulatory requirements using the same analytical engine, GUI and reporting, and its vision is that completely integrated risk solution based on common assumptions and methodologies. KRM offers, dynamic value at

risk and expected shortfall, historical value at risk measurement, Monte Carlo value at risk measurement, etc”;

- the package **NtInsight** due to Numerical Technologies Company, <http://www.numtech.com/news/basel-committee-proposes-expected-shortfall/#more-3396>. According to the authors, “Numerical Technologies understands the advantages of measuring expected shortfall. NtInsight uses massive parallel programming and applies faster codes when processing the transaction-level, 1 million Monte Carlo iterations needed to precisely capture the non-linear behavior of tail risk. It has been tested by major financial institutions in Japan where reporting expected shortfall is part of the regulatory requirement”;
- the package **G@RCH 6**, OxMetrics due to Timberlake Consultants Limited, <http://www.timberlake.co.uk/?id=6>. According to the authors, the package is “dedicated to the estimation and forecasting of univariate ARCH-type models. G@RCH provides a user-friendly interface (with rolling menus) as well as some graphical features (through the OxMetrics graphical interface). G@RCH helps the financial analysis: value-at-risk, expected shortfall, backtesting (Kupiec LRT, dynamic quantile test); forecasting, realized volatility”.

References

- [1] Acerbi, C. and Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking and Finance*, **26**, 1487-1503.
- [2] An, S., Sun, J. and Wang, Y. (2005). Research on risk management of basic social security fund in China. In: *Proceedings of the 2005 International Conference on Public Administration*, pp. 960-965.
- [3] Artzner, P., Delbaen, F., Eber, J. M. and Heath, D. (1997). Thinking coherently. *Risk*, **10**, 68-71.
- [4] Artzner, P., Delbaen, F., Eber, J. M. and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**, 203-228.
- [5] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, **12**, 171-178.
- [6] Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society, B*, **65**, 367-389.
- [7] Bae, T. and Iscoe, I. (2012). Large-sample confidence intervals for risk measures of location-scale families. *Journal of Statistical Planning and Inference*, **142**, 2032-2046.
- [8] Bernardi, M. (2012). Risk measures for skew normal mixtures. Working paper, MEMOTEF, Sapienza University of Rome.
- [9] Bi, G. and Giles, D. E. (2007). An application of extreme value analysis to US movie box office returns. In: *Proceedings of the 2007: International Congress on Modelling and Simulation: Land, Water and Environmental Management: Integrated Systems for Sustainability*, pp. 2652-2658.

- [10] Bi, G. and Giles, D. E. (2009). Modelling the financial risk associated with US movie box office earnings. *Mathematics and Computers In Simulation*, **79**, 2759-2766.
- [11] Brazauskas, V., Jones, B., Madan, L. and Zitikis, R. (2008). Estimating conditional tail expectation with actuarial application in view. *Journal of Statistical Planning and Inference*, **138**, 3590-3604.
- [12] Broda, S. A. and Paoletta, M. S. (2011). Expected shortfall for distributions in finance. In: *Statistical Tools for Finance and Insurance*, pp. 57-99, Heidelberg, Berlin.
- [13] Chen, Q., Gerlach, R. and Lu, Z. (2012). Bayesian Value-at-Risk and expected shortfall forecasting via the asymmetric Laplace distribution. *Computational Statistics and Data Analysis*, **56**, 3498-3516.
- [14] Chen, S. X. (2008). Nonparametric estimation of expected shortfall. *Journal of Financial Econometrics*, **6**, 87-107.
- [15] Cheung, K. C. (2009). Upper comonotonicity. *Insurance: Mathematics and Economics*, **45**, 35-40.
- [16] Cotter, J. and Dowd, K. (2006). Extreme spectral risk measures: An application to futures clearinghouse margin requirements. *Journal of Banking and Finance*, **30**, 3469-3485.
- [17] Embrechts, P., Kaufmann, R. and Patie, P. (2005). Strategic long-term financial risks: single risk factors. *Computational Optimization and Applications*, **32**, 61-90.
- [18] Fan, G., Wong, W. K. and Zeng, Y. (2008). Backtesting industrial market risks using expected shortfall: The case of Shanghai stock exchange. In: *Proceedings of the 2nd International Conference on Management Science and Engineering Management*, pp. 413-422.
- [19] Fan, G., Zeng, Y. and Wong, W. K. (2010). Risk contribution of different industries in China's stock market-based on VaR and expected shortfall. In: *Proceedings of the 3rd International Institute of Statistics and Management Engineering Symposium*, pp. 526-533.
- [20] Fan, G., Zeng, Y. and Wong, W. K. (2012). Decomposition of portfolio VaR and expected shortfall based on multivariate Copula simulation. *International Journal of Management Science and Engineering Management*, **7**, 153-160.
- [21] Fan, Y., Li, G. and Zhang, M. (2010). Expected shortfall and it's application in credit risk measurement. In: *Proceedings of the 2010 Third International Conference on Business Intelligence and Financial Engineering*, pp. 359-363.
- [22] Furman, E. and Landsman, Z. (2005). Risk capital decomposition for a multivariate dependent gamma portfolio. *Insurance: Mathematics and Economics*, **37**, 635-649.
- [23] Gao, L. J. and Li, J. P. (2009). The influence of IPO to the operational risk of Chinese commercial banks. *Cutting-Edge Research Topics on Multiple Criteria Decision Making Proceedings*, **35**, 486-492.
- [24] Hafner, R. (2004). *Stochastic Implied Volatility*. Springer Verlag, Berlin.
- [25] Harvey, C. R. and Siddique, A. (1999). Autoregressive conditional skewness. *Journal of Financial and Quantitative Analysis*, **34**, 465-487.

- [26] Hill, J. B. (2012). Robust expected shortfall estimation for infinite variance time series.
- [27] Hoogerheide, L. F. and van Dijk, H. K. (2008). Bayesian forecasting of value at risk and expected shortfall using adaptive importance sampling. *Tinbergen Institute Discussion Paper*, TI 2008-092/4.
- [28] Hürlimann, W. (2002). Analytical bounds for two value-at-risk functionals *Astin Bulletin*, **32**, 235-265.
- [29] Hürlimann, W. (2003). Conditional value-at-risk bounds for compound poisson risks and a normal approximation. *Journal of Applied Mathematics*, 141-153.
- [30] Inui, K. and Kijima, M. (2005). On the significance of expected shortfall as a coherent risk measure. *Journal of Banking and Finance*, **29**, 853-864.
- [31] Jadhav, D., Ramanathan, T. V. and Naik-Nimbalkar, U. V. (2009). Modified estimators of the expected shortfall. *Journal of Emerging Market Finance*, **8**, 87-107.
- [32] Johnson, N. L. (1949). System of frequency curves generated by methods of translation. *Biometrika*, **36**, 149-176.
- [33] Jones, M. C. and Faddy, M. J. (2003). A skew extension of the t distribution with applications. *Journal of the Royal Statistical Society, B*, **65**, 159-174.
- [34] Kamdem, J. S. (2005). Value-at-risk and expected shortfall for linear portfolios with elliptically distributed risk factors. *International Journal of Theoretical and Applied Finance*, **8**, 537-551.
- [35] Koenker, R. (2005). *Quantile Regression*. Cambridge University Press, Cambridge.
- [36] Koenker, R. and Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives*, **15**, 143-156.
- [37] Krehbiel, T. and Adkins, L. C. (2008). Extreme daily changes in US Dollar London inter-bank offer rates. *International Review of Economics and Finance*, **17**, 397-411.
- [38] Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1987). *Extremes and Related Properties of Random Sequences and Processes*. Springer Verlag, New York.
- [39] Lee, W. C. and Fang, C. J. (2010). The measurement of capital for operational risk in Taiwanese commercial banks. *Journal of Operational Risk*, **5**, 79-102.
- [40] Li, X. M. and Li, F. C. (2006). Extreme value theory: An empirical analysis of equity risk for Shanghai stock market. In: *Proceedings of the 2006 International Conference on Service Systems and Service Management*, pp. 1073-1077.
- [41] Lindstrom, E. and Regland, F. (2012). Modeling extreme dependence between European electricity markets. *Energy Economics*, **34**, 899-904.
- [42] Liu, X. X., Qiu, G. H. and Wu, F. X. (2008). Research of liquidity adjusted VaR and ES in China stock market based on extreme value theory. In: *Proceedings of 2008 China-Canada Industry Workshop on Enterprise Risk Management*, pp. 377-388.

- [43] Lu, Z., Huang, H. and Gerlach, R. (2010). Estimating Value at Risk: from JP Morgan's standard-EWMA to skewed-EWMA forecasting. University of Sydney Working Paper.
- [44] Magadia, J. (2011). Confidence interval for expected shortfall using bootstrap methods.
- [45] Mathai, A. M. and Moschopoulos, P. G. (1991). On a multivariate gamma. *Journal of Multivariate Analysis*, **39**, 135-153.
- [46] Mittnik, S. and Paolella, M. S. (2000). Conditional density and value-at-risk prediction of Asian currency exchange rates. *Journal of Forecasting*, **19**, 313-333.
- [47] Necir, A., Rassoul, A. and Zitikis, R. (2010). Estimating the conditional tail expectation in the case of heavy-tailed losses. *Journal of Probability and Statistics*, Article ID 596839.
- [48] Ou, S. and Yi, D. (2009). Robustness analysis and algorithm of expected shortfall based on extreme-value block minimum model. In: *Proceedings of the 2009 International Conference on Business Intelligence and Financial Engineering*, pp. 288-292.
- [49] Owen, D. B. (1956). Tables for computing bivariate normal probabilities. *Annals of Mathematical Statistics*, **27**, 1075-1090.
- [50] Pattarathammas, S., Mokkhavesa, S. and Nilla-Or, P. (2008). Value-at-risk and expected shortfall under extreme value theory framework: an empirical study on Asian markets.
- [51] Peracchi, F. and Tanase, A. V. (2008). On estimating the conditional expected shortfall. *Applied Stochastic Models in Business and Industry*, **24**, 471-493.
- [52] R Development Core Team (2011). *A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. Vienna, Austria.
- [53] Rachev, S., Jasic, T., Stoyanov, S. and Fabozzi, F. J. (2007). Momentum strategies based on reward-risk stock selection criteria. *Journal of Banking and Finance*, **31**, 2325-2346.
- [54] Richardson, L. F. (1911). The approximate arithmetical solution by finite differences of physical problems including differential equations, with an application to the stresses in a masonry dam. *Philosophical Transactions of the Royal Society of London*, A, **210**, 307357.
- [55] Richardson, L. F. and Gaunt, J. A. (1927). The deferred approach to the limit. *Philosophical Transactions of the Royal Society of London*, A, **226**, 299-349.
- [56] RiskMetrics Group (1996). *RiskMetrics-Technical Document*. Morgan J.P.
- [57] Scaillet, O. (2004). Nonparametric estimation and sensitivity analysis of expected shortfall. *Mathematical Finance*, **14**, 115-129.
- [58] Simonato, J. -G. (2011). The performance of Johnson distributions for computing value at risk and expected shortfall. *Journal of Derivatives*, **19**, 7-24.
- [59] Sims, G. V. and Kamal, I. (1996). Repowering of existing coastal stations to augment water supplies in Southern California: A study of alternatives. In: *Water Supply Puzzle: How does Desalting Fit In?*, pp. 408-429.
- [60] Song, J. S., Li, Y., Ji, F. and Peng, C. (2009). Operational risk measurement of Chinese commercial banks based on extreme value theory. *Cutting-Edge Research Topics on Multiple Criteria Decision Making Proceedings*, **35**, 531-534.

- [61] Taylor, J. W. (2008a). Estimating value at risk and expected shortfall using expectiles. *Journal of Financial Econometrics*, **6**, 231-252.
- [62] Taylor, J. W. (2008b). Using exponentially weighted quantile regression to estimate value at risk and expected shortfall. *Journal of Financial Econometrics*, **6**, 382-406.
- [63] Teng, F. and Zhang, Q. W. (2009). Cash flow risk measurement for Chinese non-life insurance industry based on estimation of ES and NRR. In: *Proceedings of the 2009 International Forum on Information Technology and Applications*, volume 2, pp. 628-631.
- [64] Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science*, **44**, 1650-1661.
- [65] Wang, Z. R. and Wu, W. T. (2008). Empirical study for exchange rate risk of CNY: Using VaR and ES based on extreme value theory. In: *Proceedings of the 2008 IEEE International Conference on Management of Innovation and Technology*, pp. 1193-1198.
- [66] Wang, Z. R., Wu, W. T., Chen, C. and Zhou, Y. J. (2010). The exchange rate risk of Chinese yuan: Using VaR and ES based on extreme value theory. *Journal of Applied Statistics*, **37**, 265-282.
- [67] Wong, C. -M. and So, M. K. P. (2010). On the estimation of multiple period expected shortfall and median shortfall using GARCH models. *Quantitative Finance*, to appear.
- [68] Wüthrich, M. V., Bühlmann, H. and Furrer, H. (2010). *Market-consistent Actuarial Valuation*, second edition. Springer, Heidelberg.
- [69] Yamai, Y. and Yoshida, T. (2002). Comparative analyses of expected shortfall and value-at-risk: Their estimation error, decomposition, and optimization. *Monetary and Economic Studies*, 87-122.
- [70] Yu, K., Ally, A. K., Yang, S. and Hand, D. J. (2010). Kernel quantile-based estimation of expected shortfall. *Journal of Risk*, **12**, 15-32.
- [71] Yu, Z. Y. and Tao, A. Y. (2008). Which one is better? - Risk measurement modeling on Chinese stock market. In: *Proceedings of the 2008 International Symposium on Financial Engineering and Risk Management*, pp. 47-51.
- [72] Zhu, D. and Galbraith, J. W. (2009). Forecasting expected shortfall with a generalized asymmetric Student-t distribution. CIRANO Working Paper.
- [73] Zhu, D. and Galbraith, J. W. (2011). Modelling and forecasting expected shortfall with the generalized asymmetric Student-t and asymmetric exponential power distributions. *Journal of Empirical Finance*, **18**, 765-778.