Addressing the potential non-robustness of sub-additive portfolio risk measures

MS18: Applied Mathematics in Industry II

John A. Dodson

University of Minnesota Center for Financial and Actuarial Mathematics / Options Clearing Corporation

SIAM AN14, Chicago, 2014-Jul-7
Introduction

The modern approach to managing financial market risk is to define an acceptance set for portfolio net asset value at some fixed time horizon according to some probability measure and to impose limits in the current period on the portfolio composition in order to ensure acceptable outcomes. Estimating these limits present practical and theoretical challenges, particularly where outcomes are highly leptokurtotic. We present a concise description of the problem and how it can be addressed.

Acknowledgments

The author wishes to acknowledge OCC colleagues Simon Babbs and Ashish Batra who contributed to this work.
Risk Measures

The general theoretical approach to financial market risk measurement (cf. Meucci) is to:

1. represent some accounting performance metric, such as mark-to-market profit on fixed holdings, $\alpha$, over the next five business days, by a random variable, say $\Psi_\alpha$;
2. design a statistical model to characterize and estimate the salient properties of the measure $\mu_{\Psi_\alpha}$;
3. choose an “index” $S(\alpha)$ whose value depends on $\mu_{\Psi_\alpha}$ (but not on $\Psi_\alpha$ itself), representing our potential (dis)satisfaction with the realized outcome of the metric.
Extreme Value Theory

An approach to step two is informed by the Pickands-Balkema-de Haan theorem, from “extreme value theory”: We can presume that $\Psi$ is (partially) defined in terms of a Generalized Pareto random variable:

$$\mu(\{\psi : \psi < \psi\}) = \begin{cases} \theta \left(1 + \xi \frac{\eta - \psi}{\beta}\right)^{-1/\xi} & \psi \leq \eta \\ ? & \text{otherwise} \end{cases}$$

for sufficiently small (fixed) mass parameter $0 < \theta < 1$, where the location parameter $\eta$ is the $\theta$-quantile of $\Psi$, $\beta > 0$ is a scale parameter, and $\xi < 1$ is the (left) “tail parameter”.

- Given a means of sampling from an approximation of $\Psi$, parameter values can be determined by the maximum likelihood estimator.
Coherent Risk Measures

In approaching step three, it has been productive to choose an index above based on adherence to a list of desired features:

- **estimable** \( S \) is non-random
- **money-equivalent** \( S \) and \( \Psi \) are in the same units
- **constant** \( \mu (\{\Psi_b = \psi_b\}) = 1 \Rightarrow S(b) = \psi_b \)
- **translation invariant** \( S(\alpha + b) = S(\alpha) + \psi_b \)
- **risk averse** \( E \Psi_f = 0 \Rightarrow S(f + b) \leq S(b) \)
- **positive homogeneous** \( \lambda \geq 0 \Rightarrow S(\lambda \alpha) = \lambda S(\alpha) \)
- **super-additive** \( S(\alpha + \beta) \geq S(\alpha) + S(\beta) \)

Indexes which satisfy these properties are termed “coherent”. Artzner et al. observe that all such indexes can be defined in terms of conditional expected values of \( \Psi \). One popular example is \( ES_c = E [\Psi | \Psi < Q(1 - c)] \) where \( \mu (\{\Psi : \Psi < Q(p)\}) = p \).
Estimating Expected Shortfall

**EVT Estimate (Embrechts–Klüppelberg-Mikosh)**

If we assume a Pareto tail, the expected shortfall at a confidence level $1 - \theta \leq c < 1$ in terms of parameter estimates for $\mu_\Psi$ is readily computable:

$$ES_c^{EVT} \triangleq \hat{\eta} - \frac{\hat{\beta}}{\hat{\xi}} \left( \frac{1}{1 - \hat{\xi}} \left( \frac{\theta}{1 - c} \right)^{\hat{\xi}} - 1 \right)$$

**Nonparametric Estimate (Chen)**

Alternatively, given an ordered sample of size $N$ of draws from $\Psi$, one can form the empirical estimator:

$$ES_c^{emp} \triangleq \frac{1}{\left\lfloor N(1-c) \right\rfloor} \sum_{i=1}^{\left\lfloor N(1-c) \right\rfloor} \Psi(i)$$
Main Result

Unfortunately, as Cont et al. point out, estimators for expected shortfall are generally not robust. For example, the standard error of Chen’s estimator involves \( \text{var} \left[ (\Psi - Q(1-c)) \mathbb{I}_{\Psi < Q(1-c)} \right] \) but:

Lemma (Babbs, 2012)

Suppose the density of the random variable \( X \) takes on the Generalized Pareto form

\[
f(x) = \frac{\theta}{\beta} \left( 1 + \frac{\xi \eta - x}{\beta} \right)^{-\frac{1}{\xi} - 1} \quad \text{for} \quad x \leq \eta
\]

for constants \( 0 < \theta < 1, \beta > 0, 0 < \xi < 1, \text{ and } \eta \geq \nu \), then \( \text{var} \left[ (X - \nu) \mathbb{I}_{X < \nu} \right] < \infty \) iff \( \xi < \frac{1}{2} \).
Main Result

Proof.
Under the assumptions of the lemma,

\[ E \left[ (X - \nu)^2 \mathbb{I}_{X<\nu} \right] = \int_{-\infty}^{\nu} \frac{\theta(x-\nu)^2}{\beta} \left( 1 + \frac{\xi (\eta - x)}{\beta} \right)^{-\frac{1}{\xi} - 1} \, dx \]

\[ = \int_{1+\xi \frac{\eta-\nu}{\beta}}^{\infty} \frac{\theta}{\xi} \left( \frac{\beta}{\xi} (1 - w) + \eta - \nu \right)^2 w^{-\frac{1}{\xi} - 1} \, dw \]

which is finite if and only if the leading power of \( w \) in the integrand is strictly less than minus unity. The result follows. □

- Many portfolios in practice have tail parameters near zero (normality), but options and credit portfolios in particular are prone to larger values.
Compensation for Estimation Risk

Since the standard error of an estimator for expected shortfall may be large, for risk management it is advisable to introduce a bias to increase the safely factor of the estimate;

▶ but for the same reason, it is inadvisable to base this on an estimate for the standard error!

As an alternative, we propose the purposely naïve:

\[
SE_{c}^{\text{comp}} \triangleq SE_{c}^{\text{EVT}} - k_c
\]

\[
\times \sqrt{\frac{1}{[N(1-c)]^2} \sum_{i=1}^{[N(1-c)]} \psi^2(i) - \frac{1}{[N(1-c)]^3} \left( \sum_{i=1}^{[N(1-c)]} \psi(i) \right)^2}
\]

inspired by the estimator for the standard error of the mean.
Testing Results

Sufficiency rates from 2,000 trials for the estimated and compensated 99% expected shortfall based on 10,000 draws:

<table>
<thead>
<tr>
<th>$k_{0.99}$</th>
<th>$\xi = 0.1$</th>
<th>$\xi = 0.2$</th>
<th>$\xi = 0.3$</th>
<th>$\xi = 0.4$</th>
<th>$\xi = 0.5$</th>
<th>$\xi = 0.6$</th>
<th>$\xi = 0.7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>51%</td>
<td>46%</td>
<td>49%</td>
<td>49%</td>
<td>49%</td>
<td>49%</td>
<td>43%</td>
</tr>
<tr>
<td>0.2</td>
<td>56%</td>
<td>53%</td>
<td>53%</td>
<td>55%</td>
<td>55%</td>
<td>54%</td>
<td>50%</td>
</tr>
<tr>
<td>0.4</td>
<td>61%</td>
<td>59%</td>
<td>59%</td>
<td>61%</td>
<td>60%</td>
<td>59%</td>
<td>54%</td>
</tr>
<tr>
<td>0.6</td>
<td>66%</td>
<td>63%</td>
<td>65%</td>
<td>65%</td>
<td>65%</td>
<td>63%</td>
<td>58%</td>
</tr>
<tr>
<td>0.8</td>
<td>71%</td>
<td>69%</td>
<td>69%</td>
<td>70%</td>
<td>69%</td>
<td>67%</td>
<td>61%</td>
</tr>
<tr>
<td>1.0</td>
<td>76%</td>
<td>74%</td>
<td>74%</td>
<td>74%</td>
<td>73%</td>
<td>70%</td>
<td>64%</td>
</tr>
<tr>
<td>1.2</td>
<td>80%</td>
<td>78%</td>
<td>77%</td>
<td>77%</td>
<td>77%</td>
<td>73%</td>
<td>67%</td>
</tr>
<tr>
<td>1.4</td>
<td>84%</td>
<td>81%</td>
<td>81%</td>
<td>80%</td>
<td>80%</td>
<td>76%</td>
<td>70%</td>
</tr>
<tr>
<td>1.6</td>
<td>86%</td>
<td>84%</td>
<td>84%</td>
<td>83%</td>
<td>82%</td>
<td>79%</td>
<td>72%</td>
</tr>
<tr>
<td>1.8</td>
<td>88%</td>
<td>87%</td>
<td>86%</td>
<td>85%</td>
<td>85%</td>
<td>81%</td>
<td>74%</td>
</tr>
<tr>
<td>2.0</td>
<td>90%</td>
<td>89%</td>
<td>88%</td>
<td>87%</td>
<td>86%</td>
<td>83%</td>
<td>76%</td>
</tr>
</tbody>
</table>

- The uncompensated ($k_{0.99} = 0$) estimate is approximately equally likely to be too large or too small.
- Adding two ‘classical’ standard errors ($k_{0.99} = 2$) increases the probability that the coverage is sufficient to 80% – 90%. 
References

P. Artzner, F. Delbaen, J. M. Eber, and D. Heath.
Coherent measures of risk.

A. A. Balkema and L. de Haan.
Residual life time at great age.

S. X. Chen.
Nonparametric estimation of expected shortfall.

R. Cont, R. Deguest, and G. Scandolo.
Robustness and sensitivity analysis of risk measurement procedures.

*Modelling extremal events for insurance and finance*.

A. Meucci.
*Risk and asset allocation*.

J. Pickands.
Statistical inference using extreme order statistics.