If Worse Comes to Worst: Systematic Stress Tests with Discrete and Other Non-normal Distributions

Thomas Breuer   Imre Csiszár

PPE Research Centre, FH Vorarlberg, Austria
Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Hungary

Quant Congress Europe, London, 3 November 2009

Outline

Purpose of Stress Testing: Complement statistical risk measurement

- Statistical risk measurements: What are probs of big losses?
  Stress Testing: Which scenarios lead to big losses?
  Derive risk reducing action.
- Statistical risk measurement: Assume fixed model.
  Stress Testing: Consider alternative risk factor distribution.
  Address model risk.

Requirements on stress scenarios (Basel II)

- Plausibility
- Severity
- Suggestive of risk reducing action

See Basel Principles of Sound Stress Testing

Framework

Reference risk factor distribution ν,
Portfolio loss function L,
both on risk factor space $\Omega \subset \mathbb{R}^n$.

First Generation Stress Tests: Hand-picked Point Scenarios

- A small number of scenarios is picked by hand, ideally involving heterogeneous groups of experts.
  $A = \{\mathbf{r}^1, \mathbf{r}^2, \dots, \} \subset \Omega$: a small set of hand-picked scenarios.
- Find worst case scenario and worst case loss in A:

$$\max_{\mathbf{r} \in A} L(\mathbf{r})$$

- Worst case loss over A is a coherent risk measure.
- Examples: SPAN rules.

Pitfalls of First Generation Stress Tests

1. Neglecting severe but plausible scenarios
2. Considering too implausible scenarios

Outline

Second Generation Stress Tests: Systematic Point Scenario Analysis

- **Measure of plausibility** for point scenarios:

$$\mathrm{Maha}(\mathbf{r}) := \sqrt{(\mathbf{r} - \mathbb{E}(\mathbf{r}))^T \cdot \Sigma^{-1} \cdot (\mathbf{r} - \mathbb{E}(\mathbf{r}))},$$

where Σ is covariance matrix of risk factor distribution ν.

- Choose as scenario set A

$$\mathrm{Ell}_h := \{\mathbf{r} : \mathrm{Maha}(\mathbf{r}) \leq h\},$$

where $h \in [0, \infty]$ is the plausibility threshold.
"Size of multivariate risk factor jump, measured in standard deviations."

Advantages of Systematic Stress Testing with Point Scenarios

All three requirements on stress testing are met:

- Do not miss plausible but severe scenarios.
- Do not consider scenarios which are too implausible.
- Worst case scenario over Ell_h gives information about portfolio structure and suggests risk reducing action.

Problems of Systematic Stress Testing with Point Scenarios

1. optimisation problem sometimes hard to solve.
2. Maha does not take into account fatness of tails.
3. How choose scenario set for non-elliptical risk factor distributions ν?
4. $\mathrm{MaxLoss}_{\mathrm{Ell}_k}$ depends on choice of coordinates.
5. $\mathrm{MaxLoss}_{\mathrm{Ell}_k}$ is not law-invariant: Portfolios L_1, L_2 might have the same profit/loss distribution but different $\mathrm{MaxLoss}_{\mathrm{Ell}_k}$.
6. Model risk is not addressed.

Outline

Generalised Scenarios

Generalised scenario: Probability distribution on $(\Omega, \mathbb{F})$.

- Interpretation 1:
  Generalisation of point scenarios,
  which have support concentrated on one point.
- Interpretation 2:
  Risk factor distributions alternative to the prior ν.
  Model risk.

Plausibility of Generalised Scenarios

- **Measure of plausibility:** relative entropy
  (I-divergence, information gain, Kullback-Leibler distance)

$$I(Q||\nu) := \begin{cases} \int \log \frac{dQ}{d\nu}(\mathbf{r}) dQ(\mathbf{r}) & \text{if } Q \ll \nu \\ +\infty & \text{if } Q \not\ll \nu \end{cases}$$

- **Scenario set:** Instead of ellipsoid take
  Kullback-Leibler sphere in the space of distributions

$$S(\nu, k) := \{Q : I(Q||\nu) \leq k\}.$$

Worst Case Scenario

- **Severity of scenarios:** Instead of $L(\mathbf{r})$ take $\mathbb{E}_Q(L)$
- Generalised MaxLoss:

$$\sup_{Q \in S(\nu,k)} \mathbb{E}_Q(L) =: \mathrm{MaxLoss}_k(L)$$

If it exists, call scenario achieving MaxLoss: $\overline{Q}$.

Advantages of Systematic Stress Testing with Generalised Scenarios

1. Relative entropy does take into account fatness of tails of ν.
2. Scenario set is naturally defined for non-elliptical risk factor distributions ν.
3. $\mathrm{MaxLoss}_k$ does not depend on choice of coordinates.
4. $\mathrm{MaxLoss}_k$ is law-invariant: Portfolios L_1, L_2 with the same profit/loss distribution have the same $\mathrm{MaxLoss}_k$.
5. Model risk is addressed:
   Generalised scenarios are alternatives to prior risk factor distribution ν.

Outline

The Basic Tool

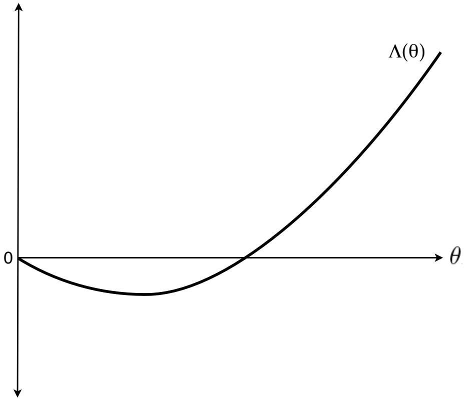
- Tool from large deviations theory for solving explicitly the optimisation problem $\sup_{Q \in S(\nu,k)} \mathbb{E}_Q(L)$:

$$\Lambda(\theta) := \log \left(\int e^{\theta L(\mathbf{r})} d\nu(\mathbf{r}) \right).$$

- $\theta_{\max} := \sup\{\theta : \Lambda(\theta) < +\infty\}$
- $\Lambda(\theta)$ is the analogue of the log-partition function $Z(\beta)$ from statistical mechanics.

Basic Properties of the Λ-function

- $\Lambda(0) = 0$.
- Λ is convex.



Solution of Worst Case: The Generic Case

Theorem

- *Except in the pathological cases (i), (ii), (iii) below, the equation*

$$\theta\Lambda'(\theta) - \Lambda(\theta) = k, \tag{1}$$

  has always a unique positive solution $\overline{\theta}$.

- *The generalised worst case scenario* $\overline{Q}$ *is the distribution with* ν*-density*

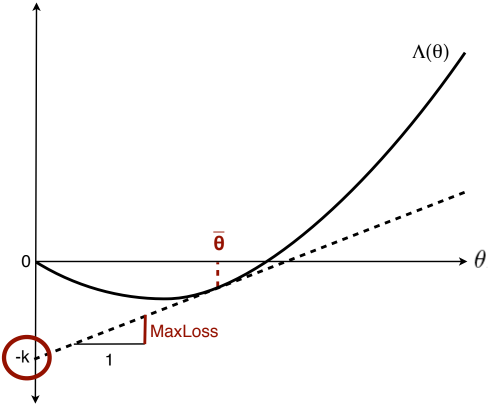
$$\frac{d\overline{Q}}{d\nu}(\mathbf{r}) := \frac{e^{\overline{\theta}L(\mathbf{r})}}{\int e^{\overline{\theta}L(\mathbf{r})}\,d\nu(\mathbf{r})} = e^{\overline{\theta}L(\mathbf{r})-\Lambda(\overline{\theta})}, \tag{2}$$

- *The generalised Maximum Loss achieved in the generalised worst case scenario* $\overline{Q}$ *is*

$$\mathbb{E}_{\overline{Q}}(L) = \Lambda'(\overline{\theta}).$$

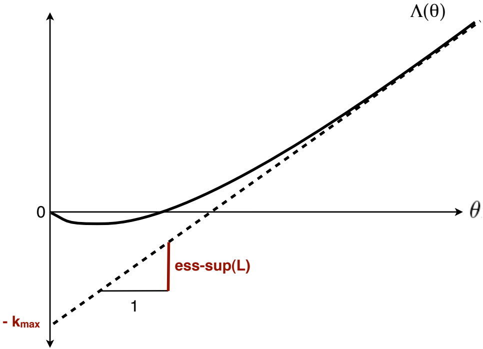
Practical Calculation of Worst Case

1. Calculate $\Lambda(\theta)$. (Evaluate n-dimensional integral.)
2. Starting from the point $(0, -k)$, lay a tangent to $\Lambda(\theta)$ curve.
3. Worst case loss is given by the slope of the tangent.
4. $\overline{\theta}$ is θ-coordinate of tangent point.
5. Worst case scenario is distribution with density $\frac{d\overline{Q}}{d\nu}(\mathbf{r}) = e^{\overline{\theta} L(\mathbf{r}) - \Lambda(\overline{\theta})}$.



Worst Case: The Pathological Cases

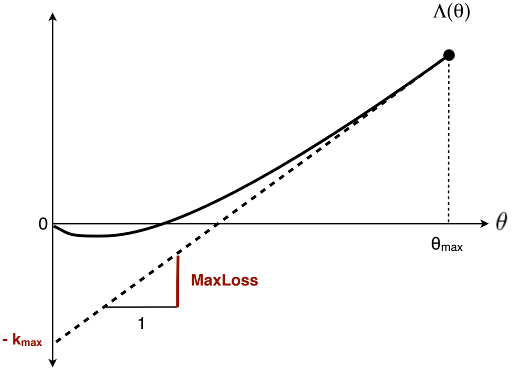
(i) If $\operatorname{ess\,sup}(L)$ is finite, and $k \geq -\log(\nu(\{\mathbf{r} : L(\mathbf{r}) = \operatorname{ess\,sup}(L)\}))$, then $\mathrm{MaxLoss}(k) = \operatorname{ess\,sup}(L)$.



(ii) If $\theta_{\max} = 0$ then $\mathrm{MaxLoss}(k) = \infty$ for all $k > 0$.

Worst Case: Pathological Cases cont'd

(iii) If $\theta_{\max}, \Lambda(\theta_{\max})$, and $\Lambda'(\theta_{\max})$ are all finite and $k > k_{\max} := \theta_{\max}\Lambda'(\theta_{\max}) - \Lambda(\theta_{\max})$, then $\mathrm{MaxLoss}(k) = (k + \Lambda(\theta_{\max}))/\theta_{\max}$.

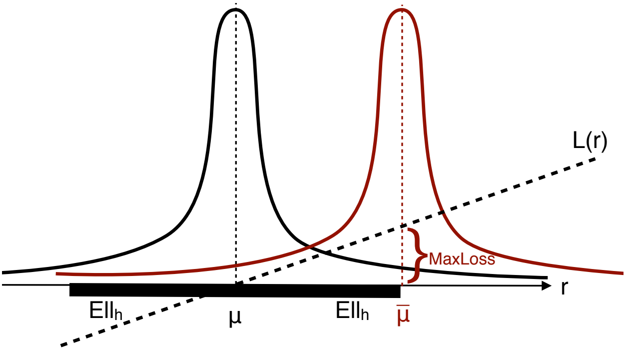


Outline

One normal risk factor, linear portfolio

Loss linear function of 1 risk factor: $L(r) = l(\mu - r)$,
$\nu \sim N(\mu, \sigma^2)$.

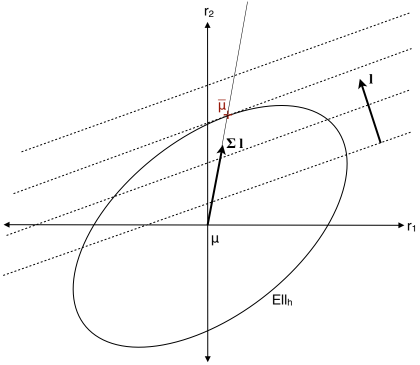
Worst case scenario: $\overline{Q} \sim N(\overline{\mu}, \sigma^2)$ where $\overline{\mu} = \mu + h\sigma \mathrm{sgn}(l)$ and $h = \sqrt{2k}$.
Worst case loss: $\mathbb{E}_{\overline{Q}}(L) = h\sigma|l| = h\sqrt{l\sigma^2 l}$.



More normal risk factors, linear portfolio

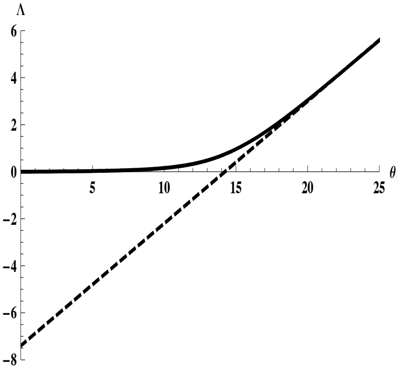
Loss linear function of n risk factors: $L(r) = \mathbf{l}^T(\mu - r)$,
$\nu \sim N(\boldsymbol{\mu}, \Sigma)$.

Λ quadratic: $\Lambda(\theta) = \mathbf{l}^T \Sigma \mathbf{l} \theta^2/2$.
Worst case scenario: $\overline{Q} \sim N(\overline{\boldsymbol{\mu}}, \Sigma)$ with $\overline{\boldsymbol{\mu}} = \boldsymbol{\mu} - \frac{h}{\sqrt{\mathbf{l}^T \Sigma \mathbf{l}}} \Sigma \mathbf{l}$ where $h = \sqrt{2k}$.
Worst case loss: $\mathbb{E}_{\overline{Q}}(L) = h\sqrt{\mathbf{l}^T \Sigma \mathbf{l}}$.



Stressed transition probabilities

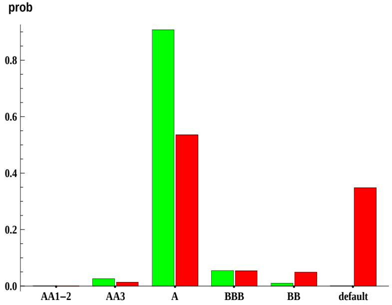
- $\Omega = \{0, 1, \ldots n\}$: rating classes.
- $\mathbf{p} = (p_1, \ldots, p_n)$: estimated transition probabilities
- $\mathbf{l} = (l_1, \ldots, l_n)$: loss caused by transitions
- $\Lambda(\theta) = \log\left(\sum_{j=1}^{n} p_j \exp(\theta l_j)\right)$.
- Worst case transition probabilities: $\overline{p}_i = \frac{p_i \exp(\overline{\theta} l_i)}{\sum_{j=1}^{n} p_j \exp(\overline{\theta} l_j)}$.



Numerical example: A-rated bond

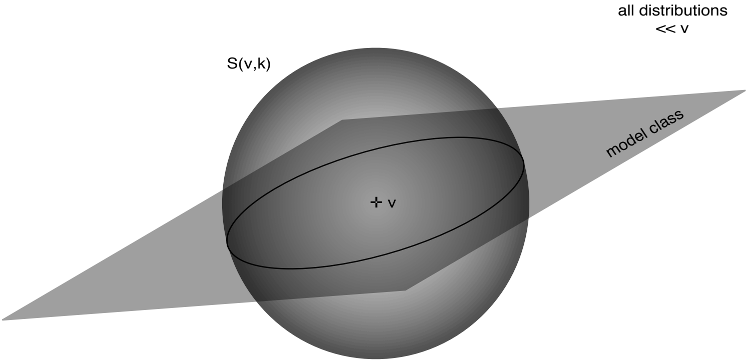
	AA1-2	AA3	A	BBB	BB	Default
loss from transitions [%]	-3.20%	-1.07%	0.00%	3.75%	15.83%	51.80%
est'd trans. prob. [%]	0.09	2.60	90.75	5.50	1.00	0.06
worst c. trans. prob. [%]	0.036	1.34	53.53	5.37	4.91	34.8

Expected loss from transitions under est'd probs: 0.37%
Expected loss from transitions under worst case probs at k=2: 19.07%



Model Risk

- estimation error: wrong distribution parameters
- model misspecification: wrong model class
- $\sup_{Q \in S(\nu,k)} \mathbb{E}_Q(L)$ quantifies effects of both on expected loss.



Summary

1. Calculate $\Lambda(\theta) := \log\left(\int e^{\theta L(\mathbf{r})} d\nu(\mathbf{r})\right)$.
2. Starting from the point $(0, -k)$, lay a tangent to $\Lambda(\theta)$ curve.
3. Worst case loss is given by the slope of the tangent.
4. $\overline{\theta}$ is θ-coordinate of tangent point.
5. Worst case scenario is distribution with density $\frac{d\overline{Q}}{d\nu}(\mathbf{r}) = e^{\overline{\theta} L(\mathbf{r}) - \Lambda(\overline{\theta})}$.