Parameter Uncertainty and Solvency Risk

Andreas Tsanakas
Cass Business School, City University London

based on joint work with R. Gerrard, V. Bignozzi, Z. Landsman

Actuarial & Financial Mathematics 2012: Theory & Applications
University of Liverpool, 7 June 2012
Motivation

• Regulation
  – Strict requirements (eg 99.5%-VaR)
  – Limited data...

• What is this talk about?
  – An idiosyncratic tour of questions that have troubled me
  – Not a comprehensive review
  – Omitting much mathematical detail
Setup (1)

• I.i.d. data $X_1, \ldots, X_n \sim F(\cdot; \theta)$

• Future loss $Y \sim F(\cdot; \theta)$, independent of $X = \{X_1, \ldots, X_n\}$

• Parameter(s) $\theta$ unknown, to be estimated by MLE $\hat{\theta} = \hat{\theta}(X)$

• MLE of distribution: $F(\cdot; \hat{\theta})$

• Predictive distribution: $\hat{F}(\cdot | X) = \int F(\cdot; t) \pi(t | X) \, dt$
Setup (2)

- Regulatory risk measure $\rho$

  - Monotonic: $Y \leq Z \implies \rho(Y) \leq \rho(Z)$

  - Translation invariant: $\rho(Y - \lambda) = \rho(Y) - \lambda$

  - Positive homogenous: For $\lambda > 0$, $\rho(\lambda Y) = \lambda \rho(Y)$

  - Law invariant: If $Y \overset{d}{=} Z$, then $\rho(Y) = \rho(Z)$.

    * So for $Y \sim F(\cdot; \theta)$, let $\rho_\theta(Y) = \rho[F(\cdot; \theta)]$

- For example

  \[
  VaR_{p, \theta}(Y) = VaR_p[F(\cdot; \theta)] = \inf \{ y \in \mathbb{R} : F(y; \theta) \geq p \}
  \]
  \[
  TVaR_{p, \theta}(Y) = TVaR_p[F(\cdot; \theta)] = \frac{1}{1 - p} \int_p^1 VaR_{\alpha, \theta}(Y) d\alpha
  \]
Parameter uncertainty

- $\rho_\theta(Y)$ is the minimum level of capital such that
  \[ \rho_\theta(Y - \rho_\theta(Y)) = 0 \]

- In reality capital held is $\rho_{\hat{\theta}}(Y) = \rho[F(\cdot; \hat{\theta})]$, such that
  \[ \rho_\theta(Y - \rho_{\hat{\theta}}(Y)) \neq 0 \]

- Randomness of capital reflects variation across time periods or firms

- When $\rho_\theta(Y - \rho_{\hat{\theta}}(Y)) \geq 0$ a positive residual risk is incurred (Bignozzi & T., 2012).
Residual Risk with $\rho \equiv \text{VaR}_p$

- For simplicity, let
  \[ \text{VaR}_{p, \theta}(Y) = F^{-1}(p; \theta) \]

- Then positive residual risk implies an increased probability of default
  \[ \text{VaR}_{p, \theta}(Y - F^{-1}(p; \hat{\theta})) \geq 0 \Leftrightarrow P_{\theta}(Y > F^{-1}(p; \hat{\theta})) \geq 1 - p \]

- VaR estimate is a *predictive limit* (Barndorff-Nielsen and Cox, 1996)

- Exponential distribution:
  \[ F(y; \theta) = 1 - \exp(-y/\theta), \quad y > 0 \]
  \[ P(Y > F^{-1}(p; \hat{\theta})) = \left(1 - \frac{1}{n} \log(1 - p)\right)^{-n} \]
Exponential example
Residual Risk with $\rho \equiv VaR_p$

- Let $Y$ be such that a known increasing transformation belongs to a location, scale, or location-scale family.
  - Normal, Log-Normal, Weibull, Log-Logistic, Exponential, Pareto...

- Then $P_\theta \left( Y > F^{-1}(p; \hat{\theta}) \right)$ does not depend on $\theta$
  - Can find $p_n^* > p$ such that $P_\theta \left( Y > F^{-1}(p_n^*; \hat{\theta}) \right) = 1 - p$

- Let $\theta = (\mu, \sigma)$ and consider probability matching prior $\pi(\mu, \sigma) = 1/\sigma$
  (Severini et al (2002); Gerrard & T. (2011))
  - Then $P_\theta(Y > \hat{F}^{-1}(p|X)) = 1 - p$
Exponential example: $\hat{\theta} = 100$, $n = 10$
Exponential example: \( \hat{\theta} = 100, \ n = 10 \)
Exponential example: $\hat{\theta} = 100$, $n = 10$
Exponential example: $\hat{\theta} = 100, \ n = 10$

$F^{-1}(p^*, \hat{\theta}) = \hat{F}^{-1}(p \mid X)$
Beyond VaR

- Let $\rho$ be some other risk measure eg $\rho \equiv TVaR_p$

- For location-scale families we find the following
  
  - One can construct another risk measure $\tilde{\rho}_n$ such that
    \[
    \rho_\theta \left( Y - \tilde{\rho}_{n,\hat{\theta}}(Y) \right) = 0
    \]
  
  - Bayesian approach
    \[
    \rho_\theta \left( Y - \rho[\tilde{F}(\cdot|X)] \right) \approx 0
    \]
  
  - Parametric bootstrap (1st order)
    \[
    \rho_\theta \left( Y - \rho\hat{\theta}(Y) - \rho\hat{\theta}(Y - \rho\hat{\theta}^*(Y)) \right) \approx 0
    \]
    * Can improve by repeating the bootstrap without increasing computational time.
Normalised residual risk for exponential / $\rho \equiv TV\alpha R$

<table>
<thead>
<tr>
<th></th>
<th>n=10</th>
<th>n=20</th>
<th>n=50</th>
<th>n=100</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MLE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p=0.95</td>
<td>0.2115</td>
<td>0.1179</td>
<td>0.0504</td>
<td>0.0261</td>
</tr>
<tr>
<td>p=0.99</td>
<td>0.2500</td>
<td>0.1442</td>
<td>0.0633</td>
<td>0.0333</td>
</tr>
<tr>
<td>p=0.995</td>
<td>0.2656</td>
<td>0.1552</td>
<td>0.0693</td>
<td>0.0355</td>
</tr>
<tr>
<td><strong>Bayesian</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p=0.95</td>
<td>-0.0181</td>
<td>-0.0088</td>
<td>-0.0024</td>
<td>-0.0015</td>
</tr>
<tr>
<td>p=0.99</td>
<td>-0.0126</td>
<td>-0.0051</td>
<td>-0.0034</td>
<td>-0.0010</td>
</tr>
<tr>
<td>p=0.995</td>
<td>-0.0108</td>
<td>-0.0053</td>
<td>-0.0031</td>
<td>0.0003</td>
</tr>
<tr>
<td><strong>Bootstrap</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p=0.95</td>
<td>0.0634</td>
<td>0.0202</td>
<td>0.0037</td>
<td>0.0010</td>
</tr>
<tr>
<td>p=0.99</td>
<td>0.0960</td>
<td>0.0332</td>
<td>0.0066</td>
<td>0.0017</td>
</tr>
<tr>
<td>p=0.995</td>
<td>0.1094</td>
<td>0.0389</td>
<td>0.0080</td>
<td>0.0021</td>
</tr>
</tbody>
</table>
Normalised residual risk for exponential / $\rho \equiv TV_\alpha R$

<table>
<thead>
<tr>
<th></th>
<th>MLE</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho = 0.95$</td>
<td>n=10</td>
<td>n=20</td>
<td>n=50</td>
<td>n=100</td>
<td></td>
</tr>
<tr>
<td>$p=0.95$</td>
<td>0.2115</td>
<td>0.1179</td>
<td>0.0504</td>
<td>0.0261</td>
<td></td>
</tr>
<tr>
<td>$p=0.99$</td>
<td>0.2500</td>
<td>0.1442</td>
<td>0.0633</td>
<td>0.0333</td>
<td></td>
</tr>
<tr>
<td>$p=0.995$</td>
<td>0.2656</td>
<td>0.1552</td>
<td>0.0693</td>
<td>0.0355</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Bayesian</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho = 0.95$</td>
<td>n=10</td>
<td>n=20</td>
<td>n=50</td>
<td>n=100</td>
<td></td>
</tr>
<tr>
<td>$p=0.95$</td>
<td>-0.0181</td>
<td>-0.0088</td>
<td>-0.0024</td>
<td>-0.0015</td>
<td></td>
</tr>
<tr>
<td>$p=0.99$</td>
<td>-0.0126</td>
<td>-0.0051</td>
<td>-0.0034</td>
<td>-0.0010</td>
<td></td>
</tr>
<tr>
<td>$p=0.995$</td>
<td>-0.0108</td>
<td>-0.0053</td>
<td>-0.0031</td>
<td>0.0003</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Bootstrap2</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho = 0.95$</td>
<td>n=10</td>
<td>n=20</td>
<td>n=50</td>
<td>n=100</td>
<td></td>
</tr>
<tr>
<td>$p=0.95$</td>
<td>0.0298</td>
<td>0.0053</td>
<td>0.0004</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>$p=0.99$</td>
<td>0.0679</td>
<td>0.0137</td>
<td>0.0012</td>
<td>0.0002</td>
<td></td>
</tr>
<tr>
<td>$p=0.995$</td>
<td>0.0900</td>
<td>0.0193</td>
<td>0.0017</td>
<td>0.0002</td>
<td></td>
</tr>
</tbody>
</table>
Something about time consistency

• Let $\theta_0$ be initial guess of parameter

$$\rho_{\theta_0} \left( Y - \tilde{\rho}_{n,\hat{\theta}}(Y) \right) = 0$$

• Implication:

$$\tilde{\rho}_{n,\hat{\theta}}(Y) \leq 0 \text{ a.s. } \implies \rho_{\theta_0}(Y) \leq 0$$
$$\tilde{\rho}_{n,\hat{\theta}}(Y) \geq 0 \text{ a.s. } \implies \rho_{\theta_0}(Y) \geq 0$$

• *Sequentially consistent* risk measurement (Roorda and Schumacher (2007); Bignozzi and T. (2012))
Something about heavy tails

- If $\rho$ is *coherent*, then finite means are required

- Let $Y \sim LN(\mu, \sigma^2)$
  
  - Predictive distribution of $Y$ is Log-t $\rightarrow$ infinite mean!

- Let $Y \sim Pareto(\theta)$ such that $P_\theta(Y > y) = y^{-\theta}$, $y \geq 1$, $\theta > 1$.
  
  - $P_\theta(\hat{\theta} < 1) > 0 \implies P_\theta(\rho_{\hat{\theta}}(Y) < \infty) < 1$

- Blame the statistics or the risk measure?

- Could follow Cont et al (2010) and use

  \[
  \rho_\theta(Y) = \frac{1}{p_2 - p_1} \int_{p_1}^{p_2} VaR_{\alpha, \theta}(Y) d\alpha
  \]
Something about pricing

- Theoretical price of layer \( l \) in excess of \( d \)

\[
\pi[F(\cdot; \theta)] = \int_d^{d+l} (1 - F(y; \theta)) \, d\theta
\]

- What should we use as premium?
  - MLE \( \pi[F(\cdot; \hat{\theta})] \) v predictive distribution \( \pi[\hat{F}(\cdot|X)] \)

- Let portfolio consist of many independent policies, with premiums calculated from different data
  - Consider \( \sum_j (Y_j - \pi[\hat{F}(\cdot|X)]) \)
  - With increasing portfolio size, premium bias becomes dominant
  - Estimation volatility \textit{diversifies away}
Relative bias of MLE of exponential tail function
Exact and approximate bias (Landsman and T., 2012)
MLE v predictive bootstrap distribution v bias-correction
References


• Bignozzi and Tsanakas (2012) Characterisation and construction of sequentially consistent risk measures.

• Landsman and Tsanakas (2012) Parameter uncertainty in exponential family tail estimation. *ASTIN Bulletin*