Solution Description

Phoenix Computing® Operational Risk ICAAP Solution
Extracting actionable intelligence from the loss and scenario data

Figure 1: 95 to 99.9 percentile QQ-plot showing the fit of the loss severity distributions to the actual loss data for the chosen unit of measure (UOM)

Introduction
Banks and other financial institutions have to prepare for an uncertain future regarding their operational risk exposure while they are unable to predict it – as really no one has a functioning crystal ball. Since the limits of human ability to forecast are well known, especially with the paucity of historical data that continues to challenge operational risk measurement at least for certain event types (e.g., EPWS or IF) in most institutions, the fallibility of such efforts is even more acute than for other measurable risks (such as market or credit risk).

The most pragmatic remedy to this situation is one of combining internal loss data (where necessary, supplemented suitably with external data, see for instance Figini, Giudici, Uberti and Sanyal (2007)\textsuperscript{1}) with internally generated scenarios data to get to possible operational risk exposures that are reflective of not only the institution’s loss experience but also sufficiently forward looking to not fall prey to the obvious limitations imposed by an insufficient data history. A solvent institution (thankfully) lacks catastrophic losses in its data history. Yet, future capital planning, at any institution, has to allow for the possibility that low frequency high impact (LFHI) losses might visit them, as they might have other peer institutions, to be considered legitimate by regulators and internal model validators alike. Business line managers, especially, have perceptions about the possible magnitude and periodicity (once in so many years…) of potential losses based on vulnerabilities that they perceive or based on considerations of failures of specific controls that might keep them up at night. The Phoenix Computing® Operational Risk ICAAP Solution provides an elegant way to harness the internal and external loss data and provides
an intuitive interface for business users to incorporate their assumptions or perceptions regarding the
range and frequency of potential future loss scenarios that might materialize outside of the realm of the
bank’s (or other financial institution’s) loss experience or might represent specific stress tests. It is
based on the LDA (loss distribution approach) described in

![Change-of-Measure Simulations with Lognormal Severity and Poisson Frequency](image)

Figure 2: Change of Measure simulations for calculating the operational value at risk (OpVaR) in closed form based on mean, variance and frequency triplets single loss approximation (SLA) in this example where the severity is lognormal and the frequency is Poisson.

Basel II under the AMA (Advanced Measurement Approach). The interface is designed, based on our experience, to minimize any cognitive bias in the enumeration and expression of such scenarios. The scenarios can be combined with internal (and/or external) loss data in a way that is robust and stable and also combines objective (historical) and subjective (scenario) data in the most optimal and easy-to-understand manner possible. To this end, we follow the Change-of-Measure methodology proposed by Dutta and Babbel (2013).

Once capital is calculated, based on historical data and pertinent scenarios and/or stress tests, there is also the question of a reasonable way to allocate it down to the legal entity or business line levels. This is to ensure that the legal entity or business line is holding an appropriate level of capital for the riskiness that it entails. These levels can then be cross-checked against existing allocations, if any. The Phoenix Computing ® Operational Risk ICAAP Solution provides complete functionality and transparency for this sort of calculation of operational risk capital and its properly risk-sensitive allocation to the satisfaction of internal and external regulators.

Advantages

1. The model allows for an effective process of exploratory loss data review for the eventual selection of the appropriate severity distribution.
2. A pragmatic selection of severity distributions is provided to make the solution affordable, understandable and self-contained without being overwhelming and/or confusing. These
include (but are not limited to) lognormal, loggamma, loglogistic, generalized Pareto (GPD), Weibull and empirical, around which best practice has coalesced\(^iv\) in the domain of LDA modeling.\(^v\) Frequency is modeled by the one parameter Poisson distribution as best practice is to use annual data (due to its stability – rather than monthly or quarterly data) and since most financial institutions only have 5-10 years-worth of loss data, it rarely makes sense to try the two parameter frequency distributions such as binomial or negative binomial as there tends not to be sufficient degrees of freedom to accommodate these. Other (more specialized) distributions may be added if the loss tail structures are complex enough to warrant their use. However, most financial institutions starting out on internal modeling for operational risk will find the selection to be quite adequate.

3. The model allows for leveraging consortium and public loss databases to facilitate scenario analysis and stress testing.

4. The analysis can be done at the level of any unit of measure (UOM) such as a Basel event type (ET) or line of business (LOB) or a product type – and, in fact, along any reporting dimension. The operational risk capital – with and without the impact of scenarios – can be computed at any level of aggregation or granularity.

5. This is not a general purpose statistical package masquerading as an OpVaR solution, as is often the case. Rather than offer a confusing universe of statistical techniques, our solution dedicated to operational risk ICAAP (or economic capital) offers a clear and prescriptive step-by-step guidance toward using the available historical and scenario data with the specific purpose of calculating a reliable operational risk capital – no more and no less. This means getting a single operational risk capital value for a given risk appetite level (as reflected by the confidence level agreed with the regulator, such as 99.5 or 99.9 percentile) with just historical (internal and/or external) loss data as well as a number inclusive of scenarios. The statistical capabilities included are selective and aligned with “best practice” and efficient LDA modeling enhanced by scenario analysis and stress testing as its goal rather than offering a plethora of general purpose statistical capabilities that would push the onus of building a statistical model for calculating OpVaR back to the business unit who rarely have the bandwidth to run a “science project;” especially since that sort of endeavor rarely produces a reasonable, stable and defensible Pillar II capital estimate.

6. Change-of-Measure or COM simulations are much faster than Monte Carlo simulations and are the mainstay of our solution. However, where necessary – for example, when historical loss data is too scarce, or the loss history related data has peculiarities that make parametric modeling untenable - empirical simulations must be conducted and in that situation we resort to Monte Carlo simulations. These run slower than the parametric COM simulations and yet are more pragmatic than waiting many more years to accumulate additional data. At least, this breaks the inertia of the classic “analysis paralysis” so that the process can get off the ground with the data available and more traditional forms of selection and estimation of parametric distributions can be undertaken at a future date when enough data has accumulated for the currently sparse ET or LOB.

7. The solution may be deployed via MS Excel™ spreadsheets or via web pages using the ASP.NET technology. Screenshots pertaining to the web-based deployment may be seen further down in the document. The Excel-based deployment has the virtue of making it easy to view all the inputs, intermediate calculations, final calculations and the OpVaR (operational value at risk)
related output. The familiar Excel GUI also makes it easier for the end-user to try any ad hoc thought experiments that might be of business interest to the business user – e.g., the increase in the capital due to the inclusion of the scenarios – individually, collectively or grouped in subsets, to just give one example. The web interface has comparable functionality and is better for staying within the interface of the existing GRC platform and for producing more official and aesthetically pleasing reports.

The Phoenix Computing® Operational Risk ICAAP solution

The Phoenix Computing® Operational Risk ICAAP Solution greatly simplifies the task of ICAAP calculation for computing the operational risk capital at a chosen risk appetite level (as defined by the agreed/prescribed percentile of the overall loss distribution). Results with both historical data (internal and/or external) and inclusive of the scenarios and/or stress tests are simultaneously reported for the chosen UOM. The former are termed the Pillar 1 OpVaR numbers and the latter the Pillar 2 OpVaR numbers as shown in the panel below for a particular event type.\textsuperscript{vi}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
\textbf{ILD only -> Pillar 1} & \textbf{Percentiles} & \textbf{UOM} \\
\hline
\textbf{Lognormal} & 9.96 & 1.50 & 129.19 & $14,409,799$ & $16,160,725$ & $21,941,639$ & \textbf{50} & $21,867$ \\
\textbf{Loglogistic} & 9.97 & 0.85 & 129.19 & $81,250,246$ & $133,588,573$ & $476,943,791$ & \textbf{75} & $57,142$ \\
\textbf{Loggamma} & 43.13 & 0.23 & 129.19 & $29,461,297$ & $37,682,301$ & $71,395,516$ & \textbf{85} & $92,301$ \\
\textbf{Weibull} & 44924.14 & 0.65 & 129.19 & $9,335,675$ & $9,501,721$ & $9,911,138$ & \textbf{90} & $123,841$ \\
\textbf{GPD} & 0.66 & 25040.40 & 129.19 & $29,510,226$ & $41,189,087$ & $101,606,597$ & \textbf{95} & $241,213$ \\
\textbf{Empirical} & & & & & & & \textbf{99} & $351,347$ \\
\hline
\textbf{ILD + Scenarios -> Pillar 2} & \textbf{Aggregate Loss Amount} & \textbf{800} & $18,785,503$ & $21,278,543$ & $29,614,378$ & $56,152,727$ & $173$ & $4,721,303$ \\
\hline
\textbf{Lognormal} & 10.11 & 1.54 & 130.28 & $18,785,503$ & $21,278,543$ & $29,614,378$ & \textbf{Min} & $70,191$ \\
\textbf{Loglogistic} & 10.07 & 0.88 & 130.49 & $119,620,208$ & $200,960,154$ & $755,749,834$ & \textbf{Max} & $4,721,303$ \\
\textbf{Loggamma} & 41.63 & 0.24 & 129.50 & $39,587,508$ & $51,608,367$ & $102,032,755$ & \textbf{Mean} & $231,942.02$ \\
\textbf{Weibull} & 53,531.91 & 0.61 & 128.41 & $12,131,709$ & $12,390,790$ & $13,036,237$ & \textbf{std} & $218.97$ \\
\textbf{GPD} & 0.72 & 27,999.12 & 131.62 & $50,249,664$ & $74,385,873$ & $209,316,662$ & \textbf{skewness} & $12.57$ \\
\textbf{Empirical} & & & & & & & \textbf{kurtosis} & $218.97$ \\
\hline
\end{tabular}
\end{table}

\textbf{Figure 3:} Pillar 1 and 2 OpVaR numbers in the two panels to the left. There is obviously a lot of variation for each percentile (99, 99.9 and 99.95) depending on the severity distribution used.

The correct distribution is chosen based on a triangulation of the industry-standard goodness-of-fit tests (see Figure 5), the QQ-plot (Figure 1) and the Dutta-Perry (2007) criteria\textsuperscript{viii} being applied to the severity loss distribution as shown in the panel to the right in the above screenshot. In this case the empirical distribution has not been run, there being enough data for parametric distributions, and hence there are no results to show in this category. On careful analysis, this sample data upholds the lognormal distribution.\textsuperscript{viii} This means that at a prescribed confidence level of 99.9 percentile, which may be determined in consultation with the regulator or is reflective of the firm’s risk appetite agreed at the board level, the Pillar 1 OpVaR is $16,160,725 and the Pillar 2 OpVaR is $21,278,543. The scenarios chosen are shown in Figure 4 below.\textsuperscript{ix}
Figure 4: Example scenarios. The second and the third columns represent lower and upper bounds of the financial impact and the third column represents the periodicity of the scenario occurring “once in so many years.” For instance, scenario 1 states that a loss of one to five million dollars can happen once in 2 years. In our experience, most participants in scenario workshops are more comfortable expressing the scenarios in this format and the resulting analysis is also more accurate. For more detail, consult Dutta-Babbel (2013). However, we can convert scenarios gathered in other formats into this format.

The impact of the scenarios is greater if their severity or frequency levels make them “fat-tailed” or, in other words, place them significantly outside of the realm of the institution’s actual loss experience. If the scenarios are more within the realm of actual loss history, they should not have too big an impact. This is due to the “residual simulation” approach detailed in the Dutta-Babbel (2013) paper and makes the model much more stable and well-behaved when scenarios are introduced or removed, unlike some earlier LDA models that would be oversensitive and unstable in this regard, not usually because the operational risk profile itself is volatile, but because of inherent stability issues in the modeling framework – e.g., in the choice of the severity distributions that are brought to bear on the model.

Figure 5: Industry standard goodness-of-fit tests – Chi-Squared, Kolmogorov-Smirnov and Anderson-Darling. A value of 1 in column H indicates that the corresponding distribution has been rejected, whereas the 0’s mean that the corresponding distributions cannot be rejected. The p-values, the QQ-plot and the Dutta-Perry criteria are then used to triangulate and hone in on the most suitable distribution.

The previous screenshots were from the MS Excel™ version of the OpVaR calculation tool. Exactly the same functionality is available on the web version of the tool that is available on the dashboard of the GRC data collection platform. Some sample screenshots are provided below – the examples are illustrative and not exhaustive. Besides, they can be customized to the client’s liking at the time of implementation.
Figure 6: Data entry and model invocation dashboard. Notice that the Data Generate Flag and the Empirical Flag checkboxes (which are mutually exclusive) – provide two pragmatic workarounds in the situation where there is a real shortage of data in a particular UOM.

The above screen is the Data Input screen in the browser version of the OpVaR model. Here a particular UOM – in this case an event type, namely, Clients, Products and Business Processes (CPBP) – can get its input data. It is also when a determination can be made if we wish to invoke the data generation process (if there is no available data for the particular UOM) using external data benchmarks that will form the body of the distribution along with internally generated scenarios. Alternatively, if there is some data but not enough for parametric estimation, the empirical distribution mode can be chosen. This kicks up the number of COM simulations from 50,000 (10,000 for each of the five parametric distributions) to 10 billion, which gives accurate results though it takes many hours to run but can be a great workaround if historical data is scant for the UOM (say, the bank has fortunately not had too many instances of internal fraud but realizes that the risk exists and wishes to model it). These alternatives are to be sparingly used but can be a great way to break the inertia of waiting around for enough data to accumulate and get started on the path to internal modeling of OpVaR.
Figure 7: Goodness-of-fit statistics on the browser version. Notice that the Weibull distribution is rejected outright. The other distributions cannot be rejected. This is where the QQ-plot at the tail and the Dutta-Perry (2007) criteria have to be invoked to choose the best fit among the remaining four.

The following screenshot shows the scenario collection and editing screen. As discussed before, in our experience, this mode of scenario data collection leads to better and more unbiased input from participants, especially if they are not quants, (which most business line managers tend not to be). In fact, research into cognitive psychology shows, that even people with substantial statistical/mathematical training can give flawed or biased estimates when the scenarios are gathered in the classic bucketed approach (“most likely” or median and “worst case” or some high percentile of severity, with corresponding probabilities)

Figure 8: Scenarios and stress-test gathering screen adapted to the Dutta-Babbel (2013) framework
**Table:**

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Parameter 1</th>
<th>Parameter 2</th>
<th>Frequency</th>
<th>Percentile 99.90</th>
<th>Percentile 99.95</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lognormal</td>
<td>9.3380</td>
<td>1.7079</td>
<td>10.0000</td>
<td>3,490,981</td>
<td>9,621,350</td>
</tr>
<tr>
<td>Logistic</td>
<td>9.2134</td>
<td>1.0416</td>
<td>10.0000</td>
<td>14,764,627</td>
<td>162,661,280</td>
</tr>
<tr>
<td>Loggamma</td>
<td>26.2265</td>
<td>0.3561</td>
<td>10.0000</td>
<td>9,294,101</td>
<td>48,671,595</td>
</tr>
<tr>
<td>Weibull</td>
<td>28,110,2919</td>
<td>0.5640</td>
<td>10.0000</td>
<td>1,279,598</td>
<td>1,855,084</td>
</tr>
<tr>
<td>Empirical</td>
<td>0.0800</td>
<td>0.0000</td>
<td>10.0000</td>
<td>2,140,498</td>
<td>2,975,699</td>
</tr>
</tbody>
</table>

**Figure 9:** Pillars I & II capital numbers

**Figure 10:** QQ-plot on the web tool
Figure 11: ICAAP tool embedded in the OneSumX GRC® (which is a product of Wolters Kluwer Financial Services) dashboard, just as an example. It can be connected in this manner to the client’s preferred GRC platform.
For further information, please contact

Aniruddho Sanyal | Principal Consultant
8 Joanne Court, Bridgewater | New Jersey | United States

Tel: +1.732.407.8966  
Email: asanyal@phoenixcomputing.net | Web: http://www.phoenixcomputing.net/home.html

---


iii This refers to using descriptive statistics on the data such as calculating the mean, standard deviation, various quantiles and percentiles, using histograms and QQ-plots to test against normality, skewness, kurtosis etc. These are computed as part of the standard output in the Parameters tab.

iv See, for instance, “Consultative Document Operational risk –Revisions to the simpler approaches,” October 2014, Bank for International Settlements, Annex 2. In trying to derive a more risk sensitive measure, an LDA approach is used to calibrate to the available industrial data. Tellingly, the distributions chosen are Pareto, lognormal, loglogistic and loggamma. We have added Weibull to this list as it baseline distribution, more thin-tailed than lognormal, that we use as a starting benchmark and an effective floor since operational loss data will not usually fit this distribution. We have also added the empirical distribution as it provides an effective workaround when there is some data for a UOM but not enough for traditional parameter-estimation. Using the empirical distribution can also help when the data does not fit standard theoretical distribution – for instance, if the loss data is bi- or multimodal.

v Due to the special characteristics or idiosyncrasies of an institution’s data, this list of distributions might have to be augmented or modified. That would constitute some customization. The determination is made at the time of the implementation, after an initial review of the data has been conducted. This step might lead to some supplemental but reasonable costs. The idea here is to provide a baseline list of the typical severity distributions that our experience shows to be the best fit for the usual data profiles that we have encountered.

vi The most common event type is EDPM (execution delivery and process management). Or, we could have chosen a particular business line, though there is evidence that it is better to start out with event types data for better homogeneity of the input and better stability of the output.


viii This stands to reason, as the sample historical data was generated here using random draws from a lognormal distribution. It also provides a ‘sanity check’ for the model since the goodness-of-fit exercise was able to correctly identify the underlying severity distribution. Since we have no way of knowing the true distribution known only to mother nature, we need to validate the model by passing it data following a distribution known to us and then checking its ability to identify it through its own triangulation process involving the goodness-of-fit tests, the QQ-plot and the Dutta-Perry (2007) qualitative criteria.

ix It is our experience that sound