Modeling Truncated Loss Data of Operational Risk in E-banking

Maryam Pirouz
Department of Computer Engineering, International Unit, University of Guilan, Rasht, Guilan, Iran
E-mail: mpirouz60@yahoo.com

Maziar Salahi
Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Guilan, Rasht, Guilan, Iran
E-mail: salahim@guilan.ac.ir

Abstract— Operational risk is an important risk component for financial institutions, especially in E-banking. Large amount of capital that are assigned to decrease this risk are evidence to this subject. One of the most important factors for modeling operational risk to estimate capital charge is loss data collections of banks. But sometimes for reasons like decreasing the costs, banks save only the losses above some determined thresholds at their database. For achieving accurate capital charge, this threshold should be considered in determining capital charge. This paper focuses on modeling truncated loss data above some given threshold. We discuss several statistical methods for modeling truncated data, and suggest the best one for modeling truncated loss data. We have tested our suggested model for some operational loss data samples. Our results indicate that our approach can be useful for increasing accuracy of estimating operational risk capital charge in E- banking.

Index Terms— E-banking, Operational Risk, Truncated Data, Basel Committee

I. Introduction

Basel Committee on Banking Supervision (BCBS) has assessed most important risks associated with banking under some main categories, namely market, credit, operational, liquidity, reputation, legal, and strategic.

By the introduction of electronic banking, the original characteristics of traditional banking are changed. Some of the observable changes including: unprecedented speed of technological changes, product and service innovation due to changing customer expectations, the ubiquitous and global nature of open electronic networks, integration of E-banking applications with legally computer systems, increasing dependence of banks on third- party service providers [1].

Due to all these developments and changes, Basel Committee obliged to research more about E-banking risks. So, they have studied traditional banking risks and E- banking capabilities. They have realized that although E-banking would not produce risk inherently, but have changed or increased some traditional risks. Their research showed that, impact is more on strategic, operational, legal and reputational risks, because of the form of financial institutions [2].

Strategic risks are mainly associated with board and management decisions. At the E-banking context, using technology when management does not adequately planed for, manage and monitor the performance of technology-related products, can lead to strategic risk. Legal risk is the risk to earnings or capital arising from violations of, or nonconformance with laws, rules, regulations, or ethical standards. Although when legal lawful and liability of two sides for one transaction would not establish well, this risk can create. Reputational risk arises from negative public opinion. E-banking services that are poorly executed or otherwise alienate customers and the public can lead damaged to a licensee's reputation [1].

There are many definition of operational risk. Basel Committee offers a more accurate definition of operational risk as "The risk of loss resulting from inadequate or failed internal processes, people and systems or from external events failures" [3].

This definition includes legal risk, but excludes strategic and reputational risk. Some analysts call it a transaction risk, security risk, or IT risk. Examples of operational risk can include internal and external fraud, technological inadequacies, human factors such as lack of training, negligence by customers and employees, product and service liability, misuse of confidential information, damage to physical assets, business disruption and system failures, failed or erroneous transaction processing, failed outsourced processes.

Complexity and structure of bank's processing environment, type of services offered, and the complexity of supporting technology also affect the level

of operational risk. The risk is heightened when the institution offers innovative services that have not been standardized.

At recent years, operational risk has transformed to the most important financial industry risk topics both at academic and practical areas. The reasons of this attention can be related to large investment at information technology and communication systems, increasing financial institution merge and growing banking industry communication. In addition, accords of Basel Committee have determined capital requirement for operational risk, and financial institutions are required to perform it [3].

Until now literature on operational risk almost focused on two topics: first, the estimation of operational risk losses [4-8] and second, application of this estimate to determine capital charge [9-13]. Estimation of capital charge for operational risk should be based on Basel Committee accords.

The most important factor for modeling operational risk and getting capital charge is the history of loss data in bank. But sometimes for some reasons these data have not been saved completely. In the best case, the data collection process should record all operational loss events, even the trivial ones. But sometimes, the data recording is related to lower recording thresholds,

so that only data more than a certain amount enter databases.

At this case, the data would be left-truncated for estimation. At operational risk modeling in E-banking to determine an exact capital charge, we must have special attention to left-truncation of the data. In other hand, we must properly address them in the estimation process.

In this paper, first we introduce truncated loss data in Section 2 and then in Section 3, we present two statistical models for truncated data. Then by comparing these two methods, we suggest one of them for modeling truncated loss data of operational risk. Also we explain the weakness of the other one. The implementation results are presented in Section 4.

II. Truncated Loss Data of Operational Risk

Basel Committee, in their Quantitative Impact Study 3 (2002 LDCE) [14], asked from 89 participant banks to report information about the minimum cutoff levels that they used when recording operational loss data. Table 1 reports the results. Most of the banks set the cutoff level at 10,000 Euro.

Table 1: Minimum cutoff thresholds used by 2002 LDCE participating banks [14]

Minimum Cutoff Threshold (in Euro)	No. of Banks
Less than 10,000 for all business lines	5
10,000:	59
10,000 For all business lines	57
10,000 For some business lines	2
More than 10,000:	13
More than 10,000 For all business lines	12
More than 10,000 For some business lines	1
No business line information reported	12
Total	**89**

There are several reasons for setting such minimum threshold [15]:

1. Data recording may be expensive. It has been proved, when the threshold is decreased linearly, the costs of recording data increases exponentially. Additionally, a large number of small losses may be recorded with mistakes that can result in extra operational losses.

2. Hidings small losses are easy, while larger loss is harder to hide. For example, a businessman who has performed a merchant mistake may counterfeit document and succeed in making trivial losses unobserved by bank management. But large amounts are much more difficult to hide.

3. In past years, incomplete ways for recording operational loss data may have been such that minor losses could be left unrecorded, but bigger losses were appropriately recorded.

The regulatory capital or capital charge is meant to assure bank's ability to cover major potential losses without causing a banking crisis. Consequently, regulatory capital management should ensure the stability of the banking sectors and protect depositors. Best value of capital charge should be extracted from suitable model for each bank [2]. Considering truncated loss data, can help to have more precise model for operational risk.

III. Modeling Truncated Loss Data of Operational Risk in E- banking

3.1 Modeling Operational Risk

Basel Committee on Banking Supervision, in its accords, has introduced the loss distribution approach (LDA) and gave enough freedom to banks to demonstrate their statistical models based on LDA. But for calculating operational risk capital charge, banks should display their presented model and evaluate measure of operational risk for one year period with 99.9% confidence level.

Under (LDA), bank's activity arranged into a 56 cell matrix included 8 Business Line (BL) × 7 Event Type (ET). For each pair, key task is estimating frequency and severity of losses. Based on these two distributions estimation, bank would calculate aggregated losses probability distribution function.

Operational capital charge is calculated as sum of VaR value (with 99.9% confidence level) in one year period for each pair BL/EV. 99.9th percentile means that capital charge totally is enough for covering losses, but with 0.1% defeat, in other words there is 0.1% probability that banks can not cover inconsistent operational losses.

Our model for operational risk has the following steps to get operational risk capital charge:

a. Calculating statistical characteristics of loss data, (we have calculated, mean, median, Skewness, Kurtosis).

b. Performing descriptive statistic analysis for loss data using different statistical plots of data, (we have used histogram of loss data).

c. Choosing appropriate statistical distributions for fitting frequency and severity of loss data.

d. Checking the fitted distribution with goodness of fit tests, (we have used Q-Q plot).

e. Computing aggregate loss distribution from combination frequency and severity distribution. To do so, we have used Monte Carlo simulation method, that has the following three steps:

- Choosing a probability model for frequency and severity of loss.

- Simulate the number of losses and individual loss amounts and then calculate the corresponding aggregate loss.

- Repeat many times (at least 5,000) to obtain an empirical aggregate loss distribution.

f. Calculating 99.9th percentile of aggregated losses distribution to estimate operational risk capital charge.

But modeling data in the presence of missing data follows two statistical methods that we explain them in the next part. Then we compare those two for choosing the best one for modeling truncated loss data of operational risk in E- banking.

3.2 Statistical Approaches for Truncated Data

First we denote the minimum threshold which data are being recorded in a database by *H*. Thus by drawing histogram of the observed data, we see the right tail of loss distribution, instead of seeing true distribution. For modeling this type of data, there are two statistical approaches, as described in Table 2.

Table 2: Statistical approaches for truncated data

Statistical approaches	Explanation
Naive approach	In this method, observed data are considered as complete data. Based on this assumption, a distribution is given to fit the data.
Conditional approach	In this method, based on the assumption that the data above the threshold H are stored, a distribution is given to fit the data.

To use conditional approach, we should accept three assumptions:

1. We do not have any information about missing data.

2. Both missing data and recorded data belong to the same group of distributions with the same parameters.

3. Frequency of data and severity of them should be considered as two independent random processes.

In conditional approach, the density of data can be represented as

$$f(x|X \geq H) = \frac{f(X \geq H)}{p(X \geq H)} I \qquad (1)$$

where

$$I = \begin{cases} 1 & \text{if } x \geq H \\ 0 & \text{if } x < H \end{cases}$$

Then, to cover the lower part of the loss, between 0 and H, we can use obtained parameters from fitting the conditional density, to extrapolate the truncated density distribution. By extrapolations, we can present the unconditional density for the complete data sample, which is collected of the missing and recorded data.

3.3 Comparison of Two Statistical Approaches for Truncated Loss Data of Operational Risk

Now by comparing characteristics of two statistical approaches, we choose conditional approach for modeling truncated loss data of operational risk.

Using naive approach for modeling truncated loss data can result to inaccurate results in computing capital charge at final operational risk model. The reason is that it connivances missing data between 0 and *H* and assumes the loss data is complete.

But, under conditional approach, we can fit the distribution for the right tail of data, based on the fact that recorded data begin from H instead of 0.

To cover the left part of loss distribution, which is between 0 and H, we can use extrapolation for parameters we obtain from conditional approach.

IV. Practical Results

In this section, first we present results of implementing our model for operational risk loss data, to get aggregated loss distribution histogram of data, and then we implement conditional approach and Naive approach for operational truncated loss data. Our comparison shows that conditional approach is a perfect method for modeling truncated loss data of operational risk.

4.1 Modeling Operational Risk for Getting Aggregated Loss Distribution Histogram of Data

In this section we implement our model to get capital charge and the aggregated loss distribution histogram of our loss data. Input data for our model is in Table 3 and we can see the results of modeling in Table 4. Fig. 1 illustrates the histogram of operational loss data that matches with information of Table 4.

Table 3: Input data for our model of operational risk

Loss Frequency	Poisson Lambda=200
Loss severity	Lognormal u=10 Sigma=2.5
Number of Simulation	100,000
Confidence level	99.9%
Iteration number of algorithm	25

Table 4: Results of our model of operational risk

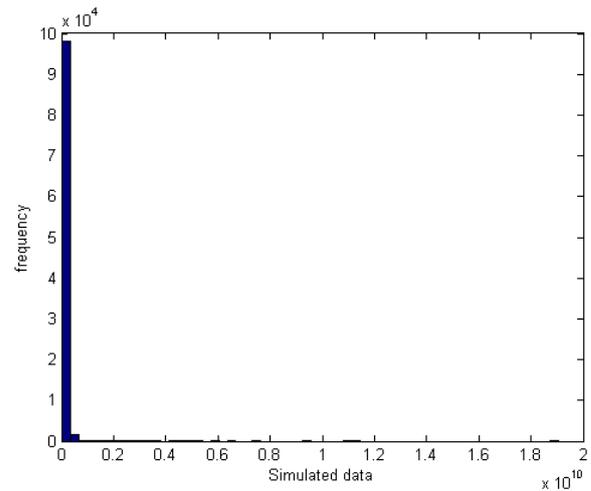| Average loss	1.0051e+008
Standard deviation	1.6253e+008
Skew	76.0204
Kurtosis	1.1576e+004
Var	1.4860e+009
CVaR	2.8436e+009
error_VaR	1.8030e+007
error_CVaR	4.8974e+007
time	110.275046 seconds

Fig. 1: Aggregated loss distribution histogram of operational losses data

4.2 Determining Minimum Threshold of Loss Data

For 1000 number of operational loss data that follow from Log Normal distributions with parameters of Table 5, we consider %10 of data as minimum threshold. This information briefly presented at Table 5.

Table 5: Information of fitted distribution and minimum threshold

Loss distribution parameters	Log Normal μ=0 σ=0.5
number of Losses	1000
Minimum Threshold	H= 0.5439

So, the minimum threshold which losses are being recorded in a bank's databases for our implementation is H= 0.5439.

With this information about operational loss data, we are going to model truncated loss data to find perfect approach for E-banking.

4.3 Modeling Complete Operational Loss Data Set

In this section, we consider complete operational loss data without minimum threshold H and we fit Log Normal distribution for this complete data set.

Fig. 2 illustrates fitting this distribution for the histogram of our complete operational loss data. As we see, fitted distribution can completely cover the histogram.

So when the data between 0 and minimum threshold H, are not absent, fitting distribution is simple and probably have no problem.

But in practice, as we said before, most of the time banks do not record loss data between 0 and H. This means that using this approach for deficient operational loss data would be impossible, but for complete loss databases would be practicable.
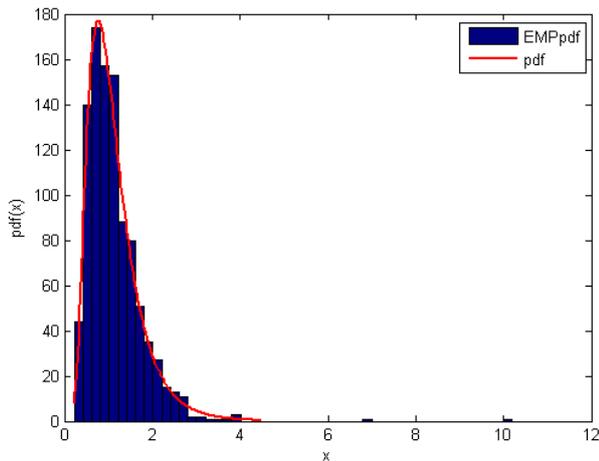
Fig. 2: Fitted Log Normal distribution for completed loss data

Fig. 3: Fitted Naive approach for truncated loss data

In the next two sections we explain two other approaches.

4.4 Modeling Truncated Operational Loss Data Set with Naive Approach

At this part, we have considered %10 of operational loss data as minimum threshold H and then used Naive approach. We have fitted Log Normal distribution for the reminding data; the right part of threshold H. Fig. 3 illustrates fitting this distribution for the histogram of truncated operational loss data.

As we see, Naive approach could not cover the histogram of operational loss data completely. The peak part of loss data is without any distribution covering of Log Normal. So as we expected earlier, for this type of data, Naive approach would not be the good choice.

4.5 Modeling Truncated Operational Loss Data Set with Conditional Approach

At the final stage, we have implemented conditional approach for truncated loss data of operational risk, that considers condition of absent data between 0 and H. Fig. 4 illustrates fitting this approach for the histogram of truncated operational loss data. As we see, this approach could properly cover the histogram, even the peak part of it.

The values of probability density function for 900 values greater than minimum threshold have reached by conditional probability law that we have explained it earlier at (1).

Implementation results of these two methods show that, the conditional approach is better than the other one for modeling truncated loss data of operational risk in E-banking. As we see, using more accurate approach for modeling truncated operational loss data in banking and other financial institutions can lead to save money and prevent many operational risk crises.
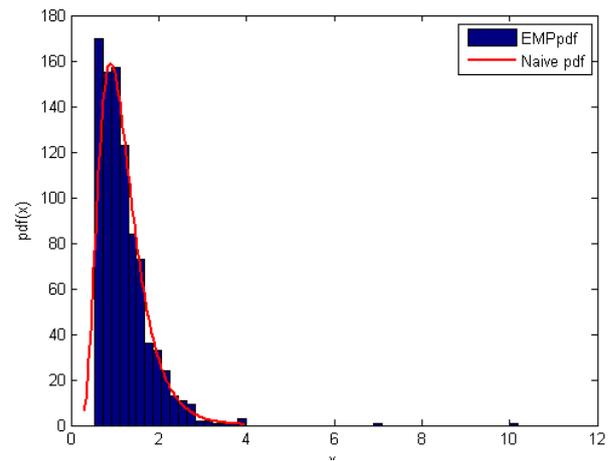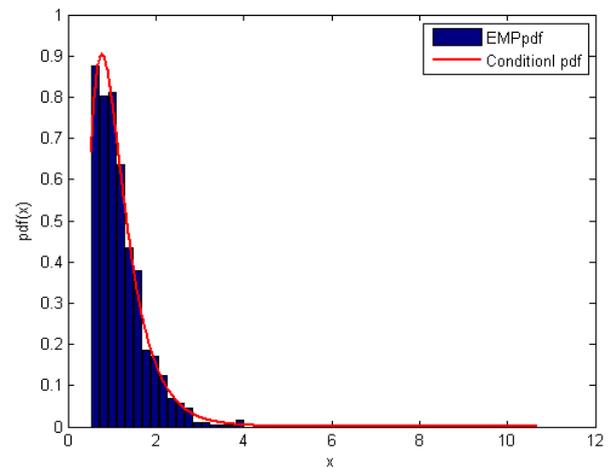

Fig. 4: Fitted conditional approach for truncated loss data

V.   Conclusion

In this paper, we have proposed a statistical approach for modeling operational risk truncated data in E-banking. Our experiments show that using the proposed approach leads to more accurate estimate of capital charge in operational risk which prevents operational risk crisis in E-banking.

References

[1] J. Kondabaghil, Risk Management in Electronic Banking, John Wiley & Sons (Asia) Pte Ltd., 2007, pp. 11-13.

[2] Basel Committee on Banking Supervision, risk management principles for Electronic banking, 2003.

[3] Basel Committee on Banking Supervision, International Convergence of Capital Measurement and Capital Standards - A Revised Framework Comprehensive Version, 2006.

[4] A. Chapelle, Y. Crama, and G. Huʾbner, Practical Methods for Measuring and Managing Operational Risk in the Financial Sector: A Clinical Study, Journal of Banking &Finance, 2008, 32, pp. 1049-1061.

[5] V. Chves-Demoulin, and P. Emberchts, Quantitative models for operational risk: Extremes, dependence and aggregation. Journal of Banking and Fiance, 2006, 30, pp. 2635-2658.

[6] R. Coleman, Operational risk modelling for extremes, Operational Risk, pp. 6–9, 2003.

[7] P. de Fontnuvelle, E. Rosenbrgren, and J. Jordan, Implications of alternative operational risk modeling techniques. Working paper, Federal Reserve Bank of Boston, 2004.

[8] P. de Fontnouvelle, J. Jordan, V. DeJesus-Ureff, and E. Rosengren, Capital and risk: New evidence onimplications of large operational losses. Working paper, Federal Reserve Bank of Boston, 2005.

[9] H. Dahen, and G. Dionne, Scaling Models for the Severity and Frequency of External Operational Loss Data, Journal of Banking & Finance, 2010, 46, pp. 3545-3568.

[10] S. Ebnother, P. Vanini, A. McNeil, and P. Antolinez-Fehr, Modeling operational risk. Working paper, ETH Zurich, 2001.

[11] P. Embrechts, and G.Puccetti, Aggregation operational risk across matrix structured loss data. Journal of Operational Risk, 2004, 3(2), pp. 29-44.

[12] F. Lindskog, and A. McNeil, Common Poisson shock models: Applications to insurance and credit risk modeling. ASTIN Bulletin, 2003, 33 (2), pp. 209–238.

[13] M. Moscadelli, "The modelling of operational risk: Experience with the analysis of the data collected by the basel committee. Technical Report 517, Banca d'Italia, 2004.

[14] BIS 2003, The 2002 Loss Data Collection Exercise for Operational Risk: Summary of the Data Collected, www.BIS.org.

[15] A.S. Chernobai,S. T. Rachev, and F. J. Fabozzi, Operational Risk, A Guide to Basel II Capital Requirements, Model and Analysis, John Wiley & Sons, 2007, pp. 183- 186.

Authors' Profiles

Maryam Pirouz: has received her Master degree in information technology from the International Unit, University of Guilan in 2011.

Maziar Salahi: is an Associate Professor at the department of applied mathematics, University of Guilan. He received his PhD from McMaster University, Canada in 2006 in the field of optimization.