Paper FS-01-2013

PROC NLMIXED and PROC IML
Mixture distribution application in Operational Risk

Sabri Guray Uner, Union Bank, N.A., Monterey Park, CA
Yunyun Pei, Union Bank, N.A., Monterey Park, CA

ABSTRACT

The main purpose of this study is to provide guidelines on how to consider mixture distributions for operational risk severity distribution modeling, with an emphasis on truncated loss data. Mixture model probability distribution function for truncated operational loss data is introduced and we presented our findings for empirical tests to estimate distribution parameters. However, this study does not intend to advocate or to propose adopting mixture forms without exploring other alternatives, but rather highlights the flexibility of the mixture models and present examples where it can serve better for some specific cases.

Keywords: Operational risk, capital model, mixture distribution, severity fitting, truncated distribution

Disclaimer: The views expressed in this paper are those of the authors and do not necessarily reflect those of the institutions they are affiliated to, such as Union Bank. The authors would like to express thanks to … for valuable feedback and comments.

INTRODUCTION

Mixture distributions can be used to model processes with samples as identified or suspected to contain a number of sub-populations. Application of finite mixture densities is most convincing for circumstances where the existence of subpopulations is strongly implied by the nature of the process.

In financial risk modeling, this can be observed due to either heterogeneous or non-stationary processes. In operational risk management, this can arise due to different factors, such as cross-sectional variation in financial institutions' risk profiles within external data (industry / consortium data) or time-variant risk factors affecting institution's own risk profile within internal loss data. In this study we discuss mixture distributions in general and possible applications in operational risk modeling as an alternative flexible distributional form to capture non-unimodal circumstances.

The main purpose of this study is to provide guidelines on how to consider mixture distributions for operational risk modeling, with an emphasis on truncated loss data. However, this study does not intend to advocate or to propose adopting mixture forms without exploring other alternatives, but rather highlights the flexibility of the mixture models and present examples where it can serve better for some specific cases.

FINITE MIXTURE DISTRIBUTIONS IN GENERAL

DEFINITION

As a simple definition, a mixture distribution is a combination of multiple distributions in a single functional form. In other words, a mixture distribution is a weighted combination of other known distributional forms. This allows for a great flexibility in statistical modeling to accommodate different multimodal shapes and allows finite mixtures applied to very different frameworks.

These building blocks are usually called as 'component distribution' of the mixture model. The number of components in a mixture form needs to be estimated or specified, and we have little theoretical guidance on this. As an example, the density function for mixture model with only two components is defined as:

$$x \sim m(.) \quad where \quad m(.) = w * f(.) + (1-w) * g(.)$$

where $f(.)$ and $g(.)$ are the density functions for each component distribution. The probability weights are simply uniform functions and equal to w and (1-w) respectively. These probabilities are simply called as 'mixture proportion' or 'mixture weight'.

These component distributions can be from the same or different distributional families. If the component distributions are from the same family, the mixture is called homogeneous. In most applications, the components are assumed to take same form and homogeneous mixtures are more commonly used (Everitt (1996)).

FLEXIBILITY OF MIXTURE DISTRIBUTIONS

Very often, the rationale of adopting a mixture model is the presence of possible sub-populations. Finite mixture distributions have many applications where the purpose is identifying and eliciting the characteristics of the heterogeneous subgroups. (Gardiner et al, (2012))

Finite mixture model provides a natural representation of heterogeneity in a finite number processes. Therefore, finite mixture models appeared to be useful where categorization is not feasible due to heterogeneity in the population. Such distributions provide an extremely flexible method of modeling unknown and multimodal distributional shapes which apparently cannot be accommodated by a single distribution.

Component distributions represent local area of support of the true distribution which may reflect the behavior of underlying process, belonging to a different state such as different regime or risk management profile. Therefore, the application of finite mixture models is most convincing in situations where the existence of separate groups of observations with differing distributions is strongly implied by the nature of the application.

MIXTURE DISTRIBUTIONS IN RISK MODELING

If the process is homogeneous and stationary (time invariant parameters and distributions) throughout the estimation period, i.e. underlying process does not change when shifted in time or space; then the historical data will exhibit desired statistical features for financial modeling. Otherwise various statistical issues will arise such as non-stationary, heterogeneity which should be accommodated properly for a robust model.

Obviously, mixture models are adopted by financial risk discipline within different modeling frameworks to accommodate these. As an example, the underlying process is usually a function of various risk factors such as macroeconomic environment, market conditions, current business profile, risk controls in place etc. Therefore the underlying process can exhibit an inherited time variant characteristics such as regime shift behavior due to factors such macroeconomic environment, risk profile and controls etc. Similarly, the data can exhibit cross-sectional heterogeneity. In financial risk this can arise due to heterogeneity across firm specific factors. D

For the cases discussed above, a mixture model can accommodate the historically observed data in that sense and offers a flexible solution for different distributional forms. Therefore it has been adopted in practice and provides an intuitive interpretation

APPLICATIONS OF MIXTURE DISTRIBUTIONS

Due to their flexibility, mixture models have been increasingly exploited as a convenient, semi-parametric way in which to model different distributional forms (McLachlan and Peel (2000)).

For example, Baixauli and Alvarez (2010) consider mixture models in credit risk context, specifically to model the market implied recovery rates for credit VaR, due to market implied recovery rates exhibiting local modes.

As Alexander (2008) explained, mixture distributions also have an intuitive interpretation in market risk context when financial markets display regime-specific behavior.

Similarly, Giacomini, Gottschling, Haefkec and White (2007) adopted mixture distributions models to forecast U.S. inflation heavier tails. Using mixture of t-distributions forecasts were more accurate, out-of-sample, than forecasts obtained using normal or standard t-distributions.

OPERATIONAL RISK MODELING AND MIXTURE MODEL

LOSS DISTRIBUTION APPROACH (LDA) IN OPERATIONAL RISK MODELING

Distributional assumptions underpin the most, if not all, operational risk modeling approaches in general. A typical LDA based approach was described by Frachot et al (2001), Aue and M. Kalkbrener (2006) and Uner (2008). In LDA, operational loss data is classified into segments with common operational risk causes, called unit of measure's (UoMs). For each unit of measure, an annual loss distribution is modeled first. For that purpose, the frequency and severity distributions are estimated separately and simply aggregated via a convolution process. Assuming that homogeneous unit of measures are specified so that losses are driven by common operational loss processes, historical data is used to calibrate the parameters for the frequency and severity distributions. These parameters together with specified distributions represent future potential losses. This way an institution can generate an annual loss distribution to estimate capital charge.

NON-STATIONARY AND HETEROGENEITY

A critical assumption here is that the underlying process is stationary and homogeneous.

Fitting a single distribution form also assumes the underlying loss process is stationary and homogeneous. However there can be deviations from these assumptions and the data can exhibit multimodality, due to non-stationary process and/or heterogeneous processes.

The presence of multimodality for internal or external loss data can be suggestive of more than one underlying loss generating process, either due to different regimes and/or different sub-group of banks. Mixture models' tractability for modeling subpopulations and flexibility enough to describe unknown and multimodal distributional shapes, which apparently cannot be modeled by single distribution, suggest them to be popular solution for this.

As Everitt (1996) suggested, mixture models are most often used in one or other of the following contexts

- Distribution to be modeled is known or consists of well-defined subpopulations but the individual class memberships are unobservable.

- Or populations are suspected to have subpopulations and mixture models are used to explore this.

There can be various reasons which will reveal non-stationary historical loss experience especially for internal data. For the underlying latent process, ie loss generating process, many different factors play role both for frequency and severity distributions, some of them being time variant in nature.

Nonetheless to say, the underlying factors change over the years. For example, the risk systems and controls in place are much more effective compared to earlier periods when operational risk management was a new practice for banks. So, the assumption of time invariant underlying process and risk factors may need to be relaxed because the underlying process can exhibit an inherited time variant characteristics.

As explained before, this regime shift behavior can be due to change in risk management practices, risk characteristics and/or macroeconomic environment. So, time variant risk factors may be a reason to suspect subpopulations with different parameters as different regimes, due to change in macroeconomic environment, risk controls or other unobservable factors.

In operational risk, subpopulations can also arise due to heterogeneity, especially for industry data. As suggested by Figini (2007), external data can exhibit a possible heterogeneity.

Two examples of increasing weights are provided in the charts below for demonstration. This way by weights adjusting to change, mixture models can also help to attain stability in parameter estimates and distributional forms, which is a challenge with single distributions.

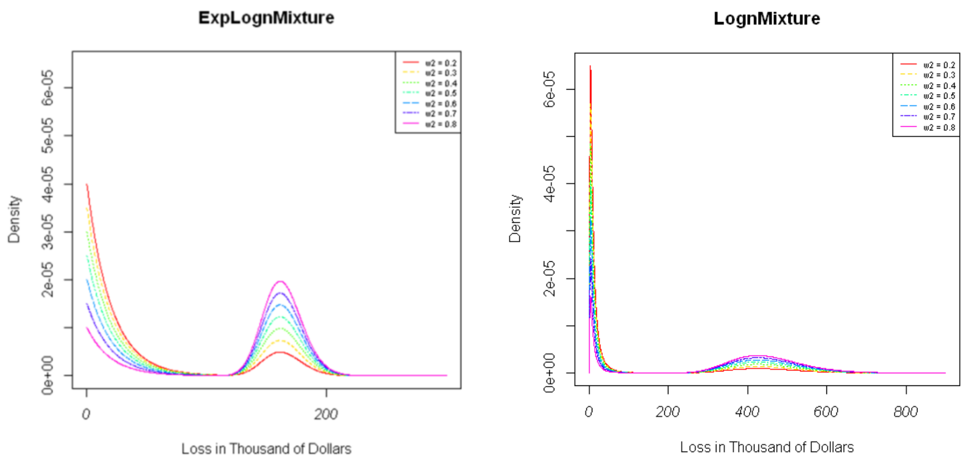

Figure 1: Examples of increasing component weight.

MIXTURE MODEL PROBABILITY DENSITY FUNCTION FOR TRUNCATED DATA

In this section, we intend to provide guideline for mixture distributions in operational risk practice, with a focus on truncated loss data. For this purpose, we present unconditional density function for mixture distribution and derive the conditional density function to estimate true parameters for the mixture distribution from a truncated data.

We consider hypothetical data to conduct empirical tests. For this, we perform empirical tests, present results for mixture model form recovering true parameters from a truncated sample or samples. For simplification, we demonstrate the two distributions example here.

Assume $X \sim f(x)$ *and* $Y \sim g(y)$ are the density functions for each component with corresponding parameters. Also assume that the probability of being generated from one component versus the other one, i.e. weight, is defined uniformly so that an observed sample will have the weights w and 1-w for the two components.

So the mixture distribution $Z \sim m(z)$ has a probability density function (pdf) defined as;

$$Z \sim m(z) \quad i.e. \;\; m(.) = p(u_{\leq w}) * f(.) + p(u_{> w}) * g(.)$$

The mixture weights are referred as prior probability weights, i.e. unconditional weights.

In operational risk management, the observed historical loss data is usually truncated. The truncation will actually make the components to be truncated and the weights for the sample to be conditional weights.

The conditional pdf for the mixture model is provided below (we will refer this as **Mixture Form**).

Mixture Form

$$\frac{m(.)}{1 - M(Threshold)} = \frac{p(u_{\leq w}) * f(.) + p(u_{> w}) * g(.)}{1 - [p(u_{\leq w}) * F(Threshold) + p(u_{> w}) * G(Threshold)]}$$

This conditional pdf will recover the unconditional (true) parameters for the mixture model using a truncated sample. By fitting this above to a truncated (conditional) loss data, we simply recover the parameters for the unconditional distribution.

Here, we show Log-Normal Mixture distribution Maximum Likelihood Estimation (MLE) fitting routine using SAS/PROC NLMIXED. Other component distributions can be easily defined for pdf and CDF functions.

```
* Distribution Fitting;
proc NLMIXED data=Input_Data  maxit = 1000 maxfu = 10000 tech = tr;

       parms
       param_1=&Param_1_init,     param_2=&Param_2_init,
       param_3=&Param_3_init,
       param_4=&Param_4_init,     param_5=&Param_5_init;

       bounds
       param_1 >= &Epsilon ,      param_2 >= &Epsilon,
       Param_3 >= &Epsilon,       Param_3 <= 1,
       Param_4 >= &Epsilon,       Param_5 >= &Epsilon;

       /*pdf for LogNormal Mixture Dist */
       pdf_LNMix =   (1-Param_3)   *pdf("LOGNORMAL", &AMOUNT, Param_1, Param_2)
              + Param_3     *pdf("LOGNORMAL", &AMOUNT, Param_4, Param_5);

       /*CDF for LogNormal Mixture Dist */
       CDF_LNMix = (1-Param_3)*cdf("LOGNORMAL",THRESHOLD, Param_1, Param_2)
              + Param_3 *cdf("LOGNORMAL",THRESHOLD, Param_4,Param_5);

       logf = log(pdf_LNMix) - log(1-CDF_LNMix) ;

       model &AMOUNT ~ General (logf);
       ods output ParameterEstimates = Distribution_Fitting;
       run;
```

Now we can derive this conditional density function for the mixture distribution.

For a truncated (i.e. conditional) sample with given threshold for the component distributions are represented as;

$$\frac{f(.)}{1 - F(Threshold)}$$

And

$$\frac{g(.)}{1 - G(Threshold)}$$

Also for the truncated sample, the weights for components are actually conditional weights.

Conditional weights are represented as

$$p(u_{\leq w|z\geq T}) \qquad \text{and} \qquad p(u_{> w|z\geq T})$$

These are posterior probability weights, ie conditional weights.

So the conditional density function for mixture distribution can be represented as;

$$\frac{m(.)}{1 - M(Threshold)} = p(u_{\leq w|z\geq T}) * \frac{f(.)}{1 - F(Threshold)} + p(u_{> w|z\geq T}) * \frac{g(.)}{1 - G(Threshold)}$$

The conditional weights $p(u_{\leq w|z\geq T})$ can be represented in terms of unconditional weight $p(u_{\leq w})$ as below.

$$p\ (u \leq w|\ z \geq Threshold) = \frac{p(u \leq w\ and\ z \geq Threshold)}{p\ (z \geq Threshold)} = \frac{p(u \leq w) * p(x \geq Threshold)}{p\ (z \geq Threshold)}$$

$$= \frac{p(u_{\leq w}) * [1 - F(Threshold)]}{1 - [p(u_{\leq w}) * F(Threshold) + p(u_{> w}) * G(Threshold)]}$$

and

$$p\ (u > w|\ z \geq Threshold) = \frac{p(u > w\ and\ z \geq Threshold)}{p\ (z \geq Threshold)} = \frac{p(u > w)\ *\ p(y \geq Threshold)}{p\ (z \geq Threshold)}$$

$$= \frac{p(u_{> w})[1 - G(Threshold)]}{1 - [p(u_{\leq w}) * F(Threshold.) + p(u_{> w}) * G(Threshold)]}$$

By substituting these weights, we derive Mixture form

$$\frac{m(.)}{1 - M(Threshold)} = \frac{p(u_{\leq w}) * f(.) + p(u_{> w}) * g(.)}{1 - [p(u_{\leq w}) * F(Threshold) + p(u_{> w}) * G(Threshold)]}$$

Or with conditional weights, we have the equation below to estimate unconditional (true) parameters and weights for the mixture model using truncated sample;

$$\frac{m(.)}{1 - M(Threshold)} = p(u_{\leq w|z\geq T}) * \frac{f(.)}{1 - F(Threshold)} + p(u_{> w|z\geq T}) * \frac{g(.)}{1 - G(Threshold)}$$

EMPIRICAL TESTS TO ESTIMATE PARAMETERS

EMPIRICAL TEST SET-UP

In this section, we conduct a few empirical tests to demonstrate how to specify the density function to estimate unconditional true parameters using Mixture Form 1.

We first generate samples with size of N=10,000 from few examples of mixture distribution with given component weights and parameters. We then truncate the samples at different thresholds to test truncated data that represents a more realistic operational loss case. We then test to recover true parameters by fitting Mixture Form 1 and Mixture Form 2.

We describe our test set-up before we move on to discuss the results.

<u>Distribution Types</u>:

For testing purpose we consider different mixture models consisting of, both from the same and different distribution families. We generally consider distributions which are commonly used in operational risk context to represent more realistic examples of loss severity distributions.

As an example we use

1. Exponential & Log-Normal mixture
2. Log-Normal & Log-Normal mixture
3. Weibull & Weibull mixture distributions.

The test can easily be extended to other mixture distributions.

Parameters:

For the mixture distributions above, we define their corresponding mixture weights and component parameters i.e. shape and scale parameters. We specified these values to have more realistic examples of mixture forms. We achieve this first by representing a severity loss range which resembles an actual loss amount, i.e. high skewness and heavy tailedness and secondly by having bimodal forms that exhibits more than one loss generating processes which demand a mixture model treatment.

Threshold:

Operational loss data is usually a truncated sample. In addition to no threshold case (ie 0), we also tested different thresholds.

Estimation method:

For simplicity, we use Maximum Likelihood Method[1] (MLE) to estimate parameters. Due to its simplicity and easy implementation, MLE serve well for our purpose to compare different functional forms. Even though our goal is not to compare performance of different methods, we also consider alternatives such as cross entropy (CE).

EMPIRICAL TEST RESULTS

CASE 1: EXPONENTIAL & LOG-NORMAL MIXTURE DISTRIBUTION

We generated a sample from Exponential & Log-Normal mixture distribution with 60% and 40% weights respectively and specified the parameters below. We defined weight and the parameters so that the sample represents bimodal form as shown in the histogram below for the sample.

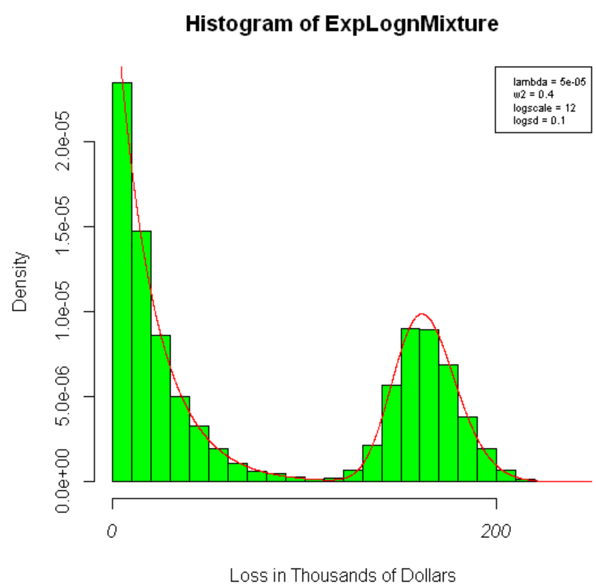

Figure 2: Exponential Log-Normal Mixture distribution probability density

Based on the parameters, Exponential and Log-Normal distributions contribute to different regions of the loss distribution domain. Exponential distribution mainly represents less severe loss amounts and Log-Normal distribution concentrates losses around $150K and contributes more at higher quantiles compared to Exponential.

[1] Shortcoming in Maximum Likelihood estimation is not the scope of this paper.

	Min	Qu.10%	Qu.20%	Qu.50%	Mean	Qu.75%	Qu.80%	Qu.90%	Max
Mixture	***9***	***3,571***	***8,078***	***35,606***	***77,476***	***157,569***	***162,503***	***174,833***	***238,239***
Exponential	9	1,955	4,378	13,935	20,030	27,348	32,185	46,275	183,689
LogNormal	114,238	142,691	149,242	162,477	163,646	174,813	178,002	186,082	238,239

Table 1: Exponential and Log-Normal Components

Since operational loss data is generally truncated, we considered different truncated samples. As shown below, since Log-Normal mainly contributes to more severe losses, increasing thresholds results in higher representation of Log-Normal (2nd component) in the truncated sample.

Threshold	N1	N2	1-w2 (cond)	w2 (cond)
0	6000	4000	60%	40%
1000	5693	4000	59%	41%
2000	5390	4000	57%	43%
3000	5158	4000	56%	44%
5000	4662	4000	54%	46%
10000	3659	4000	48%	52%

Table 2: Exponential Log-Normal Mixture Truncated Samples

We tested to recover true parameters by fitting Mixture Form. Mixture Form should recover the true parameters including unconditional weight. As expected, Mixture Form recovers the true parameters.

	True Parameters			
	Param 1	**w2**	**Scale 2**	**Shape 2**
	0.000050	40.0%	12.00	0.10

	Parameter Estimation			
	Param 1	**w2**	**Scale 2**	**Shape 2**
0	0.000050	40.1%	12.00	0.10
1000	0.000050	40.1%	12.00	0.10
2000	0.000050	40.3%	12.00	0.10
3000	0.000050	40.1%	12.00	0.10
5000	0.000050	40.1%	12.00	0.10
10000	0.000051	39.8%	12.00	0.10

Table 3: Exponential Log-Normal Mixture Parameter Estimation

Table 4 compares the recovered parameters to original values and provides absolute percentage errors. Mixture Form performs pretty well with percentage absolute errors in the range of 0%-1%.

Parameter Estimation abs% Difference from True Values				
	Param 1	w2	Scale 2	Shape 2
0	1%	0%	0%	1%
1000	0%	0%	0%	1%
2000	0%	1%	0%	1%
3000	1%	0%	0%	1%
5000	0%	0%	0%	1%
10000	1%	0%	0%	1%

Table 4: Exponential Log-Normal Mixture Parameter Estimation % Difference

CASE 2: LOG-NORMAL & LOG-NORMAL MIXTURE DISTRIBUTION

As second example, we used a mixture model with component distributions from the same family with different parameters. We generated a sample from a Log-Normal & Log-Normal mixture distribution with 75% and 25% weights respectively. We defined the parameters so that the sample represents bimodal form as shown in the histogram for the sample below.

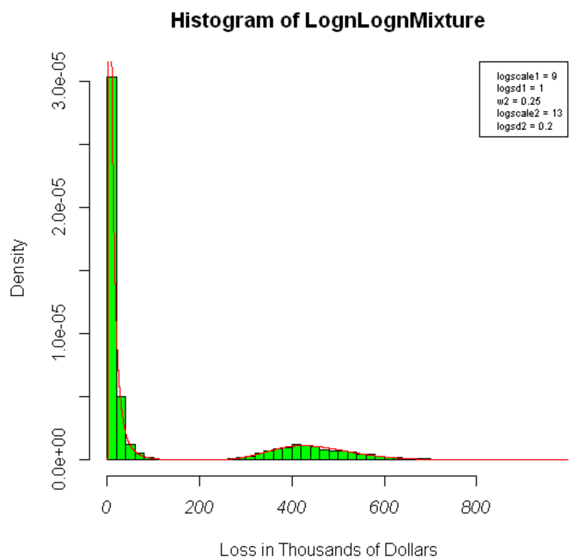

Figure 3: Log-Normal & Log-Normal Mixtures Distribution Probability Density

Based on the parameters, second Log-Normal distribution contributes with much more severe loss amounts and concentrates losses around $500K.

Quantiles for the components and the mixture distribution is provided below. Log-Normal distribution contributes to higher quantiles compared to Exponential. This imposed mixture form is also evident in the histogram.

	min	Qu.10%	Qu.20%	Qu.50%	mean	Qu.75%	Qu.80%	Qu.90%	max
Mixture	***206***	***2,600***	***4,317***	***12,512***	***122,851***	***239,457***	***370,916***	***462,997***	***947,948***
LogNormal	206	2,173	3,453	7,962	13,410	16,094	19,369	30,171	365,959
LogNormal	217,963	338,735	370,766	439,225	451,174	510,365	528,297	576,567	947,948

Table 5: Two Log-normal components

For the truncated samples below, increasing thresholds affects the observed (cond) weights significantly. Higher weight results in higher representation of second Log-Normal. For example the true weight for second Log-Normal was 25%, but at threshold of 10,000 the conditional weight is 45%. For this threshold of 10,000 the representation of components are 55% and 45% respectively compared to 75% and 25% for the original sample.

Threshold	N1	N2	1-w2 (cond)	w2 (cond)
0	7500	2500	75%	25%
1000	7333	2500	75%	25%
2000	6844	2500	73%	27%
3000	6270	2500	71%	29%
5000	5127	2500	67%	33%
10000	3091	2500	55%	45%

Table 6: Log-Normal & Log-Normal Mixture Truncated Samples

The results to recover parameters are provided below. As expected, Mixture Form recovers the true parameters.

	True Parameters				
	Scale 1	Shape 1	w2	Scale 2	Shape 2
	9.00	1.00	25%	13.00	0.20

	Parameter Estimation				
	Scale 1	Shape 1	w2	Scale 2	Shape 2
0	8.99	1.02	25%	13.00	0.20
1000	9.00	1.01	25%	13.00	0.20
2000	9.01	1.01	25%	13.00	0.20
3000	8.99	1.02	25%	13.00	0.20
5000	8.97	1.03	25%	13.00	0.20
10000	9.21	0.94	29%	13.00	0.20

Table 7: Log-Normal & Log-Normal Mixture Parameter Estimation

Table 8, compares the recovered parameters to true values and provides absolute percentage errors. Mixture form performs pretty well with percentage absolute errors in the range of 0%-6%. For truncated sample with higher thresholds, as a result of having less representation for first Log-Normal distribution due to truncation, there seems to be a slight increase in absolute percentage error for first Log-Normal parameters and weights especially at 10,000 threshold.

Parameter Estimation abs% Difference from True Values					
	Scale 1	Shape 1	w2	Scale 2	Shape 2
0	0%	2%	0%	0%	1%
1000	0%	1%	0%	0%	1%
2000	0%	1%	0%	0%	1%
3000	0%	2%	0%	0%	1%
5000	0%	3%	2%	0%	1%
10000	2%	6%	15%	0%	1%

Table 8: Log-Normal & Log-Normal Mixture Parameter Estimation % Difference

CASE 3: WEIBULL & WEIBULL MIXTURE DISTRIBUTION

As final example, we used a mixture model again with component distributions from the same model. This time we tried a Weibull & Weibull mixture distribution with equal weights of 50% but with different parameters. As before, we defined the parameters so that the sample represents bimodal form as much as we can.

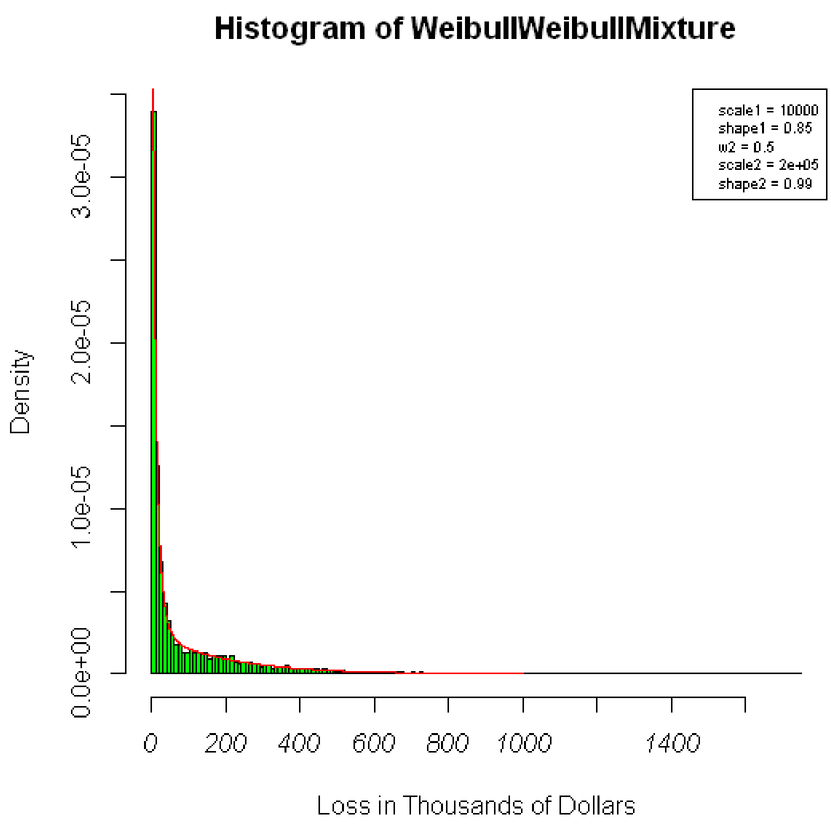

Figure 4: Weibull & Weibull Mixture Distribution Probability Density

Based on the parameters, second Weibull skewed and exhibits a much heavier tail compared to the first one. Quantiles for the components and the mixture distribution is provided below.

	min	Qu.10%	Qu.20%	Qu.50%	mean	Qu.75%	Qu.80%	Qu.90%	max
Mixture	***0***	***1,517***	***4,226***	***24,734***	***108,303***	***142,798***	***188,875***	***333,188***	***1,740,305***
Weibull	0	657	1,597	6,752	11,184	15,048	18,120	27,503	124,266
Weibull	13	19,317	41,399	142,748	205,421	284,065	333,180	476,729	1,740,305

Table 9: Two Weibull Components

For the truncated samples below, increasing thresholds eliminates observation from both components and the truncated samples contain the second Log-Normal more. At threshold of 10,000 the representation for first and second Weibull components are 28% and 72% respectively. Much less representation of first Weibull distribution at higher thresholds could be a challenge for fitting process to recover the parameters.

Threshold	N1	N2	1-w2 (cond)	w2 (cond)
0	5000	5000	50%	50%
1000	4300	4975	46%	54%
2000	3817	4948	44%	56%
3000	3466	4917	41%	59%
5000	2924	4858	38%	62%
10000	1878	4727	28%	72%

Table 10: Weibull & Weibull Mixture Truncated Samples

The results to recover parameters are provided below.

	True Parameters				
	Scale 1	**Shape 1**	**w2**	**Scale 2**	**Shape 2**
	10,000	0.850	50%	200,000	0.990

	Parameter Estimation				
	Scale 1	**Shape 1**	**w2**	**Scale 2**	**Shape 2**
0	10,223	0.832	49.8%	200,000	0.989
1000	10,529	0.844	49.9%	199,999	0.991
2000	10,783	0.885	51.1%	200,000	0.990
3000	10,709	0.883	51.1%	200,000	0.990
5000	10,265	0.847	49.5%	199,999	0.990
10000	10,436	0.846	50.0%	199,999	0.991

Table 11: Weibull & Weibull Mixture Parameter Estimation

Mixture Form recovers the true parameters. The challenge due to limited representation of the observations from first component distribution in the truncated sample does not seem to be issue.

Parameter Estimation abs% Difference from True Values					
	Scale 1	Shape 1	w2	Scale 2	Shape 2
0	2%	2%	0%	0%	0%
1000	5%	1%	0%	0%	0%
2000	8%	4%	2%	0%	0%
3000	7%	4%	2%	0%	0%
5000	3%	0%	1%	0%	0%
10000	4%	1%	0%	0%	0%

Table 12: Weibull & Weibull Mixture Parameter Estimation % Difference

Table 12, compares the recovered parameters to true values and provides absolute percentage errors. Mixture Form performs pretty well with percentage absolute errors in the range of 0%-8%. For truncated samples with higher thresholds, as a result of having less representation for first Weibull distribution due to truncation, there seems to be a slight increase in absolute percentage error.

TESTING EXTENSIONS

In this section we consider different extensions for mixture model distribution fitting. For these extensions we used the Log-Normal & Log-Normal mixture model as an example, since it is a more common Mixture Form in operational risk. The analysis can easily be extended for other mixture distribution examples.

As a first extension, we consider two samples of the mixture model with different thresholds, ie pooling two samples of a mixture distribution. This represents a mixture model assumption when both industry and bank operational data with different thresholds are used.

The second example considers a false assumption, ie mixture model form is imposed for a sample which is generated from a single distribution.

As a third extension, we demonstrate a Monte Carlo simulation using SAS/PROC IML.

EXTENSION 1: TWO TRUNCATED MIXTURE SAMPLES WITH DIFFERENT THRESHOLDS.

We consider two truncated samples of the mixture models with different thresholds. This can represent a case which two different collection thresholds are applied for two samples. It can be due to pooled industry and bank loss data with different thresholds, or due to revised collection threshold overtime for the bank's own operational loss data.

For this example we first generated two samples, ie sample A and B, from Log-Normal & Log-Normal mixture distribution used in previous section with same weight ie weight for second component being 25%. We then truncated these two samples at different thresholds.

Thresholds applied and the resulting weights in sample A, sample B and A+B combined are provided below. The conditional weights range from 30% to 40% for the A+B combined sample.

Threshold A	Threshold B	w2(cond) in A	w2(cond) in B	w2 in A+B
1000	**5000**	25%	32%	30%
3000	**8000**	29%	39%	35%
5000	**10000**	33%	43%	40%

Table 13: Log-Normal & Log-Normal Mixutre Conditional Weight with Two Truncations

Mixture Form performs well to recover true parameters for component distributions. So pooling of truncated samples with different thresholds does not seem to impose any challenge to recover the parameters.

True Parameters				
Scale 1	Shape 1	w2	Scale 2	Shape 2
9.0	1.0	25%	13.0	0.20

Parameters Estimation				
Target Function1				
Scale 1	Shape 1	w2	Scale 2	Shape 2
9.00	1.02	24.7%	12.99	0.20
8.97	1.03	24.2%	12.99	0.20
8.90	1.06	23.0%	12.99	0.20

Parameters Estimation % Difference from True Values				
Target Function1				
Scale 1	Shape 1	w2	Scale 2	Shape 2
0%	2%	1%	0%	2%
0%	3%	3%	0%	2%
1%	6%	8%	0%	2%

Table 14: Log-Normal & Log-Normal Mixture Parameter Estimation and % Difference with Two Truncations

EXTENSION 2: FALSE MIXTURE DISTRIBUTION ASSUMPTION

Even when there is no reason to believe that the underlying data is multimodal, a mixture model can be considered with the aim of testing existence of subpopulations.

In this example, we test distribution fitting with a potential false mixture model assumption. For this purpose we consider a case which a mixture model form is imposed on a sample generated from a single distribution. Again we consider sample from the original Log-Normal & Log-Normal setting before, but the sample being generated from the first component only.

Even if we had a false assumption and impose a mixture form on a sample from a single distribution, the fitting parameters will assign the weight to a single distribution and correct the assumption. For the recovered parameters, fitting performs well to recover true parameters for the distribution and assign the weight to one component only.

	True Parameter				
	Scale 1	Shape 1	w2	Scale 2	Shape 2
	9.00	1.00	0%	-	-

	Parameters Estimation				
	Scale 1	Shape 1	w2	Scale 2	Shape 2
0	8.99	1.01	0.0%	-	-
1000	9.00	1.01	0.0%	-	-
2000	9.01	1.00	0.0%	-	-
3000	8.99	1.01	0.0%	-	-
5000	8.96	1.02	0.0%	-	-
10000	9.12	0.97	0.0%	-	-

	Parameters Estimation Error				
	Scale 1	Shape 1	w2	Scale 2	Shape 2
0	0%	1%	0%	-	-
1000	0%	1%	0%	-	-
2000	0%	0%	0%	-	-
3000	0%	1%	0%	-	-
5000	0%	2%	0%	-	-
10000	1%	3%	0%	-	-

Table 15: Log-Normal & Log-Normal Mixture Parameter Estimation and % Difference with Two Truncations

EXTENSION 3: MONTE CARLO CAPITAL ESTIMATION

In this third extension, we estimate a capital estimate using Mixture Form as the true unconditional parameters. Again we consider fitted parameters from the Log-Normal & Log-Normal mixture model. For frequency, we consider the Poisson distribution as an example.

Threshold	Percentile 99.9%
0	$ 64,056,938
1000	$ 64,191,066
2000	$ 64,245,130
3000	$ 63,930,745
5000	$ 62,984,025
10000	$ 72,141,016

Table 16: Capital Impact from the Different Log-Normal & Log-Normal Mixture Parameter Estimation

We used SAS/IML Software for the following Monte Carlo simulation routine.

```
*       LN Mixture    ;
***************************************************************;
START fLNMix (lambda,param);

ALD=J(&N,2,0);
call randseed(0);
lambda1 = lambda*(1-param[,3]);
lambda2 = lambda*(param[,3]);

do i = 1 to &N;
       call randgen(Freq_ann1,"POISSON",lambda1);
```

```
            if Freq_ann1 = 0 then sev = 0;
        else do;
            norm=J(Freq_ann,1);
        call randgen(norm,"NORMAL",param[,1],param[,2]);
            end;
        sev = exp(norm)/1e6;
        sev1=sev[+,1];

        call randgen(Freq_ann2,"POISSON",lambda2);
            if Freq_ann2 = 0 then sev = 0;
        else do;
            ALD[i,1] = Freq_ann1 + Freq_ann2;
            norm=J(Freq_ann2,1);
        call randgen(norm,"NORMAL",param[,4],param[,5]);
        end;
            sev = exp(norm)/1e6;
                                        ALD[i,2]=sev1+sev[+,1];
end;

Return (ALD);
FINISH fLNMix;
***************************************************************;
```

CONCLUSION

Application of finite mixture densities is most convincing for circumstances where the existence of subpopulations is strongly implied by the nature of the application. Mixture models are valuable flexible tools to accommodate these non-unimodal processes.

In this study we provided guidelines on how to consider mixture distributions for operational risk modeling, with an emphasis on truncated loss data. We demonstrated how mixture distributions can be considered as an alternative flexible distributional form to capture non-uni-modal circumstances. For this purpose, we derived conditional probability density function; we presented results to recover true parameters from truncated data. We also considered three extension tests to demonstrate examples for two truncation levels, false mixture distribution assumption and capital simulation.

Overall, we conclude that mixture models are useful and flexible for use in operational risk modeling. The use of mixture models allows flexibility for situations where the process to be modeled is known or suspected to have subpopulations, ie non-unimodel in nature. In operational risk, this can be observed in loss data both for internal and external due to possible non-stationary and heterogeneity. In these cases, a mixture model will represent the multimodality in operational loss data statistically better than a single distribution form can.

However, we also suggest that mixture models should be taken with a grain of salt. The justification for adopting mixture distributions is critical and should be considered as a last resort due to possible over-fitting. Before considering a mixture form, there should be strong reason to believe in non-stationary and/or heterogeneity as business justification or empirical evidence showing multimodal form. Or mixture distributions should be adopted as a last resort after considering single distribution forms.

REFERENCES:

Alexander, C. (2008). Market Risk Analysis, Quantitative Methods in Finance Volume I, John Wiley & Sons, New York.

Andreas J. (2007). Operational Risk-The Sting is Still in the Tail but the Poison Depends on the Dose, *IMF Working Papers International Monetary Fund.*

Aue, F. and M. Kalkbrenner (2006). LDA at work. Risk Analytics and Instruments, *Risk and Capital Management*, Deutsche Bank, Frankfurt.

Baixauli, J. S.; Alvarez, S. (2004). Analysis of the conditional stock-return distribution under incomplete specification. *European Journal of Operational Research* 155 (2), 276-283.

Basel Committee paper, “Observed range of practice in key elements of Advanced Measurement Approaches (AMA)” October (2006).

Basel Committee paper, “Observed range of practice in key elements of Advanced Measurement Approaches (AMA)” July (2009).

Basel Committee paper, “Principles for the Sound Management of Operational Risk - final document” June (2011).

Cruz, M. (2004), Operational Risk Modelling and Analysis, *Risk Books*, London.

Cruz, M. (2012), Range of practice in operational risk measurement, *Operational Risk & Regulation*.

Dutta, K. and Perry, J. (2006). A Tale of Tails: An Empirical Analysis of Loss Distribution Models for Estimating Operational Risk Capital, *Federal Reserve Bank of Boston*

Everitt, B. S. (1988). A finite mixture model for the clustering of mixed-mode data. *Statistics and Probability Letters*, 6, 305–309.

Everitt, B. S. (1996). An Introduction to finite mixture distribution. *Statistical Methods in Medical Research*, 5, 107-127.

Everitt, B. S. and Hand, D. J. (1981). Finite Mixture Distributions, London: Chapman and Hall.

Figini, S., Giudici, P., and Uberti, P. (2007), A statistical method to optimize the combination of internal and external data in operational risk measurement, *Journal of Operational Risk* 2, 69-78

Frachot, Antoine, Georges, Pierre and Roncalli, Thierry, Loss Distribution Approach for Operational Risk (March 30, 2001).

Frühwirth-Schnatter, S. (2006), Finite Mixture and Markov Switching Models, New York: Springer-Verlag.

Frumento, P. (2009). Finite mixture models. Some computational and theoretical developments with applications. PhD Thesis, University of Florence.

Gardiner JC, Tang X, Luo Z, Ramamoorthi RV. (2012) Fitting heavy-tailed distributions to healthcare utilization by parametric and Bayesian methods. Manuscript in review.

Giacomini, R., Gottschling, A., Haefke, C., and White, H. (2007). *Mixtures of t-distributions for Finance and Forecasting,* Economics Series 216, Institute for Advanced Studies.

Joe, H. and Zhu, R. (2005), Generalized Poisson Distribution: The Property of Mixture of Poisson and Comparison with Negative Binomial Distribution, *Biometrical Journal*, 47, 219–229.

Lawrence, C. J. and Krzanowski, W. J. (1996). Mixture separation for mixed-mode data. *Statistics and Computing*, 6, 85–92.

McLachlan, G.J. & K.E. Basford (1988) Mixture models: Inference and application to clustering. New York: Marcel Dekker

McLachlan, G. and Peel, D. (2000). Finite Mixture Models, John Wiley & Sons, New York.

Pritsker, M. (2006) "The Hidden Dangers of Historical Simulation," The Journal of Banking and Finance (2006): 561-582.

Rempala, G. and Derrig. R., (2005). Modeling Hidden Exposures in Claim Severity via the EM algorithm, *North American Actuarial Journal*, 9(2)

Ripley, B. D. (1994). Neural networks and related methods for classification (with discussion). *J. Roy. Statist. Soc. Ser. B* 56, 409--456.

Roeder, K. (1994). A graphical technique for determining the number of components in a mixture of normals. *J. Amer. Statist. Assoc*. 89, 487-495.

Sattayatham, P. and Talangtam T. (2012). Fitting of Finite Mixture Distributions to Motor Insurance Claims, *Journal of Mathematics and Statistics* 8 (1): 49-56.

Teodorescu, S. (2010). Different approaches to model the loss distribution of a real data set from motor third party liability insurance, *Romanian Journal of Insurance*, 93-104,

Titterington, D.M. (1996). Mixture distributions (update). *In Encyclopedia of Statistical Sciences Update Volume 1. (S. Kotz, C.B. Read and D.L. Banks, Eds.),* 399-407. John Wiley, New York

Uner, S. (2008). SAS Global Forum 2008, "Loss Distribution Approach for the Operational Risk Economic Capital," SAS Proceedings 163-2008

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Sabri Guray Uner:
Union Bank, N.A.
Operational Risk Management Dept. V01-514,
Monterey Park, CA 91755 USA
323-720-5840
Sabri.Uner@UnionBank.com