Using Weighted Distributions to Model Operational Risk

L.B. Afonso, **P. Corte Real**

V Congresso Ibérico de Actuários 6-7 Junho 2016

This work was partially supported by the Fundação para a Ciência e a Tecnologia through the project PEst-OE/MAT/UI0297/2011

Outline

Outline

Main Problem

"Often scientists cannot select sampling units in observational studies with equal probability. Well defined sampling frames often do not exist for human, wildlife, insect, plant, or fish populations. Recorded observations on individuals in these populations are biased and will not have the original distribution unless every observation is given an equal chance of being recorded."

Lyman L. McDonald [1]

[1] The need for teaching weighted distribution theory: Illustrated with applications in environmental statistics (2010)

Main Problem

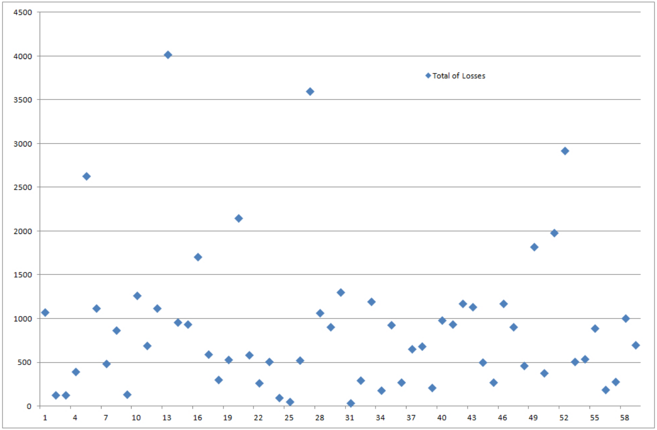

Figure: Total Losses, Exponential Distribution.

Main Problem

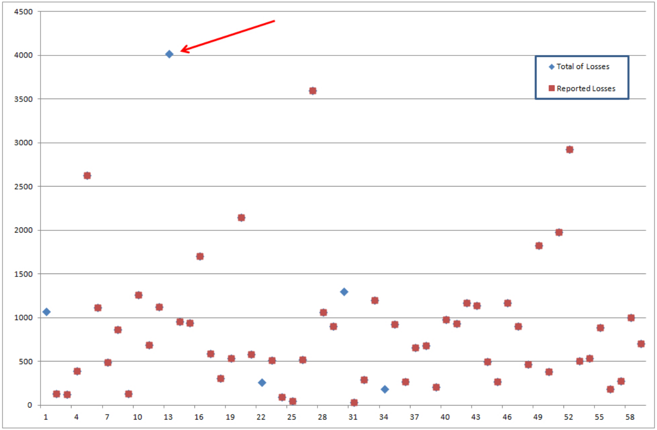

Figure: Reported Losses, Exponential Distribution.

Outline

Motivation
- The Problem

Our Contribution
- Sampling Frame and Sample
- Weighted Distributions
- Some Examples for Chosen Distributions

An Application
- Framework and Methodology

Sampling Frame and Sample

- $S_X = \{X_1, \ldots, X_N\}$ is the random sample of the operational losses occurred over a period
- X_i independent and identically distributed (*i.i.d.*) with $F_X(\cdot)$.
- Not all the observations presented in the original sample S_X, will be available to model operational losses and for statistical inference, namely, parametric estimation.

Sampling Frame and Sample

- $S_Y = \{Y_1, \ldots, Y_M\} \subseteq S_X$, with $M \leq N$ are the observations available for estimation (sample).
- To the unobservable S_X, produced by the original stochastic process, we call it sampling frame.

Sampling Frame and Sample

- Each element in the sampling frame S_X, has probability of inclusion in the sample S_Y, depending on:
  - the quality of the mechanism put in place to filter the sampling frame (ξ);
  - the size of the element, largest elements have bigger probability of inclusion, but none with probability 1, ($F_X(\cdot)$).
- The researcher of operational losses ends up with a biased sample of all the operational losses that should have been reported.
- The bias is originated due to the positive correlation between the loss amount and the probability of being reported.

Outline

Weighted Distributions: Definition

The weighted distribution model is one where the probability of including an observation in the sample is proportional to a weighting function.

Weighted Distributions: Definition

If the relative probability that x will be observed and recorded is given by $w(x) \geq 0$, and $\mathbb{E}(w(X))$ exists, then the *pdf* of the observed data is:

$$f_w(x) = \frac{w(x)}{\int_{\mathbb{R}} w(x) f_X(x) dx} f_X(x) = \frac{w(x)}{\mathbb{E}(w(X))} f_X(x), \text{where } w(x) \geq 0$$

The *pdf* $f_w(x)$ is denominated the weighted *pdf* corresponding to $f_X(x)$.

$$w(x) = x$$

A usual weight function is $w(x) = x$ (proportional to size) so that

$$f_w(x) = \frac{x}{\mathbb{E}(X)} f_X(x)$$

$$w(x) = ?$$

We propose a weight function that do not allow the recording of to many of the smaller losses and at the same time, takes in consideration the original process X and not only the size of the loss.

$w(x) = ?$

- The observations appear in the frame in a given order $\{X_1, \ldots, X_N\}$
- The sample membership indicator, $\mathbb{I}_k$, are independent with $k = 1, \ldots, N$.
- The sampling is made, naturally, without replacement.
- The sample membership indicator are distributed relating to size according to

$$\mathrm{P}(\mathbb{I}_k = 1 \mid X_k = x_k) = F_X^{\xi}(x_k), \ \xi \in [0, +\infty[.$$

$w(x) = ?$

- So, $\mathbb{I}_k \mid X_k$ has a Bernoulli distribution with probability of success $F_X^{\xi}(x)$.
-

$$
\begin{aligned}
P(\mathbb{I}_k = 1) &= \int_{\mathbb{R}} P(\mathbb{I}_k = 1 \mid X = x) P(X = x) dx \\
&= \int_{\mathbb{R}} F_X^{\xi}(x) f_X(x) dx = \mathbb{E}\left(F_X^{\xi}(X)\right) \\
&= \frac{1}{\xi + 1} \left[F_X^{\xi+1}(x)\right]_{-\infty}^{+\infty} = \frac{1}{\xi + 1}.
\end{aligned}
$$

$w(x) = ?$

- $\sharp S_Y = \sum_X \mathbb{I}_k = \sum_{i=1}^N \mathbb{I}_{X_i}$, so we have, $\mathbb{E}(\sharp S_Y \mid N) = N\frac{1}{\xi+1}$.
-

$$
\begin{aligned}
\mathrm{P}(X_j = x \mid \mathbb{I}_j = 1) &= \mathrm{P}(\mathbb{I}_j = 1 \mid X = x)\frac{\mathrm{P}(X = x)}{\mathrm{P}(\mathbb{I}_j = 1)} \\
&= F_X^{\xi}(x)\frac{f_X(x)}{\frac{1}{\xi+1}} \\
&= F_X^{\xi}(x) f_X(x)(\xi + 1).
\end{aligned}
$$

$w(x) = ?$

$$f_w(x) = \frac{F_X^{\xi}(x)}{\frac{1}{\xi+1}} f_X(x) = \frac{w(x)}{\mathbb{E}\left(w(x)\right)} f_X(x)$$

The distribution of the observations in the sample, that is, the distribution of the losses recorded, hence, the distribution of the observations available to the researcher to make inference, is a weighted distribution on $f_X(\cdot)$ with weight function

$$w(x) = F_X^{\xi}(x).$$

ξ

- ξ is as a censorship parameter (or a quality parameter).
- If $\xi = 0$ (a system so effective that all losses end up reported) we would have $\mathrm{P}(\mathbb{I}_k = 1 \mid X_k) = 1$, so that $S_Y = S_X$, and we would be in the usual situation of a random sample from $F_X(\cdot)$.
- When $\xi > 0$, we are in the presence of some degree of censorship in our sample, making more likely that big losses are included in the sample than small losses.

Important relations

$$f_w(x) = F^{\xi}(x) f(x)(\xi + 1)$$

We can write:

$$f(x) = \frac{1}{\xi + 1} F^{-\xi}(x) f_w(x)$$

$$f(x) = \frac{1}{\xi + 1} F_w^{-\frac{\xi}{\xi+1}} f_w(x)$$

and:

$$f_w(x) = (\xi + 1) F_w^{\frac{\xi}{\xi+1}} f(x)$$

Outline

Uniform model

With $Y_j \sim f_w(x)$ Uniform in $]a, b[$, we have:

$$f(x) = \frac{1}{\xi + 1} \left(\frac{x - a}{b - a} \right)^{\frac{-\xi}{\xi+1}} \frac{1}{b - a} \mathbb{I}_{]a,b[}(x)$$

with moments:

$$\begin{aligned} \mathbb{E}(X) &= \frac{b + (\xi + 1)a}{\xi + 2}, \\ \mathbb{V}(X) &= b^2 - \frac{(b - a)(\xi + 1)}{\xi + 2} - \frac{2(b - a)^2(\xi + 1)}{(\xi + 2)(2\xi + 3)} \\ &\quad - \left(\frac{b + (\xi + 1)a}{\xi + 2} \right)^2. \end{aligned}$$

These results are easily obtained integrating by parts.

Exponential model

With $Y_j \sim f_w(x)$ the Exponential distribution with $\mathbb{E}(X) = \lambda + \beta$ and $\mathbb{V}(X) = \beta^2$, we have:

$$f(x) = \frac{1}{\xi+1}\left(1 - e^{-\frac{x-\lambda}{\beta}}\right)^{-\frac{\xi}{\xi+1}} \frac{1}{\beta} e^{-\frac{x-\lambda}{\beta}} \mathbb{I}_{]\lambda,+\infty[}(x)$$

with moments:

$$\begin{aligned} \mathbb{E}(X) &= \lambda + \beta H_{\frac{1}{\xi+1}}, \\ \mathbb{V}(X) &= \beta^2 \left(\frac{\pi^2}{6} - \psi'\left(\frac{\xi+2}{\xi+1}\right)\right). \end{aligned}$$

Pareto (Type I) model

With $Y_j \sim f_w(x)$ the Pareto (α, β), we have:

$$f(x) = \frac{1}{\xi+1}\left(1-\left(\frac{\beta}{x}\right)^{\alpha}\right)^{\frac{-\xi}{\xi+1}} \frac{\alpha}{x}\left(\frac{\beta}{x}\right)^{\alpha} \mathbb{I}_{]\beta,\infty[}(x)$$

with moments:

$$\begin{aligned}
\mathbb{E}(X) &= \frac{\beta}{\xi+1} B\left(1-\frac{1}{\alpha}, \frac{1}{\xi+1}\right), \\
\mathbb{V}(X) &= \frac{\beta^2}{\xi+1} B\left(1-\frac{2}{\alpha}, \frac{1}{\xi+1}\right) - \\
& \quad \frac{\beta^2}{(\xi+1)^2}\left(B\left(1-\frac{1}{\alpha}, \frac{1}{\xi+1}\right)\right)^2.
\end{aligned}$$

Outline

Framework

- We consider that the reported losses S_Y, have a known distribution (Exponential, Pareto, Lognormal or Weibull),
- Sample Description of Table 6.3 of [Chernobai et al., 2007] based on the original data set available from [Cruz, 2002] that give us a mean value of $439.725, 99$\$ and standard deviation of $538.403, 93$\$ (in dollars),
- The Risk Department estimated a probability of $1/250$ for an operation to generate a operational loss and of 85% for the loss ending up reported and documented $\xi = (1 - 85\%)/85\% = 17,647\%$,
- So that, for every $2.941.176$ transactions made, we expect 11.765 operations originating a loss and from these, that 10.000 end up reported.

Methodology

- For each density function $f_w(x)$ for the reported losses S_Y, we generated 1.000 samples of occurred losses S_X, each with 11.765 losses with density function $f(x)$,
- For each sample, we selected the reported losses S_Y, according to our weight function $F_X^{\xi}(x_k)$, so that, the data available to make inference and take decisions have distribution $F_w(x)$. (The observed average sample size of the 1.000 reported losses is 10.020,53)

Table: Results for Exponential and Pareto examples

Occurred X	Reported Y	Exponential X	Exponential Y	Pareto X	Pareto Y
Theorical					
$E(X)$	$E(Y)$	395.056,71	439.725,99	417.366,59	439.725,19
$\sigma(X)$	$\sigma(Y)$	424.803,42	439.725,99	499.950,66	538.399,43
$V@R_{99\%}(X)$	$V@R_{99\%}(Y)$	1.953.938,46	2.025.013,02	1.723.425,99	1.849.400,38
$V@R_{99,9\%}(X)$	$V@R_{99,9\%}(Y)$	2.966.094,54	3.037.519,53	4.706.580,84	5.052.366,95
$E\left(\sum_{i=1}^{N} X_i\right)$	$E\left(\sum_{i=1}^{M} Y_i\right)$	4.647.726.036	4.397.259.900	4.910.195.194	4.397.251.910
Empirical					
$\widehat{E}(X)$	$\widehat{E}(Y)$	394.967,87	439.309,36	417.237,35	439.358,53
$\widehat{\sigma}(X)$	$\widehat{\sigma}(Y)$	424.566,34	439.141,89	432.830,30	464.561,01
$TV@R_{99\%}(X)$	$TV@R_{99\%}(Y)$	2.395.587,73	2.467.134,25	3.061.988,15	3.286.507,23
$TV@R_{99,9\%}(X)$	$TV@R_{99,9\%}(Y)$	3.408.439,46	3.479.634,14	8.316.902,89	8.912.776,77

Table: Results for Lognormal and Weibull examples

Occurred X	Reported Y	Lognormal X	Y	Weibull X	
		Theorical			
$E(X)$	$E(Y)$	396.844,27	439.714,09	391.177,00	439.725,99
$\sigma(X)$	$\sigma(Y)$	510.258,76	538.389,17	514.422,01	538.403,93
$V@R_{99\%}(X)$	$V@R_{99\%}(Y)$	2.430.724,95	2.577.549,87	2.427.557,18	2.535.410,19
$V@R_{99,9\%}(X)$	$V@R_{99,9\%}(Y)$	5.111.227,25	5.354.412,53	4.033.829,50	4.152.317,56
$E\left(\sum_{i=1}^{N} X_i\right)$	$E\left(\sum_{i=1}^{M} Y_i\right)$	4.668.756.082	4.397.140.937	4.602.082.353	4.397.259.900
		Empirical			
$\widehat{E}(X)$	$\widehat{E}(Y)$	396.697,70	439.210,72	391.046,61	439.173,74
$\widehat{\sigma}(X)$	$\widehat{\sigma}(Y)$	509.877,14	537.529,65	514.103,86	537.646,01
$TV@R_{99\%}(X)$	$TV@R_{99\%}(Y)$	3.576.296,68	3.765.141,86	3.123.561,05	3.236.677,66
$TV@R_{99,9\%}(X)$	$TV@R_{99,9\%}(Y)$	6.888.269,49	7.367.649,69	4.787.108,09	4.908.582,01

	X				
	Exponential	Pareto	Lognormal	Weibull	*Y*
K-S p-values					
min	0	0	3,9212E-02	1,5616E-04	3,4932E-04
average	4,0669E-04	4,0265E-04	6,7963E-01	1,1667E-02	5,0641E-01
max	2,3191E-01	2,2921E-01	9,9999E-01	6,4403E-02	9,9952E-01
# not reject	0	0	999	5	955
Chi2 p-values					
min	0	0	2,5390E-04	4,7621E-11	6,0648E-05
average	4,0529E-23	1,1096E-24	4,8810E-01	7,4905E-03	5,0170E-01
max	2,8235E-21	8,0041E-22	9,9846E-01	2,4295E-01	9,9934E-01
# not reject	0	0	948	44	949

Table: p-values for Kolmogorov-Smirnov and Chi2 tests

For Further Reading I

R. A. Fisher.
The effects of methods of ascertainment upon the estimation of frequencies.
Annals Eugenics, (6):13–25, 1934.

G. P. Patil and C. R. Rao.
Weighted distributions and size-biased sampling with applications to wildlife populations and human families.
Biometrics, 34(2):179–189, June 1978.

G.P. Patil and C.R. Rao.
Weighted distributions: A survey of their application.
Applications of Statistics, pages 383–405, 1977.

C.R. Rao.
On discrete distributions arising out of methods of ascertainment.
Sankhyã: The Indian Journal of Statistics, Series A, pages 311–324, 1965.

Cruz, M. G. (2002).
Modeling, Measuring and Hedging Operational Risk.
John Wiley & Sons.

Chernobai, A., Rachev, S. T., and Fabozzi, F. (2007).
Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis.
John Wiley & Sons.

Afonso, Lourdes B. and Corte Real, Pedro.
Using weighted distributions to model operational risk.
ASTIN Bulletin: The Journal of the International Actuarial Association, pages 469-485, 2016.