

Distribution analysis of the losses due to credit risk

Kamil Łyko¹

Abstract

The main purpose of this article is credit risk analysis by analyzing the distribution of losses on retail loans portfolio. This type of analysis is very important from a practical point of view, because a one of real credit risk management tools is economic capital determination. The appropriate level of economic capital is necessary to cover risks incurred by the bank. The correct determination of economic capital appropriate level depends on the goodness of it's fitting to the distribution. The distribution fitting study to empirical data will be conducted using the following tools: graphical analysis, testing empirical distribution fitting quality with goodness of fitting test to the theoretical distribution. An additional objective of the study is estimation of risks arising from the adoption of an incorrect distribution of credit risk losses. Based on the assessment's results the impact on the level of economic capital due to the wrong choice of a distribution will be determined.

Keywords: economic capital, Goodness-of-fit statistics, distribution, loss, credit risk, χ^2 Pearson's test, Anderson's-Darling's test

JEL Classification: C10

AMS Classification: 91G40

1. Literature review

A credit loss distribution fitting in credit risk management process, despite its importance isn't widely discussed in the literature. One of the few publications entirely devoted to the statistical distribution of the losses is the book written by Hogg and Klugman [3]. The authors discuss in it inter alia: the loss random variable models, testing the theoretical and empirical data distribution fitting and loss distributions modeling. Issues of credit loss distribution fitting are also raised by Andrade and Sicsú [3]. They indicate the fact, that the most commonly forms of credit loss distribution are: beta, gamma, log-normal, chi-square, weibull, exponential, log-logistic, rayleigh, johnson and extreme values distributions. The choice of these distributions is also indicated by Bluhm, Overbeck, Wagner [5] and by Gapko, Šmíd [10], and also by JP Morgan and Credit Suisse banks in their technical specifications for "CreditMetrics" [4] and "CreditRisk+" [7]. In the publications of David Vose [15] and Andrade, Sicsú [3] we can find explanations of the methodology of choosing the best fitting

¹ Wrocław University of Economics, Department of Financial Investments and Risk Management, Komandorska 118/120, 53-345 Wrocław, kamil.lyko@ue.wroc.pl

of the distribution. These authors point out two tests of fitting goodness: χ^2 Pearson's and Anderson's-Darling's.

2. Introduction to Credit Risk measurement

According to Jajuga [13] "credit risk means the possibility of default by one of the parties of the contract, which means that the party exposed to the risk will not receive specified payment value in the expected due date defined by the contract". Experience shows that even a good customer may have a tendency to default of the credit agreement due to the deterioration of financial situation. In banking, the costs of such borrower's are covered by other customers, from which the bank charges an additional fee, which is called the risk premium. Sum of these charges is collected on an internal bank account and makes capital reserve for covering costs from expected losses. In the probability theory attribute "expected" means expected or average value and in credit risk management process it is the same.

Bank calculates for each of its customers the value of probability of not paying their liabilities defined as PD (*Probability of Default*) and LGD factor (*Loss Given Default*) which inform what part of credit exposure EAD (*Exposure at Default*) will not be recovered in case of default. To determine the value of the losses we proceed according to the following formula:

$$L = EAD \cdot LGD \cdot LI, \quad (1)$$

where: L means loss value; LI is a zero-one variable, which takes a value of 1 in case of default and 0 when the terms of the agreement are met. The Loss in a case of default is determined by percentage of the exposure at risk, therefore LGD takes values in a range from 0 to 1. It is expected value of random variable which is called *severity* (SEV), that is loss in case of default:

$$LGD = EL(SEV). \quad (2)$$

PD is expected value of variable LI :

$$PD = E(LI) = 1 \cdot DP + 0 \cdot (1 - DP). \quad (3)$$

Assuming the independence of *SEV* and *LI* variables, and also taking into account the formulas (2) and (3) we obtain the formula for the expected value of a random variable which is the loss *L*:

$$EL = E(L) = EAD \cdot LGD \cdot PD. \quad (4)$$

Another important measure is the unexpected loss value given by the formula:

$$UL = EAD \sqrt{V(SEV) \cdot PD + LGD^2 \cdot PD(1 - PD)}, \quad (5)$$

where: *V* means variance [14].

3. Distribution of credit losses and economic capital

All measures of risk at the level of the credit portfolio are based on a loss variable. Not surprising therefore is the fact that the distribution of that loss in terms of credit portfolio risk management is crucial. In the Figure 1 it is shown that all credit risk measures can be identified on the chart of credit loss distribution.

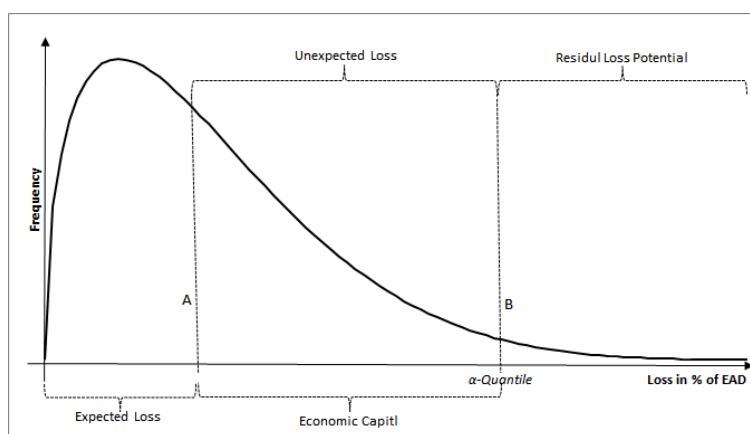


Fig. 1. Credit loss distribution.

By designating the type and parameters of distribution, it is possible to directly estimate credit risk measures. The chart is splitted into three parts: indicated by the letter A is the expected loss, and if we know distribution's formula and parameters, it is possible to mark high quantile (B). Using this quantile and expected loss it is possible to calculate unexpected loss value which is equal to the economic capital. The last part of Figure 1 – residual loss potential - are catastrophic losses occurring extremely rarely [2, 6].

Figure 1 shows that economic capital is the difference between the quantile of the loss distribution and the expected loss:

$$EC(\alpha) = q_\alpha - EL_p, \quad (6)$$

where: EC – economic capital; q_α - loss distribution quantile; EL_p - portfolio's expected loss. Economic capital value depends therefore on the parameter α , which usually takes values close to one for example 0.99. To estimate distribution quantile, it is necessary to know distribution of the losses. One of the possibilities is fitting theoretical distribution, which corresponds to the shape of the empirical distribution [14]. According to literature [5] those distributions characterizes by fat tails and they are right-slanted (see Figure 1). Additionally a distribution should be contained in a finite interval and determined by two first moments. So determined economic capital lets us recover losses exceeding expected loss.

4. Theoretical loss distribution fit

Theoretical loss distribution fitting to empirical data will be conducted by comparison using Quantile to Quantile charts (*Q-Q plot*) and Goodness-of-fit statistics. In case of Quantile to Quantile charts on the horizontal axis there are theoretical quantiles of distribution of the studied variable and on the vertical axis there are quantiles of distribution to be compared to. If studied variable distribution fits perfectly to the compared one, the quantiles create a line over 45° angle. Those charts are often used to identify outliers [11].

The second stage of verification will be studied using statistical tests of goodness of theoretical distribution and data fitting. We can therefore hypothesize that the variable follows a distribution and on the basis of randomly selected sample check whether it is necessary to reject or adopt this hypothesis. These type of tests create a general group, because they do not apply to individual parameters but to a whole distribution function. These tests are called Goodness-of-fit statistics. The test's structure requires a certain measure of distribution distance. Generally used measure is based on a cumulative distribution comparison $F_n(x)$ and theoretical distribution $F(x)$ as in the formula below [9, 12]:

$$D_n = \sup_{x \in R} |F_n(x) - F(x)|. \quad (7)$$

One of those test is χ^2 Pearson's test. This test belongs to a nonparametric tests group and its algorithm compares frequency of empirical events with theoretical. χ^2 Pearson's test is

applicable in cases of sufficiently large number of samples (over 30 observation). The distance between distributions can be calculated as follows:

$$\chi^2 = \sum_{i=1}^r \frac{(n_i - np_i)^2}{np_i}, \quad (8)$$

where: r – number of class; n_i – empirical number of class; p_i – theoretical density [8]. The next test used in the study will be Anderson's-Darling's test, which can be calculated as [1, 2]:

$$A = \frac{-\sum_{i=1}^n (2i - 1) \{ \ln F(x_{(i)}) + \ln(1 - F(x_{n+1-i})) \}}{n} - n, \quad (9)$$

where: $F(x)$ – empirical cumulative distribution. This test is more sensitive for distribution tail shape comparing to others [15].

5. Results

Based on the sample of 2000 credit loss observations, parameters of 6 theoretical distributions, such as: beta, gamma, exponential, log-normal, chi-square and weibull's were estimated. Distribution parameters were estimated by maximizing the likelihood function. The losses considered in these study are presented in percentage terms where 1 means maximum loss which is equal to credit exposure.

In the figure 2 we can compare credit loss distribution shape to theoretical distribution. On this basis we can conclude, that chi-square distribution least of all describes the tested sample. Next, to define goodness of theoretical and sample distributions fitting, the Quantile to Quantile charts method can be used.

On the basis of Figure 3, in case of log-normal and chi-square distributions we can't conclude that distribution's fitting to empirical data is correct. As we can see - quantiles of those distributions significantly differ from empirical quantiles. The best fitting into 45° line across the chart presents the weibull's distribution. But on this basis we aren't able to reject other distributions i.e.: beta, gamma and exponential distributions.

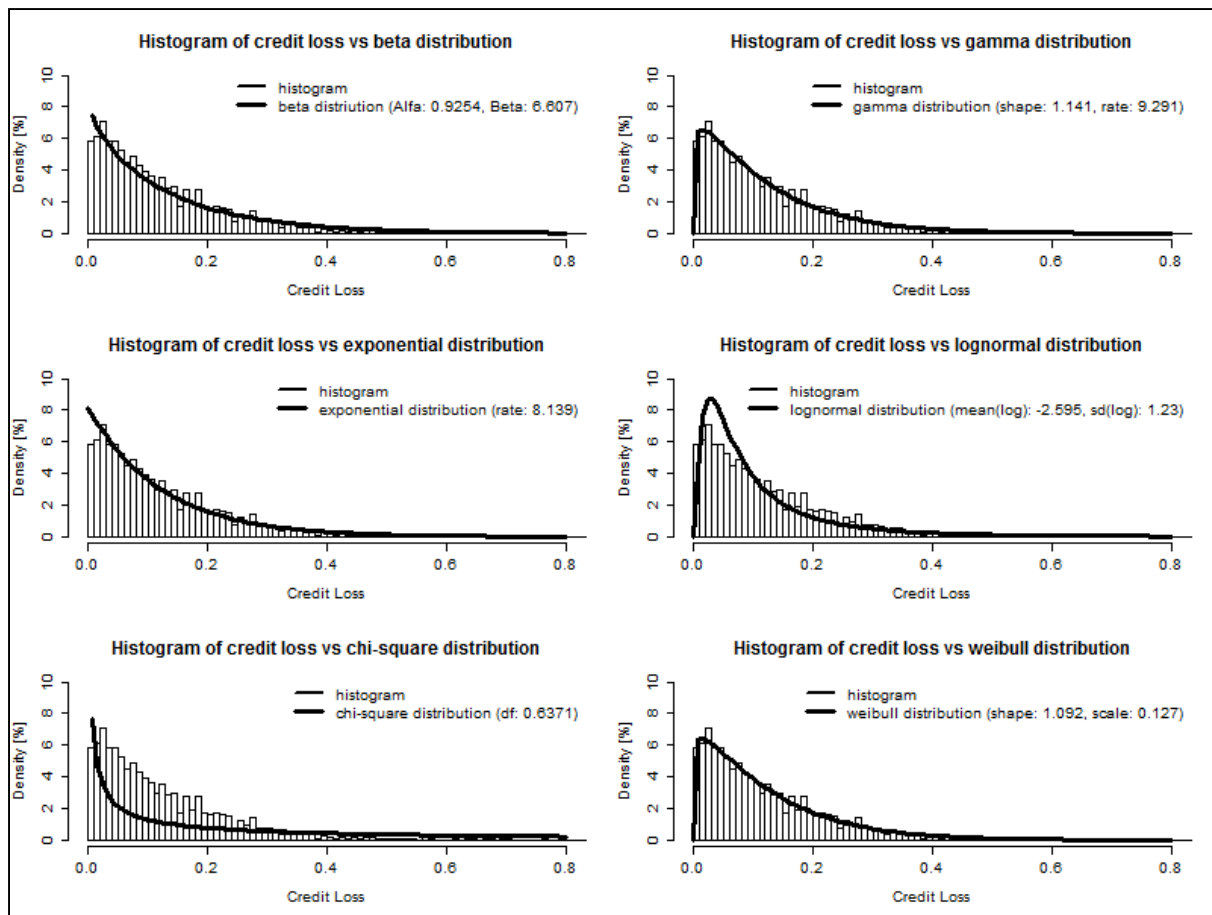


Fig. 2. Loss empirical distributions vs. theoretical distributions.

Distribution	test p-value	
	χ^2 Pearson	Anderson-Darling
Beta	0.0000	0.2607
Gamma	0.5164	0.4758
Log-normal	0.0000	0.0000
Exponential	0.4979	0.0255
Chi-square	0.0000	0.0000
Weibull	0.6025	0.4780

Table 1 Results of goodness of fit tests.

To conclude with certainty that the best fitting is represented by the Weibull’s distribution - compatibility tests were conducted: χ^2 Pearson’s and the Anderson’s-Darling’s. Results are presented in the Table 1. Both test χ^2 Pearson’s and Anderson’s-Darling’s identifies Weibull’s distribution as the best fitted to empirical data. However, it should be noted, that according to A-D test also gamma and beta distributions are well fitted, and in addition

χ^2 Pearson's test indicates an exponential distribution. Next, on this basis high distribution quantiles were estimated and we calculated economic capital value, which is necessary to cover unexpected losses.

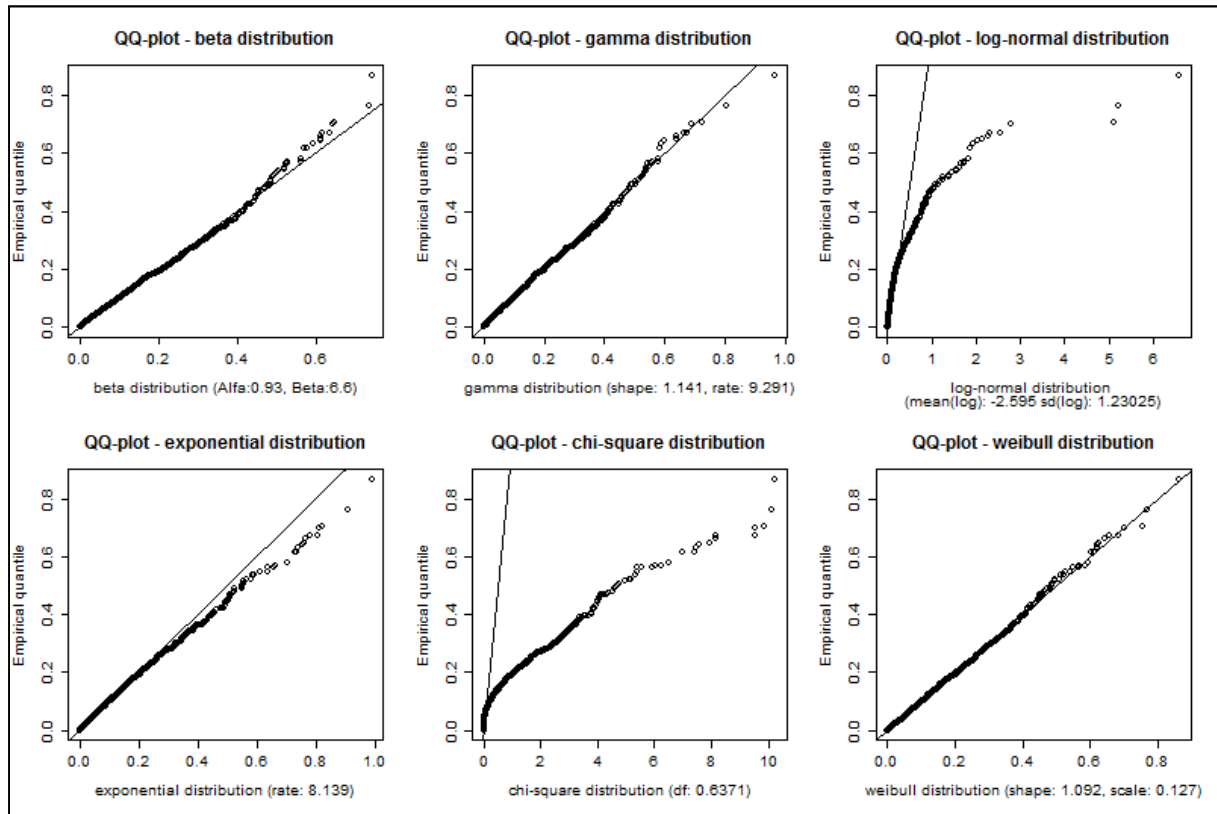


Fig. 3. Distributions Quantile to Quantile charts.

In the Table 2 we can see these calculation's results and also economic capital value comparison for each distributions. Weibull's distribution as the best fitted is a base to compare. Beta and gamma distributions as also well fitted gave results slightly different from Weibull's distribution. Less fitted or unfitting distributions gave strongly overestimated results. The economic capital calculated based on a high quantile exponential distribution is higher about 1.5 times, in cases of log-normal distribution and chi-square we are talking about several times exceeding. In addition to the above table for comparative purposes the capital calculated based on a normal distribution was added. We obtained results indicating that the expected loss is higher than the losses distribution quantile, which gave an absurd result of negative economic capital.

Distribution	Level of confidence α	α -Quantile	Economic Capital	Difference in capital (vs. weibull distribution)
Beta	0.99	0.51	5.61	-4.21
Gamma	0.99	0.54	13.71	3.90
Log-normal	0.99	1.00	126.81	116.99
Exponential	0.99	0.59	25.56	15.75
Chi-square	0.99	1.00	126.81	116.99
Weibull	0.99	0.52	9.82	0.00
Normal	0.99	0.38	-26.32	-36.13

Table 2 Economic capital for current distributions.

6. Conclusions

Based on the studies we can see how important is the proper selection of a credit loss distribution for a correct calculation of economic capital. If the bank calculated the capital on an incorrect distribution it would be exposed to the additional costs associated with maintaining excessive financial reserves to cover unexpected losses or in case of underestimation of reserves in the unexpected situation funds might be insufficient. In extreme situations irresponsible credit risk management which used an incorrect distribution to estimate the economic capital may lead to the bankruptcy of the bank due to a thin capitalization. Often, in cases of large credit portfolios, application of different types of distributions for different portfolio's parts is necessary. This prevents situations of theoretical distribution's incorrect fitting for heterogeneous portfolios. Thus, the risk of errors resulting from this type is minimized.

References

- [1] Anderson, T.W., 2010. Anderson-Darling Tests of Goodness-of-Fit. Stanford: Stanford University.
- [2] Anderson, T.W., Darling, D.A., 1954. A Test of Goodness of Fit. Journal of the American Statistical Association, December, 49(268), 765-769.
- [3] Andrade, F.W.M., Sicsú, A.L., 2008. A Credit Risk Model for Consumer Loan Portfolios. Latin American Business Review, 11 October, 8(3), 75-91.
- [4] Bhatia, M., Finger, C.C., Gupton, G.M., 1997. CreditMetrics™ - Technical Document. New York: J.P. Morgan & Co.
- [5] Bluhm, C., Overbeck, L., Wagner, C., 2010. Introduction to Credit Risk Modeling. Second Edition. Munich: CRC Press Taylor & Francis Group.
- [6] Chorafas, D.N., 2007. Stress Testing for Risk Control under Basel II. First edition. Oxford: Elsevier Ltd.

- [7] Credit Suisse First Boston International, 1997. CreditRisk+ A Credit Risk Management Framework. Boston: Credit Suisse First Boston International.
- [8] Domański, C., 1979. Statystyczne testy nieparametryczne. Warszawa: Państwowe Wydawnictwo Ekonomiczne.
- [9] Domański, C., 1990. Testy statystyczne. Warszawa: Państwowe Wydawnictwo Ekonomiczne.
- [10] Gapko, P., Šmíd, M., 2010. Modeling a Distribution of Mortgage Credit Losses. Prague.
- [11] Górecki, T., 2011. Podstawy statystyki z przykładami w R. Legionowo: Wydawnictwo BTC.
- [12] Hogg, R.V., Klugman, S.A., 1984. Loss distributions. John Wiley & Sons, Inc.
- [13] Jajuga, K., 2004. O systematyzacji modeli ryzyka kredytowego. In: Appenzeller, D., Upadłość przedsiębiorstw w Polsce w latach 1990-2003. Teoria i praktyka. Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.
- [14] Jajuga, K., 2007. Zarządzanie ryzykiem. Warszawa: Wydawnictwo Naukowe PWN.
- [15] Vose, D., 2008. Risk Analysis: a quantitative guide, third edition. John Wiley & Sons, Ltd.