УДК 330.4; 519.23/ 519.24

EXACT MAXIMUM LIKELIHOOD ESTIMATOR FOR THE PROBABILITY OF DEFAULT ON ESTIMATION PROVISION CONSUMER CREDIT PORTFOLIO OF THE BANK

ТОЧНАЯ ОЦЕНКА МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ ДЛЯ ВЕРОЯТНОСТИ ДЕФОЛТА ПРИ ОЦЕНИВАНИИ РЕЗЕРВОВ ПОТРЕБИТЕЛЬСКОГО КРЕДИТНОГО ПОРТФЕЛЯ БАНКА

©Лёвин В. В.
канд. физ.–мат. наук
Московский государственный технический университет
им. Н. Э. Баумана (национальный исследовательский университет)
г. Москва, Россия, vladimir.levin.51@mail.ru
©Levin V.
Ph.D., Bauman Moscow State Technical University
Moscow, Russia, vladimir.levin.51@mail.ru
©Хонов С. А.
канд. физ.–мат. наук
Высшая школа экономики
(национальный исследовательский университет)
г. Москва, Россия, sergey.khonov@yandex.ru
©Khonov S.
Ph.D., Higher School of Economics
Moscow, Russia, sergey.khonov@yandex.ru

Abstract. In the context of increasing competition in the banking market, increasing regulatory requirements for transparency and sound risk–creation on this basis of adequate risk provisions in the banking sector is of paramount importance. In this paper, firstly it is proposed to use for estimating credit risks the exact maximum likelihood estimators (MLE) of the structure of stratified population for any sizes of the credit portfolio. These exact MLE could be applied to estimate Basel-II risk parameter PD (Probability of Default), and could be used to optimize provisions for covering expected losses of consumer credit portfolio.

In usual banking practice for estimating risk parameter PD the frequencies (rates) of default credits of the whole consumer portfolio or of sub–portfolios of the whole consumer portfolio are usually using. But the statistical characteristics of these estimates, such as unbiased property, consistency, efficiency, exact and asymptotic distributions, usually are unknown. The new statistical estimations have derived for characteristics used in vintage analysis of consumer credit portfolio. These estimations for delinquency rates with different DPD (Days Past Due) are the exact maximum likelihood estimators (MLE) of the structure of stratified population for any sizes of the credit portfolio. These exact MLE could be applied to estimate Basel-II risk parameter PD (Probability of Default), and could be used to optimize provisions for covering expected losses of consumer credit portfolio. Making the adequate provisions to credit risks in the crisis conditions is the problem which needs to estimate risks with satisfactory accuracy.

Аннотация. В условиях роста конкуренции на рынке банковских услуг, повышения требований регулирующих органов по прозрачному и обоснованному учету рисков, создание на этой основе адекватных рискам резервов в банковском секторе приобретает первостепенное значение. В данной статье впервые предлагается использовать при оценке кредитных рисков портфеля розничных кредитов новые точные статистические оценки максимального правдоподобия (MLE) оценивания вероятности дефолта PD (Probability of

Default), параметра риска, определенного в рекомендациях Базельского комитета по банковскому надзору (Basel II). В обычной банковской практике для оценивания PD используются оценки на основе частот (долей) дефолтных кредитов во всем портфеле розничных кредитов или в суб–портфелях основного портфеля. При этом статистические свойства этих оценок, такие как несмещенность, состоятельность, эффективность, точные и асимптотические распределения и др., чаще всего неизвестны. Создание адекватных рискам резервов в условиях кризиса — задача, требующая статистических оценок параметров риска с известной точностью, обладающих оптимальными свойствами. Создание чрезмерных резервов ведет к сокращению активного капитала и сокращению прибыли, недостаточность резервов несет повышенный риск банкротства. В настоящей статье впервые предлагается использовать в классическом анализе портфеля потребительского кредитования новые точные статистические оценки максимального правдоподобия (MLE) для оценивания вероятности дефолта PD (Probability of Default), содержащейся в рекомендациях Базельского комитета по банковскому надзору. Использование предлагаемых в статье оценок риска открывает возможность получения адекватных оценок риска и резервов, соответствующих этим уровням риска.

Keywords: statistical estimation, exact maximum likelihood estimator, Bazel II, Bank for International Settlements, BIS, Banque des règlements internationaux, BRI, Basel Accords, recommendations on banking laws and regulations, Basel Committee on Banking Supervision, Probability of Default, consumer credit portfolio.

Ключевые слова: статистическое оценивание, точные оценки, максимальное правдоподобие, Базель-II, Банк международных расчетов, Базельское соглашение, рекомендации по банковскому законодательству и правилам, Базельский комитет по банковскому надзору, вероятность дефолта, потребительский кредитный портфель.

1. Risk parameters in Basel II Internal Rating Based (IRB) Approach

Basel II process has greatly increased the sophistication and profile of credit risk measurement within financial institutions. In accordance with Basel II requirements (see source 1) banks must calculate reserves for possible credit portfolio losses by the following formula (1):

$$Reserves = EAD * PD * LGD \tag{1}$$

where

EAD — the Exposure at Default, debt that should be to repay by credit obligation;

PD — Probability of Default;

LGD (Loss Given at default) — rate of non–payment of funds by credit when default occurs.

Under Basel II (see source 2, p. 30), a default event on a debt obligation is said to have occurred if:

– it is unlikely that the obligor will be able to repay its debt to the bank without giving up any pledged collateral

– the obligor is more than 90 days past due on a material credit obligation

There are challenges still exist in the development of credit models with these risk parametres, and particularly in the calculation of probability of default (PD).

The probability of default (PD) is an estimate of the likelihood that the default event will occur. It applies to a particular assessment horizon, usually one year. To get the PD estimate with good characteristics, different methods of segmentation and pooling of the credit portfolio have been used to get homogenious data for calculation of PD. Vintage analysis of the consumer credit portfolio is one of the methods of segmentation and pooling of the credit portfolio, but one of the most important.

2. Vintage analysis of the consumer credit portfolio

The term “vintage” had taken directly from the world of wine. For many years wine experts have been creating vintage tables, from which one can read a note determining the quality of a given wine from a particular year. Based on vintage table, it could be to know, whether a given wine should be stored longer in order to get the optimum taste, or if it should be drunk, or what is worse — if it should have been drunk much earlier. It is easy to notice analogy between the variable quality of the wine from a given year and variable in time quality of credit portfolio built by the bank in a given year. It turns out that loan production performed in a given time can be successfully described with the use of vintage tables.

The primary aim of vintage analysis is the presentation of the credit risk development of a given portfolio in order to enable tracking its trend of development and its further anticipation. Vintage analysis allows obtaining valuable information for:

– comparison of risk level in particular months/quarters/years,
– analysis of the influence of particular characteristic's value on the credit risk,
– analysis of the influence of the internal risk policy changes on the portfolio risk,
– forecasting the risk level in the future,
– current monitoring of the portfolio risk level.

Therefore, vintage analysis is one of the basic analyses used for measuring a risk in the process of managing it. Vintage analysis could be considered in two variants: valuable variant, when risk indicators are based on the current account rests of granted loans, and quantitative variant, when values of outstanding capital are replaced with numbers of granted loans. In this paper, we will consider usage of quantitative variant of vintage analysis, the usage of the valuable variant will be consider in a forthcoming article.

3. Vintage representation of credit portfolio

3.1. Notations

We will use of the following notations.

τ — the current time moment (in practice the last day of the calendar month is chosen usually, but the quarter or the year last day could be chosen);

Let $0 < t_1 < \cdots < t_i < \cdots < t_{\mathrm{M}}$ are the given calendar date, here the month’s last days are considered.

Indicator of risk $\boldsymbol{IR_j}$ based on the following grouping (j=0,1,2,3…,15) of the number of days of delays (DAYS PAST DUE=DPD):

0.0 days
1. from 1 up to 30 days;
2. from 31 up to 60 days;
3. from 61 up to 90 days;
4. from 91 up to 120 days;

…………………………..

11. from 301 up to 330 days;
12. from 331 up to 365 days;
13. above 365 days.

For a full reflection of the risk level, the next two values of the risk feature are also used:

14. “Repaid loans”— Number of loans with fully repaid principal.
15. “Defaults (lost loans)” — Number of loans that have not been repaid completely.

Risk Classes $\boldsymbol{RC_j}$ — sets of loans with the same $\boldsymbol{IR_j}$, j=0,1,2,3…,15

$\boldsymbol{V_i}$ ***is the vintage*** = set of loans, opened during time period $[t_{i-1}, t_i], i = 1, \dots, M$.

$\boldsymbol{V_i(\tau)}$ ***is a set of loans from vintage*** $\boldsymbol{V_i}$, which dates of closing of the credit agreement are later than τ, $i = 1, \dots, M$.

T is the credit term (in months), T=6, 12, 18, 24, … ,180.

$\boldsymbol{V_i(T) = V_i(t_i, T)}$ ***is a subvintage loans of*** $\boldsymbol{V_i}$, with the same credit term T.

$\boldsymbol{V_i(\tau, T)}$ ***is a subvintage loans of*** $\boldsymbol{V_i}$, with the same credit term T at the moment $\boldsymbol{\tau}$.

It is clear that
$V_i(\boldsymbol{\tau}) \cap V_j(\boldsymbol{\tau}) = \emptyset,\ i \neq j, V_i(\boldsymbol{\tau}) = \emptyset, if\ \boldsymbol{\tau} < t_i, i = 1, \dots, M.$
$CP(\boldsymbol{\tau}) = \bigcup_{i=1}^{M} V_i(\boldsymbol{\tau})$ is a credit portfolio (CP) at moment $\boldsymbol{\tau}$.
Also, for any T
$V_i(\boldsymbol{\tau}, T) \cap V_j(\boldsymbol{\tau}, T) = \emptyset, i \neq j, V_i(\boldsymbol{\tau}, \boldsymbol{T}) = \emptyset, if\ \boldsymbol{\tau} < t_i, i = 1, \dots, M.$
$\boldsymbol{V_i(\tau)} = \bigcup_{T=1}^{180} V_i(\boldsymbol{\tau}, \boldsymbol{T}).$
$CP(\boldsymbol{\tau}, \boldsymbol{T}) = \bigcup_{i=1}^{M} V_i(\boldsymbol{\tau}, \boldsymbol{T})$ is a credit subportfolio $CP(\boldsymbol{\tau}, \boldsymbol{T})$ with credit term T at moment $\boldsymbol{\tau}$.
$K_{ij}(\tau, T)$ — number of granted loans in $V_i(\tau, T)$ with risk indicator j,
$K_i(\tau, T) = K(V_i(\tau, T))$ – number of granted loans in $V_i(\tau, T)$,
$K_{i0}(\tau, T) + K_{i1}(\tau, T) + \dots + K_{i13}(\tau, T) = K_i(\tau, T)$
$K_{ij}(\tau)$ — number of granted loans in $V_i(\tau)$ with risk indicator j, $K_{ij}(\tau) = \sum_T K_i(\tau, T)$
$K_i = K(V_i)$ — number of granted loans in V_i, $K_i = K_i(\tau) + K_{i14}(\tau) + K_{i15}(\tau)$
$N_{ij}(\tau) = K_{ij}(\tau) / K_i(\tau)$ — rate of loans in $V_i(\tau)$ with risk indicator j

3.2. Vintage Table

Figure shows an example of a vintage table $VT(\tau, T)$ containing $K_{ij}(\tau, T)$ number of loans of vintage $V_i(\tau, T)$ with risk indicator j, i=1,2, … , M, j=0,1, … , 15

VINTAGE TABLE												
			DPD (DAYS PAST DUE)									
		Initial numbers of granted loans	0 (J=0)	1-30 (j=1)	31-60 (j=2)	61-90 (j=3)	...	301-330 (j=11)	331-365 (j=12)	>365 days (j=13)	Repaid loans (j=14)	Defaults (j=15)
Date of Granted Loans	$[t_0,t_1]$	K_1										
	$[t_1,t_2]$	K_2										
	$[t_2,t_3]$	K_3										
	$[t_3,t_4]$	K_4										
	-	-										
	-	-			$K_{ij}(t)$							
	-	-										
	$[t_{N-2},t_{N-1}]$	K_{N-1}										
	$[t_{N-1},t_N]$	K_N										
		All										

Figure. Vintage table.

Risk category j=14 is complementary to the possible situations of loans with regard to risk. After all, at the end of the portfolio life $\tau = T_{fin}$, in will be received two possible classes of risk. The first one is obviously a group of repaid loans $(RC_{14}(T_{fin}))$, and the second group is defaults (lost loans) $(RC_{15}(T_{fin}))$.

Vintage tables on Figure for different T give risk representation of the credit portfolio by DPD risk indicators at the moment τ. For the end to estimate risk parameter PD consider the following decomposition of the vintage table:

$$VT(\tau, T) = VT_{0\_13}(\tau, T) \cup VT_{14\_15}(\tau, T),$$

where $VT_{0\_13}(\tau,T)$ is the vintage table with risk indicators $j = 0,1,\dots,13$ and $VT_{14\_15}(\tau,T)$, is the vintage table with risk indicators $j = 14,15$.

Vintage table $VT_{0\_13}(\tau,T)$ consists observed data of all DPD from 0 to 365+ (365 days and more), vintage table $VT_{14\_15}(\tau,T)$ consists of observed data of repaid credits (no defaulted=ND) and observed data of defaulted credits (D). At the end of life of the credit portfolio $\tau = T_{fin}$ all credits will have distributed among categories j=14 and j=15, and the Rate of Default will can be calculated by the following way:

$$RD = \left(\sum_{i=1}^{M} K_{i15}\,(T_{fin})\,\right) \Big/ \left(\sum_{i=1}^{M} K_i\,(T_{fin})\right)$$

But before T_{fin} only part defaults are known and it is necessary the estimation of PD, which is the expected Rate of Default.

4. Exact Maximum Likelihood Estimate of the Structure of a Stratified Population

For estimation *PD* we will use the result from [1] in the simple case of a single sample without replacement of m items from the general stratified set $U = U_1 \cup U_2, U_1 \cap U_2 = \emptyset$,

with known size $N^0 = N_1 + N_2$ and unknown sizes of subsets $N_1 = |U_1|, N_2 = |U_2|$.

Let $\eta_1 -$ number of different items from subset U_1 in our sample, $\eta_1 + \eta_2 = \eta$. Then exact maximum likelihood estimate (MLE) is the

$$\widehat{N_1} = [(N^0 + 1)\eta_1/\eta],\ m < N^0, \qquad (2)$$

where [x] — the integer part of x. That is MLE proportional to the observed number of different elements of this subset

$$\widehat{N_1} \approx \eta_1 N^0/\eta,\ m < N^0.$$

However, since the expectation of the same statistics $E_{\bar{N}}\eta_1 = N_1 E_{\bar{N}}\,\eta\,/N^0$, the following equality holds $N_1 = N^0\,E_{\bar{N}}\eta_1\,/E_{\bar{N}}\,\eta$. Replacing the theoretical average corresponding observed values l и l_1, we obtain an estimate on the method of moments $\widetilde{N}_1 = N^0\,l_1/l\,,\ l = m$.

As we can see, MLE almost coincides with the estimate by the method of moments. In addition, we can always evaluate the displacement of MLE.

5. MLE for Probability of Default

We can use the estimator (2) for the estimation of PD in the following way. By the definition of default (see the point “1. Risk parameters in BaselI Internal Rating Based (IRB) Approach” of this paper), the statistics DPD 90+ for each of the time period $[t_{i-1}, t_i], i = 1,\dots,M$

$$K_{i90+}(\tau,T) = K_{i4}(\tau,T) + K_{i5}(\tau,T) + \dots + K_{i12}(\tau,T) + K_{i13}(\tau,T)$$

is considered as the number of observed defaults and the statistics DPD 0 $K_{i0}(\tau,T)$

is considered as the number of observed nondefaults.

Then for each vintage $V_i(\tau,T)$ $=V_{iD}(\tau,T) \cup V_{iND}(\tau,T)\,, i = 1,\dots,M,$ where $V_{iD}(\tau,T)$ — defaults (by Basel II (see source 2)) in vintage $V_i(\tau,T)$, and $V_{iND}(\tau,T) -$ nondefaults in vintage $V_i(\tau,T)$.

To use the result from [1], let’s denote

$$N_i(\tau,T) = K\big(V_i(\tau,T)\big) = K\big(V_{iD}(\tau,T)\big) + K\big(V_{iND}(\tau,T)\big) = N_{i1}(\tau,T) + N_{i2}(\tau,T)$$

where $N_i(\tau,T)$ is known and $N_{i1}(\tau,T)$, $N_{i2}(\tau,T)$are unknown. That is why we have $V_i(\tau,T)$ as a stratified population with к = 2 quality classes $V_{iD}(\tau,T)$ and $V_{iND}(\tau,T)$, and $K_{i90+}(\tau,T), K_{i0}(\tau,T)$ are sufficient statistics for unknown $N_{i1}(\tau,T), N_{i2}(\tau,T)$.

Denote $\widehat{N}_{il_1}(\tau,T) -$ MLE (2) for $N_{i1}(\tau,T)$, than from (2)

$$\hat{N}_{il_1}(\tau,T)=[(N_i(\tau,T)+1)l_1 / l],\hat{N}_{il_2}(\tau,T) = N_i(\tau,T) - \hat{N}_{il_1}(\tau,T), \quad (3)$$

where [x] — integer part of x and
$l_1 = K_{i90+}(\tau,T)$, $l = K_{i0}(\tau,T) + K_{i90+}(\tau,T) < N_i(\tau,T)$,
and finally we have the MLE $\widehat{PD}(\tau,T)$ for $PD(\tau,T)$ of the portfolio $CP(\boldsymbol{\tau},\boldsymbol{T})$

$$\widehat{PD}(\tau,T) = \left(\sum_{i=1}^{M} \hat{N}_{il_1}(\tau,T)\right) \Big/ \left(\sum_{i=1}^{M} N_i(\tau,T)\right) \quad (4)$$

Denominator in (4) equals to the number of the loans

$$K\big(CP(\boldsymbol{\tau},\boldsymbol{T})\big) = \sum_{i=1}^{M} K(V_i(\tau,T)) = \sum_{i=1}^{M} N_i(\tau,T)$$

in the portfolio $CP(\boldsymbol{\tau},\boldsymbol{T})$. So we get the MLE for $PD(\tau)$ of the portfolio $CP(\tau)$ for any τ

$$\widehat{PD}(\tau) = \left(\sum_T PD(\tau,T) * K\big(CP(\boldsymbol{\tau},\boldsymbol{T})\big)\right) \Big/ \left(\sum_T K\big(CP(\boldsymbol{\tau},\boldsymbol{T})\big)\right) \quad (5)$$

It is important, that $\widehat{PD}(\tau)$ is the exact MLE for PD for any time moment τ and for any fixed size of the portfolio $CP(\tau)$. So this MLE could be served as the base for calculation optimal PD, which optimize provisions for covering expected losses of consumer credit portfolio.

6. Comparison $\widehat{PD}(\tau)$ with estimator for PD based on transition probabilities

Consider the estimator of PD based on transition probabilities [2] and compare it with $PD^*(\tau)$.

Let's $\boldsymbol{P}(n) = (P_0, P_1, P_2, \dots, P_n, \dots P_{12})$, is the vector of transitions probabilities that credit having delay n at the current month will have delay $(n+1)$ at the next month. For calculating P_n the historical observed data of delays of portfolio credits for last 2–3 years are used. To calculate P_n, all data about delays of each credit at each month are pooled (grouped) at the risk classes $RC_j, j = 0,1,2, \dots, 13$. (Table).

Table.

ALGORITHM OF CALCULATION OF TRANSITION PROBABILITIES

delay (in months)	*Month $i, i = 1,2,\dots,24$ (36)*		…	*Sum over all months*		P_n
	The number of credits in the given RC at i-th month	*The number of credits from the given RC that close delay at the next (i+1)-th month*		*The number of credits in all pools with given delay (A)*	*The number of credits from RC with given delay that close delay at the next month (B)*	$1 - B/A$
0		The number of credits from the 0-th RC, which stay in the 0-th RC at the next (i+1)-th month	…		The number of credits from the 0-th RC, which stay in the 0-th RC at the next month	
1			…			
…			…			
13						

Then if N_n is the number of credits in the n-th RC, then the conditional transition probability to go from the n-th RC to the (n+1)-th RC equals

$$P_n = N_{n+1}/N_n \qquad (6)$$

The estimator PD based on transition probabilities then calculated as follows

$$PD = \prod_0^{13} P_n \qquad (7)$$

For the calculation the estimators (6) and (7) the initial data are grouped into Risk Classes (RC). So the data in the RC are not homogeneous in time (because at any RC there are initial data from current year, and also from 1, 2 and 3 years ago, with different macroeconomic conditions during the initial data were observed), and also are not homogeneous in credit characteristics (at one RC credits with different credit term are grouped). So estimators (6), (7) does not have known properties and it is difficult, if not possible, to estimate accuracy of these estimators.

On the other hand, the proposed estimators (4), (5) for the each time moment τ and for each credit term T are the exact MLE. The vintage $V_i(\tau, T)$ are the homogeneous data in the relation of macroeconomic conditions and credit characteristics. Moreover, the whole estimator (5) stay the MLE for PD(τ) for any time moment τ, using the results about asymptotic properties of MLE (5) [3], it can be calculate confidence bonds for $PD^*(\tau)$ for τ from the period last 2–3 years and it can be possible forecast PD for the future time moments.

Such developing are being planning to represent in the forthcoming articles.

Sources: / Источники:

1. Basel Committee on Banking Supervision. International Convergence of Capital Measurement and Capital Standards. 2004. Available at: http://www.bis.org/publ/bcbs107.pdf.

2. Basel Committee on Banking Supervision, Consultative Document. The Internal Ratings–Based Approach. 2001. Available at: http://seepdf.net/doc/pdf/download/www__bis__org--publ--bcbsca05.pdf.

References:

1. Ivchenko G. I., Khonov S. A., Ivanov E. A. Exact maximum likelihood estimator of the structure of a stratified population. Mathematical Notes, August 1997, Volume 62, Issue 2, pp 181–185. (In Russian).

2. Babikov V. G. Teoriya i praktika roznichnogo kreditovaniya. Upravlenie finansovymi riskami, 2014, no. 1 (37), pp. 44–61. Available at: http://www.bsc-consult.com/doc/UFR_1_2014_4.pdf. (In Russian).

3. Ivchenko G. I., Khonov S. A. An asymptotic estimate for stratified finite populations. Diskr. Mat., 1989, v. 1., no. 3, pp. 87–95. Available at: http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=dm&paperid=927&option_lang=eng. (In Russian).

Список литературы:

1. Ивченко Г. И., Хонов С. А., Иванов Е. А. Точная оценка максимального правдоподобия структуры стратифицированного населения // Математические заметки. 1997. Т. 62. №2. С. 181–185.

2. Бабиков В. Г. Теория и практика розничного кредитования // Управление финансовыми рисками. 2014. №1 (37). С. 44–61. Режим доступа: http://www.bsc-consult.com/doc/UFR_1_2014_4.pdf.

3. Ивченко Г. И., Хонов С. А. Об асимптотическом оценивании для расслоенных конечных совокупностей // Дискрет. матем. 1989. Т. 1. №3. С. 87–95. Режим доступа: http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=dm&paperid=927&option_lang=rus.

Работа поступила в редакцию 24.01.2017 г.

Принята к публикации 26.01.2017 г.

Cite as (APA):

Levin, V., & Khonov, S. (2017). Exact maximum likelihood estimator for the probability of default on estimation provision consumer credit portfolio of the bank. *Bulletin of Science and Practice,* (2), 186–193. Available at: http://www.bulletennauki.com/levin-chonov, accessed 15.02.2017.

Ссылка для цитирования:

Левин В. В., Хонов С. А. Точная оценка максимального правдоподобия для вероятности дефолта при оценивании резервов потребительского кредитного портфеля банка // Бюллетень науки и практики. Электрон. журн. 2017. №2 (15). С. 186–193. Режим доступа: http://www.bulletennauki.com/levin-chonov (дата обращения 15.02.2017). (На англ.).