

АКТУАЛЬНЫЕ ВОПРОСЫ ФИНАНСОВО-ЭКОНОМИЧЕСКОГО РАЗВИТИЯ

Международный научно-практический журнал «Вестник МИРБИС» ISSN 2411-5703 <http://journal-mirbis.ru/>
№ 1 (17) 2019 DOI: 10.25634/MIRBIS.2019.1

Ссылка для цитирования этой статьи: Шайдуллин А. И. Методика подготовки к разработке скоринговой модели для P2P-кредитования [Электронный ресурс] // Вестник МИРБИС. 2019. № 1 (17). С. 128–139.
DOI: 10.25634/MIRBIS.2019.1.16

УДК 336.77.067.31

Ансэль Шайдуллин¹

МЕТОДИКА ПОДГОТОВКИ К РАЗРАБОТКЕ СКОРИНГОВОЙ МОДЕЛИ ДЛЯ P2P-КРЕДИТОВАНИЯ

Аннотация. Актуальность выбранной темы заключается в том, что в современном мире роль технологий и альтернативных методов осуществления банковских услуг непрерывно растет. Конкуренция между различными банками обостряется, и возникает необходимость прибегать к все более рискованным методам ведения бизнеса. Цель работы — разработать принципы оценки базы заемщиков для P2P-кредитования. Данная статья направлена на раскрытие основных показателей функционирования P2P-кредитования. Ведущим подходом к исследованию данной проблемы является анализ эконометрических и финансовых параметров банков, предоставляющих услугу P2P-кредитования (на примере компании «Lending club»), что позволяет комплексно рассмотреть существующие проблемы в данной отрасли. Методология по расчету кредитного риска была адаптирована для P2P-кредитования, раскрыты основные параметры, влияющие на кредитный риск, рассчитаны существенные и несущественные показатели вероятности дефолта, предложена усовершенствованная методика по оценке качества заемщика на основе скоринговой модели. Материалы статьи представляют практическую ценность для банковских организаций и других финансовых учреждений, которые могут применять указанную модель для предоставления новой услуги.

В результате расчетов удалось определить средний показатель кредитного риска для P2P-платформы. Он оказался выше показателя традиционных банков (17 % против 5 %). Это свидетельствует о высоких рисках при предоставлении кредитов заемщикам. Для сокращения математического ожидания потерь были предложены методологические рекомендации по созданию новой усовершенствованной скоринговой модели, а также при помощи эконометрических моделей были проанализированы основные переменные, влияющие на вероятность дефолта. Самыми существенными оказались категории кредитного рейтинга заемщика, ставки процента, FICO. Именно их и необходимо учитывать при создании скоринговых моделей, предоставляя им более высокий весовой коэффициент.

Ключевые слова: P2P-кредитование, кредитный риск, вероятность дефолта, рейтинг заемщика, ожидаемые потери, скоринговые системы.

JEL: C44, G21

1 Шайдуллин Ансэль Ильгизович — Российский университет дружбы народов (РУДН), ул. Миклухо-Маклая, 6, Москва, 117198.
Email: ansel.shajdullin@yandex.ru

Введение

Современный механизм предоставления услуг кредитования подразумевает вовлечение широкого круга посредников — банков, страховых агентств и других финансовых институтов. При этом традиционные бизнес-процессы данных организаций, разработанные в условиях бумажного документооборота прошлого столетия, обладают определённой громоздкостью, и, как следствие, продолжительностью. Например, традиционная академическая схема кредитования реального сектора банком подразумевает сбор свободных накоплений граждан с целью дальнейшей выдачи этих средств в виде кредитов для компаний. Однако внедрение современных ИТ-технологий

существенно меняет привычные схемы во многих видах бизнеса, и предоставление кредитных услуг не является исключением. P2P-кредитование появилось как альтернатива традиционным кредитным организациям и расшифровывается как «peer-to-peer» («одноранговое кредитование»). Главное отличие от традиционных методов — это игнорирование финансовых институтов и предоставление кредитов на специальных онлайн-площадках. Заемщик сможет параллельно исполнять несколько ролей — быть как заемщиком, так и кредитором. P2P-кредитование — это альтернативный метод предоставления услуг, он не входит ни в один из трех категорий традиционных финансовых институтов: депозитарии, инвесто-

ры, страхователи. Но, тем не менее, P2P-банки вынуждены конкурировать именно с традиционными банками. Методы оценки деятельности P2P-и классических банков зачастую совпадают [Hulme, Wright, 2006].

При этом большинство выдаваемых кредитов в данной модели кредитования являются необеспеченными частными займами. Деньги при необеспеченных займах выдаются в ускоренном режиме, но и сроки погашения обычно бывают сокращенными (от 30 до 120 дней). У подобной формы есть свои очевидные преимущества: 1) упрощенная система для заемщика, более высокая безопасность. Многие P2P-компании не требуют от своих клиентов предоставления залога, тем самым клиент не несет имущественных обязательств и не несет риска потери имущества; 2) возможность получить деньги в любом действующем банке. Клиент сразу получает деньги на свой счет без посредников; 3) подобные займы несут нецелевой характер, то есть, клиент зачастую не указывает цель приобретения займа и может распоряжаться им по своему усмотрению. Банк часто дает рекомендации не использовать выданный заем для погашения других задолженностей; 4) более высокая клиентоориентированность. Заемщик сам выбирает сумму, срок погашения, а иногда и процентную ставку.

В случае с P2P-кредитованием невозможно обойтись без очевидных угроз, например, повышенного уровня кредитного риска и более высокого показателя ожидаемых потерь. Оценить возможности заемщика и его надежность порой бывает очень сложно в виду сокращения предоставляемых документов. Поэтому необходимо выделить те факторы, которые имеют более высокие показатели значимости и способны повлиять на вероятность дефолта. Ставки по подобным кредитам могут иметь широкую амплитуду. Кредиторы могут понизить риск невозврата средств, выдавая массу мелких кредитов разным заемщикам. Некоторые банки проводят исследования

суммы займа и вероятности дефолта в зависимости от величины займа.

Методологические основы

Анализ и оценка кредитного риска при потребительском кредитовании

В условиях осуществления и предоставления различных кредитов зачастую возникают различные риски. В широком смысле, риск — это любая вероятность наступления неблагоприятных последствий для того или иного субъекта (участника) финансовой (или иной) деятельности. Кредитный риск — риск потери финансовых средств из-за неисполнения дебитором своих обязательств, риск возникновения дефолта дебитора. Оценка кредитного риска — это методология по выявлению возможного убытка, с которым может столкнуться банк с некоторой вероятностью и в определенный промежуток времени.

Основные причины возникновения убытков:

- 1) уменьшение стоимости кредитного портфеля в связи неплатежеспособностью заемщиков;
- 2) созданием резервов под отсрочку возвращения кредитов крупными заемщиками банка;
- 3) несоблюдение заемщиком и инвестором условий договора;
- 4) внешне- и внутриэкономические факторы — кризис, нерациональное снижение/повышение процентной ставки и др. Риск — это статистическая и вероятностная величина, полностью избавиться от неопределенности невозможно, но существуют методы по снижению неблагоприятного исхода.

Любая статистическая информация имеет погрешности, она носит скорее прогнозный характер. К тому же в условиях постоянно меняющейся экономической среды оценка прежних данных не всегда дает абсолютно точный результат. Поэтому необходимо минимизировать погрешности и показатели неопределенности, например, можно применить оценку более долгой выборки лет. Существует ряд показателей, влияющих на результаты кредитного риска (табл.1).

Таблица 1. Описание основных переменных

Назв. перем.	Описание	Ед. изм.
<i>Социально-демографические</i>		
Own	Тип владения жильем	Ранг: 3 – владелец, 2 – аренда, 1 – ипотека, 0 - другое
Empl	Стаж работы	Ранг: от года до 10 лет и более
<i>Финансовые</i>		
LA	Сумма займа	\$
I	Ставка процента	%

Назв. перем.	Описание	Ед. изм.
CG	Группа кредитного рейтинга займа	Ранг: от 1 до 35 (A1 to G5)
FICO	Кредитный рейтинг заемщика	Число (300–850)
Inc	Годовой доход	\$
DTI	Долг/доход заемщика	%
RB	Суммарный баланс по возобновляемым кредитам заемщика	\$
RU	Возобновляемые долговые средства/кредитный лимит по возобновляемым кредитам	%
<i>Кредитная история</i>		
Del	Количество просроченных заемщиком платежей по займам	Число (за 2 года)
Inq	Количество заявок на кредит	Число (за полгода)
PR	Количество следующих событий: объявление о банкротстве, арест имущества и др.	Число (за полгода)
Exp	Кредитный стаж	Число лет
<i>Макроэкономические показатели</i>		
Unempl	Уровень безработицы	%
DR	Ставка рефинансирования (устанавливается Центробанком)	%
<i>Зависимая переменная</i>		
Status	Статус займа	Бинарная переменная: 0 – кредит выплачен, 1 – дефолт

Источник: аналитический портал «Мир процентов.ру». [Электронный доступ].

Режим доступа: <http://mir-procentov.ru/banks/ratings/credits-delay-part-fl.html?date1=2018-12-01&date2=2017-12-01>

Для оценки ожидаемых потерь необходимо провести ряд расчетов, которые состоят из нескольких этапов (построено на примере P2P-банка «Lending club»). Выделяют ряд показателей, которые учитываются при расчетах кредитного риска:

1) PD — вероятность дефолта. Данный параметр позволяет оценить вероятность того, что заемщик будет не в состоянии удовлетворить свои долговые обязательства.

Первый этап: разделение заемщиков-участников на «надежных» и «ненадежных». Для этого проводится анализ 3 основных групп информации о клиенте: 1) личные данные заемщика; 2) информация о кредитной истории, 3) показатели экономического окружения заемщика.

Второй этап: создание шкалы классификации заемщиков по уровню вероятности наступления дефолта. На данной ступени происходит более детальный анализ параметров, распределение различных социально-демографических, экономических факторов по группам риска неосуществления условий договора. Обычно заемщики распределяются по признаку надежности (A1 — с самой высокой, G5 — с самой низкой).

Третий этап: фиксация средней вероятности дефолта на долгосрочном среднем уровне наступления дефолта. Краткосрочная оценка (оценка на основе малой выборки лет) дает большую погрешность, более детальный анализ позволит снизить уровень погрешности.

Четвертый этап: корректировка финальной оценки исходя из «предупреждающих сигналов», помощи материнской компании и каких-либо нарушений. В целом, оценка результатов происходит по бинарному сценарию (существует два варианта итога — «выплачен» и «дефолт»). P2P-компания так же, как и традиционные создают своего рода «страховочный парашют», который в случае чего, позволит минимизировать затраты [Bachmann et al., 2016].

Для «Lending Club» процент удовлетворения требований клиента по предоставлению кредита (категория «выплачен») установился на отметке 0,47, вероятность дефолта = 17 % (0,17) (табл. 2). Со значениями каждого параметра можно ознакомиться в таблице 1.

Таблица 2. Значения описательных характеристик переменных

	Среднее значение			Максимум		Минимум		Стд. откл.	
	Выпл.	Деф.	Вер. рав-ва (%)	Выпл.	Деф.	Выпл.	Деф.	Выпл.	Деф.
DEL	0,15	0,17	17	7	4	0	0	0,49	0,47
DR	0,73	0,75	0,5	3	3	1	1	0,17	0,22
DTI	12,84	13,30	0	30	30	0	0	6,70	6,65
Exp	13,02	12,20	0	39	39	3	3	6,42	6,31
FICO	718,53	703,12	0	822	807	662	662	36,24	30,91
I	0,11	0,13	0	0,22	0,22	0	0	0,03	0,03
INC	65 281	57 064	0	159 996	157 656	6 000	9 600	30 325,4	28 080
LA	10 791	10 339	1	35 000	35 000	1 000	1 000	6 726,19	6 858
RB	10 457,47	9 767,78	0,05	39 926	39 850	0	0	9 073,55	8 618
UNEMPL	9,43	9,52	7,9	14	14	1	3	1,84	1,91

Источник: аналитический портал «Мир процентов.ру». [Электронный доступ].

Режим доступа: <http://mir-procentov.ru/banks/ratings/credits-delay-part-fl.html?date1=2018-12-01&date2=2017-12-01>

Существует огромное множество методик, использующих параметр PD, для оценки вероятности дефолта. Для правильного анализа необходимо исследовать каждый фактор отдельно, составить систему уравнений, найти общую среднюю величину с учетом стандартного отклонения. В данной работе будет проанализирован 1 показатель — DEL (количество просроченных платежей).

2) LGD — доля потерь банка в случае дефолта заемщика. LGD основан на внутреннем рейтинге заемщика, устанавливаемом самим банком. Существует два основных метода расчета доли потерь в случае дефолта: метод цепной реакции и

метод Борнхуэттера-Фергюссона. В поле первичного интереса выходит доля, которую банк (или его аналогия) может потерять из-за дефолта заемщика [Berger, Gleisner, 2009].

Теоретически, при P2P кредитовании отсутствует само понимание такого института, как банк, поэтому оценивать долю потерянного капитала довольно сложно. Зачастую информация по данному вопросу либо вовсе не публикуется, либо публикуется, но на относительно недолгий период (до 3 лет) с большой амплитудой погрешностей (табл. 3).

Таблица 3. Сводная таблица по нескольким видам предоставленных кредитов компанией «Lending club» по методике Борнхуэттера-Фергюссона

Срок (месяцы)	Заявленные претензии			Оплаченные претензии		
	CDF* до окончат.	заявлен., %	незаявлен., %	CDF* до окончат.	заявлен., %	незаявл., %
12	1.292	83.4	16.6	2.390	65.8	34.2
24	1.110	88.1	11.9	1.404	67.2	32.8
36	1.051	91.1	8.9	1.184	73.5	26.5

* CDF – фактор развития совокупных претензий

Источник: построено автором на основе материалов с официального сайта компании «Lending club». Режим доступа: <https://www.lendingclub.com/loans/resource-center/help-and-advice-on-money-and-credit> (дата обращения: 15.01.2019)

Можно отметить, что такой фактор, как развитие совокупных претензий, при P2P-кредитовании имеет более нестабильные и динамичные темпы. Это свидетельствует также о существовании некой шаткости в общем устройстве системы функционирования P2P-кредитования. Для анализа кредитного риска важно знать процент вероятности риска и процент утраченной доли. Модель Борнхуэттера-Фергюссона хорошо подходит при недостатке данных. В расчет берутся

проценты оплаченных заявленных претензий. Так как банки стремятся сохранять информацию о возможных потерях в случае дефолта, то для показателя найдем среднее арифметическое за 3 года. Тогда показатель LGD будет равен:

$$(65,8 + 67,2 + 73,5) / 3 = 68,83 \%$$

$$68,83 / 100 = 0,6883 (\sim 0,688)$$

То есть, в случае дефолта заемщика, банк теряет 68,83 % от активов, указанных в договоре. Остальная сумма либо имеет определенный за-

страхованный лимит, либо по условиям договора обладает различными защитными механизмами (гарантии, залоговые обязательства и т. д.)

3) EAD — требования под риском дефолта. Позволяет оценить, какая доля кредитных требований находится под кредитным риском на момент дефолта¹. Произведение этих трех параметров позволяет вычислить ожидаемые потери (математическое ожидание потерь) [Gobeljic, 2012]:

$$EL = PD * LGD * EAD$$

Для расчетов возьмем кредитное требование, соответствующее 100 млн рублей. Вставим полученные данные в формулу ожидаемых потерь:

$$EL = 0,17 * 0,688 * 100\ 000\ 000\ р. = 11\ 730\ 000\ р.$$

P2P-банки сокращают свои расходы благодаря отсутствию необходимости строить и развивать сети филиалов. Также здесь нет детального аппарата бюрократии и долгого процесса проверки документации заемщика. Весь процесс проходит в режиме скоринга. Таким образом, несмотря на высокие показатели ожидаемых потерь, в конечном счете, это будет нивелироваться с низкими затратами на функционирование компании и мониторинг (поиск) клиентов. Ввиду того, что существует возможность сокращения расходов, то полученный показатель потерь не совсем верно отражает действительность (имеет погрешности). Для снижения показателя суммы потерь в теории необходимо ввести поправочный коэффициент, который будет зависеть от ряда показателей: 1) отношения суммы займа к суммарному объему капитала P2P-банка; 2) динамики роста доходов и расходов компании. Самая простая формула расчета поправочного коэффициента выглядит следующим образом:

$K = \text{Затраты за определенный период} / \text{выручка за определенный период}$

1 Письмо Банка России от 29.12.2012 № 192-Т

«О Методических рекомендациях по реализации подхода к расчету кредитного риска на основе внутренних рейтингов банков»

$$K_{\text{пвр}} = 12,5 \times LGD \times \left(N \left(\frac{N^{-1}(PD) + \sqrt{R} \times N^{-1}(0,999)}{\sqrt{1-R}} \right) - PD \right) \times \frac{1 + (M - 2,5) \times b(PD)}{1 - 1,5 \times b(PD)}, \text{ где}$$

M — (maturity) срок до погашения кредитного требования (в годах), возьмем показатель, равный 3 годам;

N(x) — функция стандартного нормального распределения;

R — значение показателя корреляции, рассчитываемое по формуле:

$$R = 0,12 \times \left(\frac{1 - e^{-50 \times PD}}{1 - e^{-50}} \right) + 0,24 \times \left(1 - \frac{1 - e^{-50 \times PD}}{1 - e^{-50}} \right);$$

Умножим данную величину на показатель отношения суммы займа к суммарному объему капитала P2P-банка. Так как P2P-банкам выгоднее давать мелкие займы, то последняя величина будет иметь очень маленький показатель. Затраты на содержание банка обычно довольно низкие, выручка же может достигать огромных показателей (в отличие от прибыли). В итоге, получаем 2 предельно маленькие величины. В 2016 году объем выданных кредитов компанией «Lending club» составил более 2,75 млрд долл. США с чистой прибылью на 1 акцию 0,01 долл. У компании на 2016 год было в наличии 868 млн долл. Средняя сумма займа для P2P-банков установилась на уровне 5 000 долл. США.

Условно говоря, поправочный коэффициент будет иметь следующую формулировку:

$$K = (\text{Затраты за определенный период}) / (2,75 \text{ млрд долл. США})$$

При ежегодных затратах компании в 1 млрд, поправочный коэффициент составит 0,36. Это значительный показатель, который снижает величину ожидаемых потерь на треть. К сожалению, сейчас нет достаточных оснований, чтобы применять данный коэффициент в расчетах. Для этого необходим более детальный и точный анализ расходов компании. На деле, данная величина должна установиться на промежутке 0,8–0,9 пунктов.

Полученных данных достаточно, чтобы посчитать величину кредитного риска для всех кредитных требований, за исключением приобретенной дебиторской задолженности. Упрощенная формула имеет вид:

$$КРП = b * K_{\text{пвр}} * EAD,$$

где:

b — поправочный коэффициент (= 1,06),

$K_{\text{пвр}}$ — коэффициент риска

Формула расчета $K_{\text{пвр}}$ выглядит следующим образом:

$b(PD)$ — значение показателя корректировки на срок до погашения:

$$b(PD) = (0,11852 - 0,05478 \times \ln(PD))^2$$

1) Нахождение показателя корректировки (до погашения):

$$\ln(0,17) = -1,7719568419319 \text{ (примерно, } -1,772)$$

$$b(PD) = (0,11852 - 0,05478 * (-1,772))^2$$

$$b(PD) = (0,21559)^2$$

$$b(PD) = 0.04648$$

Нахождение показателя корреляции:
 $R = 0,12 * ((e - 50 * PD - e - 50) / (1 - e - 50) - 1)$
 $e \sim 2,7$

$R = 0,12 * ((2,7 - 50 * 0,17 - 2,7 - 50) / (1 - 2,7 - 50) - 1)$
 $R = 0,120026$
 3) Расчет $K_{пвр}$:

$$K_{пвр} = 12,5 * 0,688 * (N(N - 1(0,17) + 0,346448 * N - 1(0,999)) / 0,938) - 0,17 * (1 + (3 - 2,5) * 0,04648) / (1 - 1,5 * 0,04648)$$

$$K_{пвр} = 8,6 * (N(N - 1(0,17) + 0,346448 * N - 1(0,999)) / 0,938) - 0,17 * 1,02324 / 0,93028$$

$$K_{пвр} = 8,6 * (N(N - 1(0,17) + 0,346448 * N - 1(0,999)) / 0,938) - 0,17 * 1,099927$$

$$K_{пвр} = 9,45937 * (N(N - 1(0,17) + 0,346448 * N - 1(0,999)) / 0,938) - 0,17$$

$$K_{пвр} = 3,589447$$

4) Вычислим величину кредитного риска для всех кредитных требований:

$$КРП = 1,06 * 3,589447 * EAD$$

При двух показателях (из трех), которые являются константами, третий показатель может стать переменным. Сводная таблица по формуле кредитного риска отображена в таблице 4 (можно указать любое число для параметра EAD).

Таблица 4. Сводная таблица распределения показателя EAD и кредитного риска при данном показателе

б — поправочный коэффициент	Кпвр	EAD	Кредитный риск (КРП)
1,06	3,589447	100 000	380 481,4
1,06	3,589447	200 000	760 962,8
1,06	3,589447	500 000	1 902 406,9
1,06	3,589447	700 000	2 663 369,7
1,06	3,589447	1 000 000	3 804 813,8
1,06	3,589447	1 500 000	5 707 220,7
1,06	3,589447	2 000 000	7 609 627,6
1,06	3,589447	5 000 000	19 024 069,1
1,06	3,589447	10 000 000	38 048 138,2

Источник: построено автором на основе расчетов кредитного риска

При сумме займа (LA) более 10 000 000 руб., кредитный риск достигает «красной зоны» — почти в 4 раза превышает сумму займа. Для многих P2P-банков такой показатель представляет серьезную угрозу банкротства. Преимущественно, кредитный риск имеет способность влиять на кредитный портфель. Этот показатель является микропоказателем, то есть, банк рискует потерять свою ликвидность и снизить стоимость капитала. Кредитный риск демонстрирует, какой объем выданных средств может представлять угрозу для деятельности банка в целом. Если сумма выданного кредита превышает сумму стоимости капитала и имеющихся средств компании, то такой заём может оказаться недействительным из-за повышенного риска банкротства.

Таким образом, можно сказать, что относительно

но традиционных банков, банки, осуществляющие P2P-кредитование, имеют более высокие показатели кредитного риска (5 % — классические банки против 17 % у P2P-банков). Из-за высоких параметров кредитного риска, ожидаемые потери у подобных банков также обладают более высокими показателями, чем у традиционных банков. Данные риски покрываются экономией на содержании функционирования банка.

Реализация стратегии P2P и методология анализа данных на примере компании «Lending club»

Ипотечный кризис 2007–2008 гг. заставил все-таки задуматься о создании платформ, которые в случае возникновения экономических кризисов не будут рисковать своими материальными активами и смогут функционировать дальше без колоссальных затрат. Несколько средних традиционных банков заявили о своем банкротстве, в отличие от крупных, но все еще только начинающих P2P-банков. Уровень доверия к Интернет-займам стал медленно, но верно расти.

Существуют несколько методов анализа данных компаний:

1. Дескриптивный анализ — описательный анализ отдельных переменных. Позволяет наглядно оценить (в числовых и процентных показателях) различные параметры, на основании которых можно создать описательную модель. Согласно данным табл. 2, можно заметить, что по многим факторным переменным есть различия в математических ожиданиях по дефолтным и недефолтным заявкам. Большинство параметров обладают дескриптивной способностью, за исключением количества просроченных платежей и уровня безработицы. Уровень безработицы сложно оценить при помощи дескриптивных методик. Безусловно, учитывать его стоит хотя бы на уровне простой бинарной системы, но его компонентный вес будет незначителен, так как показатель не является параметром самого заемщика.

Стаж работы имеет параболическую зависимость с вероятностью дефолта. До опреде-

ленного момента, зависимость имеет обратную тенденцию: с ростом стажа работы, вероятность дефолта падает, но после 10 лет и более, зависимость превращается в прямую. Самой надежной является группа с опытом работы 9 лет (доля дефолтов 11 %). В целом, результаты исследования соответствуют целевой аудитории компании: лица от 25 до 32 лет со средним заработком от 1 200 до 2 300 долл. в месяц, имеющие высшее образование по гуманитарным и точным наукам. Подобный анализ позволяет также оценивать зависимость между вероятностью дефолта и положением организации в рейтинге банков и др. Дескриптивный анализ упрощает построение маркетинговой стратегии, так как дает более обширную информацию о рынке и самой компании [Cutler, Gleiser, 2014].

2. Корреляционный анализ — способ интерпретаций статистических данных путем поиска взаимосвязи между переменными. В общем виде, формула расчета корреляционной взаимосвязи выглядит следующим образом:

$$r = \frac{\sum (x_{1j} - \bar{x}_1) \cdot (x_{2j} - \bar{x}_2)}{\sqrt{\sum (x_{1j} - \bar{x}_1)^2} \cdot \sqrt{\sum (x_{2j} - \bar{x}_2)^2}}$$

В данном примере $r \in [-1, 1]$. Допустимое упущение — существуют параметры, изначально обладающие некоторой взаимосвязью, так как входят в общую систему оценки кредитного риска и вероятности дефолта. Был проведен расчет парных коэффициентов корреляции Пирсона между зависимой и объясняющей переменными (табл. 5) в условиях оценки вероятности дефолта. Ряд коэффициентов, обладающих высоким показателем значимости, имеют низкую корреляционную зависимость с вероятностью дефолта.

Таблица 5. Корреляционный анализ парных коэффициентов корреляции между зависимой и факторными переменными

	Коэффициент корреляции
CG	0,152***
Del	0,014
DR	0,028***
DTI	0,023**
Empl	0,016
FICO	-0,146***
Exp	-0,048***
I	0,166***
Inc	-0,093***
LA	-0,023**

	Коэффициент корреляции
Own	-0,035***
PR	0,063***
RB	-0,026**
Ru	0,094***
Unempl	0,017

Значимость коэффициентов: (*) 10 %, (**) 5 %, (***) 1 %
 Источник: официальный сайт компании «Lending club». Режим доступа: <https://www.lendingclub.com/loans/resource-center/help-and-advice-on-money-and-credit> (дата обращения: 15.01.2019)

Рассмотрим пример оценки корреляции между показателями группы кредитного риска (значимость коэффициента на уровне 1%) и вероятностью дефолта более подробно. При расчетах корреляции Пирсона была обнаружена особенность. Существует несколько групп заемщиков, разделенных по принципу кредитного риска: от А — самый надежный до G — наименее надежный. Каждой группе в разных количествах были выданы кредиты (оцениваются данные 2017 года). Например, для А — это 2 896 кредитов. Общая выборка представлена в табл. 6.

Таблица 6. Зависимость между группой кредитного риска и вероятностью дефолта

Группа	Количество кредитов	Доля выборки, %	Дефолт	Доля дефолтов, %
A	2 896	31,54	210	7,25
B	2 903	31,62	393	13,54
C	1 825	19,88	318	17,42
D	1 121	12,21	243	21,68
E	344	3,75	72	20,93
F	73	0,80	20	27,40
G	20	0,22	8	40,00
Итого	9 182	100,00	1 264	

Источник: Источник: построено автором при помощи аналитического сайта «Литобзор». Режим доступа: <https://lit-review.ru/biostatistika/dispersionnyjj-analiz-anova> (дата обращения: 09.03.2019)

Очевидно, что для самой ненадежной группы вероятность дефолта установится на более высоком уровне. Для группы G вероятность дефолта установилась на отметке 40 %. Конечно, логичнее всего полностью исключить данную группы из предполагаемых заемщиков вовсе, но за счёт низкой выборки (0,22 %) проблема решается сама собой и незначительно влияет на затраты компании. В данном случае, группа А (заемщики с высоким уровнем надежности) будет иметь наиболее плотные показатели FICO (с меньшей амплитудой разброса). Это связано с тем, что эта группа в даль-

нейшем будет обладать привилегированными условиями предоставления кредитов. Например, в последующие разы иметь возможность снижать процентную ставку без угроз понижения в кредитном рейтинге заемщика. Данная группа должна быть представлена меньшим количеством заемщиков. Группа А немного уступает группе В по численности предоставленных кредитов. При этом, доля дефолтов в этой группе минимальна. Наиболее небезопасной является группа G, но и количество выданных кредитов не существенно. Для построения корреляционной модели в условиях корреляционного поля построим таблицу расчетов (табл. 7).

Таблица 7. Показательные величины зависимости между группой риска заемщика (x-количество выданных кредитов) и вероятностью дефолта (y-количество запросов, по которым наблюдался дефолт)

Гр. риска	x	y	x ²	y ²	x*y
A	2 896	210	8 386 816	44 100	608 160
B	2 903	393	8 427 409	154 449	1 140 879
C	1 825	318	3 330 625	101 124	580 350
D	1 121	243	1 256 641	59 049	272 403
E	344	72	118 336	5 184	24 768
F	73	20	5 329	400	1 460
G	20	8	400	64	160
Σ	9 182	1 264	21 525 556	364 370	2 628 180

Источник: методологические рекомендации по проведению эконометрического исследования. [Электронный ресурс]. Режим доступа: <http://e.120-bal.ru/finans/39856/index.html?page=6> (дата обращения: 25.03.2019)

Для наших данных система уравнений имеет вид:

$$9a + 9\,182 \cdot b = 1\,264$$

$$9\,182 \cdot a + 21\,525\,556 \cdot b = 2\,628\,180$$

Эмпирические коэффициенты регрессии будут равны: $b = 0.1101$, $a = 28.1149$.

1) Уравнение регрессии (эмпирическое уравнение регрессии):

$$y = 0.1101x + 28.1149$$

2) Выборочные среднее:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{9182}{9} = 1020.222$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{1264}{9} = 140.444$$

$$\overline{xy} = \frac{\sum x_i y_i}{n} = \frac{2628180}{9} = 292020$$

3) Выборочные дисперсии:

$$S^2(x) = \frac{\sum x_i^2}{n} - \bar{x}^2 = \frac{21525556}{9} - 1020.222^2 = 1350875.06$$

$$S^2(y) = \frac{\sum y_i^2}{n} - \bar{y}^2 = \frac{364370}{9} - 140.444^2 = 20760.91$$

4) Среднеквадратическое отклонение

$$S(x) = \sqrt{S^2(x)} = \sqrt{1350875.06} = 1162.272$$

$$S(y) = \sqrt{S^2(y)} = \sqrt{20760.91} = 144.086$$

5) Коэффициент корреляции b можно находить по формуле, не решая систему непосредственно:

$$b = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{S^2(x)} = \frac{292020 - 1020.222 \cdot 140.444}{1350875.06} = 0.1101$$

$$a = \bar{y} - b \cdot \bar{x} = 140.444 - 0.1101 \cdot 1020.222 = 28.1149$$

6) Коэффициент корреляции.

Ковариация:

$$cov(x, y) = \overline{xy} - \bar{x} \cdot \bar{y} = 292020 - 1020.222 \cdot 140.444 = 148735.46$$

В нашем примере связь между признаком Y и фактором X высокая и прямая (по шкале Чеддока). Кроме того, коэффициент линейной парной корреляции может быть определен через коэффициент регрессии b :

$$r_{x,y} = b \cdot \frac{S(x)}{S(y)} = 0.11 \frac{1162.272}{144.086} = 0.888$$

7) Коэффициент эластичности:

$$E = \frac{\partial y}{\partial x} \frac{x}{y} = b \frac{\bar{x}}{\bar{y}}$$

$$E = 0.11 \frac{1020.222}{140.444} = 0.8$$

Коэффициент эластичности меньше 1. Следовательно, при изменении X на 1 %, Y изменится менее чем на 1 % (рис. 2).

Можно заметить, что среди самых значимых коэффициентов (10 %) — стаж работы, уровень безработицы, просроченные платежи — незначительно влияют на возникновение дефолта [Гореева, Демидова, 2015].

Самыми коррелирующими с вероятностью дефолта оказались: 1) группа кредитного риска; 2) ставка процента; 3) кредитный рейтинг займа. Некоторые проанализированные показатели имеют особенность взаимно влиять друг на друга, из-за чего нельзя утверждать о полной объективности исследования.

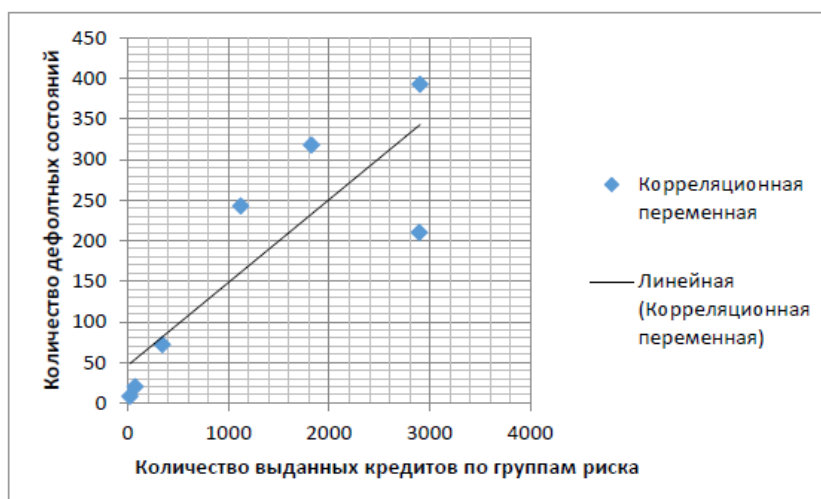


Рис. 1. Корреляционная зависимость между показателями количества дефолтных состояний и количеством выданных кредитов. Источник: построено автором

Поэтому оставшиеся факторы были проверены на мультиколлинеарность. Для этого была построена корреляционная матрица парных коэффициентов корреляции (табл. 8).

Таблица 8. Корреляционная матрица парных коэффициентов корреляции между факторными переменными

	CG	INC	DR	INQ	PR	OA	OWN	RU	TA	EXP	I	FICO
CG	1											
INC	0,027***	1										
DR	-0,015	-0,032***	1									
INQ	0,124***	0,042***	0,006	1								
PR	0,113***	-0,004	0,033***	0,036***	1							
OA	-0,056***	0,245***	-0,013	0,091***	0,018*	1						
OWN	-0,104**	0,166***	-0,006	0,048***	0,016	0,127***	1					
RU	0,484**	0,007	0,005	-0,061***	0,078***	-0,108***	-0,096***	1				
TA	-0,122***	0,376***	-0,008	0,128***	0,012	0,674***	0,226***	-0,088***	1			
EXP	-0,153***	0,286***	-0,013	0	0,067***	0,218***	0,216***	-0,042***	0,408***	1		
I	0,957***	0,012	-0,064***	0,135***	0,116***	-0,056***	-0,106***	0,491***	-0,136***	-0,167***	1	
FICO	-0,765***	0,105***	-0,038***	-0,052***	-0,156***	-0,008	0,135***	-0,558***	0,128***	0,231***	-0,785***	1

Значимость коэффициентов: (*) 10 %, (**) 5 %, (***) 1 %

Источник: Методологические рекомендации по проведению эконометрического исследования. [Электронный ресурс]. Режим доступа: <http://e.120-bal.ru/finans/39856/index.html?page=6> (дата обращения: 25.03.2019)

Категория кредитного рейтинга, ставка процента, показатель FICO — имеют сильную статистическую взаимосвязь. Эти показатели учитываются при расчетах кредитного риска, но при этом имеют способность влиять друг на друга. Действительно, группа кредитного рейтинга заемщика (FICO) тесно взаимосвязана с кредитным рейтингом займа (CG)¹, который в свою очередь строится на основе показателя ставки процента по займу (I). Группа кредитного рейтинга заемщика (FICO) рассчитывается самим банком. Обычно, принимает вид обратного «математического древа», где после оценки множества факторов полу-

чаем число от 300 до 850, что соответствует той или иной группе (от A до G). В целом, если число $a > 650$, то кредит выдается. В условиях скоринговой модели можно разработать собственные значения, соответствующие той или иной группе. Из-за подобного расклада, данные факторы необходимо: а) про дифференцировать, проверить на совместимость и влияние на другие факторы; б) исследовать показатели совместно в одной группировке с учетом фактора эндогенности остальных переменных [Пимичева, Худокормов, 2017]. Процентная ставка может предлагаться как самим заемщиком, так и банком. Слишком низкая ставка процента не позволит компании покрыть стандартные расходы по предоставлению услуги

1 Federal Reserve Discount Window (2014), Historical discount rates. URL: <http://frbdiscountwindow.org/historicalrates.cfm?hdrID=20&dtlID>

(для это сумма займа должна быть очень высокой, но при высоком показателе займа, кредитный риск также увеличивается). Высокая ставка процента ставит под вопрос актуальность займа. Скорее всего, клиент хочет воспользоваться займом для покрытия других кредитных операций, а после этого не выплачивать указанную сумму. Вероятность дефолта также возрастет.

Существуют исследования, согласно которым существует связь между полом заемщика и вероятности выдачи ему займа. Статистически мужчинам чаще отказывают в предоставлении кредитов. Объективных факторов для этого выявлено не было. Возможно, это связано с большим количеством заявителей-мужчин, чем женщин. Велико влияние и психоэмоциональных факторов. Поэтому компоновочный вес будет составлять маленькое число.

Заключение

В результате вычислений удалось определить важнейшие показатели функционирования банковских организаций: вероятность дефолта, кредитный риск, математическое ожидание потерь. Вероятность дефолта установилась на отметке 17 % (в расчете на вариацию 15–18 %), что значительно выше показателей классических банков. Доля потерь также имеет высокие параметры из-за отсутствия контролирующего органа (68,8 % против 45–50 % у традиционных банков) [Herzenstein, 2015].

P2P-банкам приходится рассчитывать на высокие математические ожидания потерь, которые составляют 11,73 % от суммы займа. Данные угрозы покрываются низкими затратами банка на содержание функционала. Процентные ставки по займам могут иметь широкую амплитуду в зависимости от предпочтений клиента (но в таком случае, слишком низкие проценты будут учтены более низким скоринговым баллом) [Тужик, Шульгина, Зайцева, 2015].

Кредитный риск, рассчитанный на показателе ПВР установил сумму, на которую можно ориен-

тироваться при предоставлении займа. Порог будет зависеть от общей суммы капитала P2P-банка. Именно поэтому банкам выгоднее давать мелкие займы большему количеству заемщиков, чем крупные, но избранному числу физических лиц.

При анализе показателей данных удалось выявить некоторые закономерности. Deskриптивный анализ показал, что к большинству параметров можно подобрать численные, ранговые, процентные значения, то есть, они обладает deskриптивной способностью. Это свидетельствует о том, что их можно применять в скоринговой оценке заемщика. Регрессионный анализ определил математическую значимость влияния показателей на вероятность возникновения дефолта. Среди значимых на уровне 10 % самыми корреляционно взаимосвязанными оказались показатели: группы кредитного риска, FICO, ставки процента. Корреляционный анализ парных коэффициентов доказал, что данные показатели обладают способностью взаимно влиять друг на друга. Для построения скоринговой модели необходимо разработать коэффициент, учитывающий и комбинирующий эти показатели. Интересно, что данные результаты коррелируют с аналогичными исследованиями при применении современных ИТ технологий в смежных отраслях делового администрирования, таких как бизнес-аналитика, маркетинг и операционный менеджмент [Ганявин, Черников, Макаров, 2018].

Ввиду отсутствия возможности проведения качественных скоринговых оценок заемщика и наличия некоторой информации, являющейся коммерческой тайной, расчет кредитного риска и математических ожиданий потерь для P2P-кредитования является довольно сложным процессом. В работе была предложена новая методология выбора коэффициентов для нахождения показателя математического ожидания потерь, а также проанализированы основные параметры вероятности дефолта.

Литература

Ганявин В. А., Черников С. Ю., Макаров А. А. RFM-сегментация и классификация клиентской базы для группы 1R алгоритмом машинного обучения [Электронный ресурс] // Вестник МИРБИС. 2018. № 2 (14). С. 20-27. DOI: 10.25634/MIRBIS.2018.2.3

Горева Н. М., Демидова Л. Н. Математические методы анализа и методология измерения рисков в банковской системе // Математическое моделирование в экономике, управлении, образовании: материалы Международной научно-практической конференции. Под редакцией Ю. А. Дробышева и И. В. Дробышевой. 2015. С. 37–51.

Пимичева Ю. Г., Худокормов В. В. Особенности применения моделей анализа экспертных оценок в ходе генерации и выбора решений технической задачи с учетом мнения потребителей // Научно-методический электронный журнал «Концепт». 2017. Т. 39. С. 971–975. URL: <http://e-koncept.ru/2017/970521.htm>.

Тужик А. М., Шуленина Д. И., Зайцева Е. В. Проблема «принципал-агент» и способы ее решения // Экономика и современный менеджмент: теория и практика: сб. ст. по матер. LIV-LV междунар. науч.-практ. конф. № 10-11(53). Часть II. Новосибирск: СибАК, 2015.

Bachmann A. et al. Online Peer-to-Peer Lending – A Literature Review // Journal of Internet Banking and Commerce, Vol. 16, No. 2. 2016.

Berger S. and Gleisner F. (2009), Emergence of Financial Intermediaries in Electronic Markets: The Case of Online P2P Lending // Business Research. Official Open Access Journal of VHB, Vol. 2, 2009, pp. 39–65.

Cutler, D. M. and Gleiser, E. L. Are ghettos good or bad? // Quarterly Journal of Economics, No. 112, 2014, pp. 827–872.

Gobeljic P. Classification of Probability of Default and Rating Philosophies. Stockholm, 2012.

Herzenstein M., et al. The Democratization of Personal Consumer Loans? Determinants of Success in Online Peer-to-Peer Lending Communities / M. Herzenstein, R. Andrews, U. M. Dholakia, and E. Lyandres. Working Paper, SSRN. 2015.

Hulme M. K., Wright, C. Internet Based Social Lending: Past, Present and Future, Social Futures Observatory: London, 2006.

ACTUAL ISSUES OF FINANCIAL AND ECONOMIC DEVELOPMENT

Ansel Shaydullin¹

METHODS OF PREPARING FOR THE DEVELOPMENT OF A SCORING MODEL FOR P2P LENDING

Abstract. *The relevance* of the chosen topic is that in the modern world the role of technologies and alternative methods of banking services is constantly growing. Competition between different banks is increasing, and there is a need to resort to increasingly risky methods of doing business. *The purpose* of the work is to develop principles for assessing the base of borrowers for P2P lending. This article is aimed at the disclosure of the main indicators of the functioning of P2P lending. The leading approach to the study of this problem is the analysis of econometric and financial parameters of banks that provide P2P lending services (for example, the company "Lending club"), which allowed to comprehensively consider the existing problems in this industry. *The methodology* for the calculation of credit risk was adapted for P2P lending, disclosed the main parameters affecting credit risk, calculated significant and insignificant indicators of the probability of default, proposed an improved method for assessing the quality of the borrower based on the scoring model. The materials of the article are of practical value for banking organizations and other financial institutions that can use this model to provide a new service.

As a result of the calculations, it was possible to determine the average credit risk for the P2P platform. It was higher than traditional banks (17% against 5%). This indicates high risks in lending to borrowers. To reduce the mathematical expectation of losses, methodological recommendations for the creation of a new improved scoring model were proposed, and the main variables affecting the probability of default were analyzed using econometric models. The most significant were the categories of the borrower's credit rating, interest rates, FICO. They should be taken into account when creating scoring models, providing them with a higher weight coefficient.

Key words: P2P loans, credit risk, probability of default, borrower rating, expected losses, scoring systems.

JEL: C44, G21

1 **Shaydullin Ansel Ilgizovich** – Peoples' Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya st., Moscow, Russia, 117198.
E-mail: ansel.shajdullin@yandex.ru

References

Ganyavin V. A., Chernikov S. Yu., Makarov A. A. RFM-segmentatsiya i klassifikatsiya kliyentskoy bazy dlya gruppy 1R algoritmom mashinnogo obucheniya [RFM-segmentation and classification of the client base for the 1R group by machine learning algorithm]. *Vestnik MIRBIS*. 2018. No. 2 (14). Pp. 20-27. DOI: 10.25634 / MIRBIS.2018.2.3 (In Russian).

Goreyeva N. M., Demidova L. N. Matematicheskiye metody analiza i metodologiya izmereniya riskov v bankovskoy sisteme [Mathematical methods of analysis and methodology for measuring risks in the banking system]. *Matematicheskoye modelirovaniye v ekonomike, upravlenii, obrazovanii: materialy Mezhdunarodnoy nauchno-prakticheskoy konferentsii* [Mathematical modeling in economics, management, education: Proceedings of the International Scientific and Practical Conference]. Edited by Yu. A. Drobyshev and I. V. Drobysheva. 2015. p. 37–51. (In Russian).

Pimicheva Yu. G., Khudokormov V. V. Osobennosti primeneniya modeley analiza ekspertnykh otsenok v khode generatsii i vybora resheniy tekhnicheskoy zadachi s uchetom mneniya potrebiteley [Features of the application of models for analyzing expert assessments during the generation and selection of solutions for a technical problem taking into account the opinion of consumers]. *Nauchno-metodicheskiy elektronnyy zhurnal "Kontsept"* [Scientific-methodical electronic journal "Concept"]. 2017. V. 39. p. 971–975. URL: <http://e-koncept.ru/2017/970521.htm>. (In Russian).

Tuzhik A. M., Shulenina D. I., Zaitseva E. V. Problema «printsipal-agent» i sposoby yeye resheniya [The problem of "principal-agent" and ways to solve it]. *Ekonomika i sovremennyy menedzhment: teoriya i praktika: sb. st. po mater. LIV-LV mezhdunar. nauch.-prakt. konf.* [Economics and modern management: theory and practice: Proceedings of the 54-55th. Intern. scientific-practical conf. No. 10–11 (53). Part II. Novosibirsk: SibAK Publ., 2015. (In Russian).

Bachmann A. et al. Online Peer-to-Peer Lending – A Literature Review, *Journal of Internet Banking and Commerce*, Vol. 16, No. 2. 2016.

Berger S. and Gleisner F. (2009), Emergence of Financial Intermediaries in Electronic Markets: The Case of Online P2P Lending, *Business Research. Official Open Access Journal of VHB*, Vol. 2, 2009, pp. 39–65.

Cutler, D. M. and Gleiser, E. L. Are ghettos good or bad? *Quarterly Journal of Economics*, No. 112, 2014, pp. 827–872.

Gobeljic P. Classification of Probability of Default and Rating Philosophies. Stockholm, 2012.

Herzenstein M., et al. *The Democratization of Personal Consumer Loans? Determinants of Success in Online Peer-to-Peer Lending Communities*. M. Herzenstein, R. Andrews, U. M. Dholakia, and E. Lyandres. Working Paper, SSRN. 2015.

Hulme M. K., Wright, C. *Internet Based Social Lending: Past, Present and Future*, Social Futures Observatory: London, 2006.