

УДК 336.71

ISSN 1812-5220
© Проблемы анализа риска, 2016

Проблемы использования внешних данных для оценки операционного риска в коммерческом банке

Л. В. Кох,**С. М. Булацкий,**Санкт-Петербургский
политехнический университет
Петра Великого**Аннотация**

В данной статье рассматриваются возможные способы решения нескольких частных задач, возникающих при формировании калибровочных выборок для моделей оценки операционного риска (ОР) в коммерческом банке. Рассматриваются аспекты разделения событий на однородные группы на основе количественных методов. Проводится качественное исследование влияния формирования выборок на итоговые оценки. В трех вариантах формулируется задача совмещения наборов данных с разными порогами отсеечения. Предлагается алгоритм для решения задачи совмещения данных с разными порогами отсеечения на основе усеченных распределений. Исследуется вопрос о недостатке данных для калибровки моделей оценки операционного риска. Обсуждаются три подхода к экстраполяции за пределы имеющихся данных с учетом потенциального наличия экстремальных потерь. В заключении статьи авторы приводят возможные пути решения проблемы малого количества данных об операционных потерях, направленные «снизу» и «сверху».

Ключевые слова: операционный риск, Базель II, внешние данные, калибровка модели LDA.

Содержание

Введение

1. Методы LDA

2. Сбор и классификация данных для калибровки модели

3. Совмещение данных с разными порогами отсеечения

4. Влияние размера выборок на точность оценки

Заключение

Литература

Введение

Согласно Базель II, операционный риск представляет собой риск убытка в результате неадекватных или ошибочных внутренних процессов, действий сотрудников и систем или внешних событий. Данное определение включает юридический риск, но исключает стратегический и репутационный риски. Это определение имеет универсальный характер, так как оно применимо для всего многообразия финансовых институтов. Кроме того, в определении сразу выделяются и источники операционного риска.

Параметры операционного риска являются важным фактором банковской деятельности. В соответствии с инструкцией Банка России от 03.12.2012 № 139-И «Об обязательных нормативах банков» нормативы достаточности базового, основного и собственного капиталов банка рассчитываются как дробь, в числителе которой находится соответствующий капитал банка, а в знаменателе — сумма, одно из слагаемых которой отражает уровень операционного риска в банке.

Следовательно, чем больше показатель операционного риска, тем жестче требования норматива к собственному капиталу банка. Порядок и методы расчета параметров операционного риска устанавливаются нормативными документами регулирующих органов. Согласно рекомендациям Базельского комитета по банковскому надзору (БКБН), а также последовавшим за ними рекомендациям Центрального Банка Российской Федерации, всем кредитным организациям следует разрабатывать основные принципы управления операционным риском, следовать им, а также совершенствовать их. Эти принципы должны являться определяющими для стратегии управления операционным риском в кредитной организации во всех аспектах, включая выявление, оценку, мониторинг и контроль [2].

Согласно Базель II, для расчета величины операционного риска следует использовать один из трех подходов.

1. Подход базового индикатора (BIA — Basic Indicator Approach).
2. Стандартизированный подход (SA — Standardized Approach).
3. Усовершенствованный (продвинутый) подход (Advanced Measurement Approach).

Согласно масштабному исследованию LDCE (Loss Data Collection Exercise), проведенному Базельским комитетом, результаты которого были опубликованы в 2009 году, во всех мировых регионах величина отчисляемого на покрытие ОР при расчете посредством VIA и SA капитала превосходит капитал, отчисляемый после расчетов по методам группы продвинутых подходов (AMA). В среднем экономия капитала достигает 2% от валового дохода банка [3]. При этом к 2009 году лишь 22% респондентов перешли на полное или частичное использование AMA. Многие кредитные организации заинтересованы в применении более сложных методов AMA, так как именно они призваны отражать реальную структуру деятельности банка, качество управления бизнес-процессами и рисками и другие особенности деятельности конкретного банка.

Одной из сложностей, с которыми сталкивается банк, желающий перейти на использование AMA, является соответствие требованиям Базель II для использования продвинутых подходов. Первое требование состоит в том, что разработанная банком система оценки уровня ОР обязана использовать четыре элемента: внутренние данные об операционных потерях, внешние данные об операционных потерях, анализ сценариев, а также факторы бизнес-окружения и внутреннего контроля банка. Также можно выделить несколько групп проблем практического применения AMA, связанных с формированием калибровочных выборок.

1. Сбор и классификация данных о потерях от ОР.
2. Совмещение данных с разным уровнем порога отсечения.
3. Масштабирование внешних данных.

В данной статье рассматриваются возможные пути решения задач первой и второй групп.

1. Методы LDA

Наибольшей популярностью среди группы AMA обладает подход на основе распределения потерь (LDA — Loss Distribution Approach) [3]. Согласно методу LDA, данные об операционных потерях классифицируются на отдельные группы. Классификация как может быть произведена согласно рекомендации Базель II (группа определяется парой «бизнес-линия/тип риска»), так и может создавать-

ся непосредственно банком в рамках собственной системы управления операционным риском. Общепринятое разделение событий по ячейкам «бизнес-линия/тип риска» формирует так называемую рисковую матрицу. Предполагается, что для группы однородных событий существует единая функция распределения потерь. Далее при помощи непараметрических или параметрических методов для каждой ячейки рискованной матрицы вычисляется мера риска при заданном уровне доверия. Агрегированная оценка по всем ячейкам определяет величину отчисляемого на покрытие операционного риска капитала.

2. Сбор и классификация данных для калибровки модели

Ясно, что в группы для последующего анализа можно объединять только однородные события. Говорится, что события операционного риска априорно однородны, когда они имеют одну природу. Например, все случаи ошибок операционистов банка при вводе данных с клавиатуры — это однородные события. При этом априорно судить о том, будут ли являться однородными, к примеру, события перебоев с подачей электричества и неполадок с интернет-соединением, нельзя. Гипотезу об однородности можно проверить в случае необходимости на этапе анализа выборок.

Большинство банков, которые в своей системе управления операционным риском применяют АМА, используют обычную рисковую матрицу, согласно которой каждое событие относится к группе, определяемой парой «тип риска/бизнес-линия». Почти три четверти банков используют не более ста однородных групп, при этом половина из них ограничиваются максимум двадцатью группами. Согласно рекомендациям БКБН, банк обязан применять качественные и количественные проверки для определения групп событий и правильной их классификации по этим группам [1, с. 54].

Для того чтобы выборка была репрезентативной, она должна иметь определенный объем, поэтому в случае если имеется несколько групп однородных событий с недостаточно репрезентативным количеством, то необходимо будет предпринять попытки для их объединения в более объемные выборки, при этом сохраняя их однородность.

В качестве критерия однородности событий в составной выборке можно использовать непараметрический критерий Колмогорова — Смирнова. Еще одним простым способом проверки однородности выборки может быть сравнение коэффициентов вариации — в том случае, когда для объединенной группы этот коэффициент будет не больше чем в каждой из исходных групп и не меньше чем коэффициент хотя бы одной исходной группы, может быть целесообразным объединение событий в единую новую группу [1, с. 47].

Грамотность формирования выборок повышает качество модели. Этап формирования выборок делает модель чувствительной к индивидуальным параметрам конкретного банка в разрезе операционного риска. В формировании выборок могут учитываться характер деятельности как всего банка, так и его подразделений, различие в причинах и последствиях событий ОР разной природы и т.п.

При этом если количество выборок сокращается, то есть если они становятся более агрегированными, то чувствительность модели снижается.

1. Неявное допущение об отсутствии взаимной связи между различными событиями.

2. В меньшей степени учитываются источники событий операционного риска, следовательно, такая модель не даст максимальной пользы для управления ОР в банке.

3. Объединение неоднородных событий в одну выборку потенциально приводит к искажению итоговых оценок уровня ОР.

С другой стороны, высокая детализация (большое количество групп, малое количество событий в группе) также имеет свои недостатки.

1. Нерепрезентативность выборки, вообще говоря, не позволяет обобщать результаты исследования на всю генеральную совокупность, из которой она была собрана.

2. Применение распределений с «тяжелым хвостом», которые должны использоваться для моделирования ОР, при условии недостаточного объема выборки для калибровки модели может очень существенно исказить итоговую оценку.

3. События операционного риска часто достаточно сложно классифицировать, и чем более детализированной будет модель, тем больше сложностей возникнет с классификацией.

Внутренние данные об операционных потерях по большей части состоят из случаев потерь небольшой величины, тогда как сама цель отчисления капитала на покрытие операционного риска состоит в предохранении банка от рискованных событий, которые имеют малую частоту, но большой размер в денежном выражении. Сбор данных о небольших потерях, вообще говоря, не всегда может быть признан эффективным для банка в плане стоимости, что приводит к отбрасыванию данных о потерях, не превосходящих некоторого значения, и, как следствие, к возникновению так называемого отклонения левого отсечения (left truncation bias). Это связано с тем, что каждый банк может самостоятельно устанавливать этот порог и в разных банках он действительно варьируется [11].

Также возможны потери, которые являются следствием одного конкретного события, но не единомоментны, а продолжаются во времени. Это может оказывать влияние на оценки частоты и величины потерь, основывающиеся на внутренних данных банка.

3. Совмещение данных с разными порогами отсечения

Как уже было упомянуто, значимым при моделировании является наличие левого порога отсечения, а также факт смещенности внешних данных: базы данных о потерях, составленные из упоминаний в открытых источниках (в том числе и Fitch), становятся жертвами так называемого феномена недопредставления, или недосказанности, т. е. очевидно, что далеко не все операционные потери в конечном итоге становятся общественным достоянием. Одним из наиболее очевидных проявлений такого феномена недосказанности является смещенность внешних данных из публичного доступа в сторону более крупных и, соответственно, более обсуждаемых потерь. В случае прямого использования внешних данных из открытых источников это явление в том числе может стать причиной увеличения итоговой оценки рискованного капитала, поэтому напрямую совмещать выборки внутренних и внешних данных недопустимо.

Эта проблема тесно связана со второй проблемой — наличием левого порога отсечения. Зная, что

так или иначе не все потери становятся достоянием гласности, поставщики внешних данных применяют на множество собранных ими данных определенное пороговое значение для отсечения слева. Естественно, порог выбирается очень большим, к примеру для базы Fitch это значение составляет 1 млн долл.

Собирая внутренние данные о потерях, каждый банк устанавливает собственное левое пороговое значение, поэтому множество внутренних данных также является отсеченным слева, но величина порога для внутренних данных обычно составляет от 0 до 20 тысяч евро [12]. В базах данных банковских консорциумов порог зафиксирован на общем для всех участников значении, к примеру для ORX он установлен на уровне 20 тысяч евро.

Что касается публично доступных данных, то в общем случае нижний порог, установленный в организации, понесшей потерю, является неизвестным (вендор не обладает этими данными) и для банка, развивающего свою модель оценки ОР, этот порог для каждой записи о внешних потерях является случайной величиной. В литературе широко обсуждается проблема совмещения выборок с различным порогом отсечения [4, 6, 7, 9].

Возможная схема совмещения внутренних и внешних данных с разным порогом отсечения представлена в работе [4]. Рассматриваются события, классифицированные как события одной однородной группы согласно Базель II. Помимо главного предположения о том, что внутренняя и внешняя выборка данных распределены одинаковым образом, предлагается рассмотреть три взаимоисключающих случая.

1. Порог отсечения выборки внешних данных H не является случайной величиной, и его значение известно.

2. Порог отсечения выборки внешних данных H не является случайной величиной, и его значение неизвестно.

3. Пороговое значение H является случайной величиной.

Пусть f — функция плотности распределения внутренних потерь ζ . Тогда для внешних данных можно записать следующее:

$$\zeta_i^* \sim f^*(\zeta, \theta),$$

так как f^* отличается от f из-за левого отсечения:

$$f^*(\zeta, \theta) = 1\{\zeta \geq H\} \cdot \frac{f(\zeta, \theta)}{1 - F(H, \theta)}.$$

Чтобы найти приближенную оценку параметров распределения θ , воспользуемся логарифмической функцией правдоподобия l . Она имеет следующий вид:

$$l(\theta) = \sum_{i=1}^n \ln f(\zeta_i, \theta) + \sum_{i=1}^{n^*} \ln f^*(\zeta_i^*, \theta).$$

Таким образом, решением задачи совмещения внутренних и внешних данных будет являться оценка параметров $\hat{\theta}$, являющаяся решением задачи максимизации:

$$\hat{\theta} = \arg \max_{\theta} l(\theta).$$

Во втором случае, если величина H полагается постоянной, но неизвестной, задача выглядит точно так же, но H теперь входит в число неизвестных параметров задачи максимизации:

$$(\hat{\theta}, \hat{H}) = \arg \max_{\theta, H} l(\theta, H).$$

В третьем случае предполагается, что H — это случайная величина, плотность распределения которой описывается некоторой невырожденной функцией g :

$$H \sim g(h, \delta),$$

тогда условная плотность вероятности внешних потерь должна иметь вид:

$$f^*(\zeta, \theta | H = h) = 1\{\zeta \geq h\} \cdot \frac{f(\zeta, \theta)}{\int_h^{\infty} f(x, \theta) dx}.$$

По формуле полной вероятности получаем:

$$f^*(\zeta, \theta, \delta) = \int_0^{\infty} f^*(\zeta, \theta | H = h) g(h, \delta) dh.$$

Задача максимизации решается при помощи численных методов. Авторами [4] рекомендуется использовать именно третий метод, и они демонстрируют, как гипотеза о некотором постоянном значении H может приводить к использованию лишь малого числа из всех имеющихся внешних данных. При предположении о стохастическом ха-

рактере порога H для оценки параметров распределения потерь будут использоваться все данные.

Открытым вопросом является выбор закона распределения величины порога H . Этот закон должен отражать то, каким образом распределены банки в контексте их возможностей по раскрытию своих внутренних операционных потерь. Одно из предложений состоит в использовании масштабированного бета-распределения $B(\alpha, \beta, H^-, H^+)$:

$$g(h, \alpha, \beta) = \frac{(H^+ - H^-)^{-1}}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1},$$

где

$$x = \frac{h - H^-}{H^+ - H^-} \wedge 1,$$

а H^- и H^+ — это наименьший порог (например, нулевой или равный значению, принятому консорциумом) и наибольший порог, относящийся к банку с наихудшими возможностями раскрытия своих данных о потерях.

Предложенный метод налагает достаточно существенные ограничения на размер выборки, так как чем меньше данных, тем более смещенные результаты даст метод максимального правдоподобия. Существуют и альтернативные подходы к оцениванию параметров распределения при совмещении внутренних и внешних данных, одним из которых является группа так называемых байесовских методов, о которых будет рассказано далее. Одним из плюсов байесовского подхода по сравнению с предложенным выше является автоматический учет экспертных оценок, которые банк обязан использовать для соответствия требованиям по использованию АМА.

4. Влияние размера выборок на точность оценки

Центральной проблемой подготовки статистической информации для использования метода LDA является недостаточное количество данных о потерях. Это отмечается в большом количестве научных работ [6]. Связь размера выборки событий операционного риска и точности оценок исследована в работе [5]. Рассмотрим непараметрическую оценку значения квантиля q_p на уровне p , основан-

ную на достаточно большом числе наблюдений n , полученных из распределения с тяжелым хвостом, имеющего плотность f . Аппроксимация стандартной ошибки q_p , согласно [8], выглядит следующим образом:

$$\sigma_{\text{aprx}} = \frac{1}{f(q_p)} \sqrt{\frac{p(1-p)}{n}}.$$

Из этой формы видно, что значение плотности распределения присутствует в знаменателе, и при значениях p , близких к единице, это будет оказывать существенный эффект на размер ошибки. Таким образом, для того чтобы получить относительную ошибку в пределах 10%, потребуются выборка приблизительно величины $n_{\pm 10\%}$, где

$$n_{\pm 10\%} = 400 \cdot \frac{p(1-p)}{(f(q_p)q_p)^2},$$

если относительная ошибка измеряется как:

$$\sigma_{\text{rel}} = \frac{2\sigma_{\text{aprx}}}{q_p}.$$

Допустим, функция распределения операционных потерь имеет форму Парето:

$$F(x) = 1 - x^{-\alpha}, \alpha > 0, x \geq 1.$$

Тогда для получения относительной ошибки в пределах $\pm 10\%$ при значении квантиля $p = 0,999$ и параметра $\alpha = 1,2$ потребуются выборка из 277 500 записей о потерях, которой не располагает ни один банк. Для сравнения, если для некоторой группы однородных событий предположить экспоненциальное распределение потерь с параметром, равным 1, то нужный размер выборки $n_{\pm 10\%}$ составит всего 8400 записей. На деле это означает, что для корректной оценки банк должен собирать данные о потерях в течение тысяч лет, тогда как на сегодня среди российских банков редким является накопление данных в течение последних десяти лет.

Для решения данной проблемы используется экстраполирование за пределы имеющихся данных. Экстраполяция возможна только на основе параметрических моделей, и то, насколько хорошо модель

работает при экстраполяции, напрямую зависит от того, насколько верно она отражает механизмы, лежащие в основе генерации операционных потерь. Существует несколько обобщенных подходов к моделированию с учетом наличия или потенциально появления экстремальных значений.

1. В основе всех наблюдаемых потерь лежит единый механизм, и предполагается, что будущие потери, величина которых выйдет за рамки наблюдаемых ранее значений, также будут результатом действия этого механизма.

2. Разные события потерь могут быть вызваны разными механизмами, при этом некоторые из них более вероятно могут привести к экстремальным потерям.

3. Экстремальные потери являются аномальными, и их появление не следует какому-либо правилу, выводимому из статистического анализа прочих данных о потерях.

В статье [10] было показано, что при наличии в выборке некоторого количества крупных потерь, они могут существенно «загрязнить» итоговые оценки в случае использования единой статистической модели. Вместо этого предлагается определить, вероятно ли повторение подобных событий в будущем. Скорее всего, факт крупной потери заставит банк пересмотреть некоторые бизнес-процессы во избежание повторения таких событий, и тогда данные о них нужно изъять из выборки. Если повторение крупных потерь, вызванных теми же причинами, все-таки представляется вероятным, то моделирование экстремальных потерь не должно учитывать другие, менее тяжелые потери, что в целом соответствует третьему подходу.

Заключение

С помощью имеющегося математического аппарата можно обоснованно совмещать данные об операционных потерях из внутренних и внешних источников. Однако большинство моделей являются достаточно требовательными к объему выборки, и это имеет особую важность для оценки операционного риска, когда речь идет об оценке очень высоких квантилей за пределами имеющихся в выборке значений.

Есть несколько шагов, которые можно предложить банкам и регуляторам для последовательного

решения проблем, касающихся недостатка данных об операционных потерях. Первый шаг — консолидация данных об операционных потерях разных финансовых институтов, что на самом деле является единственным средством для восполнения нехватки внутренних данных. Инструментами, действующими «снизу», могут быть создание или расширение консорциумов, «сверху» — создание регулятивной базы, направленной на разработку и улучшение стандартов отчетности.

Вторым шагом по преодолению проблем, вызванных недостатком данных, может быть выявление причинных связей между ошибками бизнес-процессов и операционными потерями с целью создания динамических моделей механизмов, вызывающих потери. Такие модели могли бы увеличить уверенность в допустимости экстраполяции за пределы имеющихся данных, исполнять требования надзорных органов по использованию в АМА сценарного анализа, а кроме того, помочь риск-менеджерам улучшать бизнес-процессы, чтобы уменьшить риск, а следовательно, и отчисляемый капитал.

Третьим шагом, направленным «сверху», могло бы стать ослабление надзорных стандартов уровня доверия. Расширение доверительного интервала, допустим, до 95% могло бы существенно увеличить достоверность выхода модели и позволить банку сконцентрироваться на улучшении робастности модели, используя получаемые оценки в принятии решений в ежедневной практике. Для того чтобы реально приблизиться к уровню доверия в 99,9%, возможным было бы использование некоего простого метода экстраполяции, например введение регулятором фиксированного повышающего коэффициента для результатов LDA с более низким уровнем доверия.

Решение проблемы недостатка данных требует вклада от каждого банка — только раскрывая данные о потерях, можно нарастить требуемый объем статистической информации, необходимый для получения состоятельных оценок. Задача регулятора — способствовать унификации стандартов отчетности и работать над улучшением итоговой меры оцениваемого капитала согласно текущим возможностям и потребностям банковской индустрии.

Литература

1. Золотарева Е.Л. Математическое моделирование операционного риска в коммерческом банке: дис.... канд. экон. наук. Москва, 2011.
2. О порядке расчета размера операционного риска: положение Банка России от 03.11.2009 № 346-П (ред. от 03.07.2012) // Вестник Банка России. 2009. № 77.
3. Anghelache, G.-V. Operational Risk — An Assessment at International Level / G.-V. Anghelache, B.-O. Cozmanca, C.-A. Handoreanu, C. Obreja, A.-C. Olteanu, A.-N. Radu // International journal of mathematical models and methods in applied sciences. 2011. Vol. 1(5). С. 184—192.
4. Baud N., Frachot A., Roncalli T. How to avoid overestimating capital charge for operational risk? (February 2003) // Thierry Roncalli's Home Page. URL: <http://www.thierryroncalli.com/download/oprisktech.pdf> (дата обращения: 10.02.2016)
5. Challenges in Measuring Operational Risk from Loss Data [Электронный ресурс]: офиц. сайт ORX. Цюрих, 2016. URL: https://www.orx.org/Lists/PublicDocuments/Challenges_in_Measuring_OpRisk_from_Loss_Data_9September2009.pdf (дата обращения: 15.12.2015).
6. Chernobai A., Rachev S., Fabozzi F. A Guide to Basel II Capital Requirements, Models, and Analysis. Hoboken: John Wiley & Sons, 2007. 300 p.
7. Frachot A., Roncalli T. Mixing internal and external data for managing operational risk (January 29, 2002) // Thierry Roncalli's Home Page. URL: <http://www.thierryroncalli.com/download/mixing-riskop.pdf> (дата обращения: 10.02.2016)
8. Kiefer J. On Bahadur's representation of sample quantiles // Ann. Math. Statist. 1967. Vol. 38 (5). С. 1323—1342.
9. Mignola G. Effect of a data collection threshold in the loss distribution approach / G. Mignola, R. Ugocioni // Journal of Operational Risk. 2007. Vol. 1 (4). С. 35—47.
10. Nešlehová J. Infinite-mean models and the LDA for operational risk / J. Nešlehová, P. Embrechts, V. Chavez-Demoulin // Journal of Operational Risk. 2006. Vol. 1 (1). С. 3—25.
11. Panjer H. Operational Risk: Modeling Analytics. — Hoboken: John Wiley & Sons, 2006. 431 p.
12. The 2002 Loss Data Collection Exercise for Operational Risk: Summary of the Data Collected // Basel Committee on Banking Supervision, March 2003. URL: <http://www.bis.org/bcbcs/qis/ldce2002.pdf> (дата обращения: 10.02.2016).

Сведения об авторах

Кох Лариса Вячеславовна: доктор экономических наук, профессор Международной высшей школы управления, Санкт-Петербургский государственный политехнический университет Петра Великого

Количество публикаций: 96 научных публикаций, 16 учебно-методических изданий

Область научных интересов: риск-менеджмент, банковская деятельность

Контактная информация:

Адрес: 195251, г. Санкт-Петербург, Политехническая, д. 29

Тел.: +7 (904) 338-42-37

E-mail: lkokh@mail.ru

Булацкий Станислав Михайлович: аспирант Международной высшей школы управления, Санкт-Петербургский государственный политехнический университет Петра Великого

Количество публикаций: 10 научных публикаций

Область научных интересов: банковское дело, риск-менеджмент, математические методы в экономике

Контактная информация:

Адрес: 195251, г. Санкт-Петербург, Политехническая, д. 29

Тел.: +7 (911) 919-90-92

E-mail: bulatsky@gmail.com