

УДК 330.43:336.71

ПРОГНОСТИЧЕСКИЙ И ПРЕДУПРЕДИТЕЛЬНЫЙ ПОТЕНЦИАЛ БИНАРНОЙ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ В ПРОБЛЕМЕ БАНКРОТСТВА КОММЕРЧЕСКИХ БАНКОВ

Бычков В.Е., Яшина Н.И.

Национальный исследовательский Нижегородский государственный университет им. Н.И.

Лобачевского

E-mail: bychara92@yandex.ru

В статье обосновывается использование в качестве основного инструмента прогнозирования банкротства банков метод бинарной логистической регрессии. Предлагается альтернативная последовательность этапов формирования бинарных классификаций. Рассмотрены преимущества использования ROC-анализа при оценке эффективности модели. Раскрыт потенциал модели не только как инструмента прогнозирования банкротства, но и как инструмента предупреждения банкротства. Предлагается использовать модель в качестве фундамента системы нормативов финансовой устойчивости.

Ключевые слова: прогнозирование банкротства банков, бинарная логистическая регрессия, финансовая устойчивость, ROC-анализ.

PREDICTED AND PREVENTIVE POTENTIAL OF BINARY LOGISTIC REGRESSION IN THE PROBLEM OF COMMERCIAL BANKS' BANKRUPTCY

Bychkov V.E., Yashina N.I.

This article gives reason to use binary logistic regression method in the function of the main tools of bank bankruptcy prediction. The article suggests an alternative order of formation stages in binary classification. The advantages of using ROC-analysis in determining the level of efficiency of the model are considered as well. The article also shows that the model potential could be used as means to prevent bankruptcy, not only to predict it. The model is suggested to be used as the basis for criteria system of financial soundness.

Keywords: bank bankruptcy prediction, binary logistic regression, financial soundness, ROC-analysis.

За последние несколько лет проблема прогнозирования банкротства банков (в данном случае под банкротством в узком смысле понимается отзыв лицензии на осуществление банковских операций) стала особенно острой как для внутреннего менеджмента самих банков, так и для их

клиентов и кредиторов. Только за 2014-2016 годы Банк России отозвал лицензии у 282 банков. Основными причинами этому послужили нарушение банковского законодательства (около 85% банков), в том числе легализация доходов преступных путем, а также неудовлетворительное финансовое положение (около 30% банков) [1]. В результате чего около 30% требований кредиторов 1 и 2 очереди и около 95% требований кредиторов 3 очереди были не удовлетворены, что обусловлено низким качеством активов банков, у которых была отозвана лицензия. Таким образом, для обеспечения эффективного функционирования банка, а также для принятия решения о его кредитовании необходимо не только осуществлять анализ его финансовой устойчивости, но и проводить раннюю диагностику его возможного банкротства в будущем.

На сегодняшний день методология прогнозирования банкротства обширна и включает большое количество математических и статистических методов, позволяющих разрабатывать модели прогнозирования банкротства компании или банка. Однако среди многообразия существующих зарубежных и отечественных моделей нет модели, успешно применимой к современным российским условиям. Первые в силу различий законодательства, форм отчетностей банков, макроэкономических ситуаций в странах [3]. Вторые являются либо неудачно адаптированными зарубежными моделями [2], либо успешно применяются только при определенных условиях, например, только в одном субъекте РФ или только во время финансового кризиса, или только в рамках самой обучающей выборки.

При разработке моделей прогнозирования банкротства чаще всего используется множественный дискриминантный анализ, однако наиболее целесообразно использовать бинарную логистическую регрессию, т.к. результаты, полученные при помощи такой модели проще интерпретировать. Наиболее важным критерием является возможность использования модели в дальнейшем для разработки нормативов показателей, т.к. в данном случае модель имеет не только прогностический потенциал, но и возможность выполнять предупреждающую банкротство функцию. Преимущество бинарной логистической регрессии здесь обусловлено простой интерпретацией получаемых результатов, обеспечивающих наибольшую четкость, понятность и прозрачность разрабатываемых нормативов. Также данная модель была признана наиболее подходящей для стран с развивающейся экономикой [6].

Логистическая регрессия является разновидностью множественной регрессии, и позволяет находить взаимосвязь между независимыми переменными и зависимой переменной [8]. Бинарная логистическая регрессия широко используется в медицине, банковском скоринге и идеально подходит для прогнозирования банкротства банков, т.к. зависимая переменная может принимать только два значения – банкрот/не банкрот.

Для обоснования использования метода бинарной логистической регрессии необходимо сравнить его с наиболее популярными методами по ключевым параметрам, отражающим не только качество модели, но и в перспективе использования получаемых результатов в дальнейшем (табл. 1).

Таблица 1. Сравнительный анализ метода бинарной логистической регрессии с наиболее популярными методами

Критерий сравнения	Множественный дискриминантный анализ	Нейронные сети	Бинарная логистическая регрессия
Точность модели	Высокая	Самая высокая	Высокая
Степень сложности интерпретации результатов	Высокая	Самая высокая	Низкая
Степень сложности использования модели	Низкая	Высокая	Низкая
Возможность использования модели в дальнейшем для разработки нормативов показателей	Есть	Нет	Есть

Необходимость использования бинарной логистической регрессии объясняется тем, что применение множественной линейной регрессии (1) приведет к любым значениям зависимой переменной, в то время, когда необходимо получить результат 1 (банкрот) или 0 (не банкрот). Для этого с помощью логит-преобразования (2) ставится задача спрогнозировать не саму бинарную переменную, а непрерывную переменную со значениями на отрезке $[0;1]$. В таком случае можно оценить вероятность банкротства банка в заданный промежуток времени.

$$y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n, \quad (1)$$

где y – уравнение регрессии;

b_j – коэффициенты регрессии;

x_j – независимые переменные;

n – количество независимых переменных.

$$p = \frac{1}{1+e^{-y}}, \quad (2)$$

где p – вероятность наступления события;

e – основание натурального логарифма; $e \sim 2,718$;

y – стандартное уравнение регрессии.

Процесс построения бинарной классификации, т.е. получения итогового уравнения регрессии, традиционно включает в себя 4 этапа [2, 3]:

1. Формирование обучающей статистической выборки;
2. Отбор независимых переменных;
3. Построение самой модели;
4. Апробация модели.

Однако на первом этапе построения бинарной классификации целесообразно производить отбор предикторов, поскольку в ином случае теряется логика построения модели. Если изначально сформировать выборку данных, это может привести к некорректному подбору нужных предикторов, исходя из имеющихся данных, или их использованию в недостаточном или излишнем количестве для правильной интерпретации результатов модели и модель будет низкоэффективной. В модель должны быть включены факторы, наиболее точно отражающие финансовую устойчивость банка, такие как, например, банковские риски (кредитный, операционный, рыночный и т.д.), качество собственного капитала банка, качество активов банка, ликвидность, рентабельность и т.п. В качестве фундаментального примера можно использовать зарубежные блоки независимых переменных CAMELS [9] или 6i [10].

Большинство существующих моделей прогнозирования банкротства банков ограничиваются набором предикторов, включающих либо только нормативные показатели, используемые Банком России, либо отражающие только состояние ликвидности и капитала, либо вообще не учитывают финансовые показатели. Для построения успешной модели, которая будет работать не только в определенных условиях (в кризис или в отдельном регионе) необходимо учитывать по возможности все факторы, отражающие деятельность коммерческого банка, которые условно можно разделить на 3 группы: финансовые предикторы, нефинансовые предикторы и макроэкономические условия (рис. 1).

Далее используя, например, метод 20 EPV [7] необходимо сформировать выборку, достаточную для разработки бинарной классификации. Для определения необходимого количества наблюдений можно воспользоваться авторской формулой (3).

$$\text{НКН} = \text{КФ} * 20 * 1,3, \quad (3)$$

где НКН – необходимое количество наблюдений для построения эффективной модели прогнозирования банкротства банков;

- КФ – количество факторов, наиболее полно отражающих финансовую устойчивость банка;
 - 20 – по методу 20 EPV на 1 предиктор должно быть не менее 20 наблюдений;
 - 1,3 – дополнительные 30% наблюдений потребуются для апробации модели.
-

Для достижения высокой эффективности модели необходимо использовать в обучающей выборке ($KФ*20$) равное количество положительных и отрицательных наблюдений, т.е. банков банкротов и нет. Тестовая выборка используется после построения модели для проверки ее достоверности [4].



Рисунок 1. Набор факторов, необходимых для построения успешной модели прогнозирования банкротства банков

На третьем этапе формируется непосредственно сама бинарная классификация. На четвертом этапе проводится апробация модели и оценивается ее эффективность, например, с помощью ROC-анализа [5]. ROC-анализ также позволит скорректировать модель под определенную цель, например минимизировать риск потерь вследствие открытия депозитного счета в банке посредством ограждения компании от банков-банкротов. В таком случае модель с большей точностью будет предсказывать возможность банкротства банка, но может неправильно классифицировать банки с хорошей финансовой устойчивостью как будущих банкротов. И наоборот, если необходимо найти банк, с хорошей финансовой устойчивостью в будущем, модель

может ложно классифицировать банк с плохой финансовой устойчивостью как надежный. На практике используется некоторый баланс соотношения истинно положительных и истинно отрицательных случаев, классифицируемых моделью.

Итоговая модель может использоваться на практике внешними субъектами (предприятиями и банками – для принятия управленческих решений, аналитиками и проч.) и менеджерами самого банка в прогностических целях. Однако конечной целью можно представить не только раскрытие прогностического потенциала модели, но и предупредительного. По типу существующих систем нормативов, которые используются для оценки финансовой устойчивости коммерческого банка, на основе модели логично разработать систему нормативов, отражающую «реальное» состояние кредитной организации, в которую войдут показатели, по сути являющиеся преобразованными предикторами модели в нормативные показатели, с заданными интервалами устойчивости, весом в системе и т.д. Такая система способна отражать не только финансовое состояние банка, но и выявлять угрозы, которые могут привести к банкротству, включая изменение макроэкономических условий, сигналы о «bank run» и проч. Система также может использоваться как внешними аналитиками, так и менеджерами банка для принятия управленческих решений.

Важно отметить, что на практике ни одна методика прогнозирования банкротства не может обеспечить точность 100%, и модель считается хорошей, если достигает 80-90% точности. При разработке бинарной классификации можно столкнуться с универсальными проблемами прогнозирования, сложностью создания выборок, определения факторов, охватывающих все аспекты деятельности банка и т.д. Также следует учитывать, что помимо анализа финансовых показателей, которые могут быть неполными, неточными, искаженными и т.п., необходимо учитывать нефинансовые факторы, приводящие к банкротству банков, такие как нарушение законодательства, фиктивное банкротство и т.п. Их практически невозможно идентифицировать и использовать в модели, в то время когда они во многих случаях являются решающими.

Список литературы

1. О ликвидации кредитных организаций (по состоянию на 01.01.2016). – URL: http://www.cbr.ru/credit/likvidbase/information_01012016.pdf
 2. Буздалин А.В. Особенности анализа надежности банка в долгосрочном периоде // Бюллетень финансовой информации, 2004. №8.
 3. Плещицер М.В. Комплексная методика прогнозирования банкротства банков // Финансовая аналитика: проблемы и решения. 2011. №14.
 4. Сорокин Александр Сергеевич Построение скоринговых карт с использованием модели логистической регрессии // Науковедение. 2014. №2 (21).
-

5. Bradley, A. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30, 1145–1159.
 6. Ciotina D., Ciotina I. M. (2013). Symptoms of Bankruptcy and Prediction Models of Bankruptcy Risk. *Anale. Seria Stiinte Economice*. – Timisoara, 2013, vol. XIX – p. 114-121
 7. Harrell, Frank. (2001). *Regression modeling strategies*. [Text] – NY: Springer. – 608 p. – ISBN 0387952322, 9780387952321
 8. Hosmer D., Lemeshow S. (1989, 2000, 2013). *Applied logistic regression*. [Text] – New York: John Wiley and Sons. – 528 p. – 3rd ed. – ISBN 0470582472, 9780470582473
 9. Gamze Özel. (2013). Probabilistic Prediction of Bank Failures with Financial Ratios: An Empirical Study on Turkish Banks – *Pak.j.stat.oper.res*. Vol.IX No.4 2013 – p. 407-426
 10. Vaithilingam S., Nair M., Samudra M. (2006). Key drivers for soundness of the banking sector: lessons for developing countries – *Journal of Global Business and Technology*, Volume 2, Number 1, Spring 2006 – p. 1-11
-