

АПРОБАЦИЯ РЕЗУЛЬТАТОВ НЕЛИНЕЙНОЙ РЕГРЕССИОННОЙ ЛОГИТ-МОДЕЛИ ПРОГНОЗИРОВАНИЯ РИСКА БАНКРОТСТВА ПРЕДПРИЯТИЙ И ОПРЕДЕЛЕНИЕ ЕЕ ОПТИМАЛЬНЫХ ПОРОГОВЫХ ЗНАЧЕНИЙ



Т.С. Космыкова

Главный специалист ОАО «Банк БелВЭБ», заместитель декана инженерно-экономического факультета по научно-исследовательской работе студентов БГУИР, ассистент кафедры экономической информатики, магистр экономических наук, магистр технических наук

Белорусский государственный университет информатики и радиоэлектроники, Республика Беларусь
E-mail: t.kosmykova@gmail.com

Abstract. This article is about the models of binary choice, that can be used to predict the risk of bankruptcy. There is some results of constructing models of binary choice in this article. This scientific material presents information about these models and their predictive ability, and also it includes the stages of model valuing. This article is focus on the critical values for the model for bankruptcy risk prediction and their determination. It is noted that the model is good for the bankruptcy risk prediction.

В качестве аппарата для моделирования риска банкротства предприятий реального сектора экономики целесообразно использовать нелинейную логистическую регрессионную модель.

В статье «Моделирование риска банкротства предприятий реального сектора экономики Республики Беларусь» [1] рассмотрен процесс построения данного типа моделей. Наилучший результат приобрела модель следующего вида:

$$\varphi(z) = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}}, \quad (1)$$

где z – основание нелинейной логит-модели, представленное в виде пятифакторной регрессионной модели со следующей спецификацией:

$$z=7,88+5,78x_1+8,75x_2+60,75x_3+17,96x_4+3,51x_5, \quad (2)$$

где x_1 – коэффициент обеспеченности собственными оборотными средствами,

x_2 – коэффициент финансовой независимости (автономии),

x_3 – коэффициент абсолютной ликвидности,

x_4 – темп прироста выручки,

x_5 – показатель качества кредитной истории предприятия.

Прежде, чем приступать к определению пороговых значений, представляется целесообразным провести анализ распределения итоговых значений модели прогнозирования риска банкротства на обучающей выборке, на основании которой происходило ее построение.

Диапазон изменения результирующего показателя находится в пределах от 0 до 1. Чем

ближе расчетное значение к 0, тем ближе организация к состоянию банкротства, чем ближе расчетное значение к 1, тем ближе организация к состоянию финансовой устойчивости.

Обучающая выборка была сформирована из 219 предприятий: 65 относящихся к категории «банкрот» и 154 организаций без такого признака.

При этом, если значение результирующего показателя для *i*-ой организации было менее или равное 0,55, то значение классифицировалось как стремящееся к 0, а значение более 0,56 – как стремящееся к 1. Выявим ошибки 1-го и 2-го рода для обучающей выборки. Результаты содержатся в таблице 1.

Таблица 1. Результаты ошибок 1-го и 2-го рода для обучающей выборки

	Всего значений	Значение 1	Значение 0	Процент верно предсказанных значений
Значение 1	154	149	5	96,64%
Значение 0	65	5	60	92,31%
	219	Итого верных предсказаний: 209	Итого неверных предсказаний: 10	95,43%

Анализ данных таблицы показал, что на обучающей выборке модель дала высокие результаты и показала высокую прогностическую способность.

Проведем анализ качества предсказательной способности модели на выборках последующих периодов:

По состоянию на 01.01.2016 выявим ошибки 1-го и 2-го рода. Результаты приведены в таблице 2.

Таблица 2. Результаты ошибок 1-го и 2-го рода для выборки предприятий по состоянию на 01.01.2016

	Всего значений	Значение 1	Значение 0	Процент верно предсказанных значений
Значение 1	1556	1519	37	97,6%
Значение 0	81	5	76	93,8%
	1637	Итого верных предсказаний: 1595	Итого неверных предсказаний: 42	97,4%

На выборке по состоянию на 01.01.2016 модель дала высокие результаты и высокую прогностическую способность.

По состоянию на 01.04.2016 ошибки 1-го и 2-го рода следующие. Результаты приведены в таблице 3.

Таблица 3. Результаты ошибок 1-го и 2-го рода для выборки предприятий по состоянию на 01.04.2016

	Всего значений	Значение 1	Значение 0	Процент верно предсказанных значений
Значение 1	1317	1173	144	89,1%
Значение 0	104	8	96	92,3%
	1421	Итого верных предсказаний: 1269	Итого неверных предсказаний: 152	89,3%

На выборке по состоянию на 01.04.2016 модель также дала высокие результаты и хорошую прогностическую способность.

Снижение процента верно предсказанных показателей по сравнению с 01.01.2016 вызвано «выравниванием» показателей клиентов, стремящихся наилучшим образом «закрыть»

год.

По состоянию на 01.07.2016 выявим ошибки 1 и 2-го рода. Результаты приведены в таблице 4.

Таблица 4. Результаты ошибок 1-го и 2-го рода для выборки предприятий по состоянию на 01.07.2016

	Всего значений	Значение 1	Значение 0	Процент верно предсказанных значений
Значение 1	1248	1139	109	91,3%
Значение 0	98	9	89	90,8%
	1346	Итого верных предсказаний: 1228	Итого неверных предсказаний: 118	91,2%

На выборке по состоянию на 01.07.2016 модель дала хорошие результаты и хорошую прогностическую способность.

По состоянию на 01.10.2016 выявим ошибки 1-го и 2-го рода. Результаты приведены в таблице 5.

Таблица 5. Результаты ошибок 1-го и 2-го рода для выборки предприятий по состоянию на 01.10.2016

	Всего значений	Значение 1	Значение 0	Процент верно предсказанных значений
Значение 1	1171	1097	74	93,7%
Значение 0	101	9	92	91,1%
	1272	Итого верных предсказаний: 1189	Итого неверных предсказаний: 83	93,5%

На выборке по состоянию на 01.10.2016 модель дала хорошие результаты и высокую прогностическую способность.

По состоянию на 01.01.2017 выявим ошибки 1-го и 2-го рода. Результаты приведены в таблице 6.

Таблица 6. Результаты ошибок 1-го и 2-го рода для выборки предприятий по состоянию на 01.01.2017

	Всего значений	Значение 1	Значение 0	Процент верно предсказанных значений
Значение 1	918	882	36	96,1%
Значение 0	87	5	82	94,3%
	1005	Итого верных предсказаний: 964	Итого неверных предсказаний: 41	95,9%

На выборке по состоянию на 01.01.2017 модель дала высокие результаты и высокую прогностическую способность.

Увеличение процента верно предсказанных показателей по сравнению с 01.10.2016 вызвано «выравниванием» показателей клиентов, стремящихся наилучшим образом «закрыть» год.

Анализ полученных результатов показал, что модель в целом пригодна для прогнозирования риска банкротства предприятий.

Теперь представляется целесообразным произвести градирование итоговых значений интегрального показателя модели, расклассифицировав организации на большее количество групп, посредством установления пороговых значений для модели.

Для этого проанализируем обучающую выборку. Проведем графический анализ и построим диаграмму рассеивания значений, полученных в результате апробации модели на обучающей выборке (приведена на рисунке 1).

Из полученной диаграммы видно, что имеется возможность в зависимости от диапазона, в который попадает расчетное значение конкретного наблюдения, разделить полученные данные не менее, чем на 4 группы.



Рис. 1. Диаграмма рассеивания значений, полученных в результате апробации модели на обучающей выборке

Для более точного определения «оптимального» числа интервалов воспользуемся формулой Старджесса [2]:

$$k = \log_2 N + 1 = 3,322 \lg N + 1, \quad (3)$$

где N – количество встречающихся в обучающей выборке повторяющихся значений результирующего показателя.

Исходя из полученного интегрального показателя модели, предприятия распределились следующим образом (приведено в таблице 7):

Таблица 7. Распределение предприятий в зависимости от значений интегрального показателя модели

Значение интегрального показателя	0,00	0,04	0,07	0,08	0,13	0,15	0,19
Количество предприятий	35	8	2	1	3	4	1
Значение интегрального показателя	0,20	0,25	0,31	0,37	0,41	0,48	0,51
Количество предприятий	1	1	1	1	1	1	1
Значение интегрального показателя	0,53	0,54	0,57	0,69	0,86	1,00	Всего
Количество предприятий	1	1	10	20	78	47	219

При $N=20$, значение $k=5,29$ (или 5 при округлении). Таким образом, для модели можно выделить 5 интервалов пороговых значений.

Определим шаг изменения интервалов по формуле:

$$h = \frac{n_{\max} - n_{\min}}{k}, \quad (4)$$

где n_{\max} – максимальное значение результирующего показателя,
 n_{\min} – минимальное значение результирующего показателя,
 k – оптимальное количество интервалов.

Следовательно, для анализируемых данных, шаг изменения интервалов h равен 0,19.

Таким образом, получаем следующие интервальные значения модели:

- от 0,00 до 0,19, в данный интервал входят организации категории «банкрот»,
- от 0,20 до 0,39, к данному интервалу относятся организации, близкие к банкротству
- от 0,40 до 0,59, в интервал входят предприятия, имеющие признаки финансовой неустойчивости,

- от 0,60 до 0,79, интервал стабильных организаций,
- от 0,80 до 1,00, интервал финансово устойчивых организаций.

При этом, для простоты отнесения предприятий к финансовой устойчивым предприятиям, предприятиям с признаками финансовой неустойчивости и предприятиям-банкротам представляется целесообразным указанные интервалы укрупнить следующим образом:

- от 0,00 до 0,39 – предприятия-банкроты,
- от 0,40 до 0,59 – предприятия, с признаками финансовой неустойчивости,
- от 0,60 до 1,00 – финансово устойчивые предприятия.

Литература

[1]. Космыкова Т.С. Моделирование риска банкротства предприятий реального сектора экономики Республики Беларусь / Т.С. Космыкова // Материалы XXIII Междунар. науч.-практ. конф. «BIG DATA and Advanced Analytics. Conference and EXPO», 3-4 мая 2017 / г. Минск, Республика Беларусь – 2017.

[2]. Выбор числа интервалов [Электронный ресурс], режим доступа: https://www.ami.nstu.ru/~headrd/seminar/xi_square/28.htm. – М., 2017.