

Олейник Елена Борисовна, д-р экон. наук, доцент
Дальневосточный федеральный университет, г. Владивосток
e-mail: oleinik.elena@gmail.com
Ли Анна Владимировна, магистрант
Дальневосточный федеральный университет, г. Владивосток
e-mail: anna_li_b1104@mail.ru

ПРИМЕНЕНИЕ ЛОГИСТИЧЕСКОЙ РЕГРЕССИИ ДЛЯ ОЦЕНКИ КРЕДИТНОГО РИСКА¹

Аннотация. Рассматривается количественная оценка кредитного риска коммерческого банка на основе модели логистической регрессии. Дается обоснование применения метода максимального правдоподобия для оценки коэффициентов регрессии. Предложен набор факторов и интерпретация коэффициентов модели.

Ключевые слова: кредитный риск, логистическая регрессия, кредитоспособность, вероятность дефолта, случайная компонента.

Основным назначением банка является осуществление операций по кредитованию. Поэтому основным риском, влияющим на деятельность коммерческого банка, является кредитный риск.

Проведение оценки кредитного риска, то есть определение максимально возможного убытка, который может быть получен банком с заданной вероятностью в течение определенного периода времени, является актуальной задачей в деятельности коммерческого банка. С количественной точки зрения кредитный риск представляет собой функцию параметров займа и заемщика. Степень риска, связанного с определенным заемщиком и видом кредита, базируется на оценке различных видов риска, которые возникают для банка при предоставлении кредита [1, 68].

В целях минимизации кредитных рисков необходимо давать оценку кредитоспособности заемщика, которая основана на прогнозировании банкротства кредитруемого лица. Рассмотрим logit-модели, которые относятся к группе моделей бинарного выбора для оценки вероятности дефолта. Математический аппарат логистической регрессии предназначен для решения задач предсказания значения непрерывной зависимой переменной при условии, что эта зависимая переменная может принимать значения от 0 до 1. В силу такой специфики её часто используют для предсказания вероятности наступления некоторого события в зависимости от значений некоторого числа признаков. Основная идея метода заключается в том, чтобы на первом этапе выявить влияющие на кредитоспособность заемщика факторы, а затем на основе их значений оценить вероятность возможного дефолта.

Пусть имеется выборка объема n наблюдений (x_i, y_i) , $i = 1, \dots, n$, где $x_i^T = (1, x_{i1}, \dots, x_{ik})^T$, y_i – зависимая переменная, которая может принимать только два значения: ноль и единица. Рассмотрим стандартную модель линейной регрессии: $y_i = \beta^T x_i + \varepsilon_i$, где $\beta^T = (\beta_0, \beta_1, \dots, \beta_k)$ – вектор неизвестных параметров, ε_i – случайная компонента.

В регрессионном анализе случайная компонента подчиняется нормальному закону распределения с нулевым математическим ожиданием. Учитывая это, получаем: $E y_i = \beta^T x_i$. Т.к. y принимает только два значения, то математическое ожидание имеет вид:

$$E y_i = 1 \times P\{y_i = 1\} + 0 \times P\{y_i = 0\} = P\{y_i = 1\}.$$

Таким образом, получаем модель линейной вероятности:

$$P\{y_i = 1\} = \beta^T x_i.$$

¹ Работа выполнена при поддержке Министерства образования и науки РФ. Госзадание №1398.2014

В данном случае нельзя успешно применить метод наименьших квадратов для оценки параметров β и построения прогнозов, так как в этом случае прогнозные значения вероятности $P\{y_i = 1\}$ могут принимать как отрицательные значения, так и значения больше единицы. Поэтому для оценки коэффициентов модели используют метод максимального правдоподобия, который заключается в оценивании параметров путём максимизации функции правдоподобия, которая имеет вид:

$$L = L(Y_1, \dots, Y_n) = \prod_{i: Y_i = 0} (1 - F(X_i^T \beta)) \prod_{i: Y_i = 1} F(X_i^T \beta)$$

Поскольку Y_i является бинарной переменной, то вид функции упрощается и принимает вид:

$$L = L(Y_1, Y_2) = \prod_{i=1}^n [F(X_i \beta^T)]^{y_i} [1 - F(X_i \beta^T)]^{1-y_i}$$

Logit-модель включает два типа переменных: зависимую бинарную переменную вида $y = \{1, 0\}$, которая принимает данные значения в следующих случаях:

$$y = \begin{cases} 1, & \text{заемщик не кредитоспособен} \\ 0, & \text{в противном случае} \end{cases}$$

и совокупность объясняющих переменных, которые образуют вектор

$$X_i = \{X_{i1}, X_{i2}, \dots, X_{in}\}^T.$$

Таким образом, вероятность дефолта i -го заемщика равна вероятности того, что $y = 1$:

$$p_i = P(y_i = 1) = \frac{1}{1 + e^{-z}}$$

Где $P(y_i = 1)$ – вероятность дефолта i -го заемщика, z – управляющий параметр (интегральный показатель), вычисляемый в зависимости от разработанной модели по формуле:

$$z = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_n X_{in},$$

где X_{ij} – значение j -го признака для i -го заемщика, β_j – коэффициент регрессии для j -го признака.

Положительный коэффициент говорит о том, что данный фактор увеличивает общий риск, т.е. повышает вероятность анализируемого исхода. Отрицательный коэффициент означает, что данный фактор уменьшает риск, т.е. понижает вероятность наступления исхода. Результат можно представить в виде таблицы 1. В таблице 1 a_{ij} ($i=0,1; j=0,1$) – это количество значений отклика, для которых наблюдалось значение i , а предсказано значение j ; – количество наблюдаемых значений i в исходной выборке ($i=0,1$); – количество предсказанных по модели значений i ($i=0,1$).

Таблица 1 – расчет процентов правильных предсказаний

| | | Предсказано | | |
|----------|-------|-------------|----------|-------|
| | | 0 | 1 | Сумма |
| Получено | 0 | a_{00} | a_{01} | |
| | 1 | a_{10} | a_{11} | |
| | Сумма | | | S |

Обозначим процент правильных предсказаний за φ , тогда:

$$\varphi_0 = \frac{a_{00}}{S_{0E}} \quad \text{и} \quad \varphi_1 = \frac{a_{11}}{S_{1E}},$$

где φ_0 – процент правильных предсказаний $Y = 0$,

φ_1 – процент правильных предсказаний $Y = 1$.

«Хорошая» модель должна давать высокий процент правильных предсказаний, $\varphi \in [60\%, 75\%]$ обозначает низкий уровень качества модели, $\varphi \in [75\%, 80\%]$ – средний уровень качества, $\varphi \in [80\%, 100\%]$ – высокий уровень [2, 6].

В качестве факторов для модели выбираются финансовые коэффициенты, которые входят в профессиональное суждение для определения финансового состояния заемщика: X_1 – коэффициент финансовой независимости, X_2 – коэффициент обеспеченности собственными средствами, X_3 – коэффициент текущей ликвидности, X_4 – коэффициент срочной ликвидности, X_5 – оборачиваемость оборотных активов.

Список литературы:

- 1 Матигорова, И. Ю. Характеристика основных подходов к оценке кредитного риска / И.Ю. Матигорова // Экономическая наука и практика: материалы междунар. науч. конф. (г. Чита, февраль 2012 г.). — 2012. — С. 68 – 69.
- 2 Тимофеев, В.С. Модели бинарного выбора: оценка качества и интерпретация коэффициентов / В.С. Тимофеев // Сборник трудов НГТУ. – 2005. – №5. – С. 2 – 8.