Статистические оценки показателей валидности внутренних рейтингов банков для случая малых выборок

© 2015 Гуськов Сергей Юрьевич
экономист
Банк ЗЕНИТ
129110, г. Москва, Банный пер., д. 9
© 2015 Л□вин Владимир Владимирович
кандидат физико-математических наук, доцент
Московский государственный технический университет им. Н.Э. Баумана
105005, г. Москва, 2-я Бауманская ул., д. 5, стр. 1
E-mail: OET2004@yandex.ru

Предложены статистические оценки показателей валидности внутренних рейтингов банков на основе группированных данных. Для интервальных доверительных оценок уровни значимости определены на основе точных распределений соответствующих статистик и могут быть использованы для случая малых выборок, когда асимптотические распределения не применимы.

Ключевые слова: показатели валидности, бинарный классификатор, внутренний рейтинг банка, скоринг, функция распределения, доверительный интервал, кумулятивный профиль точности.

Введение

Впервые идея использования внутренних рейтингов банка (IRB approach) в регулировании достаточности капитала была предложена в 1999 г. в документе “Совершенствование корпоративного управления в кредитных организациях”[1]. Данный подход является:

1) более точным в оценках кредитного риска;

2) стимулирующим дальнейшее совершенствование внутрибанковских систем рейтинговой оценки.

Анализ таких систем, проведенный Базельским комитетом[2], свидетельствует, что большинство западных коммерческих банков с успехом рассчитывают показатели кредитоспособности и кредитных рисков на основе внутренних систем оценки. Согласно документу[3] рейтинг заемщика определяется как оценка риска на основе конкретных и четких рейтинговых критериев, из которых выводятся оценки вероятности дефолта (PD).

Под дефолтом понимают невозможность выполнения заемщиком взятых на себя обязательств по кредитному договору в сроки и объемы, установленные договором. В качестве дефолтных на практике используют кредиты, по которым была допущена непрерывная просрочка свыше 90 дней (просрочка 90+). Очевидно, что множество настоящих дефолтов, по которым клиент был признан безнадежным должником, является подмножеством множества кредитов, по которым была зафиксирована непрерывная просрочка 90+.

В целях получения модели внутреннего рейтинга, в частности скоринга высокого качества, необходимо наличие не менее 300 дефолтных кредитов. При недостаточности количества дефолтных кредитов используются другие определения дефолтности, например, множество кредитов с зафиксированной просрочкой 60- или множество кредитов с суммарной накопленной просрочкой 90+, 60+ и т.д.

В материалах Базельского комитета[4] было предложено два IRB-подхода: базовый и усовершенствованный. Базовый подход (F-IRB the Foundation Internal Ratings-Based Approach) предполагает, что банки сами производят расчет вероятности дефолта PD для своих заемщиков. Усовершенствованный подход (A-IRB the Advanced Internal Ratings-Based Approach) заключается в том, что банкам предоставляется возможность использовать свои собственные оценки по всем четырем переменным:

1) PD (probability of default) - вероятность дефолта;

2) LGD (loss given default) - доля потерь по кредиту при наступлении дефолта;

3) EAD (exposure at default) - общие кредитные потери в момент дефолта;

4) M-эффективный срок погашения.

Базельский комитет своим документом “Базель 2”[5] определил порядок использования внутренних рейтингов как основы для рационального формирования резервов в соответствии с кредит-

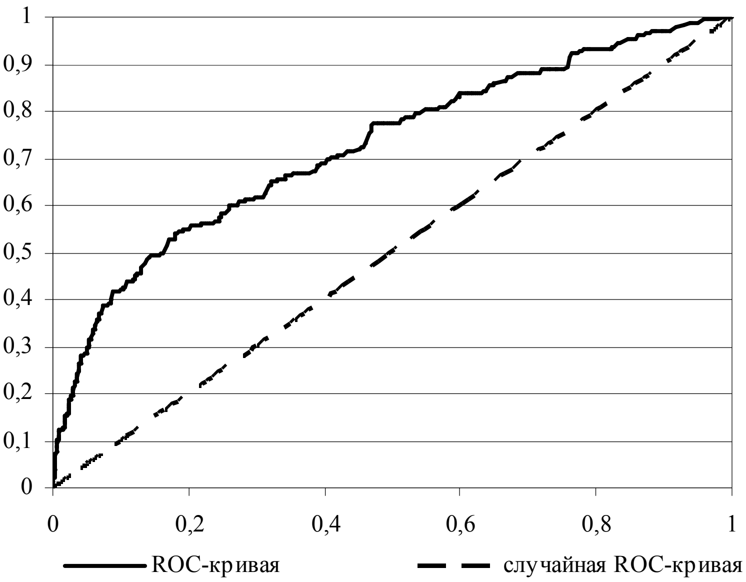

Рис. 1. **Общий вид ROC-кривой**

ным риском. Это ставит перед банками и контролирующими органами задачу разработки статистических инструментов для оценки качества (валидности) моделей внутреннего рейтинга. Важность качества моделей валидности для рейтинговых систем определяется тем, что модели низкого качества могут привести к неоптимальному размещению капитала, в частности к недооценке резервов в случае кризисной ситуации. В своих нормативных документах Базельский комитет отмечает, что область построения адекватных моделей валидности является одним из главных вопросов для финансовых институтов и контролирующих органов.

В статье представлен подход к построению доверительных границ для показателей качества бинарных классификаторов - ROC-кривых и связанных с ними характеристик (AUC, AR, CAP)[6] для случая малых выборок. В банках в качестве бинарного классификатора используются скоринг заявки, поведенческий скоринг и другие внутренние рейтинги, оценивающие вероятность дефолта клиента в течение определенного периода времени. Под дефолтом принято понимать невозможность выполнения обязательств заемщиком. В качестве дефолтных на практике используют кредиты, по которым была допущена непрерывная просрочка свыше 90 дней.

Предложены оценки доверительных границ для эмпирических функций распределения на основе группированных данных с использованием точных доверительных интервалов для пуассоновского распределения[7]. На основе этих оценок построены доверительные границы для ROC-кривых и связанных с ними характеристик[8].

ROC-кривая представляет собой графическую характеристику качества бинарного классификатора, зависимость доли верных положительных классификаций от доли ложных положительных классификаций при варьировании порога решающего правила (рис. 1).

Множество классов $Y = \{-1, +1\}$, здесь классификацию " +1" получают объекты, которые по своим характеристикам могут быть отнесены к положительным объектам, соответственно, классификацию "-1" получают объекты с противоположными характеристиками. В зависимости от содержательной интерпретации задачи к положительным объектам относят объекты, выявление которых необходимо обеспечить в первую очередь, чтобы максимально избежать ошибки неправильной классификации положительных объектов (ошибка I рода).

В медицинской диагностике положительные объекты - это пациенты с признаками заболевания, которые необходимо выявить на ранней стадии заболевания с использованием диагностического теста. В приемочном контроле качества продукции это дефектные изделия, которые необходимо отсеять с помощью процедур приемки качества, в банковском деле - клиенты, у которых высок риск невозврата долга и предоставление кредита которым нежелательно.

Для определенности далее положительные объекты будем обозначать как "плохие" (BAD), соответственно, отрицательные - как "хорошие" (GOOD).

Правило классификации задается в виде функции

$$\alpha(\mathrm{x}) = sign(f(x, w) - w_0),$$

где x - классифицируемый объект,

$f(x, w)$ - дискриминантная функция;

w - вектор параметров, определяемый по обучающей выборке;

w_0 - порог отсечения (cut-off-value), $0 < w_0 < 1$.

Для оценки качества классификаторов рассматриваются следующие абсолютные и относительные показатели[9]. Абсолютные показатели качества различения сопряженности (ошибок) (confusion matrix) определяются для заданного уровня отсечения w_0 (см. таблицу).

Специфичность (Specificity) - доля истинно отрицательных случаев, которые были правильно идентифицированы моделью:

$$\mathrm{Sp} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \cdot 100\ \%.$$

$$FPR = 100\text{-}Sp.$$

$Se = TPR$ - ф.р. "плохих" (BAD) объектов в зависимости от уровня отсечения.

Абсолютные показатели качества различения сопряженности (ошибок) (confusion matrix) (определяются для заданного уровня отсечения w_0)

Модель	Фактически	
	Положительно	Отрицательно
Положительно	*TP*	*FP*
Отрицательно	*FN*	*TN*

В таблице использованы обозначения:

TP (*True Positives*) - верно классифицированные положительные примеры (так называемые истинно положительные случаи);

TN (*True Negatives*) - верно классифицированные отрицательные примеры (истинно отрицательные случаи);

FN (*False Negatives*) - положительные примеры, классифицированные как отрицательные (ошибка I рода). Это так называемый "ложный пропуск" - когда интересующее нас событие ошибочно не обнаруживается (ложно отрицательные примеры);

FP (*False Positives*) - отрицательные примеры, классифицированные как положительные (ошибка II рода); это "ложное обнаружение", так как при отсутствии события ошибочно выносится решение о его присутствии (ложно положительные случаи).

Относительные показатели качества различения (определяются для заданного уровня отсечения w_0). К относительным показателям качества относятся следующие коэффициенты.

Доля истинно положительных примеров (True Positives Rate):

$$TPR = \frac{TP}{TP + FN} \cdot 100\%.$$

Доля ложно положительных примеров (False Positives Rate):

$$FPR = \frac{FP}{TN + FP} \cdot 100\ \%.$$

Чувствительность (Sensitivity) - это доля истинно положительных случаев:

$$Se = TPR = \frac{TP}{TP + FN} \cdot 100\ \%.$$

Все возможные ROC-кривые могут располагаться между ROC-кривой, соответствующей отсутствию различающей способности у классификатора, когда ROC-кривая совпадает с графиком функции $У = X$, и ROC-кривой, соответствующей идеальному классификатору, безошибочно различающему "плохих" от "хороших", у которого ROC-кривая совпадает с графиком зависимости $У = 1$ (рис. 2).

Количественной характеристикой классификатора является показатель AUC (Area Under Curve), равный площади под ROC-кривой. Возможные значения AUC варьируются в интервале [0 %; 100 %]. Приемлемыми считаются классификаторы с AUC $\geq$ 70 %.

Функции распределения доли BAD (GOOD) объектов обозначим $F_B(z)(F_G(z))$. Если они непрерывные для $0 \leqslant z \leqslant 1$, то для ROC-кривой справедливо представление[10]:

$$ROC(z) = F_B\left(F_G^{-1}(z)\right).$$

Соответственно, для AUC справедливо представление[11]:

$$AUC = \int_0^1 TPR(FPR)d(FPR). \qquad (1)$$

Так как показатели ROC-кривая и AUC строятся по результатам наблюдений, для практического использования получаемых показателей необходимо иметь интервальные доверительные оценки для ROC-кривой и AUC, что особенно необходимо для случая малых выборок (число наблюдений результатов классификации n = 50 - 100), когда асимптотические оценки имеют большую погрешность.

Для построения интервальных доверительных оценок функций распределения F_G, F_B используется группировка наблюдений.

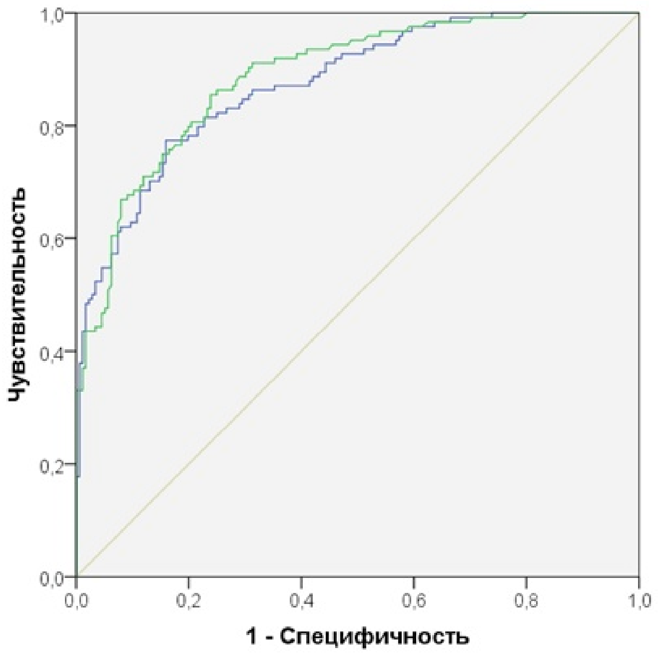

Рис. 2. **ROC-кривая, построенная по параметрам матрицы ошибок по всем возможным уровням отсечения w_i**

Группировка наблюдений

Интервал значений [0, 1] разбивается на N подинтервалов $A_{iN} = \left[\frac{i-1}{N}; \frac{i}{N}\right], i = \overline{1,n}$, $\sum_{i=1}^{N} A_{iN} = A = [0;1]$.

Функции $F_{G(B)N}(z)$ (оценки F_G, F_B по группированным данным) определяются как:

$$F_{G(B)N} = \begin{cases} 0, z \le 0 \\ \frac{1}{n}\sum_{j=1}^{[Nz]+1} \nu_{jN}^{G(B)}, 0 < z \le 1, \\ 1, z > 1 \end{cases}$$

где $F_{G(B)N}$ функции распределения для “хороших” (GOOD) и “плохих” (BAD), соответственно, построенные по группированным данным;
n - количество наблюдаемых данных;
N - количество интервалов разбиения отрезка [0,1].

Здесь $\overline{\nu_N} = \left(\nu_{1N}^{G(B)}, ..., \nu_{NN}^{G(B)}\right)$ - случайный вектор частот попаданий “хороших” (“плохих”) наблюдений в соответствующие интервалы.

Случайный вектор $\overline{\nu_N}$ имеет полиномиальное распределение $M(n; p_{1N}, ..., p_{NN})$:

$$P(\nu_{1N} = k_1, ..., \nu_{NN} = k_N) =$$

$$= \begin{cases} \frac{n!}{k_1!...k_N!} p_{1N}^{k_1} \cdots p_{NN}^{k_N}, & \sum k_i = n \\ 0, & \sum k_i \ne n \end{cases}.$$

Известно, что полиномиальное распределение является совместным распределением независимых пуассоновских случайных величин при условии, что их сумма фиксирована:

если $\xi_{1N}, ..., \xi_{NN}$ - независимые пуассоновские случайные величины

$$p(\xi_{jN} = k) = \frac{(np_{jN})^k}{k!} e^{n\lambda_i}$$

$$i = 1, 2, ..., N,$$

то $p(\nu_{1N} = k_1, ..., \nu_{NN} = k_N) =$

$$= p\left(\xi_{1N} = k_1, ..., \xi_{NN} = k_N \middle| \sum_{j=1}^{N} \xi_{jN} = n\right). \qquad (2)$$

Доверительные интервалы для пуассоновских случайных величин

Для отдельной пуассоновской случайной величины ξ_{jN} для среднего значения np_{jN} можно указать точный доверительный интервал с уровнем доверия $\gamma_j = 1 - \varepsilon_j$, где ε_j - заданный уровень значимости. Если k - это наблюдавшееся значение случайной величины ξ_{jN}, имеющей пуассоновское распределение $\Pi(np_{jN})$, то для среднего значения np_{jN} точный доверительный интервал имеет вид

$$P\left\{\frac{1}{2}\chi^2_{2k_j, \frac{\varepsilon_j}{2}} \leqslant np_{jN} \leqslant \frac{1}{2}\chi^2_{2k_j+2, 1-\frac{\varepsilon_j}{2}}\right\} = \gamma_j = 1 - \varepsilon_j. \qquad (3)$$

Здесь $\chi^2_{m,\varepsilon}$ - квантиль распределения Хи-квадрат с m степенями свободы, уровнем значимости ε и уровнем доверия γ.

Таким образом, для неизвестного параметра p_{jN} имеем доверительный интервал

$$P\left\{\frac{1}{2n}\chi^2_{2k_j,\frac{\varepsilon_j}{2}} \leqslant p_{jN} \leqslant \frac{1}{2n}\chi^2_{2k_j+2,1-\frac{\varepsilon_j}{2}}\right\} = \gamma_j = 1 - \varepsilon_j.$$

Для независимых $\xi_{1N}, \xi_{2N}, \dots, \xi_{NN}$, где $\xi_{jN} \sim \Pi(np_{jN})$, $j = 1, 2, \dots, N$, получаем доверительный параллелепипед

$$D_N(p_{1N}, \dots, p_{NN}) =$$

$$= \left\{(p_{1N}, \dots, p_{NN}) \left(\frac{1}{2n}\chi^2_{2k_j,\frac{\varepsilon_j}{2}} \leqslant p_{jN} \leqslant \right.\right.$$

$$\left.\left. \leqslant \frac{1}{2n}\chi^2_{2k_j+2,1-\frac{\varepsilon_j}{2}} \right), j = 1, 2, \dots, N \right\}.$$

$$P\left\{\left(\frac{1}{2n}\chi^2_{2k_j,\frac{\varepsilon_j}{2}} \leqslant p_{jN} \leqslant \frac{1}{2n}\chi^2_{2k_j+2,1-\frac{\varepsilon_j}{2}}\right),\right.$$

$$\left. j = 1, 2, \dots, N \right\} = \prod_{j=1}^{N} P\left\{\left(\frac{1}{2n}\chi^2_{2k_j,\frac{\varepsilon_j}{2}} \leqslant p_{jN} \leqslant \right.\right.$$

$$\left.\left. \leqslant \frac{1}{2n}\chi^2_{2k_j+2_j,1-\frac{\varepsilon_j}{2}} \right)\right\} = \prod_{j=1}^{N}(1 - \varepsilon_j) = \bar{\gamma} = 1 - \bar{\varepsilon}, \quad (4)$$

где ε_j - уровень значимости интервала для ξ_{jN};

$\bar{\gamma}$ - уровень доверия;

$\bar{\varepsilon} = \prod_{j=1}^{N}(1 - \varepsilon_j)$ - уровень значимости для параллелепипеда $D_N(p_1, \dots, p_N)$.

Здесь $(k_1, k_2, \dots, k_N)$ - наблюдавшиеся значения случайной величины $\xi_{1N}, \dots, \xi_{NN}$.

Используя связь между полиномиальным и пуассоновским распределениями, получаем следующее утверждение.

Утверждение 1

Для полиномиального случайного вектора $\bar{\nu}_N = (\nu_{1N}, \dots, \nu_{NN}) \sim M(n, p_{1N}, \dots, p_{NN})$, распределение которого является совместным распределением независимых пуассоновских случайных величин при условии, что их сумма фиксирована, с неизвестными $p_{1N}, \dots, p_{NN}$, множество $D_N(p_{1N}, p_{2N}, \dots, p_{NN})$ является точным доверительным множеством с уровнем доверия

$$\bar{\gamma_1} = 1 - \bar{\varepsilon_1} = \frac{(1-\bar{\varepsilon})}{(1-\varepsilon_N)} n^{-n} n! e^n,$$

где ε_1 - уровень значимости $D_N(p_{1N}, p_{2N}, \dots, p_{NN})$.

Доказательство:

В силу соотношения (2)

$$B = P\left\{\frac{1}{2n}\chi^2_{2\nu_{jN},\varepsilon_j/2} \le p_{jN} \le \frac{1}{2n}\chi^2_{2\nu_{jN}+2,1-\varepsilon_j/2},\right.$$
$$\left. j = 1, 2, \dots, N\right\} =$$

$$= \frac{P\left\{\frac{1}{2n}\chi^2_{2\xi_{jN},\varepsilon_j/2} \le p_{jN} \le \frac{1}{2n}\chi^2_{2\xi_{jN}+2,1-\varepsilon_j/2}, \; j = 1, 2, \dots, N-1, \sum_{j=1}^{N-1}\xi_{jN} = n\right\}}{P\left\{\sum_{j=1}^{N}\xi_{jN} = n\right\}} =$$

$$= \frac{P\left\{\frac{1}{2n}\chi^2_{2\xi_{jN},\varepsilon_j/2} \le p_{jN} \le \frac{1}{2n}\chi^2_{2\xi_{jN}+2,1-\varepsilon_j/2}, \; j = 1, 2, \dots, N-1\right\}}{P\left\{\sum_{j=1}^{N}\xi_{jN} = n\right\}} = B^*$$

$$P\left\{\sum_{j=1}^{N}\xi_{jN} = n\right\} = \frac{n^n e^{-n}}{n!},$$

так как сумма $\sum_{j=1}^{N}\xi_{jN}$ независимых пуассоновских случайных величин $\xi_{jN} \sim \Pi(np_{jN})$ также имеет распределение Пуассона с параметром $\sum_{j=1}^{N} np_{jN} = n$.

В то же время

$$B^* = \frac{\prod_{j=1}^{N-1}(1-\varepsilon_j)}{p\left\{\sum \xi_{jN} = n\right\}} = \prod_{j=1}^{N-1}(1-\varepsilon_j) n! e^n n^{-n} = \gamma_1.$$

Что и требовалось доказать (Ч.Т.Д.).

Примечание 1

Необходимо отметить, что случайный вектор $\frac{\overline{\nu_N}}{n} = \left(\frac{\overline{\nu_{1N}}}{n}, ..., \frac{\overline{\nu_{NN}}}{n}\right)$ представляет собой гистограмму наблюдений, построенную на группированных наблюдениях. Таким образом, множество $D_N = \left(p_{1N}, p_{2N}, ..., P_{NN}\right)$ является точным доверительным множеством для гистограммы, которая является оценкой плотности распределения, при этом уровень доверия для параллелепипеда $D_N = \left(p_{1N}, p_{2N}, ..., P_{NN}\right)$ равен γ_1, соответственно, уровень значимости $\overline{\varepsilon_1} = 1 - \gamma_1$.

Примечание 2

В разделе "Доверительные интервалы для пуассоновских случайных величин" для каждого p_{jN} задан свой уровень значимости ε_j и получен соответствующий доверительный интервал для p_{jN} с уровнем доверия $\gamma_j = 1 - \varepsilon_j, j = 1,2,...,N$.

Утверждение 2

Оценки $F_{GN}(z)$, $F_{BN}(z)$ при $N \to \infty$ и $0 < c_1 \le \frac{n}{N} \le c_2 < \infty$ при некоторых константах c_1, c_2 являются асимптотически не смещенными и асимптотически состоятельными оценками для $F_B(z)$, $F_G(z)$. Доказательство данного факта приведено в работе[12].

Примечание 3

Пусть число n - число наблюдений, задано, а N - число интервалов группировки. В случае, когда $\varepsilon_1 = \varepsilon_2 = ... = \varepsilon_{N-1} = \varepsilon$ - уровни значимости одинаковы для всех $\nu_{1N}, ..., \nu_{N-1,N}$ при заданных n, ε, ε_1 подходящее N можно выбрать по формуле

$$N = \frac{ln(1-\varepsilon_1) + n \cdot ln(n) - n - ln(n!)}{ln(1-\varepsilon)} + 1.$$

Определим верхнюю и нижнюю граничные функции.

$F^{v}_{G(B)N}(z), F^{L}_{G(B)N}(z)$, соответственно:

$$F^{v}_{G(B)N}(z) = \begin{cases} 0, & z \le 0 \\ \frac{1}{2n}\sum\limits_{j=1}^{[Nz]+1} \chi^2_{2*\nu_{jN}G(B)+2,1-\varepsilon_j/2} & 0 < z \le 1 \\ 1, & z > 1 \end{cases}$$

$$F^{L}_{G(B)N}(z) = \begin{cases} 0, & z \le 0 \\ \frac{1}{2n}\sum\limits_{j=1}^{[Nz]+1} \chi^2_{2*\nu_{jN}G(B),\varepsilon_j/2} & 0 < z \le 1 \\ 1, & z > 1 \end{cases}.$$

Утверждение 3

Функции $F^{v}_{G(B)N}(z), F^{L}_{G(B)N}(z)$ являются точными доверительными границами для $FG(B)N(z)$ с уровнем доверия γ_1.

Доказательство:

Утверждение следует из того факта, что для заданных k_1, k_2, ... , k_N случайное событие $\{\xi_{1N} = k_1, \xi_{2N} = k_2, ..., \xi_{NN} = k_N\}$ эквивалентно событию $\{\xi_{1N} = k_1, \xi_{1N} + \xi_{2N} = k_1 + k_2, ..., \xi_{1N} + ... + \xi_{NN} = k_1 + + k_N\}$ и, следовательно, их вероятности совпадают.

Ч.Т.Д.

Примечание 4

Функции $F^{L|V|}_{G(B)N}$ составляют доверительную полосу для всей функции $F_{G(B)N}(z)$ с заданным уровнем доверия γ_1. При этом, как следует из утверждения об асимптотической состоятельности (при $N \to \infty$) $F_{G(B)N}(z)$ для оценки истинной функции распределения $F_{G(B)}(z)$, максимальное расхождение между $F_{G(B)N}(z)$ и $F_{G(B)}(z)$ имеет порядок $1 / N$. Поэтому для оценки $F_{G(B)}(z)$ доверительным интервалом с уровнем доверия γ_1 является $\left(F^{v}_{G(B)N}(z) + \frac{1}{N}, F^{L}_{G(B)N}(z) - \frac{1}{N}\right)$.

$$\overline{\gamma_1} = 1 - \overline{\varepsilon_1} = \frac{(1-\overline{\varepsilon})}{(1-\varepsilon_N)} n^{-n} n! e^n.$$

Дополнительно к оценкам $F^{L|V|}_{G(B)N}$ рассматривались поточечные оценки:

$$F^{v}_{G(B)N}(z) = \begin{cases} 0, & z \le 0 \\ \frac{1}{2n}\chi^2_{2*\left(\sum\limits_{j=1}^{[Nz]+1} \nu_{jN}G(B)\right),1-\varepsilon_j/2} & 0 < z \le 1 \\ 1, & z > 1 \end{cases}$$

$$F^{L}_{G(B)N}(z) = \begin{cases} 0, & z \le 0 \\ \frac{1}{2n}\chi^2_{2*\left(\sum\limits_{j=1}^{[Nz]+1} \nu_{jN}G(B)+2\right),\varepsilon_j/2} & 0 < z \le 1 \\ 1, & z > 1 \end{cases}$$

Эти оценки для каждого z дают доверительный интервал для $F_{G(B)N}(z)$ с уровнем доверия $\gamma_j = 1 - \varepsilon_j$, $j = [Nz] + 1$. При этом, повторяя рассуждения в доказательстве утверждений 1 и 3, можно показать: доверительная вероятность того, что при всех z $0 \le z \le 1$ $F_{G(B)N}(z)$ принадлежит к доверительному интервалу, соответствующему z, равна $\overline{\gamma_1} = 1 - \overline{\varepsilon_1}$.

С помощью $F_{G(B)N}(z)$ строится оценка

$$ROC_N(z) = G_{BN}\left(F_{GN}^{-1}(z)\right),$$

а с помощью $F_{G(B)N}^{L|V|}$ верхняя и нижняя доверительные оценки ROC_N^v, ROC_N^L в соответствии с формулой (1):

$$ROC_N^L(z) = G_{GN}^L\left(\left(F_{BN}^L\right)^{-1}(z)\right),$$
$$ROC_N^V(z) = G_{GN}^V\left(\left(F_{BN}^V\right)^{-1}(z)\right),$$

которые составляют доверительную полосу для функции $ROC(z)$ с заданным уровнем доверия γ_1.

Аналогично по оценке $ROC_N(z)$ строится оценка AUC_N и по ROC_N^V, ROC_N^L строятся AUC_N^V, AUC_N^L - верхняя и нижняя границы доверительного интервала для AUC.

$$AUC_N^L(z) = \int_0^1 \left(G_{GN}^L\left(\left(F_{BN}^L\right)^{-1}(z)\right)\right)dz.$$
$$AUC_N^V(z) = \int_0^1 \left(G_{GN}^V\left(\left(F_{BN}^V\right)^{-1}(z)\right)\right)dz.$$

Выводы

В статье предложены статистические оценки показателей мощности (валидности) внутренних рейтингов банков на основе группированных данных. Также выведены зависимости для расчета точных верхних и нижних границ для эмпирических функций распределения “плохих” и “хороших” клиентов с заданным уровнем доверия для малых выборок. На основе этих оценок построены доверительные интервалы с заданным уровнем доверия для показателей валидности внутренних рейтингов банков таких показателей мощности рейтингов, как ROC-кривая и Area Under Curve (AUC). Точность полученных доверительных интервалов показывает, что значение доверительной вероятности устанавливается точным, а не асимптотическим распределением статистик, определяющих доверительный интервал. Полученные доверительные интервалы могут быть использованы для малых объемов наблюдений, когда асимптотические формулы имеют большую погрешность.

В дальнейшем предполагается получить оптимальные доверительные интервалы с выбором числа интервалов N, минимизирующим погрешность от использования группировки наблюдений. Также планируется распространить этот подход к построению доверительных оценок для показателей валидности небинарных классификаторов. Данный инструмент может использоваться для получения показателя ценности банковских услуг[13].

---

[1] Совершенствование корпоративного управления в кредитных организациях : рекомендация Базельского комитета по банковскому надзору. Базель. 1999. Сентябрь.

[2] Там же.

[3] Базельский комитет по банковскому надзору: Международная конвергенция измерения капитала и стандартов капитала: Уточненные рамочные подходы // Банк международных расчетов. 2004. Июнь. С. 64, п. 246, 247.

[4] Совершенствование...

[5] Базельский комитет...

[6] См.: *Engelmann B., Hayden E., Dirk Tasche.* Testing rating accuracy // RISK. 2003. Vol. 16. P. 82-86; *Stein R.M.* Benchmarking default prediction models pitfalls and remedies in model validation // Journal of Risk Model Validation. 2007. Vol. 1. □ 1. P. 77-113.

[7] См.: *Большев Л.Н.* Сравнение интенсивностей простейших потоков // Теория вероятностей и ее применения. Москва, 1962. Т. 7. С. 353-355; *Его же.* О построении доверительных пределов // Теория вероятностей и ее применения”; *Garwood F.* Fiducial limits for Poisson distribution // Biometrica. 1936. Vol. 28. P. 437-442; *Stevens W.L.* Fiducial limits of the parameter of discontinuous distribution // Biometrica. 1950. Vol. 37. P. 117-129.

[8] См.: *Sofus A.* Macskassy and Foster Provost, Confidence Bands for ROC Curves: Methods and an Empirical Study // CeDER Working Paper. 02-04, Stern School of Business, New York University, NY, NY 10012. Jan 2004; *Alicja Jokiel-Rokita, Micha□ Pulit.* Nonparametric estimation of the ROC curve based on smoothed empirical distribution functions // Statistical Computing. 2013. □ 23. P. 703-712; *Ayman Baklizi.* A Simple Method for Finding Emperical Liklihood Type Intervals for the ROC Curve // Journal of Modern Applied Statistical Methods. 2007. Vol. 6, issue 2. P. 589-595.

[9] *Engelmann B., Hayden E., Dirk Tasche.* Op. cit.

[10] *Alicja Jokiel-Rokita, Michal Pulit.* Op. cit.

[11] *Engelmann B., Hayden E., Dirk Tasche.* Op. cit.

[12] *Гуськов С.Ю. Л□вин В.В.* Интервальные доверительные оценки для показателей качества бинарных классификаторов - ROC-кривых, AUC для случая малых выборок // Инженерный журнал: наука и инновации. 2015. Вып. 3. URL: http://engjournal.ru/catalog/mesc/idme/1376.html.

[13] *Гуськова М.Ф.* Взаимосвязь экономических теорий полезности и ценности : автореф. дис. ... д-ра экон. наук. Москва, 2009. С. 15.

Поступила в редакцию 04.08.2015 г.