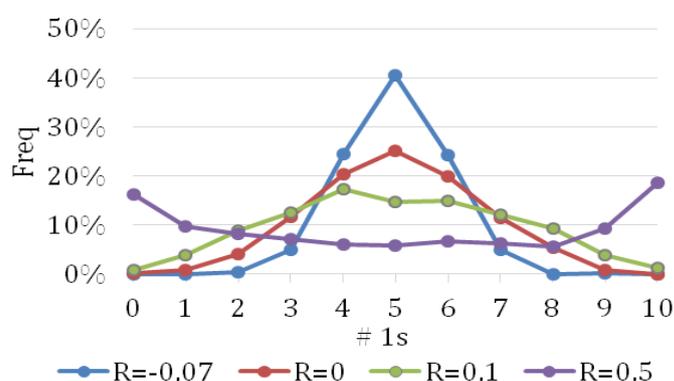


ВАЛИДАЦИЯ ТОЧНОСТИ МОДЕЛЕЙ БИНАРНОГО ВЫБОРА В СЛУЧАЕ КОРРЕЛИРОВАННЫХ ИСХОДОВ В ПРИЛОЖЕНИИ К ЗАДАЧАМ ОЦЕНКИ КРЕДИТНОГО РИСКА

Борзых Д., Пеникас Г. (Москва)ⁱ

Оценка вероятности невозврата (probability of default, PD) кредита лежит в основе регулирования кредитного риска в рамках подхода на основе внутренних рейтингов соглашений Базель II-III. При разработке моделей PD регулятор требует, чтобы они обладали достаточной точностью, что можно проверить биномиальным тестом. Но он предполагает независимость случайных величин. (Blochwitz, Martin, & Wehn, 2006) делают попытку скорректировать тест для коррелированных исходов через расширение доверительного интервала (ДИ), что неверно, как будет показано в докладе.

Доверительные интервалы для проверки точности моделей бинарного выбора с коррелированными бинарными исходами будут получены на основе генераций соответствующих распределений на основе непараметрического ("генетического") алгоритма (Курра, Lerenies, & Jung, 2018), поскольку он позволяет дополнительно получить явление с отрицательно коррелированными исходами, что недоступно в параметрическом подходе (Lunn & Davies, 1998) (результат см. пример на рисунке).



1s – число единичных исходов;
Freq. – частота исходов;
R – коэффициент корреляции между каждой парой бинарных случайных величин (с.в.);
P = 50% – вероятность единичного исхода для отдельно взятой случайной величины.

Тогда для случай существенно положительных коррелированных бинарных с.в. при $p=50\%$ и $R=50\%$ традиционный биномиальный тест предлагал бы использовать доверительный интервал (ДИ) примерно [2; 8] (как для случая $R=0\%$), хотя по факту все реализации доли единичных исходов сконцентрированы либо около нуля, либо единицы, т.е. нужен двойной ДИ вида [0; 3] и [7; 10]. Если бы корреляция была отрицательной ($R=-7\%$), то доверительный интервал сузился бы до [3; 7].

Научная новизна предложенной статистической процедуры проверки точности моделей бинарного выбора, когда имеют место коррелированные исходы, включает:

1) При наличии коррелированных бинарных исходов и частной (предельной) вероятности исхода, отличной от 50%, ДИ должен быть смещен относительно центра, используемого в биномиальном тесте и интервале Джефриса, а не просто расширяться, как предлагает (Blochwitz, Martin, & Wehn, 2006, p. 295).

2) При высоких значениях положительной корреляции (как минимум, более 50%) более точным (более узким) является ДИ, состоящий из двух частей (отрезков), центрированных относительно появляющихся мод бимодального распределения бинарных исходов (долей исходов). Данный результат распада ДИ на две части никогда до этого не обсуждался в литературе, хотя само понятие бимодальности встречалось ранее (например, в (Witt, 2014), (Ермолова & Пеникас, 2018)).

Литература

- Ермолова М.Д., Пеникас Г.И. Биномиальный тест для коррелированных бинарных случайных величин для проверки точности рейтинговой модели. Управление финансовыми рисками. 2018. 03 (55). С. 174–190.
- Blochwitz S., Martin M., Wehn C. XIII. Statistical Approaches to PD Validation. В. Engelmann, & R. Rauhmeier (Ред.), The Basel II Risk Parameters. Estimation, Validation, Stress-Testing – with Applications to Loan Risk Management (pp. 289–306). London: Springer. 2006.
- Kruppa J., Lepenies B. & Jung K. A genetic algorithm for simulating correlated binary data from biomedical research. Computers in Biology and Medicine. 2018. 92, pp. 1–8.
- Lunn A., Davies S. A note on generating correlated binary variables. Biometrika. 1998. 2 (85), pp. 487–490.
- Witt G. A simple distribution for the sum of correlated, exchangeable binary data. Communications in Statistics – Theory and Methods. 2014. 43, pp. 4265–4280.

ⁱ Борзых Дмитрий Александрович – НИУ ВШЭ, dborzykh@hse.ru;
Пеникас Генрих Иозович – Банк России, penikas@gmail.com