

Практическое применение критерия Манна-Уитни-Уилкоксона в оценочной деятельности

К. А. Мурашев

14 мая 2022 г.

В своей практике оценщики часто сталкиваются с необходимостью учёта различий количественных характеристик объектов. В частности, одной из стандартных задач является установление признаков, влияющих на стоимость (т. н. ценообразующих факторов) и их отделение от признаков, влияние которых на стоимость отсутствует либо не может быть установлено. В практике оценки широкое распространение получил субъективный отбор признаков, учитываемых при определении стоимости. При этом конкретные количественные показатели влияния этих признаков на стоимость зачастую берутся из т. н. «справочников». Не отказывая такому подходу в быстроте и невысокой стоимости его реализации, нельзя не признать, что только данные, непосредственно наблюдаемые на открытом рынке, являются надёжной основой суждения о стоимости. Приоритет таких данных над прочими, в частности, полученными путём опроса экспертов, закреплён, в том числе в Стандартах оценки RICS [3], Международных стандартах оценки 2022 [4], а также МСФО 13 «Оценка справедливой стоимости» [2]. Поэтому можно говорить о том, что математические методы анализа данных, полученных на открытом рынке, являются наиболее надёжным средством интерпретации рыночной информации, применяемой при исследованиях рынка и предсказании стоимости конкретных объектов. В данном материале будут рассмотрены основные теоретические вопросы, касающиеся теста Манна-Уитни-Уилкоксона (далее U-тест), а также проведён пошаговый разбор применения данного теста к конкретным данным. Материал содержит строки кода, необходимые для проведения U-теста с использованием языков программирования Python и R, а также приложение в виде электронной таблицы, содержащей формулы для проведения данного теста и полностью готовой для её применения на любых иных данных. Данный материал и все приложения к нему распространяются на условиях лицензии cc-by-sa-4.0 [6].

1. Технические данные

Данный материал, а также приложения к нему доступны по постоянной ссылке [5]. Исходный код данной работы был создан с использованием языка \TeX [14] с набором макрорасширений \LaTeX [15], дистрибутива TeXLive [16] и редактора TeXstudio [19]. Расчёт в форме электронной таблицы был выполнен с помощью LibreOffice Calc [8] (Version: 7.3.3.2, Ubuntu package version: 1:7.3.3 rc2-0ubuntu0.20.04.1 lo1 Calc: threaded). Расчёт на языке R [17] (version 4.2.0 (2022-04-22) – "Vigorous Calisthenics") был выполнен с использованием IDE RStudio (RStudio 2022.02.2+485 "Prairie Trillium" Release (8acbd38b0d4ca3c86c570cf4112a8180c48cc6fb, 2022-04-19) for Ubuntu Bionic Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) QtWebEngine/5.12.8 Chrome/69.0.3497.128 Safari/537.36) [13]. Расчёт на языке Python [7] был выполнен с использованием среды разработки Jupyter Lab [9].

2. Предмет исследования

В случае работы с рыночными данными перед оценщиком часто встаёт задача проверки гипотезы о существенности влияния того или иного признака, измеренного в количественной или порядковой шкале, на стоимость. Аналогичная задача возникает у аналитиков рынка недвижимости, специалистов компаний-застройщиков, риелторов. При этом зачастую отсутствует возможность сбора больших массивов данных, позволяющих применить широкий спектр методов машинного обучения. В ряде случаев оценщики осознанно сужают область сбора данных до узкого сегмента рынка, в результате чего в их распоряжении оказываются лишь сверхмалые выборки объёмом менее тридцати наблюдений. При этом, ценовые данные чаще всего имеют распределение отличное от нормального. В данном случае рациональным решением является применение U-теста. Сформулируем задачу:

- предположим, что у нас существуют две выборки удельных цен коммерческих помещений, часть из которых обладает некоторым признаком (например, имеет отдельный вход), часть — нет;
- необходимо установить: оказывает ли наличие этого признака существенное влияние на удельную стоимость недвижимости данного типа или нет.

На первый взгляд, согласно сложившейся практике, оценщик может просто субъективно признать те или иные признаки значимыми, а прочие нет, после чего принять значения корректировок на различия в этих признаках из справочников. Однако, как было сказано выше, такой подход вряд ли может считаться лучшей практикой, поскольку в этом случае отсутствует какой-либо серьёзный анализ рынка. Кроме того, в таком случае вряд ли можно говорить о какой-либо ценности такой работы в принципе.

Вместо этого возможно использовать случайные выборки рыночных данных и применять к ним математические методы анализа, позволяющие делать доказательные с научной точки зрения выводы о значимости влияния того или иного признака

на стоимость. Данные, используемые в настоящей работе, являются вымышленными и служат для учебных целей. Прилагаемая электронная таблица настроена таким образом, что исходные данные могут быть сгенерированы случайным образом, что позволяет лучше понять поведение теста.

3. Основные сведения о тесте

В первую очередь необходимо сказать, что, несмотря на заявленное общее название, правильное всё же говорить о двух тестах:

- двухвыборочный критерий Уилкоксона, разработанный Фрэнком Уилкоксоном в 1945 году [11];
- U-критерий Манна-Уитни, являющийся дальнейшим развитием вышеуказанного критерия, разработанный Генри Манном и Дональдом Уитни в 1947 году [10].

Забегая вперёд, можно сказать о том, что статистики данных критериев линейно связаны, а сами p -значения практически одинаковы, что с практической точки зрения позволяет скорее говорить о вариациях одного теста, а не о двух отдельных [11]. В данной работе по всему тексту используется общее название, а также его сокращённый вариант — U-тест, исторический относимый к критерию Манна-Уитни. Некоторые авторы [1] рекомендуют использовать двухвыборочный критерий Уилкоксона в случаях, когда нет предположений о дисперсиях, а в случае равных дисперсий применять U-критерий Манна-Уитни. Однако экспериментальные данные указывают, что p -значения критериев Уилкоксона и Манна-Уитни практически совпадают, в том числе и в случае, когда дисперсии выборок существенно различаются. Придерживаясь принципа KISS [18], лежащего в основе всего данного цикла публикаций, автор приходит к выводу о возможности применения единого подхода.

Также следует помнить о том, что существует Критерий Уилкоксона для связанных выборок [12], представляющий собой отдельный тест, предназначенный для анализа различий между связанными выборками, тогда как рассматриваемый в данной работе U-тест предназначен для работы с двумя независимыми выборками.

Предположим, что заданы две выборки:

$$x^m = (x_1, x_2, \dots, x_m), x_i \in \mathbb{R}; \quad y^n = (y_1, y_2, \dots, y_n), y_i \in \mathbb{R} \quad |m \leq n.$$

- Обе выборки являются простыми, объединённая выборка независима.
- Выборки взяты из неизвестных непрерывных распределений $F(x)$ и $G(y)$ соответственно.

Простая выборка — это случайная, однородная, независимая выборка. Эквивалентное определение: выборка $x^m = (x_1, x_2, \dots, x_m)$ является простой, если значения (x_1, x_2, \dots, x_m) являются реализациями m независимых одинаково распределённых случайных величин. Иными словами, отбор наблюдений является

не только случайным, но и не предполагает наличия каких-либо специальных правил (например, выбор каждого 10-го наблюдения).

U-тест — это непараметрический тест для проверки нулевой гипотезы, заключающейся в том, что для случайно выбранных из двух выборок наблюдений $x, x \in X$ и $y, y \in Y$ вероятность того, что x больше y , равна вероятности того, что y больше x . На математическом языке запись нулевой гипотезы выглядит следующим образом:

$$H_0 : P\{x < y\} = \frac{1}{2}. \quad (1)$$

Для целостности теста требуется альтернативная гипотеза, которая заключается в том, что вероятность того, что значение признака наблюдения из выборки X превышает его у наблюдения из выборки Y , отличается (больше или меньше) от вероятности того, что значение признака у наблюдения из Y превышает значение у наблюдения из X . На математическом языке запись альтернативной гипотезы выглядит следующим образом:

$$H_1 : P\{x < y\} \neq P\{y < x\} \vee P\{x < y\} + 0.5 \cdot P\{x = y\} \neq 0.5. \quad (2)$$

Согласно базовой концепции U-теста, при справедливости нулевой гипотезы распределение двух выборок непрерывно, при справедливости альтернативной распределение одной из них стохастически больше распределения другой. При этом, можно сформулировать целый ряд нулевых и альтернативных гипотез, для которых данный тест будет давать корректный результат. Его самое широкое обобщение заключается в следующих предположениях:

- наблюдения в обеих выборках независимы;
- тип данных является как минимум ранговым, т.е. в отношении любых двух наблюдений можно сказать, какое из них больше;
- нулевая гипотеза предполагает, что распределения двух выборок равны;
- альтернативная гипотеза предполагает, что распределения двух выборок не равны.

Простыми словами, суть U-теста заключается в том, что он позволяет ответить на вопрос, является ли существенным различие значения количественного признака двух выборок. Применительно к оценке можно сказать, что применение данного теста помогает ответить на вопрос, является ли необходимым учёт того или иного признака в качестве ценообразующего фактора. Из сказанного выше следует, что речь идёт о двухстороннем тесте. На практике это означает, что тест не даёт прямой ответ, например на такой вопрос: «имеет ли место значимое превышение удельной стоимости помещений, имеющих отдельный вход, относительно помещений, не обладающих им». Вместо этого корректно говорить о том, «существует ли

существенное различие в значении стоимости между помещениями двух типов: с отдельным входом и без такового».

Условиями применения U-теста помимо вышеуказанных требований к самим выборкам являются:

- распределение значений количественного признака выборок отлично от нормального (в противном случае целесообразно использование параметрического t-критерия Стьюдента для независимых выборок);
- не менее трёх значений признака в каждой выборке, допускается наличие двух значений в одной из выборок, при условии наличия в другой не менее пяти;

4. Практическая реализация

5. Выводы

Источники информации

- [1] А. И. Кобзарь. *Прикладная математическая статистика*. 2006.
- [2] Министерство финансов России. *Международный стандарт финансовой отчётности (IFRS) 13 «Оценка справедливой стоимости»*. с изменениями на 11 июля 2016 г. Russian. Russia, Moscow: Минфин России, 2015-12-28. URL: <https://normativ.kontur.ru/document?moduleId=1&documentId=326168#10> (дата обр. 10.06.2020).
- [3] Royal Institution of Chartered Surveyors (RICS). *RICS Valuation — Global Standards*. English. UK, London: RICS, 2021-11-30. URL: <https://www.rics.org/uk/upholding-professional-standards/sector-standards/valuation/red-book/red-book-global/> (дата обр. 11.05.2022).
- [4] International Valuation Standards Council. *International Valuation Standards*. 2022-01-31. URL: <https://www.rics.org/uk/upholding-professional-standards/sector-standards/valuation/red-book/international-valuation-standards/>.
- [5] Kirill A. Murashev. «Practical application of the Mann-Whitney-Wilcoxon test (U-test) in valuation». English, Spanish, Russian, Interslavic. B: (2022-05-15). URL: https://github.com/Kirill-Murashev/AI_for_valuers_book/tree/main/Parts&Chapters/Mann-Whitney-Wilcoxon (дата обр. 15.05.2022).
- [6] Creative Commons. *cc-by-sa-4.0*. URL: <https://creativecommons.org/licenses/by-sa/4.0/> (дата обр. 27.01.2021).
- [7] Python Software Foundation. *Python site*. Английский. Python Software Foundation. URL: <https://www.python.org/> (дата обр. 17.08.2021).
- [8] The Document Foundation. *LibreOffice Calc*. Английский. URL: <https://www.libreoffice.org/discover/calc/> (дата обр. 20.08.2021).

- [9] *Jupyter site*. URL: <https://jupyter.org> (дата обр. 13.05.2022).
- [10] Machinelearning.ru. *U-критерий Манна-Уитни*. Russian. URL: http://www.machinelearning.ru/wiki/index.php?title=%D0%9A%D1%80%D0%B8%D1%82%D0%B5%D1%80%D0%B8%D0%B9_%D0%A3%D0%B8%D0%BB%D0%BA%D0%BE%D0%BA%D1%81%D0%BE%D0%BD%D0%B0-%D0%9C%D0%B0%D0%BD%D0%BD%D0%B0-%D0%A3%D0%B8%D1%82%D0%BD%D0%B8 (дата обр. 14.05.2022).
- [11] Machinelearning.ru. *Критерий Уилкоксона двухвыборочный*. Russian. URL: http://www.machinelearning.ru/wiki/index.php?title=%D0%9A%D1%80%D0%B8%D1%82%D0%B5%D1%80%D0%B8%D0%B9_%D0%A3%D0%B8%D0%BB%D0%BA%D0%BE%D0%BA%D1%81%D0%BE%D0%BD%D0%B0_%D0%B4%D0%B2%D1%83%D1%85%D0%B2%D1%8B%D0%B1%D0%BE%D1%80%D0%BE%D1%87%D0%BD%D1%8B%D0%B9 (дата обр. 14.05.2022).
- [12] Machinelearning.ru. *Критерий Уилкоксона для связанных выборок*. Russian. URL: http://www.machinelearning.ru/wiki/index.php?title=%D0%9A%D1%80%D0%B8%D1%82%D0%B5%D1%80%D0%B8%D0%B9_%D0%A3%D0%B8%D0%BB%D0%BA%D0%BE%D0%BA%D1%81%D0%BE%D0%BD%D0%B0_%D0%B4%D0%BB%D1%8F_%D1%81%D0%B2%D1%8F%D0%B7%D0%BD%D1%8B%D1%85_%D0%B2%D1%8B%D0%B1%D0%BE%D1%80%D0%BE%D0%BA (дата обр. 14.05.2022).
- [13] PBC Studio. *RStudio official site*. Английский. URL: <https://www.rstudio.com/> (дата обр. 19.08.2021).
- [14] CTAN team. *TeX official site*. English. CTAN Team. URL: <https://www.ctan.org/> (дата обр. 15.11.2020).
- [15] LaTeX team. *LaTeX official site*. English. URL: <https://www.latex-project.org/> (дата обр. 15.11.2020).
- [16] *TeXLive official site*. URL: <https://www.tug.org/texlive/> (дата обр. 15.11.2020).
- [17] The R Foundation. *The R Project for Statistical Computing*. Английский. The R Foundation. URL: <https://www.r-project.org/> (дата обр. 17.08.2021).
- [18] Wikipedia. *KISS principle*. URL: https://en.wikipedia.org/wiki/KISS_principle (дата обр. 06.11.2020).
- [19] Benito van der Zander. *TeXstudio official site*. URL: <https://www.texstudio.org/> (дата обр. 15.11.2020).