# Практическое применение критерия Манна-Уитни-Уилкоксона в оценочной деятельности

К. А. Мурашев

19 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 реальным 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 тестовые данные и формулы для проведения рассматриваемого 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].

## Содержание



## 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 [15] с набором макрорасширений LATEX [16], дистрибутива TeXLive [17] и редактора TeXstudio [24]. 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 [18] (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) [14]. Расчёт на языке Python (Version 3.8.10) [7] был выполнен с использованием среды разработки Jupyter Lab (Version 3.4.2) [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 при проведении U-теста средствами Python и R, представляют собой реальные рыночные данные, часть из которых была собрана автором путём парсинга, часть — предоставлена коллегами для анализа. Прилагаемая электронная таблица настроена таким образом, что исходные данные могут быть сгенерированы случайным образом.

## 3. Основные сведения о тесте

### 3.1. Предпосылки и формализация гипотез

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

- двухвыборочный критерий Уилкоксона, разработанный Фрэнком Уилкоксоном в 1945 году [12];
- U-критерий Манна-Уитни, являющийся дальнейшим развитием вышеуказанном 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 [12]. 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 [20], 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 связных выборок [13], 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:

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

- 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 $x^m = (x_1, x_2, \ldots, x_m)$ является простой, если значения $(x_1, x_2, \ldots, x_m)$ 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 математической языке запись нулевой гипотезы выглядит следующим образом:

$$H_0 : P\{x < y = \frac{1}{2}\}. \tag{1}$$

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:

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

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.

В случае более строгого набора допущений, чем приведённые выше, например, в случае допущения о том, что распределение двух выборок в случае справедливости нулевой гипотезы непрерывно, альтернативной — имеет сдвиг расположения двух распределений, т. е. $f_1 x = f_2(x + \sigma)$, можно сказать, что U-тест представляет собой тест на проверку гипотезы о равенстве медиан. В этом случае, U-тест можно интерпретировать как проверку того, отличается ли от нуля оценка Ходжеса-Лемана разницы значений мер центральной тенденции. В данной ситуации оценка Ходжеса-Лемана представляет собой медиану всех возможных значений различий между наблюдениями в первой и второй выборках. Вместе с тем, если и дисперсии, и формы распределения обеих выборок различаются, U-тест не может корректно проверить медианы. Можно показать примеры, когда медианы численно равны, при этом тест отвергает нулевую гипотезу с вследствие малого p-значения.

Таким образом, более корректной интерпретацией U-теста является его использование для проверки именно гипотезы сдвига [11].

**Гипотеза сдвига** — статистическая гипотеза, часто рассматривающаяся как альтернатива гипотезе о полной однородности выборок. Пусть даны две выборки данных. Пусть также даны две случайные величины $X$ и $Y$, которые распределены как элементы этих выборок и имеют функции распределения *F(x)* и *G(y)* соответственно. В этих терминах гипотезу сдвига можно записать следующим образом:

$$H : F(x) = G(x + \sigma) \quad |\forall x,\, \sigma \neq 0. \tag{3}$$

В этом случае U-критерий является состоятельным независимо от особенностей выборок.

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.

Подытоживая вышесказанное, можно сказать, что существуют три варианта нулевой гипотезы, в зависимости от уровня строгости.

Таблица 1. Варианты нулевой гипотезы при использовании U-теста при оценке стоимости

| Тип гипотезы | Формулировка |
|---|---|
| Научная | Наблюдения из двух выборок полностью однородны, т. е. принадлежат одному распределению, сдвиг отсутствует, оценка, сделанная для первой выборки, является несмещённой и для второй |
| Практическая | Медианы двух выборок равны между собой |
| Изложенная в терминах оценки | Различие признака между двумя выборками объектов-аналогов не является существенным, его учёт не требуется, данный признак не является ценообразующим фактором |

## 3.2. Реализация теста

### 3.2.1. Статистика критерия

Допустим, что элементы $x_1, \ldots, x_n$ представляют собой простую независимую выборку из множества $X \in \mathbb{R}$, а элементы $y_1, \ldots, y_n$ представляют собой простую независимую выборку из множества $Y \in \mathbb{R}$, при этом выборки являются независимыми относительно друг друга. Тогда соответствующая U-статистика определяется следующим образом:

$$U = \sum_{i=1}^{m} \sum_{j=1}^{n} S(x_i, y_j),$$
$$\text{при} \tag{4}$$
$$S(x,y) = \begin{cases} 1, & \text{если } x > y, \\ \frac{1}{2}, & \text{если } x = y, \\ 0, & \text{если } x < y. \end{cases}$$

### 3.2.2. Методы вычисления

Тест предполагает вычисление статистики, обычно называемой U-статистикой, распределение которой известно в случае справедливости нулевой гипотезы. При работе со сверхмалыми выборками распределение задаётся таблично, при размерах выборки более двадцати наблюдений оно достаточно хорошо аппроксимируется нормальным распределением. Существуют два методы вычисления U-статистики: подсчёт вручную по формуле 4, применение специального алгоритма. Первый способ подходит только для сверхмалых выборок в силу трудоёмкости. Второй способ может быть формализован в виде пошагового набора инструкций и будет описан далее.

1) Необходимо построить общий вариационный ряд для двух выборок, а затем присвоить каждому наблюдению ранг, начиная с 1 для наименьшего из них. В случае наличия связок, т. е. групп повторяющихся значений (такой группой могут являться в т. ч. только два равных значения), каждому наблюдению из такой группы присваивается значение, равное медиане значений рангов группы до корректировки (например, в случае вариационного ряда (*3, 5, 5, 5, 5, 8*) ранги до корректировки имеют вид (*1, 2, 3, 4, 5, 6*) после — (*1, 3.5, 3.5, 3.5, 3.5, 6*)).

2) Необходимо провести подсчёт сумм рангов наблюдений каждой из выборок, обозначаемых как $R_1$, $R_2$ соответственно. При этом, общая сумма рангов $R$ может быть вычислена по формуле

$$R = \frac{N(N+1)}{2}, \tag{5}$$

где $N$ — общее число наблюдений в обеих выборках.

3) Далее вычисляем U-значение для первой выборки:

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2}, \tag{6}$$

где $R_1$ — сумма рангов первой выборки, $n_1$ — число наблюдений в первой выборке.

Аналогичным образом вычисляется U-значения для второй выборки:

$$U_2 = R_2 - \frac{n_2(n_2+1)}{2}, \tag{7}$$

где $R_2$ — сумма рангов второй выборки, $n_2$ — число наблюдений во второй выборке.

Из вышеприведённых формул следует, что

$$U_1 + U_2 = R_1 - \frac{n_1(n_1+1)}{2} + R_2 - \frac{n_2(n_2+1)}{2}. \tag{8}$$

Также известно, что

$$\begin{cases} R_1 + R_2 = \dfrac{N(N+1)}{2} \\ N = n_1 + n_2. \end{cases} \tag{9}$$

Тогда

$$U_1 + U_2 = n_1 n2. \tag{10}$$

Использование данной формулы в качестве контрольного соотношения может быть полезно для проверки корректности вычислений при расчёте в табличном процессоре.

4) Из двух значений $U_1$, $U_2$ во всех случаях выбираем меньшее, которое и будет являться U-статистикой и использоваться в дальнейших расчётах. Обозначим его как $U$.

### 3.2.3. Интерпретация результата

Для корректной интерпретации результата теста необходимо указать:

- размеры выборок;
- значения меры центральной тенденции для каждой выборки (с учётом непараметрического характера теста, подходящей мерой центральной тенденции представляется медиана);
- значение самой U-статистики;
- показатель CLES [19];
- рангово-бисериальный коэффициент корреляции (RBC) [21];
- принятый уровень значимости (как правило 0.05).

Понятие U-статистики было рассмотрено ранее, большинство других показателей широко известны и не требуют какого-либо отдельного рассмотрения. Остановимся на показателях CLES и RBC.

#### 3.2.3.1. Показатель CLES

**Common language effect size (CLES)** — вероятность того, что значение случайно выбранного наблюдения из первой группы больше значения случайно выбранного наблюдения из второй группы. Данный показатель вычисляется по формуле

$$CLES = \frac{U_1}{n_1 n_2}. \tag{11}$$

Вместо обозначения *CLES* часто используется обозначение *f (favorable)*. Данное выборочное значение является несмещённой оценкой значения для всей совокупности объектов, принадлежащих множеству.

Следует отметить, что значение и смысл данного показателя эквивалентны значению и смыслу показателя AUC[22]. Таким образом, можно говорить о том, что

$$CLES = f = AUC_1 = f = \frac{U_1}{n_1 n_2}. \tag{12}$$

**3.2.3.2. Рангово-бисериальная корреляция** Метод представления степени влияния для U-теста заключается в использовании меры ранговой корреляции, известной как рангово-бисериальная корреляция. Как и в случае с иными мерами корреляции значение коэффициента рангово-бисериальной корреляции может принимать значения в диапазоне $[-1; 1]$, при этом нулевое значение означает отсутствие какой-либо связи. Коэффициент рангово-бисериальной корреляции обычно обозначает как $r$. Для его вычисления используется простая формула, основанная на значении CLES. Выдвинем гипотезу о том, что в паре случайных наблюдений, одно из которых взято из первой выборки, другое — из второй, значение первого больше. Запишем её на математическом языке:

$$H : x_i > yj \quad x\ inX,\, y \in Y. \tag{13}$$

Тогда значение коэффициента рангово-бисериальной корреляции представляет собой разницу между долей случайных пар наблюдений, удовлетворяющей (favorable) гипотезе — $f$, и комплементарной ей доле случайных пар, не удовлетворяющих (unfavorable) гипотезе — $u$. По сути, данная формула представляет собой формулу разности между показателями CLES для каждой из групп.

$$r = f - u = CLES_1 - CLES_2 = f - (1 - f) \tag{14}$$

Существует также ряд альтернативных формул, дающих идентичный результат:

$$r = 2f - 1 = \frac{2U_1}{n_1 n_2} - 1 = 1 - \frac{2U_2}{n_1 n_2}. \tag{15}$$

### 3.2.4. Вычисление p-значения и итоговая проверка нулевой гипотезы

При достаточном большом числе наблюдений в каждой выборке, значение U-статистики имеет приблизительно нормальное распределение. Тогда её стандартизированное значение (z-метка, z-score) [23] может быть вычислено по формуле

$$z = \frac{U - m_U}{\sigma_U}, \tag{16}$$

где $m_U$ — среднее арифметическое $U$, $\sigma_U$ — её стандартное отклонение. Визуализация понятия стандартизированное значения для нормального распределения приведена на рисунке 1. Среднее для $U$ вычисляется по формуле

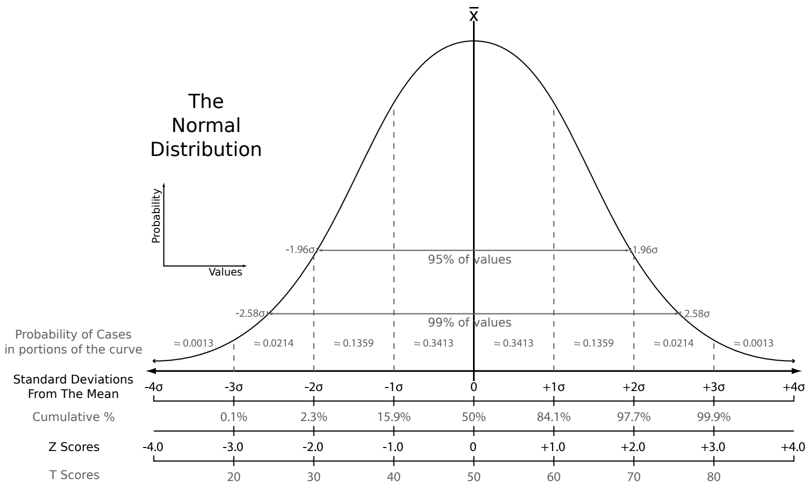


Рис. 1. Визуализация понятия стандартизированного значения (z-score) для нормального распределения [23]

$$m_U = \frac{n_1 n_2}{2}. \tag{17}$$

Формула стандартного отклонения в случае отсутствия связок выглядит следующим образом:

$$\sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}. \tag{18}$$

В случае наличия связок используется другая формула:

$$\sigma_{U_{ties}} = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12} - \frac{n_1 n_2 \sum_{k=1}^{K} ({t_k}^3 - t_k)}{12n(n-1)}} = \sqrt{\frac{n_1 n_2}{12} \left( (n+1) - \frac{\sum_{k=1}^{K} ({t_k}^3 - t_k)}{n(n-1)} \right)}, \tag{19}$$

где $t_k$ — количество наблюдений, имеющих ранг $k$, $K$ — общее число рангов, имеющих связки. Далее, получив стандартизированное значение (z-score), и используя аппроксимацию стандартного нормального распределения, вычисляется *p-значение* для заданного уровня значимости (как правило 0.05). Интерпретация результата осуществляется следующим образом:

$$\begin{aligned} &p < 0.05 \Rightarrow \text{нулевая гипотеза отклоняется} \\ &p \geq \Rightarrow \text{нулевая гипотеза не может быть отклонена.} \end{aligned} \tag{20}$$

## 4. Практическая реализация

### 4.1. Реализация в табличном процессоре LibreOffice Calc

На данный момент можно с уверенностью сказать, что табличные процессоры являются стандартом для расчётов оценщиков. Проникновение средств разработки, например на языке Python либо R, в профессиональную деятельность оценщиков идёт достаточно медленно. Кроме того, самостоятельный пошаговый расчёт позволяет лучше понять методику U-теста. Поэтому было принято решение создать пошаговую инструкцию для проведения U-теста в электронной таблице. Был использован программный продукт LibreOffice Calc (Version: 7.3.3.2, Ubuntu package version: 1:7.3.3 rc2-0ubuntu0.20.04.1 lo1 Calc: threaded), существенная часть функционала которого имеется также и в наиболее распространённом приложении такого рода Microsoft Excel. Отсутствуют основания полагать, что сделанные расчёты не будут корректно работать в приложениях, отличных от LibreOffice Calc. Однако гарантировать это также невозможно. Для однозначно корректного проведения теста рекомендуется использовать именно данное приложение, имеющее версии для всех основных операционных систем. Актуальная версия файла U-test.ods находится в репозитории вместе с остальными материалами данной работы.

Данные, рассматриваемые в данной подсекции, являются вымышленными и были созданы алгоритмом генерации псевдослучайных чисел LibreOffice Calc. Для повторной генерации необходимо использовать сочетание клавиш *ctrl+shift+F9*.

## 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#l0` (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. В: (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. *Гипотеза сдвига*. URL: `http://www.machinelearning.ru/wiki/index.php?title=%D0%93%D0%B8%D0%BF%D0%BE%D1%82%D0%B5%D0%B7%D0%B0_%D1%81%D0%B4%D0%B2%D0%B8%D0%B3%D0%B0` (дата обр. 15.05.2022).

[12] 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).

[13] Machineleraning.ru. *Критерий Уилкоксона для связных 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).

[14] PBC Studio. *RStudio official site*. Английский. URL: https://www.rstudio.com/ (дата обр. 19.08.2021).

[15] CTAN team. *TeX official site*. English. CTAN Team. URL: https://www.ctan.org/ (дата обр. 15.11.2020).

[16] LaTeX team. *LaTeX official site*. English. URL: https://www.latex-project.org/ (дата обр. 15.11.2020).

[17] *TeXLive official site*. URL: https://www.tug.org/texlive/ (дата обр. 15.11.2020).

[18] The R Foundation. *The R Project for Statistical Computing*. Английский. The R Foundation. URL: https://www.r-project.org/ (дата обр. 17.08.2021).

[19] Wikipedia. *Commom Language Effect Size*. English. URL: https://en.wikipedia.org/wiki/Effect_size#Common_language_effect_size (дата обр. 16.05.2022).

[20] Wikipedia. *KISS principle*. URL: https://en.wikipedia.org/wiki/KISS_principle (дата обр. 06.11.2020).

[21] Wikipedia. *Rank-biserial correlation*. English. URL: https://en.wikipedia.org/wiki/Effect_size#Rank-biserial_correlation (дата обр. 16.05.2022).

[22] Wikipedia. *Receiver operating characteristic*. URL: https://en.wikipedia.org/wiki/Receiver_operating_characteristic (дата обр. 17.05.2022).

[23] Wikipedia. *Standard score*. URL: https://en.wikipedia.org/wiki/Standard_score (дата обр. 17.05.2022).

[24] Benito van der Zander. *TeXstusio official site*. URL: https://www.texstudio.org/ (дата обр. 15.11.2020).