

Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ  
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

кафедра социальной информатики

**ОСОБЕННОСТИ ИСПОЛЬЗОВАНИЯ  
НЕПАРАМЕТРИЧЕСКИХ ТЕСТОВ В АНАЛИЗЕ  
СОЦИОЛОГИЧЕСКИХ ДАННЫХ**

(автореферат бакалаврской работы)

студентки 4 курса 451 группы  
направления 09.03.03 - Прикладная информатика  
профиль прикладная информатика в социологии  
Социологического факультета  
Селивановой Юлии Владимировны

Научный руководитель

кандидат социологических наук, доцент \_\_\_\_\_ С.В.Ситникова

подпись, дата

Зав. кафедрой

кандидат социологических наук, доцент \_\_\_\_\_ И.Г.Малинский

подпись, дата

Саратов 2017

## Введение

**Актуальность темы исследования.** Информатизация общества в социальной реальности заключается в воздействии, которое распространяет информационные технологии в обществе и тем самым ускоряет происходящие в нем процессы. Существует проблема, заключающаяся в противоречии между необходимостью разработки новых подходов к социологическому анализу социально-информационной реальности и их отсутствием, что приводит к невозможности прогноза и управления аспектами информатизации.<sup>1</sup>

На сегодняшний день главными задачами социологов являются измерение и анализ социологической информации. Ю. М. Плотинский считает, что социальная информатика должна развиваться путем соединения математики и информатики в социологии. Проблема ориентации социологов в пространстве информации, в то время как есть свободный доступ к любой информации, решается путем развития когнитивных методов в социологии. Данный подход является наиболее подходящим подходом социальной информатики.<sup>2</sup>

**Объект:** эвристический потенциал непараметрических критериев в анализе социологических данных.

**Предмет:** специфика и особенности применения непараметрических тестов в анализе социальных аспектов значимости мобильных телефонов в повседневной жизни студентов.

**Цель** выпускной квалификационной работы состоит в том, что бы раскрыть эффективность, специфику и аналитические возможности применения непараметрических критериев в анализе социологических данных.

Для достижения этой цели необходимо выполнить следующие **задачи**:

- представить теоретические основы аналитической модели на базе непараметрических критериев;

---

<sup>1</sup> Соколова И. В. Информатизация общества как объект социологического исследования/ Соколова И. В.// Ученые записки Российского государственного социального университета, 2008, № 4. С. 20-26.

<sup>2</sup> Соколова И.В. Социальная информатика и социология: проблемы и перспективы взаимосвязи - М.: Союз, 1999. - 228 с.

- раскрыть особенности алгоритмов аналитической работы с применением непараметрических тестов;

- раскрыть и показать эффективность применения непараметрических критериев в анализе социальных аспектов значимости мобильных телефонов в жизни студентов.

**Эмпирическую базу** данной работы составили вторичные данные прикладного исследования.<sup>1</sup> В основу данной работы положены результаты авторского исследования «Мобильный телефон – предмет необходимости или роскоши».<sup>2</sup>

**Практическая значимость** исследования заключается в возможности использования основных положений и выводов квалификационной работы для дальнейшего изучения практик применения непараметрических критериев в анализе социологических данных.

**Структура работы.** Квалификационная работа состоит из введения, трех разделов, заключения, списка литературы и приложений.

### **Основное содержание работы**

**В первом разделе** я рассматривала практику применения непараметрических критериев. Статистика – это наука, которая изучает процессы сбора, измерения и анализа статистических данных.<sup>3</sup>

Статистический критерий - важнейшее правило, которое принимает истинные и отвергает ложные гипотезы с высокой вероятностью.

Критерии подразделяются на 2 вида: параметрические и непараметрические.

---

<sup>1</sup> Грановская, М. Мобильный телефон в жизни студента: роскошь или средство связи? / М. Грановская, А. Лиджиева, Е. Селезнева. // Российская социология завтрашнего дня: сб. студенческих работ / Гос. ун-т - Высш. шк. экономики. - Москва : ТЕИС, 2006. –С. 340 - 446.

<sup>2</sup> Авторское исследование проводилось в 2014 году. Объем выборочной совокупности составил 90 человек.

<sup>3</sup> Википедия [Электронный ресурс]: свободная энциклопедия / текст доступен по лицензии Creative Commons Attribution-ShareAlike; Wikimedia Foundation, Inc, некоммерческой организации. Электрон. дан. (1393089 статей, 5242152 страниц, 198652 загруженных файлов). Wikipedia®, 2001- . URL: <http://ru.wikipedia.org/wiki/Статистика> (дата обращения: 08.04.2017)

Параметрическим называют критерий, построенный на конкретном типе распределения генеральной совокупности. Так как в этих формулах находятся одна или несколько других формул и расчет по ним является более сложным, их еще называют сложными критериями.

Непараметрическим называют критерий, если он не базируется на предположении о типе распределения генеральной совокупности и не использует параметры этой совокупности. Именно поэтому этот критерий так же называют простым. Кроме того расчет по нему не предполагает дополнительных вычислений.

Первым кто использовал непараметрический критерий согласия стал Колмогоров. А. Н, использовавший его в своей работе примерно 80 лет назад. После этого была создана серия непараметрических критериев, уже ставших классическими. Их свойство “свобода от распределения” при проверке простых гипотез нашло широкое применение при решении задач статистического анализа.<sup>1</sup> При нормальном распределении параметрический критерий обладает большей мощностью.<sup>2</sup>

При выборе между параметрическим и не параметрическим критерием нужно принимать во внимание, что наиболее статистической мощностью обладают параметрические критерии. Поэтому в том случае, когда имеется вариационный ряд количественных показателей без явных признаков асимметричности, нужно начать обработку с помощью параметрических критериев. Если она даст результаты, которые далеки от границы значений критериев, то на этом можно остановиться. Если же результаты находятся на пределе значений достоверности, следует с помощью непараметрических критериев, проверить имеется ли достоверность различия. Как раз именно

---

<sup>1</sup> Falunina E. Qualitative and quantitative methods of psychological and pedagogical studies North Carolina - 145 с.

<sup>2</sup> Ермолаев О.Ю. Математическая статистика для психологов Учебник /О.Ю. Ермолаева – 2 е изд. испр. – М.: Московский психолого социальный институт Флинта 2003 – 336 с.

поэтому о непараметрическом критерии говорят как о «вспомогательном критерии».<sup>1</sup>

Непараметрические критерии применяются для следующих переменных:

- для количественных переменных, распределение у которых не подчиняется нормальному закону распределения;
- для переменных, которые измеряются в порядковой шкале;
- для переменных, которые измеряются в номинальной шкале.

Непараметрические критерии так же может применяться в ситуациях при нормальном распределении, однако, в этом случае они будут иметь 95% эффективность в отличие от параметрических критериев.<sup>2</sup>

Существует множество разнообразных непараметрических тестов. Их можно разделить на группы:

- одновыборочные критерии (критерий  $\chi^2$ , критерий Колмогорова - Смирнова, критерий серий и др.);
- критерий для независимых выборок (Критерий Манна-Уитни, критерий Колмогорова – Смирнова для двух выборок, медианный критерий и др.);
- критерий для зависимых или связанных выборок (критерий Кохрана, критерий согласия Кендалла, непараметрический дисперсионный анализ Фридмана и др.).

Одновыборочные критерии проверяют гипотезу о равенстве распределения выборки заданному. Критерии для независимых выборок проверяют гипотезу о совпадении распределений в выборах. Критерии для зависимых выборок проверяют гипотезу о совпадении распределений вы выборках в одной и той же группе распределений.<sup>3</sup>

---

<sup>1</sup> Ашихмина Т.В. Методы математической статистики обработки результатов выпускной квалификационной работы: учебно-методическое пособие. – Киров: Изд-во ВятГТУ, 2014. – 113 с.

<sup>2</sup> [Электронный ресурс]:. <http://www.statmethods.ru/konsalting/statistics-metody/108-neparametricheskie-testy.html>( дата обращения 12.04.2017)

<sup>3</sup> [Электронный ресурс]:. <http://www.statmethods.ru/konsalting/statistics-metody/108-neparametricheskie-testy.html>( дата обращения 12.04.2017)

Существует восемь непараметрических критериев наиболее употребляемых. Все они перечислены ниже:

Критерий Манна - Уитни, сравнивающий две независимые выборки. Этот критерий определяет различия между двумя независимыми выборками по уровню выраженности порядковой переменной.

Критерий знаков и Критерий Уилкоксона сравнивают, сопоставляют две связанные выборки. Однако между ними есть различие. Критерий знаков вычисляет число отрицательных и положительных разностей между повторными измерениями. Критерий Уилкоксона кроме этого так же учитывает величину разности.

Критерий серий распознает, какая последовательность бинарных величин является случайной, а какая упорядоченной.

Биномиальный критерий устанавливает, если отличия между распределением дихотомической величины и заданным соотношением.

Критерий Колмогорова – Смирнова для одной выборки находит разницу между распределением переменной и нормальным распределением, таким как равномерное, Пуассона и др.

Критерий хи-квадрат для одной выборки распознает степень отличия наблюдаемого распределения частот переменной от ожидаемого распределения.

Критерий Н Краскела – Уоллеса сравнивает  $K$  независимых выборок. Этот критерий определяет степень различия между тремя или более независимыми выборками по уровню выраженности порядковой переменной.

Критерий Фридмана сравнивает  $K$  зависимых или как их еще называют связанных выборок. Этот критерий, так же как и предыдущий определяет степень различия по уровню выраженности порядковой переменной только в отличие от критерия Н Краскела – Уоллеса между тремя или более зависимыми выборками.<sup>1</sup>

---

<sup>1</sup> Наследов А. IBM SPSS Statistics 20 и AMOS: профессиональный статистический анализ данных. – СПб.: Питер, 2013. – 416 с.

Таким образом, применение непараметрического критерия не требует знания законов распределения изучаемой совокупности, он может быть применим к совокупностям не только количественным, но и качественным.

**Во второй главе** я рассматривала применение непараметрических критериев в гуманитарных науках. В гуманитарных областях практики применения непараметрических критериев широко используются в социально – статистических исследованиях.

В первом примере для решения задач типа сравнение двух рядов чисел психолог применяет парный критерий Т – Вилкоксона, так как данный критерий используется для оценки различий экспериментальных данных, которые получаются из двух различных условиях, но на одних испытуемых. Плюсом данного критерия является то, что он мало того что обнаруживает оценки различий экспериментальных данных, он еще также определяет насколько этот сдвиг произошел, а также сравнивает сдвиг показателей различных направлениях.

Критерий Т – Вилкоксона построен на ранжирование абсолютными величин разности между двумя рядами выборочных значений в первом и во втором эксперименте. Ранжирование абсолютных величин подразумевает, что знаки разностей не принимают во внимание. Тем не менее, впоследствии вместе с общей суммой рангов, так же еще присутствует самостоятельно сумма рангов для положительных сдвигов и сумма рангов для отрицательных сдвигов. Сдвиг называют типичным, когда интенсивность сдвига в этом направлении и сумма рангов больше, чем в другом направлении. Нетипичным называют сдвиг с меньшей суммой рангов. Типичный и нетипичный сдвиг взаимодополнительны. Однако критерий Т – Вилкоксона в большей степени ориентирован на величине нетипичного сдвига.

Психолог вместе с младшими школьниками проводит коррекционную работу по формированию навыков внимания. Для того что бы дать оценку полученным результатам применяется корректурная проба. Цель состоит в том,

чтобы проверить, изменится ли количество ошибок у младших школьников после коррекционных упражнений.

Для того чтобы добиться этой цели психолог устанавливает число ошибок при выполнении корректурной пробы до и после коррекционных упражнений у 19 детей.

По итогам получилось, что применение коррекционных упражнений способствует возрастанию точности выполнения корректурной пробы, а значит, оказывает позитивное воздействие на улучшение внимания школьников.

Для использования критерия  $T$  – Вилкоксона требуется исполнять ряд условий.

- 1) Измерять можно все шкалы, исключая, конечно, номинальные.
- 2) Выборка должна быть связанной.
- 3) Количество элементов в сравнительных выборках должно быть одинаковым.
- 4) Критерий  $T$  – Вилкоксона может быть использован только в случае если численность выборки будет находиться в пределах от 5 до 50, так как на большее количество таблица достоверности не рассчитана.

Во втором примере используется критерий  $U$  Манна – Уитни для несвязанных выборок. Несвязанные выборки возникают, когда для эксперимента сравниваются данные двух или более выборок, которые получены из разных генеральных совокупностей. Из чего можно заключить, что в экспериментах с несвязанными выборками участвуют разные испытуемые.

Для того чтобы оценить достоверность различий между несвязанными выборками применяется ряд непараметрических критериев. Наиболее популярным считается критерий  $U$ , который используют для оценки различий по уровню выраженности признака для двух независимых выборок. Следует также отметить, что количество испытуемых входящих в выборки может отличаться. Благодаря этой особенности данный критерий очень уместен в ситуации, когда количество испытуемых не очень много, то есть в обеих



выборках их количество не больше 20. Не смотря на это, таблицы критических значений рассчитаны для выборок, у которых количество испытуемых не превышает 60.

Две группы испытуемых с разной численностью решают техническую задачу. Показателем ее успешного выполнения считается время выполнения. Группа испытуемых с меньшей численностью получает денежное вознаграждение в качестве дополнительного стимула. Психолог задается вопросом, который звучит следующим образом. «Оказывает ли влияние вознаграждение на время решения задачи?»

В результате получилось, что вознаграждение не приводит к значительному увеличению эффективности решения технической задачи.

При использовании критерия U требуется исполнять ряд условий:

1) Измерение может быть проведено только в шкале интервалов и отношений.

2) Допускается использование только несвязанных выборок.

3) Нижняя граница применимости критерия должны быть  $n_1 \geq 3$  и  $n_2 \geq 3$

4) Верхняя граница применимости критерия должны быть  $n_1$  и  $n_2 \leq 60$ .<sup>1</sup>

В первом примере психолог используя критерий T – Вилкоксона, сравнивал данные полученные из 2 различных условиях, но на одних испытуемых. В результате получилось, что применение упражнений способствует возрастанию точности выполнения корректурной пробы.

Во втором примере применялся критерий U Манна – Уитни, данный критерий сравнивает выборки, полученные из разных генеральных совокупностей. По итогам вознаграждение не повлияло на эффективность решения задачи.

---

<sup>1</sup> Нижегородцева Н. В. Методические рекомендации по написанию и оформлению курсовой и выпускной квалификационной работы по психологии и конфликтологии. – ГОУ ВПО «Ярославский государственный педагогический университет. К. Д. Ушинского», 2006

**В третьем разделе** эффективность использования непараметрических критериев я рассматривала на основе моего исследования, которое проводилось в 2014 году.

Я начала с того что проверила все переменные с помощью  $Z$  – стандартизации на нормальность, поскольку непараметрические критерии могут применяться только для переменных с ненормальным распределением.

Расчет стандартизированной величины  $Z_{xi}$  осуществляется по формуле

$$Z_{xi} = \frac{x_i - \bar{x}}{S}$$

где  $x_i$  - значение переменной для  $r^i$  респондента;

$\bar{x}$  - среднее значение переменной  $x$ ;

$S$  – стандартное отклонение для переменной  $x$ .

Из свойства стандартного нормального распределения видно, что 68%  $Z_{xi}$  должны лежать в интервале от -1 до 1, 95% – в интервале от -2 до 2.<sup>1</sup>

Непараметрические тесты можно использовать как для количественных и порядковых переменных. В моем исследовании нашлось два таких вопроса.

Первый количественный вопрос - Сколько телефонов вы сменили?

Второй вопрос - Сколько рублей вы примерно тратите в месяц на оплату услуг сотовой связи?

Так как обе переменные имеют ненормальное распределение, применение непараметрических критериев возможно.

В авторском исследовании я использовала  $U$ -критерий Манна — Уитни, так как этот критерий будет наиболее приемлем для независимых выборок с малым объемом.

$U$ -критерий Манна — Уитни является статистическим критерием, который применяется для оценки различий между малыми выборками по уровню какого

---

<sup>1</sup> А. О. Крыштановский. Анализ социологических данных с помощью пакета SPSS: учеб. пособие для вузов / А. О. Крыштановский; Гос. ун-т – высшая школа экономики. – М. : Изд. Дом ГУ ВШЭ, 2006. – 281с.

- либо признака, измеренного количественно. Критерий Манна — Уитни выявляет различия в значении параметра между малыми выборками.<sup>1</sup>

Критерий Манна – Уитни или как его еще называют U-критерий, по назначению похож на t-критерий для независимых выборок. Отличие состоит в том, что t-критерий ориентирован на нормальные и близкие к ним распределения, в то время как критерий Манна – Уитни – на распределения отличных от нормальных. Существуют случаи, когда критерий Манна – Уитни можно применять для нормально распределенных данных, только в этом случае данный критерий менее чувствителен к различиям, чем t-критерий.

Наша первая гипотеза предполагала, что количество сменившихся телефонов у студентов зависит от их формы обучения.

Я проверила предполагаемую гипотезу. В результате выяснилось, что она является верной. Таким образом, можно сделать вывод о том, что студенты коммерческой формы обучения чаще, чем студенты бюджетной меняют мобильные телефоны.

Вторая гипотеза о зависимости количества сменившихся телефонов от пола студента.

По результату проверки предполагаемая гипотеза оказалась неверной. То есть студенты обоих полов примерно одинаковое количество раз меняют свои телефоны.

Третья гипотеза звучит следующим образом – студенты коммерческой формы обучения тратят на услуги связи больше чем студенты бюджетной формы обучения.

После проверки гипотезы с помощью критерий Манна — Уитни оказалось, что гипотеза является неверной. Отсюда следует, что студенты обеих форм

---

<sup>1</sup> Википедия [Электронный ресурс]: свободная энциклопедия / текст доступен по лицензии Creative Commons Attribution-ShareAlike; Wikimedia Foundation, Inc, некоммерческой организации. Электрон. дан. (1393089 статей, 5242152 страниц, 198652 загруженных файлов). Wikipedia®, 2001-. URL: [http://ru.wikipedia.org/wiki/U-критерий\\_Манна\\_—\\_Уитни](http://ru.wikipedia.org/wiki/U-критерий_Манна_—_Уитни) (дата обращения: 16.082016)

обучения тратят примерно одинаковое количество средств на оплату мобильных услуг.

Четвертая гипотеза, заключается в том, что пол студента влияет на количество сменных им телефонов.

По результатам проверки стало очевидным, что данная гипотеза не подтверждается. Из чего можно заключить, что студенты обоих полов в месяц тратят на оплату мобильного телефона примерно одинаковое количество средств.

Непараметрический критерий имеет ряд достоинств:

- производит поиск качественно – количественных связей,
- с его помощью можно подтвердить или опровергнуть проверяемую гипотезу
- применяется при небольшом объеме выборок
- расчет прост, но при этом он обладает достаточной мощностью.

Все это на практике позволило провести корреляционный анализ статистических данных, для которых применение параметрических критериев было некорректно.

**В заключении** я сделала общие выводы. В наши дни услугами сотовой связи пользуется практически каждый человек. Отдельного упоминания требуют студенты. Кажется, что сейчас было бы невероятно трудно отыскать студента, у которого нет мобильного телефона.

Рынок мобильных товаров и услуг значительно вырос за последние несколько лет. Производители разрабатывают новые модели телефонов, снабженные огромным количеством функций, активно работают над дизайном продукта. То или иное техническое усовершенствование (часто незначительное) специально подчеркивается, создается реклама, пытающаяся убедить приобрести новую модель с усовершенствованным дизайном или дополнительной функцией. В результате у потребителя появляется желание сменить мобильный телефон, купить более «продвинутой», даже если в этом нет необходимости и прежний еще вполне пригоден для использования.

Практическая значимость моего исследования состоит в применении непараметрических тестов на примере анализа значимости мобильного телефона для студентов.

Непараметрические критерии – группа статистических критериев, основанных на оперировании частотами или рангами. Непараметрический критерий в отличие от параметрического критерия не основывается на оценке параметров при описании выборочного распределения. Именно поэтому критерий еще называют свободным от параметров.

В своем исследовании я использовала U-критерий Манна — Уитни, так как этот критерий подходит для выявления различий между малыми выборками, а также для оценки различий между выборками по уровню определенного признака.

В ходе проведения исследования доказана гипотеза о том, что форма обучения влияет на количество сменившихся телефонов. Среди студентов коммерческого отделения больше вероятность отыскать тех, у кого имеющийся мобильный телефон далеко не первый.

Вторая гипотеза о том, что пол влияет на количество сменившихся телефонов, не подтвердилась, так как и юноши и девушки в среднем меняют телефоны одинаковое количество раз.

Третья гипотеза, состоящая в том, что студенты коммерческой формы обучения больше чем студенты бюджетной формы обучения тратят средства на услуги связи в месяц, не подтвердилась.

Четвертая гипотеза, заключающаяся в том, что пол студента влияет на количество сменившихся им телефонов, также не подтвердилась.

Непараметрический критерий обладает множеством преимуществ. Он применяется для количественных, порядковых переменных. Так как непараметрический критерий не использует параметры генеральной совокупности, он достаточно прост в использовании, однако при этом обладает достаточной мощностью. В отличие от параметрического критерия используется для небольших выборок.