


См. обсуждения, статистику и профили авторов для этой публикации по адресу: <https://www.researchgate.net/publication/324703583>.

Тестирование на отсу ст вие свя зи между к а т егориаль ными переменными

Статья в The Quantitative Methods for Psychology - апрель 2018 г.  
DOI: 10.20982/tqmp.14.3.p167

Ц И Т А Т Ы  
8

З авт ора, в т о м ч исле:




Линд а Фармус

Йорк ск ий университет

20 ПУБЛИК АЦ ИИ132 Ц ИТ АТ А

СМОТ Р ЕТЬ ПРОВИТЬ

Ч ИТ АЕТ  
402



Роберт К. рэббит

Йорк ск ий университет

114 ПУБЛИК АЦ ИИ3505 Ц ИТ АТ ОВ

СМОТ Р ЕТЬ ПРОВИТЬ

Некоторые авторы этой публикации также работают над следующими родственными проектами:

- Project

Healthy Coping Project [Посмотреть проект](#)
- Project

Образец проекта "Планирование размера"

Тест на отсутствие отношений

Средикатегориальных переменных

Таня Шишкина (orcid.org/0000-0002-1500-5520)

Линда Фармус (orcid.org/0000-0002-5303-6408)

Роберт А. Крибби (orcid.org/0000-0002-9247-497X)

Программа количественных методов

Кафедра психологии

Йоркский университет

## Абстрактный

Часто бывает сложно определить отсу́тствие связи между двумя или более категориальными переменными.

необходимы психологические проекты, таких как сравнительный анализ результатов, оценок групповых

эквивалентности на исходном уровне и оценок результатов терапии. Несмотря на это, литература

редко предлагает информацию или технические рекомендации от носителей соответствующих

статистической методологии, которая будет использоваться для выполнения этой задачи. В этой статье исследуются две эквивалентности

тесты для категориальных переменных, введенные Роджерсом, Ховардом и Весси (1993) и

еще один Wellek (2010), а также предложенная стратегия, основанная на модели Крамера V. Моделирование

было проведено исследование, чтобы изучить и сравнить ошибки типа I и мощности, связанные с

эти тесты. Результаты показывают, что V-процентура Крамера, основанная на эквивалентности, является наиболее

подходящим методом для определения отсу́тствия связи между категориальными переменными в двух

способностях.

Ключевые слова: проверка эквивалентности, категориальные переменные, табличная частота.

Тест на отсуствие от ношений

Средикатегориальных переменных

В психологических исследованиях исследователи часто заинтересованы в подтверждении отсуствия связи между двумя или более категориальными переменными. Это часто имеет место для конструкций, так как как сравнительный анализ исходных оценок групповой эквивалентности по категориальным переменным на базовом уровне. Например, исследователи могут захотеть продемонстрировать, что два или более лечения равны в зависимости от их терапевтических результатов, так как возвращение к работе, снижение долга, приверженность лечению или воздержание от азартных игр (Chen, Tsong, & Kang, 2000). В Кроме того, экспериментальные группы часто оцениваются в начале исследования, чтобы уловить эквивалентность демографических и/или клинических переменных (Rogers, Howard, & Vessey, 1993). В этих обстоятельствах для исследователей крайне важно использовать соответствующие методологию, чтобы сделать вывод, что категориальные переменные эквивалентны в группах.

Буоли, Кумерлато Мелтер, Калдиероли и Альтамора (2015) представляют собой практический пример статистических усилий по установлению эквивалентности. Авторы исследовали эффективность различных фармакологических классов антидепрессантов при длительном лечении большого депрессивного расстройства. На исходном уровне исследователям необходимо было продемонстрировать эквивалентность экспериментальных групп по демографическим характеристикам (например, работе и семейному положению) и клиническим характеристикам (например, семейный анамнез психических проблем, количество попыток суицида и участие в злоупотреблении психоактивными веществами). Авторы применили критерий независимости хи-квадрат для определения были ли экспериментальные группы эквивалентны по этим переменным. Цель состояла в том, чтобы найти и не значительный результат, чтобы сохранить нулевую гипотезу о том, что группы показали одинаковые результаты.

Другой пример взят из исследования Bailine et al. (2010), которые стремились оценить, одинаково ли пациенты с биполярной и униполярной депрессией реагировали на

ПРОВЕРКА ЭКВИВАЛЕНТНОСТИ ДЛЯ АНАЛИЗА ЧАСТИ

электропроводная терапия. На исходном уровне эквивалентности необходимо было доказать

равенство однополярных и биполярных субъектов по демографическим признакам (таким как

пол, раса, образование и семейное положение) и клинические переменные (таким как наличие или

отсутствия психоза). Более того, для оценки гипотезы о том, что электропроводная

терапия повлияет на эти группы в равной степени, авторы сравнили две группы точкой зрения

их реакции на лечение, определяя снижение на 50% по сравнению с исходным уровнем на

Рейтинговая шкала Гамильтона для депрессии как положительный результат. Соответственно, ч

квадратные тесты независимости были проведены как для базовых сравнений, так и для

ответа на гипотезы лечения, с целью демонстрации отсутствия статистической

значимости (т.е. радиационная нулевая гипотеза не отвергается).

Эти два примера подчеркивают исследования, в которых исследователи изучают отсутствие

взаимосвязи между категориальными переменными; они также иллюстрируют неправильное использование радиационного нулевого значения.

проверка гипотез для выявления так называемого отсутствия. Важно сразу заявить, что мы не

занимались критическими решениями, принятыми авторами в рассмотренных выше примерах;

действительно, эти два исследования были выбраны из бесчисленного множества других, проводившихся аналогично.

процедуры Скорее, мы хотим подчеркнуть, что подходящие тесты для оценки отсутствия родства

средних категориальных переменных в настоящее время широко неизвестны и они не всегда доступны

пакетов статистических программ. Кроме того, открытое обсуждение ограничений радиационных

методов оценки эквивалентности средних категориальных переменных немногочисленны. Также, пока эти

примеры иллюстрируют ситуацию, когда исследуют однородность (т.е. эквивалентность

пропорций по группам), те же проблемы возникают при исследовании независимости. Эта бумага

решает от носительную нехватку информации о недостатках обычных

ПРОВЕРКА ИЭК ВИВАЛЕНТНОСТИ ДЛЯ АНАЛИЗА ЧАСТИ

методы оценки отсутствия взаимосвязи между категориальными переменными и предлагает

более надежные методологические инструменты, которые подходят для этой задачи.

Введение в тестирование эквивалентности

Как показывают предыдущие примеры, исследователи обычно пытаются вывести эквивалентность

групп или установить отсуствие связи между переменными на основании отсуствия

существенных различий или ассоциаций. Однако этот метод не подходит для некоторых

причин. Во-первых, как отмечает Quertemont (2011), незначительные результаты могут возникнуть из-за недостаточного

статистической мощности. Таким образом, ранее статистически незначимые различия между группами могут

на самом деле становиться значимыми после достаточного увеличения размера выборки. Во-вторых, неспособность

отклонения нулевой гипотезы означает, что нулевая гипотеза верна; это просто означает, что есть

в настоящее время недостаточное доказательство, чтобы сделать вывод о том, что это неверно (Walker & Nowacki, 2010).

Теоретическая формулировка нулевой гипотезы для тестов эквивалентности прямо противоположна

утверждению нулевой гипотезы для традиционных тестов на основе различий (Крибби, Арпин-Крибби и

Груман, 2009). Такая динамика свидетельствует о необходимости статистических процедур, специально предназначенных для

проверки отсутствия взаимосвязи между переменными, что и является целью эквивалентности

тестирования.

Чтобы проверить отсуствие взаимосвязи между переменными, исследователи начинают с

выбора наименее подходящей ассоциации, которую исследование призвано практиковать

значительный. На практике ошибка выборки приводит к нулевым связям (например, идентичные средние, нулевые значения).

корреляции) невозможно (Counsell & Cribbie, 2015). Цель проверки эквивалентности состоит не в том, чтобы

тестировать полное отсуствие взаимосвязи между переменными, а скорее исследовать, есть ли различия

обнаруженные актуальны (Cribbie, Gruman, & Arpin-Cribbie, 2004). Чтобы выполнить эту задачу,

исследователи должны дать количественную оценку своей окончательной нерелевантной разницы, приняв решение о критическом

диапазон значений, называемый интервалом эквивалентности, часть обозначаемый симметрично с помощью  $(-\delta, \delta)$ ;  $\delta$  может

представлять собой интересующий эффект, такой как отсутствующие корреляции или нерелевантная разница в

пропорциях. Интервал эквивалентности обычно имеет как верхний, так и нижний пределы при этом

определенный диапазон, представляющий наименьшую ассоциацию (например, разница в популяции

пропорциях), которую в рамках исследования будут считать значимыми.

Нулевая гипотеза проверки эквивалентности утверждает, что взаимосвязь между

переменных не меньше, чем эффект, указанный исследователем через эквивалентность

интервал. И наоборот, альтернативная гипотеза утверждает, что отклонения между

переменных меньше заданного через интервал эквивалентности. Эквивалентность или отсутствия

связи устанавливается, когда данные представляют достаточное доказательство, чтобы заключить, что

величина отклонения попадает в интервал эквивалентности (Schuirmann, 1987; Walker &

Новацкий, 2011). Не существует фиксированных правил установления пределов эквивалентности; их оправдание

сильно зависит от характера исследования, интересующей переменной исхода, предыдущих результатов

в конкретных областях исследований, а также суждения соответствующих экспертов о риске/пользе (Комитет по

Лекарственные препараты для человека, 2006 г.). Например, O'Reilly et al. (2007) проверил

эквивалентность телепсихиатрии и очной психиатрии с помощью консультации. Один из исходов

измерялась доля участников, госпитализированных в психиатрическую больницу в течение двенадцати

месяцев после первоначальной оценки. Следователи, посоветовавшись с психиатрами, решили

что разница в пропорциях 10% между группами будет наименьшей клинической

значительная разница, приводящая к ненаправленному запасу эквивалентности  $(-\delta, \delta) = (-0,10, 0,10)$ .

Эквивалентность есть для отклонений между категориальными переменными

В этом проекте мы рассмотрим релятивный подход к тестированию отсуствия связи между

категориальными переменными. Первый, описанный Rogers et al. (1993), представляет собой модифицированную версию

двух пропорций (TOST; Schuirmann, 1987) процедура, цель которой является проверка

эквивалентности двух пропорций, обозначаемых как  $p_1$  и  $p_2$ . Тест основан на нормальном

аппроксимация разницы между двумя пропорциями. Мы будем ссылаться на этот тест

эквивалентности двух пропорций как тест «ЕР». Хотя он оформлен как тест на однородность, он может

также использоваться для оценки независимости двух категориальных переменных. Например, исследователи

можно использовать этот тест, чтобы проверить, что пропорция самцов и самок в контроле и

экспериментальной группы похожи или что пол (мужской/женский) минимально связан с выбором боли

лекарства (препарат А против препарата Б).

Первая нулевая гипотеза,  $H_0: p_1 - p_2 = \delta$ , отвергается, если  $z_1 - z_2$  адекватная,  $H_0: p_1 -$

$p_2 = \delta$ , отклоняется, если  $z_2 - z_1$  где:

$$z_1 = \frac{(\hat{p}_1 - \hat{p}_2) - \delta}{\sqrt{\hat{p}_1(1-\hat{p}_1) + \hat{p}_2(1-\hat{p}_2)}}, \quad z_2 = \frac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}_1(1-\hat{p}_1) + \hat{p}_2(1-\hat{p}_2)}},$$

•

— доля выборки, а  $z_1 - \alpha$  и  $z_2 - \alpha$  — значения стандартного нормального распределения,

отсекать нижние доли распределения  $1 - \alpha$  и соответственно. Заметим, что  $(-1)z_2 = z_1 - \alpha$

Стандартную ошибку разницы между двумя пропорциями можно рассчитать, используя:

$$\hat{p}_1, \hat{p}_2 = \{[\hat{p}_1(1-\hat{p}_1)/n_1] + [\hat{p}_2(1-\hat{p}_2)/n_2]\}^{1/2},$$

где  $n_1$  и  $n_2$  — размеры выборок для первой и второй групп соответственно. Когда оба нулевые

гипотезы отвергнуты, исследователи могут отклонить нулевую гипотезу о том, что

разница в пропорциях больше, чем  $\delta$  ( $-\delta < p_1 - p_2 < \delta$ ; разница в пропорциях

попадает в интервал эквивалентности), или, другими словами, что между

двух дихотомических переменных. Процедура ЭП функционально сравнима с простой

подход с асимптотическим интервалом (SAI) (Баркер, Ролка, Ролка и Браун, 2001). Согласно

Метод SAI, если  $\hat{p}_1 - \hat{p}_2 \pm z_{1-\alpha} \sqrt{\hat{p}_1(1-\hat{p}_1) + \hat{p}_2(1-\hat{p}_2)}$  доверительный интервал (ДИ) для  $p_1 - p_2$  находится в пределах эквивалентности



интервал  $(-\delta, \delta)$ , то  $H_0$  и  $H_1$  отбрасываются на заданном уровне  $\alpha$ . Важно

Ограничение теста EP состоит в том, что он применим только к конструкции  $m \times 2$ .

Другая процедура проверки иэковалентности, описанная Welles (2010), основана на

Евклидово расстояние (т.е. расстояние между двумя точками в евклидовом пространстве). Мы будем ссылаться на него

как процедура «ЭД». Нулевая гипотеза для этого теста утверждает, что сумма квадратов

расстояний  $(D^2)$  между наблюдаемыми вероятностями чек, обозначенными как  $p_i$  и ожидаемыми чек-ами

вероятностей (произведение предельных сумм), обозначаемое как  $g(p)$ , в популяции, по крайней мере, как

более критическое расстояние. Welles предложил  $\epsilon = 0,15$  как наибольшее приемлемое расстояние между

$p$  и  $g(p)$  (т.е. между вектором наблюдаемых вероятностей и вектором ожидаемых

вероятностей), однако исследователи должны рассмотреть, какое значение  $\epsilon$  является наиболее подходящим, учитывая

характеристики исследователя (итребуются дополнительные исследования для понимания величины в

чтобы помочь исследователям установить соответствующие значения для  $\epsilon$ ). Оять же, этот тест может быть использован для

исследования либо независимости, либо однородности групповых пропорций. Таким образом,  $H_0$ :  $\epsilon^2 = 2$

отклоняется, если:

$$D^2 + 1 / \epsilon^2 < 2,$$

где  $\epsilon$  — стандартная ошибка:

$$D^2 = \sum_{i=1}^n (p_i - g_i)^2 / g_i, \quad \text{знак равно} \quad \sum_{i=1}^n p_i^2 + \sum_{i=1}^n g_i^2,$$

где  $i, j$  задает строку и столбец соответственно,  $r$  — количество строк,  $s$  — количество

столбцов.  $\sum_{i=1}^n p_i^2$  представляет собой сумму наблюдаемых вероятностей для строк  $i$ , и  $\sum_{i=1}^n g_i^2$  представляет собой сумму

наблюдаемых вероятностей для столбца  $j$ . Другой способ формулировать тест ED состоит в том, что нулевая гипотеза

важного отклонения между двумя категориальными переменными можно отклонить, если верхний

предел CI для  $D^2$  меньше 2. Дисперсия  $D^2$  может быть выражена как:

$$2 = \begin{matrix} \diamond & \diamond & \wedge \\ =1 & =1 \end{matrix} \quad 2 \quad \begin{matrix} \diamond & \diamond & \diamond & \wedge \\ 2=1 & 2=1 & 1=1 & 1=1 \end{matrix} \quad \begin{matrix} \wedge \\ 11 \quad 221122 \end{matrix},$$

где:

$$\begin{matrix} \wedge \\ = 2 \diamond \diamond \quad + + \diamond \quad \diamond \diamond \quad + + \diamond + \diamond \quad [(\quad =1 \quad + + \quad) + \diamond.] \end{matrix}$$

Наконец, мы предлагаем подход, основанный на  $V$ -критерии (Cramer, 1946), упомянутом в этой статье.

бумага как резюме.  $V$ -критерий, мера размера эффекта для ассоциаций между категориальными переменными,

учитывает размеры таблиц частот, из чего следует, что  $V$  для таблиц разных

размеров можно осмысленно сравнивать (Smithson, 2003). Таким образом,  $V$ -критерий можно использовать для

определить отсуствие связи между категориальными переменными в общем двустороннем (или выше)

таблицы, поскольку  $V$  находится в диапазоне от 0 до 1, задача нахождения подходящего интервала эквивалентности

сделано проще.

Чтобы вернуться к нулевой гипотезе о неэквивалентности для  $CV$ -подхода ( $H_0: \quad$ ),

верхний предел критерия  $V$  CI должен лежать ниже заданной границы эквивалентности  $\delta$ .

$V$ -критерий вычисляется как:

$$= \frac{2}{[(\quad - 1)]},$$

где  $k$  — меньшее из чисел строк  $r$  или столбцов  $s$ . Следуя Смитсону (2003),  $K$  и

для критерия  $V$  рассчитывается как:

$$= \frac{(\quad + )}{[(\quad - 1)]}, \quad \text{и}$$

$$= \frac{(\quad + )}{[(\quad - 1)]}$$

где  $m = (s - 1)(r - 1)$ , а  $\Delta L$  и  $\Delta U$  представляют собой нижний и верхний доверительные пределы

параметра неэквивалентности для неэквивалентности  $\chi^2$  (критерий определяется итерацией, см. Signorell,

2017). Мы включили или прикладной пример для всех трех методов (см. Приложение А).

Метод

Исследование методом Монте-Карло было проведено для оценки ошибок типа I и мощности генератора.

Процедуры проверки эквивалентности EP, ED и CV для категориальных переменных в 2 x 2 (EP, ED, CV)

и 2 x 4 (ED, CV). Мы использовали  $\alpha = 0,05$  и выполнили 5000 симуляций для каждого условия.

с помощью статистического программного обеспечения R с открытым исходным кодом (R Development Core Team, 2016).

Управляемые переменными были размеры выборок и степень ассоциации (для условия мощности) и

дизайн исследования (конкретные условия можно найти в таблице 1). Размеры выборок 50, 100, 200 и

1000 были исследованы потому что они обычно используются в психологических исследованиях. Мы сосредоточились

только в случае равных суммарок для плана исследования 2 x 2, в то время как в плане исследования 2 x 4

суммарок не обязательно были равными. Обратите внимание, что Wellek (2010) признает либеральный характер

Процедуры ED (т.е. эмпирическая частота ошибок типа I может превышать  $\alpha$ ) и, следовательно, рекомендует

использованием скорректированного номинального уровня  $\alpha^* < \alpha$  «всякий раз, когда требуется строгое поддержание заданного уровня».

считалось необходимым требованием» (стр. 277). Таким образом, был использован номинальный уровень  $\alpha 0,05$ .

для всех тестов, кроме ED, где мы использовали как  $\alpha = 0,05$ , так и  $\alpha^* = 0,025$ .

Когда степень связи между переменными точно соответствует эквивалентности

интервала, ожидается, что эмпирическая частота ошибок типа I будет равна номинальной частоте ошибок типа I ( $\alpha$ ). Мы

использовали свободные границы Брэдли (1978),  $\alpha \pm 0,5\alpha$  в качестве критерия наличия удовлетворительной ошибки типа I.

контрроль. Таким образом, при  $\alpha = 0,05$  частота ошибок типа I приемлема, если она находится между 0,025 и 0,075.

Поскольку в тестах EP, ED и CV используются разные шкалы процедуры определения

интервал эквивалентности также различались. Следуя рекомендации Wellek (2010) по

В тесте ED интервал эквивалентности был установлен как  $\epsilon = (-,15, 0,15)$ . Таким образом, для ошибок I рода

условия мы выбрали значения долей леток популяции, которые дали евклидову

расстояние  $d = 0,15$  (конкретные случаи 2 x 4 фактически были выбраны из тех, которые были получены Веллеком,

2010, с. 276). Эти пропорции клеток затем использовались для расчета сопоставимой эквивалентности.

интервалы для пропорций  $EP$  и  $CV$  (т.е. мы вычислили разницу в пропорциях и

ВК-размера для пропорций клеток популяции, которые дают  $d = 0,15$ ). В результате в типе I

В условиях ошибок и оценок эквивалентности для теста ED всегда была одинаковой,  $\epsilon = 0,15$ , в то время как для

интервал  $EP$  и  $CV$  менялся для разных матриц населения (см. табл. 1).

Для условий мощность и сила связаны между двумя дихотомическими

переменные, измеряемые евклидовым расстоянием, принимались равными 0, 0,02 и 0,10. Для

В условиях мощность и верхний предел границ эквивалентности составлял  $\epsilon = 0,15$  для теста ED, 0,30 для

теста EP и 0,40 для теста CV. Как указывалось ранее, Wellek (2010) рекомендует использовать  $\epsilon = 0,15$ .

с тестом ED, а значения для EP и CV были сопоставимы по силе с значением для  $\epsilon = 0,15$ .

Мы также провели тест независимости хи-квадрат, чтобы сравнить производимость

традиционный подход с тестом эквивалентности. Важно подчеркнуть, что, поскольку

цель состоит в том, чтобы продемонстрировать отсутствия ассоциаций (независимость/однородность), этот тест не будет

уместно, поскольку цель состоит в том, чтобы отвергнуть нулевую гипотезу; он включен, хотя, поскольку

этот метод часто используется исследователями для демонстрации отсуствия ассоциаций. Чтобы измерить

сопоставимый результат для каждого теста, переменная результат была пропорцией симуляций в

вывод, касающийся «отсутствия связи». Для тестов эквивалентности это означает отклонение

$H_0$ , но для традиционного критерия независимости хи-квадрат это означает отвергнуть  $H_0$ . Примечание,

следовательно, в условиях ошибок типа I и условиях ненулевого эффекта мощность для

теста эквивалентности по-прежнему, что заявленные показатели для традиционного теста хи-квадрат являются ошибками типа II, и

для условия мощность с нулевым эффектом для тестов эквивалентности это переводится в показатели правильных

неотклонения для теста хи-квадрат (с ожидаемой долей 1- $\alpha$ ).

Полученные результаты

2 x 2 Дизайн

Частота ошибок типа I. Доля случаев, в которых отсутствует связь между двумя дихотомическими переменными делают ложный вывод для тестов EP, ED и CV (частота ошибок типа I), поскольку так же вероятность ошибок типа II для критерия независимости хи-квадрат представлены в

Рисунок 1.

Как EP, так и CV-процедуры имели частоту ошибок типа I, которая соответствовала ограничениям Брэдли (0,025–0,075). В отличие от EP и CV, размер выборки значительно повлиял на тип I частоту ошибок теста ED; при небольшом размере выборки ( $N = 50$ ) показатели типа I были в два раза выше по сравнению с  $N = 1000$  при  $\alpha = 0,05$  и в среднем в три раза больше при  $N = 50$  по сравнению с  $N = 1000$ , когда  $\alpha = 0,025$ . Только при  $N = 1000$  эмпирическая частота ошибок типа I ED падает в рамках Брэдли.

Доля неотклонений нулевой гипотезы для теста хи-квадрат, когда в тот факт, что это неверно (частота ошибок типа II), представлен здесь для сравнения с тестами эквивалентности. Как и ожидалось, частота ошибок типа II для критерия хи-квадрат сильно связана с размером выборки. Таким образом, для небольших размеров выборки, таких как  $N = 50$ , критерий хи-квадрат имеет наибольшую вероятность обьявление эквивалентности по сравнению с остальными тестами эквивалентности, рассмотренными в этой статье, тогда как он имеет наименьшую вероятность обьявления эквивалентности, когда размер выборки увеличивается до  $n = 1000$ . Оба этих результата являются ожидаемыми, учитывая то, что регистрируются ошибки типа II.

Тарифная мощность. Вероятности правильного заключения об эквивалентности для ED, EP и Процедур CV (мощности) представлены на рисунке 2. Все тесты эквивалентности, рассмотренные в этом документе, имеют сходные модели точечных показателей мощности, т.е., как и ожидалось, они увеличиваются по мере того, как образец увеличивается. ED с  $\alpha = 0,025$  и тесты CV имеют одинаковые значения мощности в разных диапазонах. Степень ассоциации. Процедура EP показывает меньшую мощность, чем другие тесты эквивалентности, когда размер выборки  $N = 50$ . Однако эта разница в мощности исчезает, когда выборка

размер увеличивается до  $N = 200$ . Как и ожидалось, степень связи между категориальными переменными

оказывает существенное влияние на мощность, при этом  $d = 0$  дает заметно более высокие вероятности, чем  $d =$

.1. Структура пропорций членов населения умеренно влияет на мощность процедуры EP (например,

0,250, 0,250, 0,250, 0,250 против 0,100, 0,400, 0,100, 0,400), но не для тестов ED или CV.

В условиях мощности, когда  $d = 0,02$  или  $d = 0,10$ , заявленные скорости для радиационной ц

кватратным критерием независимости являются вероятности не отвергнуть  $H_0$  (ошибка второго рода). Как показано на рисунке 1

и 2, как и ожидалось, показатели больше (а не меньше) евклидово расстояние, что отражает

степень связи между категориальными переменными - приводит к меньшей частоте ошибок типа II.

2 x 4 Дизайн

Частота ошибок типа I. На рис. 3 показаны эмпирические частоты ошибок типа I для ED и

Тесты CV и вероятности ошибок типа II для критерия независимости хи-квадрат в 2 x 4

Дизайн исследования. Тест ED с  $\alpha = 0,05$  выявляет несоответствующую эмпирическую частоту ошибок типа I.

(преимущественно для условия 1) для размеров выборки менее  $N = 200$ . Когда  $\alpha = 0,025$ , тест ED

Частота ошибок типа I остается в пределах пределов Брэдли для всех условий, кроме одного. Хотя слегка

консервативно, частота ошибок типа I для процедуры CV всегда находилась в пределах установленных Брэдли для

прочность. Подходы ED и CV были чувствительны к пропорциям клеток популяции.

при этом частота ошибок типа I незначительно различается в зависимости от условий. Все тесты вивалентности

учитываются для дизайна исследования 2 x 4, зависящего от размера выборки с меньшей ошибкой типа I.

скорости, происходящие с большими выборками. Как и в случае 2 x 2, доля случаев, в которых

радиационный критерий независимости хи-квадрат дал ложный вывод о том, что ассоциация была нулевой.

умеренный для условия  $N = 50$ , но равный или близкий к нулю для всех  $Ns > 50$ .

Тарифы на мощность. Изучение значений на рисунке 4 показывает, что ED и CV

тесты показывают аналогичные показатели мощности, когда  $d = 0$  (условия 1 и 2) и  $d = 0,02$  (условия 3 и 4),

хотя тест ЭД в целом был мощнее. При  $d = 0,1$  (условия 5 и 6) мощность

показатели снижались, особенно для теста CV, потому что степень ассоциации очень высока.

Близко к интервалу эквивалентности. Как ожидалось, по мере увеличения размера выборки показатели мощности для

оба теста эквивалентности растут. Для традиционного теста, когда ассоциация равна нулю записанное

скорости оказались, как ожидалось, примерно равными 1 -  $\alpha$ . Когда ассоциация была больше нуля

но в пределах интервала эквивалентности частота ошибок типа II уменьшилась для традиционного хи.

квадратный критерий независимости по мере увеличения размера выборки. Однако эти показатели подчеркивают

проблема с использованием теста на основе различий для оценки эквивалентности; при небольшом размере выборки

Тест CS частично неправильно заключал эквивалентность, тогда как при большем размере выборки эквивалентность была

редко или никогда не завершается. Другими словами, верооятность необоснованного объявления об отсуствии

связи между переменными уменьшалась, а не увеличивалась по мере увеличения размера выборки.

### Обсуждение

Многие психологические исследования явно нацелены на то, чтобы показать, что между

исследуемые категориальные переменные. Частота исследований хочет показать, что, прежде чем

В начале исследования некоторые ключевые характеристики (также как этническая принадлежность, статус работы, состояние здоровья и

образовательный статус) одинаковы среди различных групп. Другим типичным примером является

исследователя, который хочет продемонстрировать, что разные подходы к лечению дают сходные результаты

частоты для разных групп (например, мужчины и женщины, городские и сельские жители).

Несмотря на общность этих обстоятельств, попытки доказать эквивалентность

группы или методы лечения обычно страдают от двух основных проблем. Во-первых, тестирование на отсуствие

Ассоциация частично сопряжена с трудностями, связанными с выбором подходящего

статистического метода. Во-вторых, возникла трудность с адекватным определением «эквивалентности».

Связано с понятием границы эквивалентности.

Что касается вопроса выбора подходящего статистического теста, в этом исследовании статистические свойства трех различных подходов для проверки отсуствия связи между категориальными переменными в планах 2 x 2 и 2 x 4. Кроме того, мы исследовали взаимосвязь эти тесты на эквивалентность традиционному критерия независимости хи-квадрат с целью предоставить рекомендации исследователям поведения относительно их пригодности и практичности. Несколько ключевых отличий отличают различные тесты эквивалентности, рассмотренные в этой статье. Бумага. Процедура ED основана на разнице между наблюдаемой и ожидаемой частотами, и его логика близка к логике традиционного критерия независимости хи-квадрат. ЭПтест это коренится в различных пропорциях. Подход CV использует метрику корреляции, и так образом, потенциальные значения варьируются от 0 до 1. Чтобы сравнить статистические свойства этих трех проверок эквивалентности, необходимо было определить, как изображения их интервалов эквивалентности могут быть приравнены (имея в виду, что каждый из них измеряется по разным шкалам). Так образом интервал эквивалентности для теста ED соответствует рекомендуемому евклидову расстоянию  $d = 0,15$ . (Wellek, 2010), что дает набор двусторонних таблиц частот, которые позволили нам изучить Type I частоты ошибок и мощность (см. Таблицу 1). Учитывая эти производные таблицы частоты населения, затем были определены интервалы эквивалентности для тестов EP и CV, чтобы соответствовать ассоциации населения (условия ошибок типа I) или быть пропорциональными евклидову расстоянию мера, используя тест ED. Наши результаты показывают, что предлагаемый тест, основанный на V Крамера обеспечивает наилучший баланс между количеством ошибок типа I и мощностью юидоступен как для 2 x 2 и более крупных двусторонних конструкций.

Что касается вопроса выбора подходящего интервала эквивалентности, некоторые важные вопросы возникли в этом исследовании. Как отмечено в таблицах результатов, евклидово расстояние 0,15 относится к различным пропорциям и корреляциям, которые теоретически спорны. Хотя это было



ПРОВЕРКА ИЭК ВИВАЛЕНТНОСТИ ДЛЯ АНАЛИЗА ЧАСТИ ОТ

необходимо для целей данного исследования приравнять интервалы для тестов EP и CV с интервалами

теста ЭД, что возможно было провести и честное сравнение между процедурами, многие возразят.

что различия в пропорциях и корреляциях, превышающие 0,3, слишком велики, что обуславливает отступы

ассоциация между переменными. Например, рассмотрим интервалы эквивалентности, используемые Роджерсом и др.

др. (1993) с тестом EP. Авторы использовали процедуру EP для сравнения двадцати семи

исходных характеристик и между двумя группами женщин, одна из которых носила свою беременность до срока

и еще один прерванный. Исследователи указали на различия в пропорциях и эквивалентности

интервалов для каждой из измеряемой характеристик. Важно отметить, что интервалы эквивалентности и их

использованное исследование (20% значения контрольных групп) варьировалось от 0,001 до 0,199. Таким образом,

по сравнению с работой Rogers et al. (1993), интервал эквивалентности, используемый для теста EP в нашей

исследование (т.е. 3) кажется заметным либеральным. По этой причине исследователи должны решить, основываясь на

метрике, выбранная для исследования отступы ассоциации, наименьшее значение, к которому они рассматривали

осмысленный.

Рекомендации

Баркер и др. (2001) предполагают, что рекомендации от носителя того, какой вид эквивалентности

тест, который следует использовать, должен быть основан на соотношении между эмпирическим и номинальным

Частота ошибок типа I, а также мощность рассматриваемого теста. В дизайне исследования 2 x 2

эмпирическая частота ошибок типа I теста ED значительно выше, чем номинальная частота. Это

исходно исключает процедуру ЭД из дальнейшего рассмотрения, несмотря на то, что показывает

лучшие показатели мощности, чем другие процедуры во многих условиях. Тесты CV и EP

производят жизнеспособные коэффициенты ошибок типа I для всего диапазона исследованных условий с CV

подход несколько более консервативен, чем EP. Таким образом, мы рекомендуем использовать резюме или

тесты EP для дизайна исследования 2 x 2 и подход CV для дизайна исследования 2 x 4.

Подводя итог, тесты эквивалентности позволяют исследователям оценить эффективность гипотезы.

Что две категориальные переменные пренебрежимо связаны? Установление «минимально важного отклонения» (или интервала эквивалентности) является сложным и субъективным аспектом проверки на отклонение ассоциации. Есть надежда, что в будущих дискуссиях будут освещены вопросы, связанные с определением соответствующего интервала и сделать его менее сложной задачей для исследователей. Однако даже допущение некоторой субъективности при установлении области эквивалентности лучше, чем ненадлежащим образом использовать незначительность традиционного критерия независимости хи-квадрат для исследовать отклонение связи между категориальными переменными.

Приложение

Прикладной пример с использованием основанной на эквивалентности версии К-рамера V, метода ЕР и метода ЭД.

Исследователь хочет продемонстрировать отсутствие связи между полом ребенка и имеет ли он или она высокие баллы по показателю синдрома дефицита внимания (СДВ) в 8-й класс. Данные (процентная часть от общего числа в скобках) были следующими (обратите внимание, что данные были сгенерированы для целей этой демонстрации):

Секс	Высокий балл по СДВ	Невысокий балл по СДВ 25 (7%)	Всего
Мальчик	32 (9%)	172 (46%)	315
Девочка	57	143 (38%)	204
Общий			372

К-рамер V

"=" \_\_\_\_\_

11 " = "  $\frac{168(57)}{372} = 25,74$

12 " = "  $\frac{168(315)}{372} = 142,25$

21 " = "  $\frac{204(57)}{372} = 31,25$

22 " = "  $\frac{204(315)}{372} = 172,74$

$$= \frac{25,74}{31,25} \frac{142,25}{172,74}$$

$$= \frac{(25,74)^2}{31,25^2} + \frac{(142,25)^2}{172,74^2} + \frac{(31,25)^2}{172,74^2} = 0,046$$

$$= \frac{(25,74)^2}{31,25^2} + \frac{(142,25)^2}{172,74^2} + \frac{(31,25)^2}{172,74^2} = 0,046$$

$$= 0,046 / [372(2 - 1)] = 0,0111262$$

$$= \frac{(25,74 + 142,25 + 31,25)^2}{372(2 - 1)} = \frac{(199,24)^2}{372(2 - 1)}$$

$$= \frac{(199,24)^2}{372(2 - 1)} = \frac{39696,0176}{372} = 106,71$$

Исследователи оценивают нулевую гипотезу  $H_0$ : по сравнению альтернативной гипотезой о том, что  $H_1$ :

<. Учитывая теоретическую основу границ эквивалентности была установлена на уровне  $\delta = 0,3$ , что

было найдено приблизительно значение, при котором корреляция между переменными становится

осмысленные (Берисский, 2018). Напомним, что  $\Delta L$  и  $\Delta U$  представляют собой нижний и верхний КИ для

параметры энтропии и распределения  $\chi^2$ . Не существует прямого метода анализа  $\Delta L$  и

$\Delta U$ , поэтому мы используем в R функцию `CramerV` из пакета `DescTools`.

(Синьорелли и др., 2017). Эта функция использует итеративный подход для определения значений  $\Delta L$  и

$\Delta U$  (см. Smithson, 2003). Например, код `DescTools::CramerV(mat)`, где 'mat' — это

матрица наблюдаемых частот возвращает 90% ДИ (0,000, 0,082). Поскольку верхняя граница

90% ДИ для  $V$  К-мера (0,082) падает ниже границы эквивалентности  $\delta = 0,3$ , мы отвергаем  $H_0$ :

и сделаем вывод, что существует значимая связь между полом ребенка и высоким баллом

на ДСП. Значение К-мера  $V$  (0,011) можно использовать в качестве величины эффекта, поскольку его можно

интерпретировать в соответствии с положительной корреляцией, мы могли бы сказать, что эффект очень мал.

Эквивалентность двух пропорций (EP)

$H_{01}: p_1 - p_2 = \delta$  и  $H_{02}: p_1 - p_2 = \delta$  отклоняются, если 90% ДИ для  $\hat{p}_1 - \hat{p}_2$  падает полностью

в интервале эквивалентности (в данном случае установленном при  $(-\delta, \delta) = (-.10, .10)$ )

$$\hat{p}_1 = \frac{25}{168} = 0,149$$

$$\hat{p}_2 = \frac{32}{204} = 0,157$$

$$\hat{p}_1 - \hat{p}_2 = 0,149 - 0,157 = -0,008$$

Стандартная ошибка разницы между двумя пропорциями рассчитывается по формуле:

$$s_{\hat{p}_1 - \hat{p}_2} = \sqrt{\left\{ \hat{p}_1(1 - \hat{p}_1)/n_1 + \hat{p}_2(1 - \hat{p}_2)/n_2 \right\}}$$

$$= \sqrt{\{[0,1488(1 - 0,1488)/168] + [0,1569(1 - 0,1569)/204]\}} = 0,037$$

$$90\% \text{ ДИ} = \hat{p}_1 - \hat{p}_2 \pm z_{1-\alpha/2} s_{\hat{p}_1 - \hat{p}_2} = -0,008 \pm 1,65(0,037) = (-0,070, 0,054)$$

Поскольку КИ для  $\hat{p}_1 - \hat{p}_2$  полностью попадает в интервал эквивалентности  $(-\delta, \delta = -.10, .10)$  как

$H_{01}$  и  $H_{02}$  отбрасываются, и мы заключаем, что пропорции для мальчиков и девочек эквивалентны

Евклидово расстояние (ED)

Вслед за Веллеком (2010),  $H_0: \chi^2_{D2}$  отвергается, если верхний предел 1- $\alpha$  ДИ для  $D2$

падает ниже  $\epsilon^2$ , где верхний предел для  $D2$  рассчитывается как:

$$\chi^2_{D2} + \chi^2_{1-\alpha}$$

где  $\chi^2_{D2} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$

$$\chi^2_{D2} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$$

Где  $\bar{x}$  представляет наблюдаемые вероятности и  $\bar{x}$  представляет собой ожидаемое

вероятности.

$$\chi^2_{11} = \frac{25}{372} = 0,067, \quad \chi^2_{12} = \frac{143}{372} = .384, \quad \chi^2_{21} = \frac{32}{372} = .086, \quad \chi^2_{22} = \frac{172}{372} = .462$$



# ПРОВЕРКА ИЭК ВИВАЛЕНТНОСТИ ИДЛЯ АНАЛИЗА ЧАСТИ ОТ

22

Расчет <sup>2</sup> чрезвычайно громоздка, даже для этой простой матрицы 2 x 2. Таким образом, используя

функция `gofind_t` из пакета `EQUIVNONINF`,  $\alpha = .003$ . Наконец, чтобы определить, является ли

верхняя граница доверительного интервала для  $D2 < 2$ ,

$$1 - (2) = \frac{2}{1} + \frac{1}{1} = 0,00002 + 1,65 = 1,652 \quad \frac{0,003}{372}$$

С  $1 - (2)(0,003) < \frac{2}{1} (0,152 = 0,0225)$  ноль  $H_0$ :  $\chi^2$  от вергается и отсутствие

ассоциация заключена.

## Рекомендации

Бейлин, С., Финк, М., Кнапп, Р., Петридис, Г., Хусейн, М.М., Расмуссен, К., ... и Келлер, С.Н.

(2010). Электрофизиологическая терапия одинаково эффективна при униполярном и биполярном

депрессия. *Acta Psychiatrica Scandinavica*, 121(6), 431-436. doi: 10.1111/j.1600-

0447.2009.01493.x

Баркер, Л., Ролк, А., Ролк, А.Д. и Браун, К. (2001). Проверка эквивалентности для биномиального случайного

переменного: как это использовать? *Американский статистик*, 55(4), 279-287. doi:

[10.1198/000313001753272213](https://doi.org/10.1198/000313001753272213)

Бернштейн, Н., Дэвидсон, Х. и Кристи, Р. (представлено). Что является наименьшим значащим

отклонением между двумя переменными? *Текучая психология*.

Брэдли, С.П. (1978). Прочность? *Британский математический и статистический журнал*

*Психология*, 31(2), 144-152. doi:10.1111/j.2044-8317.1978.tb00581.x

Буоли, М., Кумерлато, Меллер, К., Калдиероли, А., и Альтамора, А. (2015). Антидепрессанты

столь же эффективны при длительном лечении большого депрессивного расстройства? *Человек*

*Психофармакология: клиническая и экспериментальная*, 30(1), 21-27. doi: 10.1002/hup.2447

Чен, Дж. Дж., Цон, Ю. и Канг, С.Х. (2000). Тесты на эквивалентность или наименьшей эффективности между

двумя пропорциями. *Журнал информации о наркотиках*, 34(2), 569-578.

doi: 10.1177/009286150003400225

Комитет по лекарственным средствам для человека. (2006). Комитет по лекарственным средствам

руководство для использования человеком (CHMP) по выбору предела наименьшей эффективности. *Статистика в*

*Медицина*, 25(10), 1628-1638. doi: 10.1002/sim.2584

Кунселл, А., и Кристи, Р.А. (2015). Эквивалентные тесты для сравнения корреляции и

коэффициент регрессии. *Британский математический и статистический журнал*

*Психология*, 68(2), 292-309. doi: 10.1111/bmsp.12045



Кремер, Х. (1946). Математический метод истинности. Принстон: Издательство Принстонского университета.

doi: [10.1515/9781400883868](https://doi.org/10.1515/9781400883868)

Крибб, Р.А., Грuman, Дж.А., и Арпин-Крибб, Калифорния (2004). Рекомендации по применению

теста вивалентности. Журнал клинической психологии, 60 (1), 1-10. doi: [10.1002/jclp.10217](https://doi.org/10.1002/jclp.10217)

Крибб, Р.А., Арпин-Крибб, Калифорния, и Грuman, Дж.А. (2009). Тест вивалентности для односторонних

проективных групп. Журнал экспериментального образования, 78 (1), 1-13.

doi: [10.1080/00220970903224552](https://doi.org/10.1080/00220970903224552)

Орейли, Р., Бишоп, Дж., Мэддокс, К., Хатчинсон, Л., Фисман, М., и Тахар, Дж. (2007). Является

телепсихиатрическая вивалентная психиатрия лицом к лицу? Результаты рандомизированного

испытания контрлируемой вивалентности. Психиатрические службы, 58(6), 836-843.

doi: [10.1176/appi.ps.58.6.836](https://doi.org/10.1176/appi.ps.58.6.836)

Квертемонт, Э. (2011). Как статистический показатель отстоит от эффекта. Психология

Бельгия, 51(2), 109-127. doi: [10.5334/pb-51-2-109](https://doi.org/10.5334/pb-51-2-109)

Роджерс, Дж. Л., Ховард, К. И. и Весси, Дж. Т. (1993). Использование тестов значимости для оценки

иввалентности между двумя экспериментальными группами. Психологический бюллетень, 113 (3), 553-565.

doi: [10.1037//0033-2909.113.3.553](https://doi.org/10.1037//0033-2909.113.3.553)

Шуирманн, ди-джей (1987). Сравнение процедур односторонних испытаний и мощности

подход для оценки вивалентности средней биодоступности. Журнал

Фармакокинетики и биофармацевтики, 15, 657-680. doi: [10.1007/bf01068419](https://doi.org/10.1007/bf01068419)

Синьорелл, А. и др. (2017). DescTools: инструмент описательной статистики. Версия пакета R 0.99.23.

Смитсон, М. (2003). Доверительные интервалы № 140. Ключевые приложения в социальной

Серия наук. Бельмонт, Калифорния: Сейдж. doi: [10.4135/9781412983761](https://doi.org/10.4135/9781412983761)

Уокер, Э., и Новак, А.С. (2011). Понимание эквивалентности и не меньшей эффективности

тестирования. Журнал общей внутренней медицины 26 (2), 192-196. doi:[10.1007/s11606-010-](https://doi.org/10.1007/s11606-010-1513-8)

[1513-8](https://doi.org/10.1007/s11606-010-1513-8)

Веллек, С. (2010). Проверка статистических гипотез эквивалентности и не меньшей эффективности. Бок

Ратон, Флорида: CRC Press. doi: [10.1201/ebk1439808184](https://doi.org/10.1201/ebk1439808184)

Таблица 1.

Условия для исследования методом Монте-Карло.

2 x 2 Дизайн															
Состояние	a11	a12	a21	a22	d	EBED	PD	EBEP	EBCV						
Условия ошибки типа I	1 2 3														
	.175	.325	.325	.175	.150	.150	.300	.300	.298	.250	.250	.100	.400	.150	.300
	.300	.312	.050	.450	.200	.300	.150	.150	.300	.300	.344				

Условия питания 1

2 3 4 5 6	.250	.250	.250	.250	0	.300	.400	.100	.400	.100	.400	.100	.400	.100	.400	.270	.230	.250
	.250	.020	.150	.040	.300	.400	.260	.240	.240	.260	.020	.150	.040	.300	.400	.270	.230	.250
	.100	.150	.200	.300	.400	.200	.300	.300	.200	.100	.150	.200	.300	.400				

2 x 4 Дизайн																				
Состояние	a11	a12	a13	a14	a21	a22	a23	a24	d	ЭБЭД ЭБК В										
Условия ошибки типа I 1 2																				
3 4	.050	.050	.050	.050	.659	.050	.050	.041	.150	.150	.507	.100	.050	.250	.150	.181	.100	.050	.119	.150
	.150	.410	.200	.050	.050	.050	.166	.300	.150	.034	.150	.150	.399	.150	.150	.150	.150	.253	.050	.050
	.047	.150	.150	.382																

Условия питания 1																							
2	3	4	5	6	.125	.125	.125	.125	.125	.125	.125	.125	0	.120	.120	.120	.120	.130	.130	.130	.15	.4	
					.130	0	.145	.105	.125	.125	.125	.125	.125	.125	.02	.149	.120	.120	.120	.130	.15	.4	
					.130	.130	.101	.02	.200	.070	.190	.220	.080	.090	.125	.025	.1	.100	.150	.134	.100	.15	.4
					.150	.120	.046	.200	.1												.15	.4	
																					.15	.4	
																					.15	.4	

Примечание.  $a_{ij}$  = пропорция чек в таблицах 2 x 2 и 2 x 4; d = евклидово расстояние; EBED = граница эквивалентности для теста ED; ПД =  $p_1 - p_2$ ; EBEP = верхняя граница интервала эквивалентности для теста EP; EBCV = граница эквивалентности для V-критерия Крамера.

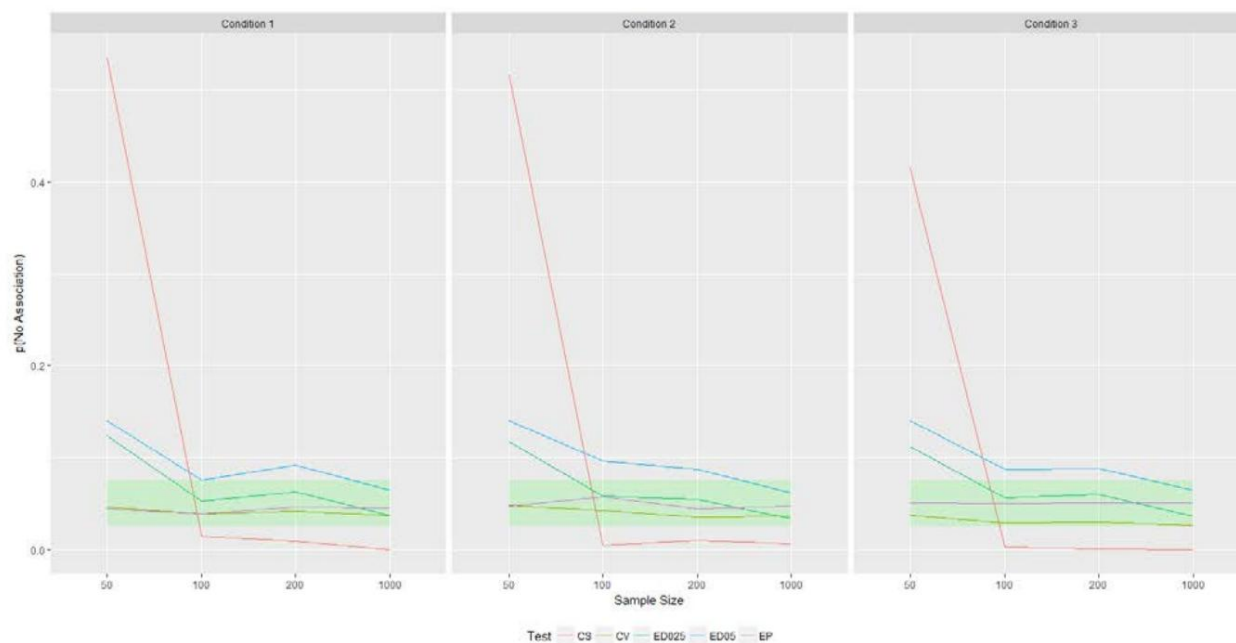


Рисунок 1. Коэффициент ошибки типа I для тестов на основе эквивалентности, коэффициент ошибки типа II для традиционного критерия независимости  $\chi^2$  в схеме 2 x 2. CS = традиционный критерий независимости  $\chi^2$ , CV = V Крамера, ED025 = критерий евклидова расстояния ( $\alpha = 0,025$ ), ED05 = критерий евклидова расстояния ( $\alpha = 0,05$ ), EP = критерий эквивалентности пропорций; выделенная зеленая область = либерализованная область Брэдли (0,025–0,075). См. Таблицу 1 для информации о составлении.

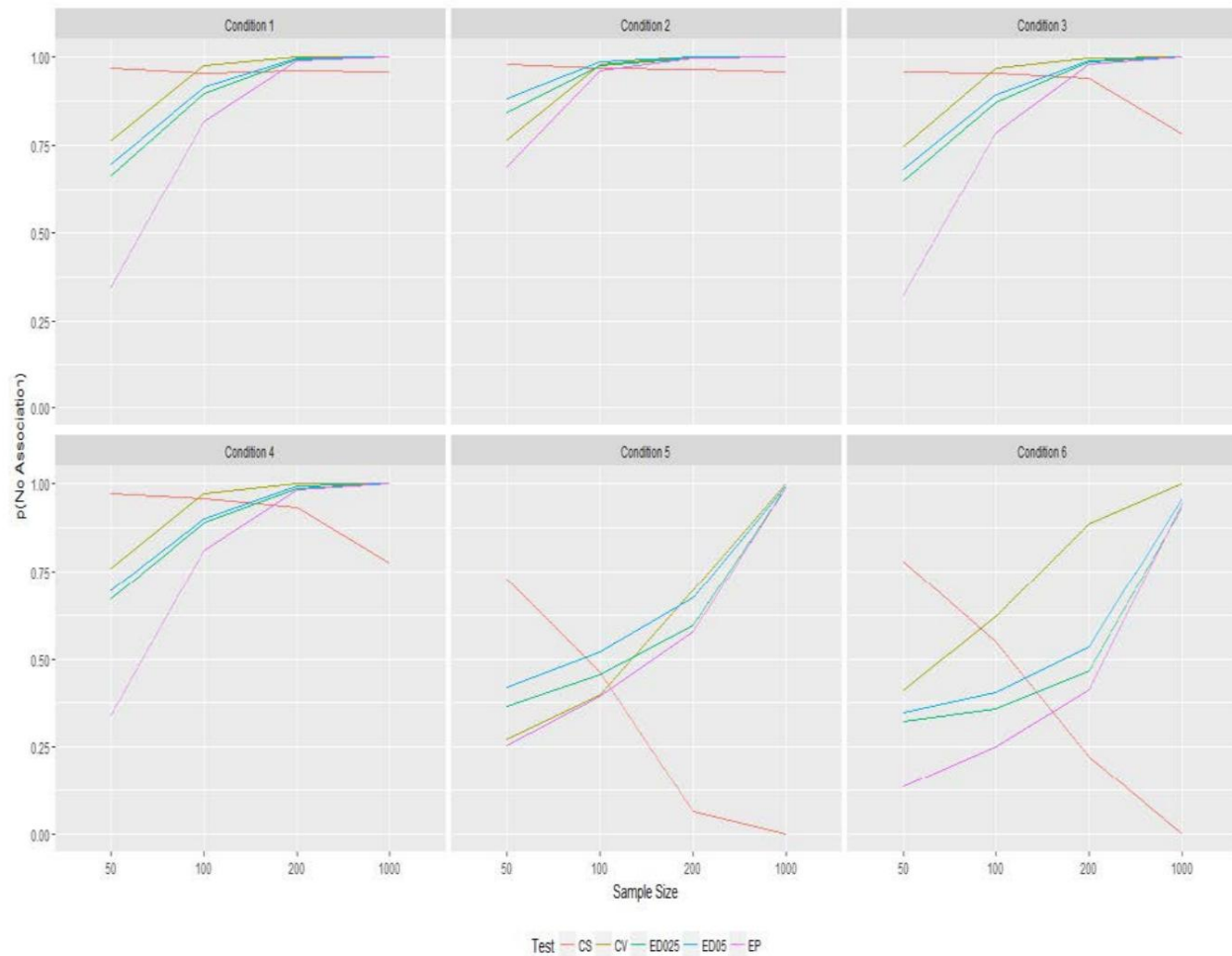


Рисунок 2. Показатели мощности и для тестов на основе эквивалентности, правильное решение (условия 1, 2)/ коэффициент ошибки типа II (условия 3-6) для традиционного теста независимости  $\chi^2$  в схеме 2 x 2. CS = традиционный критерий независимости  $\chi^2$ , CV = V Крамера, ED025 = критерий эквивалентности расстановки ( $\alpha = 0,025$ ), ED05 = критерий эквивалентности расстановки ( $\alpha = 0,05$ ), EP = критерий эквивалентности пропорций. См. Таблицу 1 для информации о состоянии.

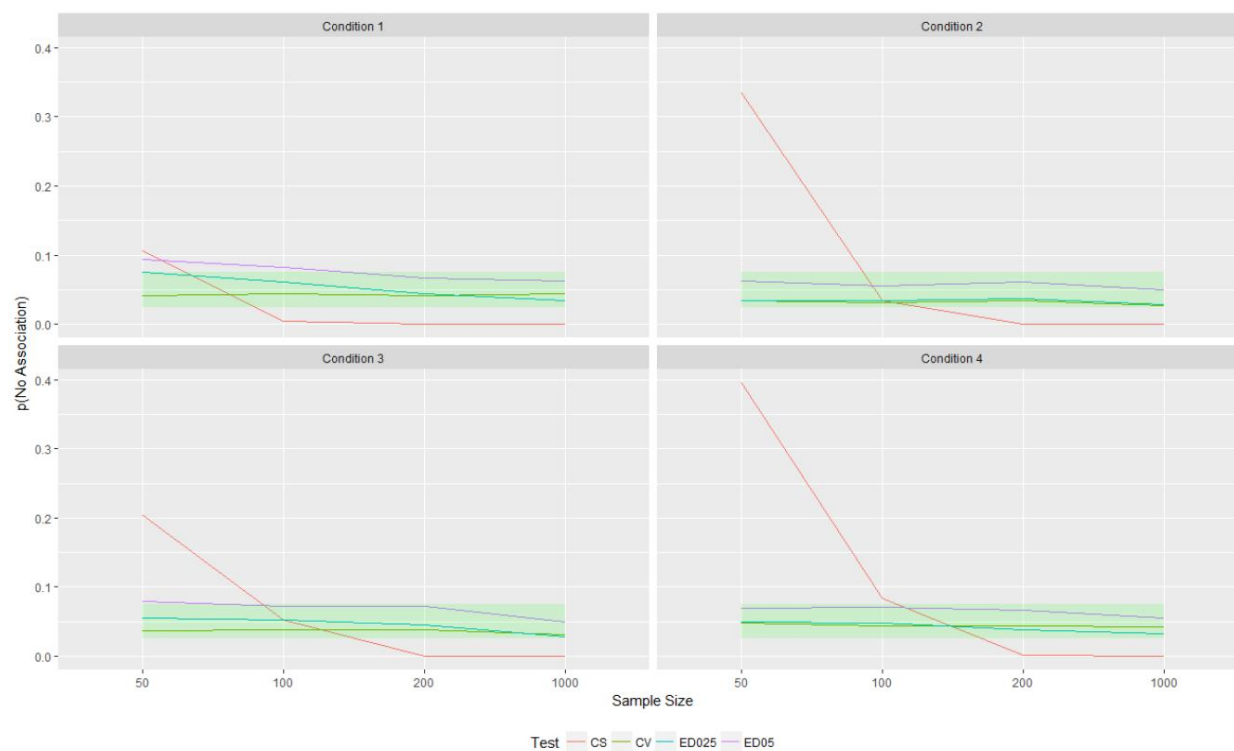


Рисунок 3. Коэффициент ошибок типа I для тестов на основе эквивалентности, коэффициент ошибок типа II для традиционного критерия независимости  $\chi^2$  в схеме  $2 \times 4$ . CS = традиционный критерий независимости  $\chi^2$ , CV = V Крамера, ED025 = критерий евклидова расстояния ( $\alpha = 0,025$ ), ED05 = критерий евклидова расстояния ( $\alpha = 0,05$ ), выделенная зеленая область = либеральные пределы Брэдли (0,025–0,075).). См. Таблицу 1 для информации о состояниях.

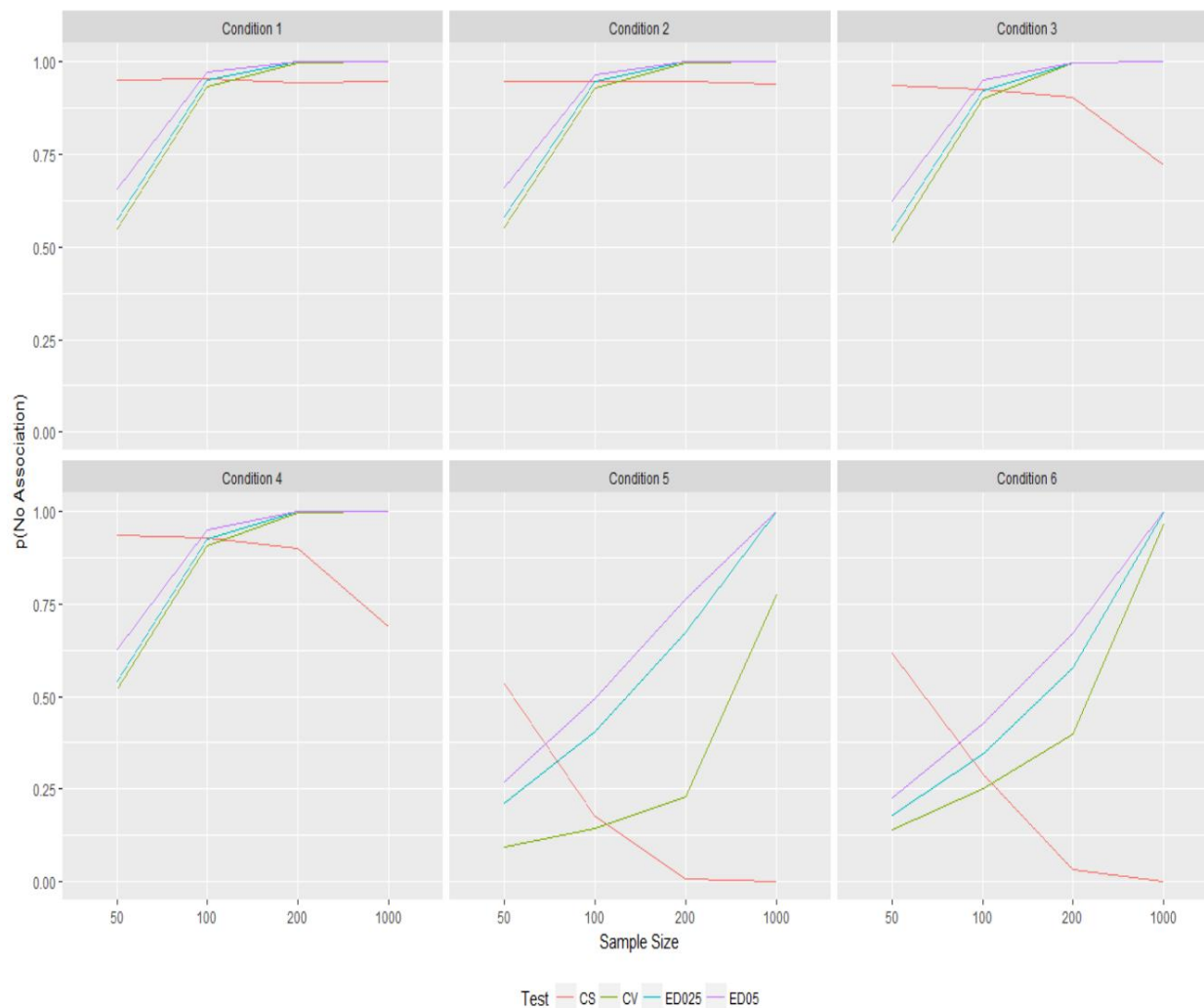


Рисунок 4. Мощность и для тестов на основе эквивалентности, правильное решение (условия 1, 2)/ коэффициент ошибки типа II (условия 3-6) для традиционного теста независимости  $\chi^2$  в схеме  $2 \times 4$ . CS = традиционный критерий независимости  $\chi^2$ , CV = V Крамера, ED025 = критерий евклидова расстояния ( $\alpha = 0,025$ ), ED05 = критерий евклидова расстояния ( $\alpha = 0,05$ ). См. Таблицу 1 для информации о состоянии.