

Copyright © Британское психологическое общество

Воспроизведение в любой форме (включая Интернет) запрещено без предварительного разрешения Общества.

527

Британский журнал математической и статистической психологии (2010 г.), 63, 527-537, кв.
2010 г. Британское психологическое общество



www.bpsjournals.co.uk

Обнаружение отсутствия ассоциации: подход к проверке эквивалентности

Джейсон Р. Герцен и Роберт А. Крибби* Йоркский университет, Торонто, Онтарио, Канада

Исследователи часто проверяют отсутствие связи между переменными. Отсутствие ассоциации – это обычно устанавливается путем демонстрации незначительной связи с традиционным тестом (например, r Пирсона). Однако как по логическим, так и по статистическим причинам такие выводы проблематичны. В этой статье мы обсуждаем и сравниваем эмпирическую ошибку типа I и мощность трех тестов на отсутствие ассоциации. Результаты показывают, что для того, чтобы статистические данные теста были подходящими, требуются большие, а иногда и очень большие размеры выборки. Что особенно проблематично, так это то, что требуемые размеры выборки могут превышать практически достижимые для условий, которые, как ожидается, будут распространены среди исследователей в области психологии. В этой статье подчеркивается важность использования доступных тестов отсутствия ассоциации вместо традиционных тестов ассоциации для демонстрации независимости переменных и квалифицируются условия, при которых эти тесты подходят.

1. Введение

Исследователи в области психологии часто интересуются проверкой отсутствия связи между переменными. Например, исследователи, использующие линейные модели с несколькими ковариатами, могут быть заинтересованы в ограничении числа ковариат путем удаления тех, которые не связаны с переменной результата. Другие исследователи проверяют отсутствие ассоциации в качестве основной исследовательской гипотезы. Например, Wheadon et al. (1992) исследовали потенциальное отсутствие связи между суицидальными наклонностями и применением флуоксетина при лечении нервной булимии. Однако основная проблема с тестированием на отсутствие связи между переменными заключается в том, что отсутствие связи обычно демонстрируется незначимой тестовой статистикой (например, r Пирсона). Эта стратегия не подходит, потому что исследовательская гипотеза (т. е. отсутствие связи) согласуется с нулевой гипотезой (например, $H_0: r = 0$), а не с альтернативной гипотезой. Другими словами, традиционный тест корреляции был разработан для проверки наличия ассоциации, а не ее отсутствия. Короче говоря, если исследователи хотят надлежащим образом обнаружить отсутствие ассоциации, они должны использовать тесты на отсутствие ассоциации.

* Корреспонденцию следует направлять доктору Роберту А. Крибби, кафедра психологии, Йоркский университет, Торонто, Онтарио, Канада, М3J 1P3 (электронная почта: cribbie@yorku.ca).

Copyright © Британское психологическое общество

Воспроизведение в любой форме (включая Интернет) запрещено без предварительного разрешения Общества.

528 Джейсон Р. Герцен и Роберт А. Крибби

Цели настоящей статьи: (1) обозначить некоторые проблемы, с которыми сталкиваются исследователи при попытке обнаружить отсутствие ассоциации; и (2) предложить методы обнаружения отсутствия связи. Кроме того, мы надеемся повысить осведомленность о том, как в психологии в настоящее время принимаются решения об отсутствии ассоциации и как можно исправить эту проблемную тактику. Насколько нам известно, настоящая статья является первым исследованием, посвященным тестам на отсутствие ассоциаций в области психологии.

2. Проблемы, связанные с текущим подходом к обнаружению отсутствия ассоциации

Как указано во введении, в настоящее время наиболее распространенным методом обнаружения отсутствия связи между двумя переменными является поиск незначительной связи с помощью традиционной корреляционной или регрессионной статистики. Однако этот метод не подходит по двум основным причинам: (1) отказ от отклонения нулевой гипотезы об отсутствии ассоциации (т. е. $H_0: r = 0$) означает, что мы принимаем нулевую гипотезу; и (2) незначимый критерий корреляции имеет плохие асимптотические свойства для демонстрации отсутствия связи, потому что вероятность объявления отсутствия связи (т. е. не отвергнуть $H_0: r = 0$) уменьшается (а не увеличивается) по мере увеличения размера выборки. увеличивать. Первый пункт просто напоминает исследователям, что мы не можем «принять» нулевую гипотезу, когда наша тестовая статистика незначительна, поскольку связь может присутствовать, но наша тестовая статистика может просто не обнаружить ее (возможно, из-за недостатка мощности). Второй момент подчеркивает, что, когда исследователь проводит традиционный тест на корреляцию, вероятность обнаружения значимой связи увеличивается по мере увеличения размера нашей выборки. Другими словами, чем больше размер выборки, тем больше возможностей для обнаружения ассоциации в популяции. И наоборот, вероятность обнаружения незначительной взаимосвязи уменьшается по мере увеличения размера выборки, и поэтому это не лучший метод для обнаружения отсутствия связи. Другими словами, исследователи, пытающиеся найти отсутствие связи между переменными с помощью традиционного теста на связь, повысят вероятность обнаружения отсутствия связи (т. е. недостоверной статистики теста) за счет использования выборки меньшего размера, и исследователи будут объявить эквивалентность почти в 100% случаев с $N = 3$ (наименьшее значение N , для которого степени свободы положительны).

В целом, эти две основные проблемы представляют собой проблемы, которые необходимо решить путем разработки новых статистических тестов или принятия альтернативных методов. Что касается альтернативных методов, одной из потенциальных альтернатив является байесовский анализ. Байесовские подходы обращаются к вероятности гипотезы с учетом данных, позиция, которая может быть полезна для оценки отсутствия связи путем определения вероятности того, что нулевая гипотеза в традиционном тесте на связь (т. е. $H_0: r = 0$) окажется истинной. Однако, хотя байесовский и другие подходы могут предложить уникальные решения проблемы обнаружения отсутствия ассоциации, основное внимание в этом исследовании уделяется разработке процедур проверки нулевой гипотезы (т. е. частотного анализа) на отсутствие ассоциации.

3. Еще одна важная проблема: распределение r Пирсона

Приступая к разработке

соответствующего теста для демонстрации отсутствия связи, важно иметь представление о распределении r Пирсона, когда между интересующими переменными нет базовой связи. среди населения

Copyright © Британское психологическое общество

Воспроизведение в любой форме (включая Интернет) запрещено без предварительного разрешения Общества.

Отсутствие ассоциации 529

(т. е. $r = 0$). Например, мы смоделировали 5000 r -статистик Пирсона, когда $r = 0$ для каждого из семи различных условий размера выборки ($N = 10, 15, 20, 25, 50, 100$ и 200). Таблица 1 представляет долю выборочных корреляций каждой величины для каждого из выборочных условия размера. Важно отметить, что мы не фокусируемся на статистических данных. значение в этом примере, но вместо этого на величину корреляции, поскольку мы ожидать, что это будет наиболее важным фактором в попытке определить, является ли связь существует между двумя переменными. Из таблицы видно, что это будет чрезвычайно трудно «доказать» отсутствие связи, когда размеры выборки малы до умеренный. При размере выборки $N = 10$ примерно 75% выборочной корреляции значения (по абсолютной величине) превышают $r = .1$, 40 % превышают $r = .3$ и даже 15 % превышают $r = .5$. При больших размерах выборки по-прежнему сохраняется значительная вариабельность корреляции. коэффициенты. При размере выборки $N = 20$ примерно 40 % выборочной корреляции значения превышают $r = .1$, а 20% превышают $r = .3$. Даже при размере выборки $N = 100$, более 30% выборочных значений корреляции превышают $r = .1$.

Таблица 1. Доля выборочных корреляций, превышающих установленные значения r , в зависимости от размер выборки, когда $r = 0$

N	Величина выборочной корреляции (r) (абсолютное значение)								
	.8	.7	.6	.5	.4	.3	.2	.1	.0
10	0,001	0,007		0,067	.140	.252	.398	.578	.780
15	0	0,025	0,001 0,005	0,019	0,060	.139	.273	.465	.720
20	0	0,001		0,005	0,025	.082	.199	.400	.687
25	0	0	.001 .002		0,012	.046	.145	.334	.628
50	0	0	0	0	.001	0,004	.036	.163	.492
100	0	0	0	0	0	0	.003	.048	.323
200	0	0	0	0	0	0	0	0,005	.159

Почему это такая проблема? Учтите, что вы пытаетесь продемонстрировать отсутствие связь между двумя переменными; из этих результатов следует, что более Для надлежащей демонстрации отсутствие какой-либо значимой связи, учитывая крайнюю изменчивость выборки r статистика. Другими словами, попытка продемонстрировать отсутствие ассоциации является равносильно попытке определить, что величина эффекта пренебрежимо мала и, как ожидать, это очень трудно достичь. Мы расширяем эту тему ниже, когда мы обсудить установление интервала незначительной связи. Далее, в результате трудности с демонстрацией отсутствия связи с малыми и средними размерами выборки, одной из целей настоящей статьи является определение размера выборки, ассоциативные тесты потребуются для получения приемлемых результатов.

4. Тест на отсутствие ассоциации

Первый тест на отсутствие ассоциации (equiv_r), который мы предлагаем, основан на «двух односторонних тестах». подход к тестам, распространенный при тестировании эквивалентности среднего (например, Schuirmann, 1981, 1987 год; Westlake, 1981) и аналогичен регрессионным тестам эквивалентности, предложенным Диксон и Пехманн (2005 г.) и Робинсон, Дуурсма и Маршалл (2005 г.); несмотря на то что

Copyright © Британское психологическое общество

Воспроизведение в любой форме (включая Интернет) запрещено без предварительного разрешения Общества.

530 Джейсон Р. Герцен и Роберт А. Крибби

последний ориентирован на логарифмическую модель, которая не так широко применима в психологии. Составные нулевые гипотезы, $H_{01} : r = 0$ и $H_{02} : r = 2r$, отбрасываются, если $t_1 \neq t_{\alpha/2; N-2}$ и $t_2 \neq t_{\alpha/2; N-2}$, где

$$t_1 = \frac{r - 2r}{\sqrt{\frac{12p^2}{N-2}}}, \quad t_2 = \frac{r - 2r - 2r}{\sqrt{\frac{12p^2}{N-2}}}$$

r^* представляет собой интервал отсутствия ассоциации (т.е. $(2r^*, r^*)$), N представляет собой размер выборки, $t_{\alpha/2; N-2}$ представляет собой критическое значение уровня α из распределения t с $N-2$ степенями свободы, а r представляет собой выборочное значение корреляции. Одновременное отбрасывание H_{01} и H_{02} означает, что корреляция населения находится в пределах от $2r^*$ до r^* . Как видно из приведенных выше уравнений, отсутствие тестов ассоциации делает потенциальное наличие ассоциации нулевой гипотезой, а альтернативной гипотезой становится $H_1 : 2r < r < r$.

Важным преимуществом тестов на отсутствие ассоциации по сравнению с использованием традиционных тестов, предназначенных для обнаружения ассоциации, является то, что исследователи вынуждены указывать интервал отсутствия ассоциации ($2r^*, r^*$). Отсутствие интервала ассоциации определяет границы, для которых корреляция считается бессмысленной (т.е. практически равной нулю). Например, исследователь может установить $r = .1$ (и, следовательно, интервал $(2.1, .1)$), где в этом случае любая корреляция меньше .1 (по абсолютной величине) будет считаться незначимой в рамках исследования. Для исследователей важно учитывать характер исследования при установке r^* , поскольку то, что считается практически незначительным, может значительно варьироваться от исследования к исследованию.

5. Возможные проблемы с тестом на отсутствие ассоциации

Достаточно хорошо известно, что формула t для проверки нулевой гипотезы с r Пирсона необъективна, поскольку стандартная ошибка тестовой статистики связана с r (Bond & Richardson, 2004; Fisher, 1915). Другими словами, t -статистика

$$t = \frac{r}{\sqrt{\frac{12p^2}{N-2}}}$$

обычно не подходит для проверки гипотез с помощью r Пирсона. Однако эта формула является корнем теста на отсутствие ассоциации, предложенного выше. Наиболее распространенной коррекцией смещения является z -преобразование Фишера (1915), которое преобразует r в z , а тестовая статистика (F_z) равна z_0 / s_{z0} , где:

$$F_z = \frac{z_0}{s_{z0}} = \frac{\frac{r}{\sqrt{\frac{12p^2}{N-2}}}}{\sqrt{\frac{12p^2}{N-2}}}$$

При этом тесте $H_0 : r = 0$ отбрасывается, если $|F_z| \geq z_{\alpha}$, где z_{α} — критическое значение уровня α из стандартного нормального распределения. Таким образом, потенциально улучшенный тест на отсутствие ассоциации (equiv_z) может быть результатом использования z -преобразования Фишера с ранее предложенным тестом на отсутствие ассоциации (т.е. equiv_r). Полученный тест будет

... в любой форме (включая Интернет) запрещено без предварительного разрешения

$$Fz_LA1 \quad \frac{1}{4} \quad \frac{r^0 \quad 2 \quad z\ddot{o}r \quad \flat}{\text{размер 0}} \quad \frac{\text{журнал} \quad \frac{1 \quad \text{pys}}{12p}}{2} \quad \frac{\text{журнал} \quad \frac{1p}{12p}}{2} \quad \frac{\text{тестирование}}{N23 \quad p}$$

Третий подход к созданию беспристрастного теста на отсутствие ассоциации заключается в

(с заменой) N парных точек данных из исходных N парных наблюдений

(эквивалент). Этот подход аналогичен процедуре повторной выборки, р

Робинсон и др. (2005) для оценки эквивалентности в задачах на основе регрессии. Это

между парными данными. Таким образом, эмпирическое выборочное распредел

Результаты тестирования на нормальность распределения для выборок, сформированных из выборочных данных. Результирующий тест отклонил бы $H_0: F$, если

выборочные значения корреляции, вычисленные из повторно выбранных парных данных. R.

р. Мы рекомендуем провести не менее 10 000 повторных замеров, чтобы обеспечить

соответствующую тестовую статистику.

Основная цель настоящих

вышеупомянутые тесты для обнаружения отсутствия связи (т. е., equiv_r, equiv_fz и

проведение тестов на отсутствие ассоциации. В частности, авторов интересует

рекомендации относительно соответствующих размеров выборки для достижения со

достичь надлежащих результатов с наименьшим размером выборки, что может быть важным

Исследования

предлагаемые тесты (equiv_r, equiv_fz и equiv_rs). Было использовано десять тысяч ресэмплров.

в этом исследовании манипулировали переменными: (1) отсутствие интервала ассоциации; (2)

корреляция; и (3) размер выборки. r^* было установлено равным 0,05, 0,1, 0,15, 0,2, 0,25 или 0,3, а n установлено равным 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 430, 440, 450, 460, 470, 480, 490, 500, 510, 520, 530, 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700, 710, 720, 730, 740, 750, 760, 770, 780, 790, 800, 810, 820, 830, 840, 850, 860, 870, 880, 890, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, 1200, 1210, 1220, 1230, 1240, 1250, 1260, 1270, 1280, 1290, 1300, 1310, 1320, 1330, 1340, 1350, 1360, 1370, 1380, 1390, 1400, 1410, 1420, 1430, 1440, 1450, 1460, 1470, 1480, 1490, 1500, 1510, 1520, 1530, 1540, 1550, 1560, 1570, 1580, 1590, 1600, 1610, 1620, 1630, 1640, 1650, 1660, 1670, 1680, 1690, 1700, 1710, 1720, 1730, 1740, 1750, 1760, 1770, 1780, 1790, 1800, 1810, 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980, 1990, 2000, 2010, 2020, 2030, 2040, 2050, 2060, 2070, 2080, 2090, 2100, 2110, 2120, 2130, 2140, 2150, 2160, 2170, 2180, 2190, 2200, 2210, 2220, 2230, 2240, 2250, 2260, 2270, 2280, 2290, 2300, 2310, 2320, 2330, 2340, 2350, 2360, 2370, 2380, 2390, 2400, 2410, 2420, 2430, 2440, 2450, 2460, 2470, 2480, 2490, 2500, 2510, 2520, 2530, 2540, 2550, 2560, 2570, 2580, 2590, 2600, 2610, 2620, 2630, 2640, 2650, 2660, 2670, 2680, 2690, 2700, 2710, 2720, 2730, 2740, 2750, 2760, 2770, 2780, 2790, 2800, 2810, 2820, 2830, 2840, 2850, 2860, 2870, 2880, 2890, 2900, 2910, 2920, 2930, 2940, 2950, 2960, 2970, 2980, 2990, 3000, 3010, 3020, 3030, 3040, 3050, 3060, 3070, 3080, 3090, 3100, 3110, 3120, 3130, 3140, 3150, 3160, 3170, 3180, 3190, 3200, 3210, 3220, 3230, 3240, 3250, 3260, 3270, 3280, 3290, 3300, 3310, 3320, 3330, 3340, 3350, 3360, 3370, 3380, 3390, 3400, 3410, 3420, 3430, 3440, 3450, 3460, 3470, 3480, 3490, 3500, 3510, 3520, 3530, 3540, 3550, 3560, 3570, 3580, 3590, 3600, 3610, 3620, 3630, 3640, 3650, 3660, 3670, 3680, 3690, 3700, 3710, 3720, 3730, 3740, 3750, 3760, 3770, 3780, 3790, 3800, 3810, 3820, 3830, 3840, 3850, 3860, 3870, 3880, 3890, 3900, 3910, 3920, 3930, 3940, 3950, 3960, 3970, 3980, 3990, 4000, 4010, 4020, 4030, 4040, 4050, 4060, 4070, 4080, 4090, 4100, 4110, 4120, 4130, 4140, 4150, 4160, 4170, 4180, 4190, 4200, 4210, 4220, 4230, 4240, 4250, 4260, 4270, 4280, 4290, 4300, 4310, 4320, 4330, 4340, 4350, 4360, 4370, 4380, 4390, 4400, 4410, 4420, 4430, 4440, 4450, 4460, 4470, 4480, 4490, 4500, 4510, 4520, 4530, 4540, 4550, 4560, 4570, 4580, 4590, 4600, 4610, 4620, 4630, 4640, 4650, 4660, 4670, 4680, 4690, 4700, 4710, 4720, 4730, 4740, 4750, 4760, 4770, 4780, 4790, 4800, 4810, 4820, 4830, 4840, 4850, 4860, 4870, 4880, 4890, 4900, 4910, 4920, 4930, 4940, 4950, 4960, 4970, 4980, 4990, 5000, 5010, 5020, 5030, 5040, 5050, 5060, 5070, 5080, 5090, 5100, 5110, 5120, 5130, 5140, 5150, 5160, 5170, 5180, 5190, 5200, 5210, 5220, 5230, 5240, 5250, 5260, 5270, 5280, 5290, 5300, 5310, 5320, 5330, 5340, 5350, 5360, 5370, 5380, 5390, 5400, 5410, 5420, 5430, 5440, 5450, 5460, 5470, 5480, 5490, 5500, 5510, 5520, 5530, 5540, 5550, 5560, 5570, 5580, 5590, 5600, 5610, 5620, 5630, 5640, 5650, 5660, 5670, 5680, 5690, 5700, 5710, 5720, 5730, 5740, 5750, 5760, 5770, 5780, 5790, 5800, 5810, 5820, 5830, 5840, 5850, 5860, 5870, 5880, 5890, 5900, 5910, 5920, 5930, 5940, 5950, 5960, 5970, 5980, 5990, 6000, 6010, 6020, 6030, 6040, 6050, 6060, 6070, 6080, 6090, 6100, 6110, 6120, 6130, 6140, 6150, 6160, 6170, 6180, 6190, 6200, 6210, 6220, 6230, 6240, 6250, 6260, 6270, 6280, 6290, 6300, 6310, 6320, 6330, 6340, 6350, 6360, 6370, 6380, 6390, 6400, 6410, 6420, 6430, 6440, 6450, 6460, 6470, 6480, 6490, 6500, 6510, 6520, 6530, 6540, 6550, 6560, 6570, 6580, 6590, 6600, 6610, 6620, 6630, 6640, 6650, 6660, 6670, 6680, 6690, 6700, 6710, 6720, 6730, 6740, 6750, 6760, 6770, 6780, 6790, 6800, 6810, 6820, 6830, 6840, 6850, 6860, 6870, 6880, 6890,

при оценке мощности с $\alpha = .05$ условие $\alpha = .05$ не исследовалось

100, 500 и 1000.

Copyright © Британское психологическое общество

Воспроизведение в любой форме (включая Интернет) запрещено без предварительного разрешения Общества.

532 Джейсон Р. Герцен и Роберт А. Крибби

Для каждого моделирования были сгенерированы две случайные нормальные переменные (X, Y), где $X = a_1 + e_1$ и $Y = a_2 + e_2$. X и Y были сгенерированы, чтобы иметь население корреляция r ($r \neq 0$), где a_1 , e_1 и e_2 — случайные нормальные переменные ($m = 0$, $s = 1$) которые были сгенерированы с помощью R-генератора «rnorm» (R Development Core Team, 2005 г.) и

$$b = \frac{rs_{e1} - rs_{e2}}{2s_{a1}^2 - 2s_{a2}^2}$$

Каждое из условий пересекалось, и для каждого было рассчитано 5000 симуляций. условие с номинальной частотой ошибок типа I 1/4:05. Программа моделирования была написано на R (R Development Core Team, 2005).

9. Результаты

9.1. Частота ошибок типа I

Коэффициенты ошибок типа I для процедур equiv_r, equiv_fz и equiv_rs в каждом из исследованные условия представлены в таблице 2. Результаты показывают, что тип I частота ошибок для теста на отсутствие ассоциации на основе повторной выборки (equiv_rs) была самой высокой. точным, за которым следует процедура преобразования Фишера по оси z (equiv_fz), и исходный тест на отсутствие ассоциации (equiv_r). Однако в целом результаты показывают, что ни одна из процедур не имела точной частоты ошибок типа I, когда оба размера выборки и отсутствие ассоциативных интервалов были небольшими. Более конкретно, все процедуры были чрезвычайно консервативны, когда и размер выборки, и отсутствие ассоциативные интервалы были небольшими, а процедура, основанная на повторной выборке, была наименьшей. консервативный. Например, при размере выборки N = 100 отсутствие связи интервал не менее 2,2–0,2 потребует для проведения соответствующего испытания гипотезы (т. е. эмпирическая частота ошибок типа I не менее 0,04). Даже с размер выборки N = 500, отсутствие интервала ассоциации, по крайней мере, от 2,15 до 0,15 будет требуется для того, чтобы эмпирическая частота ошибок типа I была близка к номинальной оценивать. Особо следует отметить, что при установлении интервала ассоциации от 2,05 до .05, не было ложных отказов (т.е. заявлений об отсутствии ассоциации) для любого из трех тестов, даже когда N = 1000.

Таблица 2. Частота ошибок типа I для тестов на отсутствие ассоциации в зависимости от размера выборки и r*

	N = 50			N = 100			N = 500			N = 1000		
r*	eq_r	eq_fz	eq_rs	eq_r	eq_fz	eq_rs	eq_r	eq_fz	eq_rs	eq_r	eq_fz	eq_rs
0,05	0	0	0	0	0	0	0	0	0	0	0	0
.10	0	0	0	0	0	0	.050	.051	.052	.002	.047	.049
.15	0	0	0	0	0	0	.047	.005	.030	.034	.042	.046
.20	0	0	.052	.25	.009	.017	.035	.039	.048	.051	.043	.049
.041	.052	.043	.050	.052	.044	.050	.052	.040	.048			
												0,047

Примечание. r* представляет интервал отсутствия ассоциации; eq_r — исходный тест на отсутствие ассоциации; eq_fz — модифицированный тест на отсутствие ассоциации, основанный на z-преобразовании Фишера; соответствует тест отсутствия ассоциации на основе повторной выборки.

Copyright © Британское психологическое общество

Воспроизведение в любой форме (включая Интернет) запрещено без предварительного разрешения Общества.

Отсутствие ассоциации 533

9.2. Сила

Норма мощности для процедур equiv_r, equiv_fz и equiv_rs для каждого из исследованных условий, представлены в таблице 3. Результаты для $r = 0$ были очень похожими. к таковым для $r = .05$, и поэтому следующее обсуждение будет иметь дело одновременно с оба набора результатов. Результаты мощности очень точно отражают результаты для ошибки типа I. ставки, при этом все процедуры чрезвычайно консервативны, когда как размеры выборки, так и отсутствие ассоциативных интервалов было небольшим. В частности, когда размеры выборки были умеренная (например, $N = 100$), удовлетворительная мощность (например, $.8$) может быть достигнута только при отсутствии принимается интервал ассоциации не менее $2.3-0.3$. Даже при размере выборки 500 , исследователям потребуется использовать интервал отсутствия ассоциации, по крайней мере, от 2.15 до 0.15 в чтобы иметь приемлемую мощность. Опять же, процедура на основе повторной выборки (equiv_rs) была равномерно самый мощный; тем не менее, преимущество в мощности над equiv_r и equiv_fz процедуры в большинстве случаев были невелики.

Таблица 3. Показатели мощности тестов на отсутствие ассоциации в зависимости от размера выборки, r и r^*

	N = 50				N = 100				N = 500				N = 1000			
r^*	eq_r	eq_fz	eq_rs	eq_r	eq_fz	eq_rs	eq_r	eq_fz	eq_rs	eq_r	eq_fz	eq_rs	eq_r	eq_fz	eq_rs	
$r = 0$																
0.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.10$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.15$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.20$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.25$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.3$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$r = .05$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.10$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.15$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.20$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.25$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.3$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$r = .10$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.15$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.20$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.25$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.3$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$r = .15$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.20$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.25$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.3$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$r = .20$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.25$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.3$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$r = .25$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.3$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$r = .3$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$.3$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Примечание. r^* представляет интервал отсутствия ассоциации; eq_r — исходный тест на отсутствие ассоциации; eq_fz, модифицированный тест на отсутствие ассоциации, основанный на z-преобразовании Фишера; eq_rs, основанный на передискретизации тест на отсутствие ассоциации.

10. Обсуждение

По мере того, как исследователи узнают о доступности тестов на отсутствие ассоциации, важно, чтобы были доступны рекомендации по проведению этих тестов в порядке, соответствующем. Основная цель этого исследования состояла в том, чтобы оценить и сравнить производительность трех тестов на отсутствие ассоциации: equiv_r, equiv_fz и equiv_rs. С что касается сравнения производительности трех тестов на отсутствие ассоциации, процедура на основе повторной выборки (equiv_rs) имела более точную частоту ошибок типа I и равномерно большая мощность. Тем не менее, результатом, который выделяется в этом исследовании, является обнаружение что очень большие размеры выборки (например, $n > 500$; см. что тесты на отсутствие ассоциации имеют достаточную мощность при отсутствии интервала ассоциации уже от 2.25 до 0.25 . Это очень важно, потому что теоретически становится очень сложно утверждать, что две переменные не связаны с отсутствием ассоциативного интервала шириной от 2.25 до 0.25 (или шире).

Copyright © Британское психологическое общество

Воспроизведение в любой форме (включая Интернет) запрещено без предварительного разрешения Общества.

534 Джейсон Р. Герцен и Роберт А. Крибби

Другой способ сформулировать эти результаты — принять во внимание, что для большинства психологических исследований выборочная корреляция, приближающаяся к 0,3, будет считаться практически, а также статистически значимой в большинстве случаев. Однако при попытке провести тест на отсутствие ассоциации с интервалом отсутствия ассоциации от 2,25 до 0,25 или меньше размеры выборки, необходимые для получения надлежащих результатов, очень велики (практически, для многих исследователей психологии). Поскольку мы ожидаем, что распространенные во многих психологических исследованиях интервалы отсутствия ассоциации будут меньше или равны от 2,20 до 0,20, исследователи должны знать, что размеры выборки, необходимые для получения соответствующих результатов с этими интервалами, могут быть непомерно большими (т. е. N \$ 500).).

Также важно подчеркнуть, что, хотя доступные тесты на отсутствие ассоциации не оптимальны (т.е. чрезмерно консервативны), возврат к традиционным тестам на ассоциацию (например, r Пирсона) нецелесообразен. Например, представьте, что кто-то хочет определить, находится ли ассоциация между двумя переменными в пределах интервала ассоциации $r = .10$. Мы смоделировали статистику r Пирсона для $r = .2$ и обнаружили, что при $N = 50$ 71,4% тестовой статистики не были значимыми. Другими словами, исследователь, неправильно использовавший традиционный ассоциативный тест для демонстрации отсутствия связи, объявил бы переменные независимыми примерно в 71% случаев, даже если корреляция населения выходит за пределы интервала ассоциации (обратите внимание, что интервал ассоциации не даже играют роль в оценке отсутствия ассоциации с помощью традиционного теста ассоциации). Даже при $N = 100$ исследователи будут объявлять переменные независимыми примерно в 50% случаев. Поэтому традиционные тесты на ассоциацию крайне непригодны для оценки отсутствия ассоциации, особенно при небольших размерах выборки. На самом деле частота ошибок типа I для оценки отсутствия связи с традиционными тестами ассоциации будет приближаться к единице по мере уменьшения размера выборки. Чтобы упростить использование тестов отсутствия ассоциации, мы включили в Приложение функцию R, которая будет генерировать статистику тестов для всех методов, описанных в этой статье.

В результате крайней консервативности статистики тестов на отсутствие ассоциаций с небольшими и умеренными размерами выборки мы предполагаем, что текущее исследование должно стать отправной точкой для исследования тестов на отсутствие ассоциаций в психологии. Результаты настоящего исследования представляют собой современные передовые методы получения надлежащих результатов, которые, что важно, включают в себя призыв немедленно прекратить выводы об отсутствии ассоциации, сделанные в результате неудачного теста на ассоциацию. Однако, следуя этим передовым методам, многие исследователи могут обнаружить, что они не могут провести желаемый тест из-за размера выборки, необходимого для текущих процедур. Поэтому мы рекомендуем провести дополнительные исследования тестов на отсутствие ассоциации; в частности, результаты этого исследования подчеркивают важность разработки тестов на отсутствие ассоциации, которые могут обеспечить приемлемую мощность при малых и средних размерах выборки.

использованная литература

Бонд, К.Ф., и Ричардсон, К. (2004). Наблюдение за z-преобразованием Фишера. Психометрика, 69, 291–303.

Диксон, П.М., и Пехманн, Дж.Х.К. (2005). Статистический тест, показывающий незначительную тенденцию. Экология, 86, 1751–1756.

Фишер, Р.А. (1915). Частотное распределение значений коэффициента корреляции в выборках неопределенно большой популяции. Биометрика, 10, 507–521.

Основная группа разработки R (2005 г.). R: Язык и среда для статистических вычислений.

R Foundation for Statistical Computing: Вена, Австрия. Получено с <http://www.R-project.org>.

Copyright © Британское психологическое общество

Воспроизведение в любой форме (включая Интернет) запрещено без предварительного разрешения Общества.

Отсутствие ассоциации 535

Робинсон, А.П., Дуурсма, Р.А., и Маршалл, Д.Д. (2005). Регрессионный тест эквивалентности для проверки модели: перенос бремени доказательства. Физиология деревьев, 25, 903–913.

Шуирманн, ди-джей (1981). При проверке гипотез, чтобы определить, является ли среднее значение нормального распределения содержится в известном интервале. Биометрия, 37, 617.

Шуирманн, ди-джей (1987). Сравнение двух процедур односторонних тестов и мощностного подхода для оценки эквивалентности средней биодоступности. Журнал фармакокинетики и биофармацевтики, 15, 657–680.

Вестлейк, WJ (1981). Ответ TBL Kirkwood: Тестирование биоэквивалентности – необходимость переосмысления. Биометрия, 37, 589–594.

Уэдон, Д.Е., Рэмпи, А.Х., Томпсон, В.Л., Потвин, Дж.Х., Масика, Д.Н., и Бизли, К.М.

(1992). Отсутствие связи между флуоксетином и суицидальными наклонностями при нервной булимии. Журнал клинической психиатрии, 53, 235–241.

Поступило 27 января 2009 г.; исправленная версия получена 2 сентября 2009 г.

Приложение

Следующая функция была создана для выполнения традиционного теста корреляции Пирсона, теста отсутствия связи на основе эквивалентности, теста отсутствия связи на основе эквивалентности с использованием z-преобразования Фишера и теста отсутствия связи на основе эквивалентности с ресемплинг в программном пакете R. Электронную копию функции можно найти по адресу: <http://www.psych.yorku.ca/cribbie/loa.html>. R — это программное обеспечение с открытым исходным кодом, доступное на сайте www.r-project.org. Чтобы использовать эту функцию, сначала запустите всю функцию в R. Другими словами, скопируйте и вставьте приведенный ниже синтаксис в R. Программа запустится, но вы не увидите никаких результатов в это время; все, что вы сделали, это определили функцию. На этом этапе вам нужно будет открыть свой набор данных (или иным образом определить свои переменные) в R. Если вам нужна помощь в открытии набора данных или определении переменных в R, см. справочные файлы и руководства, доступные на указанном выше веб-сайте R. Затем в командной строке R запустите функцию `equiv_corr`, набрав (без кавычек) `'equiv_corr(v1,v2,equivint)'`. Вы должны заменить имена ваших двух переменных на «v1» и «v2», а ваш интервал эквивалентности на «эквивинт». Вы также можете изменить альфа-уровень, изменив четвертый аргумент в функции. Например, после того, как вы выполнили всю функцию в R, вы можете ввести (без круглых скобок) `«equiv_corr(x,y, .2, .1)»` в подсказке в R, чтобы запустить статистические тесты для переменных x и y с интервалом эквивалентности 0,2 и альфа-уровнем 0,10.

```
equiv_corr ,-функция (var1, var2, эквивалент, альфа = 0,05, na.rm
¼ ИСТИНА, ::: ) {
  if (na.rm) x ,x[!is.na(var1)] if (na.rm) y ,y[!
is.na(var2)] corxy ,cor(var1,var2) n ,length(var1)
  nresamples ,-10000 ##### Запуск
  традиционного теста t для определения
  значимости корреляции ##### t ,corxy/
  (sqrt((1 2 corxy^2)/(n 2 2 ))) pvalue_tradt ,-1 2 pt(abs(t),n 2 2) ifelse
  (pvalue_tradt ,¼ альфа,
```

Copyright © Британское психологическое общество

Воспроизведение в любой форме (включая Интернет) запрещено без предварительного разрешения Общества.

536 Джейсон Р. Герцен и Роберт А. Крибби

```
decis_tradt, "Нулевая гипотеза об отсутствии корреляции между x и y может быть  
отвергнута.", decis_tradt, "Нулевая гипотеза об отсутствии корреляции между x и y  
не может быть отвергнута").
```

```
#### Запуск исходной процедуры двух t-тестов на эквивалентность  
#####
```

```
equivt1, -(corxy 2 эквивалент)/sqrt((1 2 corxy^2)/(n 2 2)) pvalue1_equivt, -pt(equivt1, n 2 2) equivt2, -  
(corxy + эквивалент)/sqrt((1 2 corxy^2)/(n 2 2)) pvalue2_equivt, -1 2 pt(equivt2, n 2 2) ifelse  
(pvalue1_equivt, 1/4 alpha & pvalue2_equivt, 1/4 alpha, decis_equivt, " Нулевая гипотеза о том, что  
корреляция между var1 и var2 выходит за пределы интервала эквивалентности, может быть  
отвергнута.", decis_equivt, " «Нулевая гипотеза о том, что корреляция между var1 и var2 выходит  
за пределы интервала эквивалентности, не может быть отвергнута».)
```

```
##### Запустите процедуру двух t-тестов для эквивалентности с критерием Фишера.  
z преобразование #####
```

```
zei, -log((1 + эквивинт)/(1 2 эквивинт))/2 zcorxy, -log((1 + corxy)/(1 - corxy))/  
2 equivt1_fz, -(zcorxy 2 zei)/(1/ sqrt(n 2 3)) pvalue1_fz, -pnorm(equivt1_fz)  
equivt2_fz, -(zcorxy 2 zei)/(1/ sqrt(n 2 3)) pvalue2_fz, -1 2 pnorm(equivt2_fz)  
ifelse (pvalue1_fz, 1/4 альфа и pvalue2_fz, 1/4 альфа,
```

```
decis_fz, " Нулевая гипотеза о том, что корреляция  
между var1 и var2 выходит за пределы интервала эквивалентности, может быть  
отклонено.»,
```

```
decis_fz, " Нулевая гипотеза о том, что корреляция  
между var1 и var2 выходит за пределы интервала эквивалентности, не может быть  
отклонено».)
```

```
#### Запустить версию процедуры двух t-тестов для эквивалентности с повторной выборкой #####  
resamp, function(x, m = 10000, theta, conf.level = 0.95, ...)
```

```
{ n, длина (x)
```

```
Данные, матрица(выборка(x, размер 1/4 n*m, заменить 1/4 T), nrow 1/4 m) thetastar, apply(Data, 1,  
theta, ...)
```

```
M, -среднее (тетазвезда)
```

```
S, -sd(thetastar) alpha, -1 2
```

```
conf.level CI, -quantile(thetastar, c(alpha/
```

```
2, 1 2 alpha/2)) return(list(ThetaStar = thetastar, Mean.ThetaStar = M,
```

```
SEThetaStar = S, Percentile.CI = CI)) } matr, -cbind(var1, var2) mat, -  
as.matrix(matr)
```

Copyright © Британское психологическое общество

Воспроизведение в любой форме (включая Интернет) запрещено без предварительного разрешения Общества.

Отсутствие ассоциации 537

```
тета-функция (x, mat) { кор (mat [x,
1], mat [x, 2])}

результаты ,-resamp(x = 1:n,m = nresamples, тета = тета,
mat ¼ mat)
q1 ,-quantile(results$ThetaStar, alpha) q2 ,-
quantile(results$ThetaStar,1 2 alpha) q1negei ,-q1 2 эквивалент
q2negei ,-q2 2 эквивалент q1posei ,-q1 + эквивалент q2posei ,-q2 þ
эквивалентно ifelse (q2negei , 0 и q1posei . 0,

decis_rs ,- "Нулевая гипотеза о том, что корреляция
между var1 и var2 выходит за пределы интервала эквивалентности, может быть
отклонено.»,
decis_rs ,- "Нулевая гипотеза о том, что корреляция
между var1 и var2 выходит за пределы интервала эквивалентности, не может
быть отклонено.».)
#### Резюме ####
title1 ,-«Традиционный тест корреляции, Но: rho = 0» title2 ,-«Тест отсутствия
ассоциации на основе эквивалентности» title3 ,-«Тест отсутствия ассоциации на
основе эквивалентности с критерием Фишера трансформация z»

title4 ,-«Эквивалентный тест на отсутствие ассоциации с
Передискретизация»
stats_tradt,-c(corxy,t,n 2 2,pvalue_tradt,decis_tradt) name(stats_tradt),-c("Пирсон r", "t-
статистика", "df",
«р-значение», «Решение»)
stats_equivt,-c(corxy,equivint,equivt1,pvalue1_equivt,
equivt2,pvalue2_equivt,n 2 2,decis_equivt) name(stats_equivt) ,-
c("Pearson r", "Equivalence Interval", "t-stat 1", "pval_t1", "t-stat 2", "pval_t2", «дф», «Решение»)

stats_fz ,-c(corxy,equivint,equivt1_fz,pvalue1_fz,equivt2_fz,pvalue2_fz,decis_fz)

имена (stats_fz), -c («Пирсон r», «Интервал эквивалентности»,
«z-стат 1», «pval_z1», «z-стат 2», «pval_z2», «Решение»)
stats_rs,-c(corxy,equivint,nresamples,q1,q2,decis_rs) name(stats_rs),-c("Pearson r",
"Equivalence Interval", "# of Resamples", "100(alpha) Percentile", " 100 (1-альфа)
Процентиль», «Решение»)
out ,-list (title1,stats_tradt,title2,stats_equivt,title3,
stats_fz, title4, stats_rs)
вне
}
```