

Критерий Мак-Немара и разность долей зависимых выборок

	До лечения	После лечения
До лечения	23	5
После лечения	0	14

Без поправки: $\chi^2 = \frac{(5-0)^2}{(5+0)} = 5$; $df=1$, $P= 0,025347$.

С поправкой Йейтса на непрерывность: $\chi^2 = \frac{(5-1)^2}{(5+0)} = 3,2$; $df=1$, $P= 0,073638$.

StatXact 7:

McNemar's Test for Two Related Binomial Proportions (Conditional)

mcnemars (method = exact, time_limit = none);

Data File:

Number of Rows: 2

Number of Columns: 2

Summary of the Test Statistic:

Binomial Proportions [COLUMN]: piHat_1	0.5476
Binomial Proportions [ROW]: piHat_2	0.6667
Difference of Proportion: piHat_2 - piHat_1	0.119
Std. Error: (pooled estimate of stdev of piHat_2-piHat_1)	0.05324
Standardized Difference: (piHat_2-piHat_1)/Std. Error	2.236

Inference:

Type	P-Value			
	Tail	1-Sided	Tail	2-Sided
Asymptotic	.GE.	0.01267	.GE.	0.02535
Exact	.GE.	0.03125	.GE.	0.0625

Elapsed time: 0:0:0.00

AtteStat 12.5:

23	5
0	14

Критерий Мак-Немара

Асимптотика хи-квадрат

Статистика, Р-значение (одностороннее), Р-значение (двустороннее)

5 0,025347 0,050695

Асимптотика с поправкой Йейтса

Статистика, Р-значение (одностороннее), Р-значение (двустороннее)

3,2 0,073638 0,147277

Точный условный

Статистика, Р-значение (одностороннее), Р-значение (двустороннее)

5 0,03125 0,0625

В добротном учебнике Закса Л. Статистическое оценивание на стр. 339 со ссылкой на 3 источника 1969-1970 гг приводятся другие формулы:

Макнимар смог показать, что изменения такого рода (если $(b + c)/2 \geq 4$) могут проверяться на основании статистики

$$\hat{\chi}_1^2 = \frac{(b-c)^2}{b+c+1} \quad (4.40)$$

(см. [Bennett, Underwood, 1970]), а также [Gart, 1969] и [Maxwell, 1970], а если $(b + c) < 30$, то с поправкой на непрерывность — с помощью статистики

$$\hat{\chi}_1^2 = \frac{(|b-c|-1)^2}{b+c+1} \quad (4.40a)$$

при одной степени свободы.

В википедии (http://en.wikipedia.org/wiki/McNemar%27s_test) со ссылкой на Флейса (Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: [John Wiley](#). p. 114.) приводится другая формула:

The McNemar test statistic with Yates' correction for continuity^[3] is given by:^[4]

$$\chi^2 = \frac{(|b-c|-0.5)^2}{b+c}.$$

Предварительные выводы:

1. Существует несколько вариантов формул для расчёта как собственно критерия Мак-Немара, так и для введения поправки на непрерывность. Все они дают разные значения P . Возможно это связано с разными подходами к учёту степени отклонения статистики от распределения хи-квадрат в зависимости от насыщенности четырёхпольной таблицы.
2. В результате анализа четырёхпольной таблицы получается двухсторонняя оценка. При односторонней гипотезе полученное значение P нужно разделить на 2. В программе AtteStat для асимптотики в качестве односторонней оценки приводится двухсторонняя, а в качестве двухсторонней - удвоенная двухсторонняя.
3. Точные значения P для критерия Мак-Немара в пакетах StatXact 7 и AtteStat 12.5 идентичны.

Разность долей:

Вероятность до лечения – Вероятность после лечения =

$$= \left(\frac{23+5}{42} \right) - \left(\frac{23+0}{42} \right) = \frac{5}{42} = 0,119047619 \text{ или округлённо } 0,119.$$

StatXact 7:

Confidence Interval on Difference of Two Related Binomial Proportions Based on the Standardized Statistic and Inverting Two 1-Sided Tests

binomial (type = related, test_type = difference, compute = ci, method = exact, time_limit = none);

Data File:

Number of Rows: 2

Number of Columns: 2

Summary of the Test Statistic:

Binomial Proportions [COLUMN]: piHat_1	0.5476
Binomial Proportions [ROW]: piHat_2	0.6667
Difference of Proportions: piHat_2 - piHat_1	0.119
Std. Error: (pooled estimate of stdev of piHat_2-piHat_1)	0.05324
Standardized Difference: (piHat_2-piHat_1)/Std. Error	2.236

Inference:

Type	P-Value			95.00% Conf. Interval for (pi_2-pi_1)	
	Tail	1-Sided	2*1-Sided	Lower Limit	Upper Limit
Asymptotic	.LE.	0.01267	0.02535	0.02527	0.25
Exact	.LE.	0.0128	0.02561	0.01754	0.2564

Elapsed time: 0:0:7.25

Калькулятор плава:

Биномиальные доверительные интервалы

Доверительный интервал

95%

95%ДИ

N	k	p	3,98%	25,63%	Метод Клоппера-Пирсона
42	5	11,90%	2,11%	21,70%	Нормальная аппроксимация (Вальд)
			5,19%	25,00%	Метод Вилсона
			3,87%	27,78%	По распределению Пуассона (через хи2)
			4,73%	25,46%	Откорректированный метод Вальда

Предварительные выводы:

1. Все приведённые в калькуляторе методы дают для этого примера более высокую оценку нижней границы доверительного интервала.
2. Оценка значимости двусторонней гипотезы через точный критерий Мак-Немара и через разность зависимых долей различна: если строго придерживаться порогового значения в 0,05, то Мак-Немар «говорит» о неэффективности лечения ($P=0,0625$), а разность долей – об эффективности ($P=0,02561$). Или здесь правильнее использовать одностороннюю гипотезу?