СИЛА НА ЕФЕКТА ПРИ ДИХОТОМНИ ПРОМЕНЛИВИ

Стефан Матеев

Въведение

Непосредствен повод за написването на настоящия текст бе дискусията в сайта http://www.nature.com/news/psychology-journal-bans-p-values-1.17001?WT.mc_id=FBK_NatureNews относно изискването, поставено от редакторите на списанието Basic and Applied Social Psychology (BASP). Според тях, интерпретацията на емпиричните данни трябва да става само въз основа на индекси за сила на ефекта и доверителни интервали. Не се допускат за публикуване статии, в които психологическата интерпретация е въз основа на пресмятане на известните ни p-стойности и „значимости". Цитираното по-горе списание не е единственото, което поставя такива изисквания към тези, които желаят да публикуват в него. Подобни текстове четем и в „Напътствия към авторите" на авторитетни психологически списания, като например Psychological Science.

Тези изисквания не са плод на редакторски капризи. С течение на времето все повече професионалисти – не само психолози, но и клиницисти, епидемиолози и др., се убеждават, че анализът на данни, получени от групи лица, е ненадежден, ако е извършен само въз основа на „значимости". Изглежда, че броят на тези хора вече наближава критичния. Това налага значително преосмисляне на знанията, които може да придобием от „класическите" учебници и наръчници по статистика. Все повече се осъзнава, че описанието и интерпретацията на данни и на резултати от изследвания трябва да става чрез индекси за т.н. *сила на ефекта* (effect size). Накратко, става въпрос за клас от индекси, които (1) са безразмерни, т.е. не се изразяват в единиците на скалата на измерване, (2) чиято величина не се определя от общия брой на изследваните лица, (3) които описват данните, като дават информация за съществеността (а не на „значимостта") на получените резултати, и (4) дават възможност за количествени сравнения между данни, получени от различни изследвания на различни автори и с различни инструменти.

Понастоящем все още не може да се каже, че индекси за сила на ефекта са разработени за всички възможни типове данни и изследователски планове. Предложени са добри индекси за описание на данни, които се измерват в интервална скала или скала на отношения. Полезни източници за философията и пресмятанията на силата на ефекта при такива случаи са Rosnow & Rosenthal (2003), Kirk (2003), Kline (2004), Henson, (2006). Също така, препоръчителни четива на български са Василев (2014) и Матеев (2014, 2014a). Настоящата статия е посветена на индексите за описание на данни, които са получени в рамките на дихотомна скала.

§1. Подреждане на дихотомни данни: 2x2-матрици

Нека си представим, че сме избрали група лица, и сме им предложили въпросник, в който те трябва да отбележат какъв им е пола, и дали имат или нямат права за управление на автомобил. Съответно в бланките на въпросника получаваме отговори, кодирани като М (мъжки) и Ж (женски), както и „да“ (имам шофьорска книжка) и „не“ (нямам шофьорска книжка). Отговорите носят информация за определени характеристики на лицата, които наричаме променливи. Чрез въпросника ние измерваме променливите „пол“ и „права за правоуправление“ за всяко лице, като данните от измерването са в т.н. номинална скала. Всяка променлива се състои от две нива. Такива променливи се наричат дихотомни. Тъй като променливите са две, и възможните отговори са също два, от въпросника ще получим 2x2 = 4 възможни и изхода. Такива данни е удобно да се подредят в матрица от следния вид (Табл. 1)

Таблица 1. Типично представяне на данни, получени от измерване на дихотомни променливи. Таблицата се нарича 2x2-матрица, contingency table. Обозначаването на клетките със символите a, b, c, d е обичайно в литературата.

		ПОЛ		
		М	Ж	
ПРАВА	да	a	b	a+b
	не	c	d	c+d
		a+c	b+d	N

Символите a, b, c и d в Табл.1 представляват честотите (или просто бройките) на лицата, които се характеризират с всеки от четирите изхода:
a е честотата на лицата, които са и мъже, и имат шофьорски права
b е честотата на лицата, които са и жени, и имат шофьорски права
c е честотата на лицата, които са и мъже, и нямат шофьорски права
d е честотата на лицата, които са и жени, и нямат шофьорски права

Съответно
a + b е броят на шофьорите в групата,
c +d е броят на лицата в групата, които нямат права,
a + c е броят на мъжете, и
b + d е броят на жените в групата.
Тези числа се наричат странични честоти (marginal frequencies)
Общият брой на лицата в групата е N = a + b + c + d

От така подредените данни е лесно да се извлече различен вид количествена информация. Например, веднага се вижда броят (b + d) на дамите в групата, ако трябва да им се поднасят цветя. Относителни дялове, или проценти, също могат да се пресмятат от данните в матрицата. Например, относителният дял на шофьорите в групата е (a+ b)/N. Този относителен дял ни дава вероятността, ако изберем случайно едно лице от групата, то да се окаже шофьор.

Някои относителни дялове представляват особен интерес. Такива например, са: a/(a + c) – това е делът на шофьорите от мъжете, или b/(b + d), делът на шофьорите от жените. Такива относителни дялове служат за пресмятането на т.н. *условни вероятности*. Например, вероятността едно лице да се окаже шофьор, *при условие,* че то е мъж, се дава именно от a/(a + c). Съответно и вероятността едно лице да се окаже шофьор, *при условие,* че то е жена, се дава от b/(b + d). Защо тези условни вероятности представляват интерес?

Работата е там, че подобни проучвания не се правят просто така. Обикновено изследователят има предварителна работна хипотеза за причина и следствие. Той може да смята, че едната променлива, например „пол“, оказва влияние, или *ефект* върху другата, в случая „права за правоуправление“. Емпиричните данни, представени в 2x2 матрица, могат да подкрепят неговата хипотеза силно, или не много силно, или никак да не я подкрепят, в зависимост от *силата на ефекта* (effect size). Но за тази цел силата на ефекта трябва да се определи количествено с подходящ индекс.

При други обстоятелства, работна хипотеза може и да не е формулирана предварително, но данните да показват, че двете променливи не са независими, или, че те са свързани. В такива случаи се говори обикновено за наличие на *връзка* (association) – силна, средна или слаба, между променливите. Вместо „връзка“, в статистиката се използва и думата *корелация*. Понятието „ефект“ не изглежда особено подходящо при описание на такива данни, тъй като установяването на наличие на връзка между променливите не е основание за формулиране на причина и следствие. Все пак, изследователят често се старае да обясни наблюдаваната връзка между променливите чрез някаква хипотеза, и тогава понятието „ефект“ придобива смисъл. То формално навлиза все повече в психологическата литература, като сила на ефекта и сила на връзката са еквивалентни понятия и индексите, които ги описват, се пресмятат по един и същ начин.

Именно двете условни вероятности, a/(a + c) и b/(b + d), могат да ни информират за наличието или липсата на връзка между променливите. Ако е изпълнено равенството

$$a/(a + c) = b/(b + d) \qquad (1)$$

връзка между променливите *няма*. С други думи, ако относителните честоти, с които се пресмятат двете условни вероятности, са равни, то променливата „пол“ няма връзка с променливата „права за управление“. Колкото е вероятно мъж да има шофьорска книжка, толкова вероятно е и жена да има книжка. Полът не играе роля, или не оказва ефект, върху придобиването на права за правоуправление.

Връзка се наблюдава, ако например се окаже, че
$a/(a + c) > b/(b + d)$.

Това би означавало, че процентът на мъжете-шофьори е по-голям от процента на жените-шофьори в групата. Такива данни могат да дадат храна за размисъл например относно ефекта, който полът оказва върху желанието да се изкарат курсове за кормуване, или да окажат подкрепа на някоя друго разсъждение в термини на причина и следствие. Но преди това следва да се установи наличието на връзка между двете променливи, и това най-лесно става чрез сравнението на вероятностите в равенство (1). Формален начин за извеждане на равенство (1) е даден в Допълнение А.

Тъй като разглеждаме дихотомни променливи, няма значение дали променливата „пол“ е нанесена по колоните на матрицата (както е на Табл. 1), или по редовете. Няма значение и реда, в който са нанесени редовете и колоните. При различно подреждане на нивата на двете променливи, вербалната изказване на вероятностите в равенство (1) ще се променя, те ще се четат по друг начин, но правилото за липса на връзка между променливите ще се запазва.

Съответно, ако равенство (1) е изпълнено, ще бъдат изпълнени и следните равенства

$c/(a + c) = d/(b + d)$ (2)

$a/(a + b) = c/(c + d)$ (3)

$b/(a + b) = d/(c + d)$ (4)

Всяко от равенствата (1) – (4) означава, че няма връзка между променливите. При нарушаването на кое да е от тях се нарушават и останалите; това означава, че има връзка между променливите. Числени примери за пресмятане на вероятностите в равенства (1) – (4) са представени в Табл. 2 и 3, като променливите са отново пол и права за кормуване..

Таблица 2. Пример за данни, получени от N = 175 лица, от които 75 жени. Данните показват липса на връзка между променливите пол и права. Равенствата (1) – (4) са:
(1): $80/100 = 60/75$, (2): $20/100=15/75$, (3): $80/140=20/35$, (4): $60/140 = 15/35$

		ПОЛ		
		М	Ж	
ПРАВА	да	80	60	140
	не	20	15	35
		100	75	175

Таблица 3. Пример за данни, получени от N = 175 лица, от които 75 жени. Данните показват наличие на връзка между променливите пол и права. Равенствата (1) – (4) стават неравенства:
(1): $90/100 \neq 40/75$, (2): $10/100\neq35/75$, (3): $90/130\neq10/45$, (4): $40/130 \neq 35/45$

		ПОЛ		
		М	Ж	
ПРАВА	да	90	40	130
	не	10	35	45
		100	75	175

Когато променливите са в интервална скала, връзката между тях може да се опише с думите „положителна“ или „отрицателна“. При дихотомните променливи понятията „по-голямо“ и „по-малко“ нямат смисъл, и съответно връзката не може да бъде нито положителна, нито отрицателна. Затова при тях, вербално, или в текст, твърдението за връзка следва да се описва с данни. Най-простият начин е да се съобщят двете вероятности в кое да е от равенствата (1) – (4). Например, данните от Табл. 2 позволяват твърдението: „Връзка между променливите пол и права не се установява, 80% както от мъжете, така и от жените имат права за кормуване“. От Табл. 3 може да се каже, че „установява се наличие на връзка между променливите пол и права – 90% от мъжете имат права за кормуване, докато такива права имат около 53% (40/75) от жените“. Тези твърдения се основават на равенство (1) в примерите, но в зависимост от контекста могат да се съобщават и вероятностите от другите равенства (2) – (4).

От тези разглеждания следва една проста поука – за наличие или липса на връзка може да се говори, ако са налице данни за поне две вероятности, или два процента! Съобщаването само на един процент не дава сериозни основания за извод за връзка. Да вземем, например, изказването „75% от болните от рак на белия дроб са пушачи!“, (дочуто от телевизионно предаване) което има за цел да илюстрира вредата от цигарите. Без да отричаме тази вреда, отново следва да определим това твърдение като твърде непрофесионално. В него се разчита на наивната вяра на слушателя, че колкото е по-висок процентът, толкова е по-опасно пушенето. Това може да е така, но може и да не е. Например, сигурно е, че не 75%, а даже 99% (да речем) от болните от рак на белия дроб ядат краставици, но това съвсем не означава, че яденето на краставици има връзка със заболяването. Препоръчваме на читателя да има пред вид, че зад такива проценти и твърдения всъщност е „скрита“ 2x2 матрица от типа на тази на Табл. 1. Без да се даде повече информация за съдържанието на матрицата, не може да се установи наличието на връзка и нейната сила.

Да разгледаме още веднъж равенство (1). Имаме
$a/(a + c) = b/(b + d)$
С малко аритметика това равенство се преобразува в
$a*(b + d) = b*(a + c)$
и като се разкрият скобите и се съкрати a*b от двете страни, то става
$a*d = b*c$
или
$$a*d - b*c = 0 \qquad (5)$$
също и
$$a*d/b*c = 1 \qquad (6)$$

Равенствата (5) и (6) могат да се изведат и от кое да е от равенствата (2) – (4). Те представляват, макар и по-неочевидни, правила за установяване на липса на връзка между променливите, които ще използваме в по-нататъшните разглеждания.

Нека да навлезем по-навътре в изкуството да извличаме връзка от 2x2 матриците. Вече определихме, че за наличие връзка между дихотомните променливи може да се говори, ако се нарушават равенствата (1) – (6). Но това все още не е достатъчно за добър извод. Следва да се отговори на въпроса: колко силна е тази връзка, и представлява ли тя някакъв практически, или теоретически интерес? Читателите, които са изучавали статистика, ще кажат веднага: ами трябва да се провери дали връзката е „значима“! Ако е значима, тя представлява интерес, ако не е значима, можем просто да забравим за нея. Както отбелязахме във въведението, именно този вид разсъждение се критикува все повече в специализираната литература през последните години. В резултат на тази критика, водещите списания по психология започват да представят изисквания към авторите, да не интерпретират резултатите си въз основа на „значимости“. Вместо това те трябва преди всичко да опишат данните си в термини на сила на връзката или еквивалентното понятие сила на ефекта.

По-нататък разглеждаме четири широко използвани индекса за описание на силата на връзката между две дихотомни променливи, които се използват не само в психологията, но и в медицината, особено в епидемиологичните изследвания. Това са:
- разлика между рискове
- относителен риск
- φ-коефициент (фи-индекс)
- отношение на шансове
Ще разгледаме техните преимущества и недостатъци. Преди това обаче, ще разгледаме понятието риск (risk). То се използва съществено при два от индексите за описание на сила на връзка между променливи.

§2. Понятието „риск“

В ежедневието използваме думата риск, когато искаме да информираме за някаква опасност или за нежелано събитие. Ако в зоопарка влезнете в клетката на лъва, съществува риск да бъдете изяден. Но може и да ви се размине, ако лъвът е вече сит. Следователно, риск означава вероятна, а не сигурна опасност, или нежелано събитие. Може би именно по тази причина думата риск в много области, най-вече в клиничните специалности, се използва вместо термина вероятност. Риск може да се пресмята точно, така както и вероятността.

Например, можем да определим риска пушач да пострада от дадено заболяване. За тази цел може да подберем група от N здрави лица, които пушат, да проследяваме съдбата на всяко лице в период от, например десет години, и да регистрираме честотата n на случаите на поява на заболяването в групата. Тогава рискът се пресмята като n/N, може и в проценти, и представлява смислена оценка на вероятността белият дроб на пушача да пострада.

Примерът, който разгледахме, представлява *проспективно* (prospective) изследване. Подбираме здрави лица и проследяваме какво е бъдещето им. В този случай пресмятането на риска като относителен дял или процент е логически правомерно. Смислено е същото изследване и пресмятане да се проведе с група от непушачи, и след това резултатите от двете изследвания да се сравнят по някой от начините, които ще разгледаме по-долу.

Проспективните изследвания са „златен стандарт“ при психологически и клинични изследвания. Те обаче, често не са възможни поради финансови, технически и етични причини. Затова се прилагат и т.н. *ретроспективни* (retrospective) изследвания, наричани още case-control studies. При тях подбираме група лица, които вече страдат от заболяването, и ги разпитваме за навиците им на живот до момента на изследването, т.е. за миналото им. Установяваме броя на болните лица, които пушат или са пушили. За съжаление, риск не може да се пресметне смислено от тези данни. Можем да определим вероятността, едно лице от болните да се окаже пушач. Но това не е „риск“. Ние знаем колко е броят на болните в групата, но не знаем колко е броят на пушачите, от които е извлечена групата от болни, за да пресметнем риска (или вероятността) пушач да се разболее.

От тези разглеждания става ясно, че за да говорим за „риск“, са необходими някои предпоставки. Първата е, че макар и имплицитно, трябва да предположим някаква причинно-следствена връзка. Предполагаме, че пушенето води до заболяване, а не обратното. Наличието на статистическа връзка би подкрепило предположението за причинно-следствена връзка. Втората е, че трябва да се съобразяваме с това, по какъв начин са събрани данните, и тогава да използваме подходящ индекс за наличие и сила на връзка. Понятието „риск“ не е подходящо за всеки изследователски план, или дизайн.

В настоящия текст нямаме за цел да разглеждаме различните изследователски подходи, които са повече на брой, отколкото разгледаните два екстремални примера. Нататък в текста разглеждаме четирите индекса за сила на връзката чрез примери, в които данните са получени от проспективни изследвания, и ще споменаваме ретроспективните само при необходимост.

§ 3. Разлика между рискове, risk difference, RD

Възможен и често употребяван индекс за определяне на силата на връзката между две дихотомни променливи, е разликата между два риска, или, което е същото, между две вероятности. Нека припомним равенство (1), което информира за липса на връзка между променливите:
$a/(a + c) = b/(b + d)$
Нека използваме примера от Табл. 1, като въведем за удобство означенията
$p(да|М) = a/(a + c)$, - вероятността за отговор „да, имам права за управление“ от мъжете, и
$p(да|Ж) = b/(b + d)$, вероятността за „да, имам права“ от жените

От числата в Табл. 3 пресмятаме $p(да|М) = 90/100 = 0.90$ и $p(да|Ж) = 40/75 = 0.53$. Абсолютната стойност на разликата между тези две вероятности

$|p(да|М) - p(да|Ж)| = |0.9 – 0.53| = 0.37$

може да се използва за индекс, който да информира относно силата на връзката между двете променливи, пол и права. Тази разлика е нула при отсъствие на връзка, и теоретически не може да надхвърли стойността 1. Стойността единица определя и максимално силна, или перфектна връзка, тя се наблюдава ако едната вероятност е равна на 1 и другата е нула. От матрицата това се установява „на око“ когато две от диагоналните клетки, например c и b, са нули.

Предимно в клиничните изследвания индексът се нарича „разлика между рисковете“ (risk difference, RD). Той е прост за разбиране, илюстрира се на графика обикновено с две различни по височина стълбчета и масово се използва в студентски разработки. Съществува тест за проверка на „значимост“ на разликата, и стандартна формула за пресмятане на доверителния интервал на разликата (Калинов, 2013). Както тестът, така и методът за пресмятане на доверителния интервал за приблизителни, те „не работят“

когато двете вероятности са твърде близки до нула или единица, и едновременно с това броят на лицата, мъже и жени, от които се пресмятат вероятностите, са твърде малки.

Индексът разлика между вероятностите или рисковете RD не е без недостатъци. Преди всичко, изследователят трябва да има някаква представа за това, кое е причина и кое е следствие, когато представя двете вероятности. Той може да се изкуши да представи разликата между вероятностите p(М|да) - p(Ж|да) вместо p(да|М) - p(да|Ж) с надеждата, че е все едно с коя разлика ще опише силата на връзката. В общия случай не е все едно. Предлагаме на читателя да провери, че p(М|да) - p(Ж|да) = 0.47, докато по-горе пресметнахме, че p(да|М) - p(да|Ж) = 0.37. Разликите не са еднакви, което може да доведе до недоразумения при описанието на данните. Разликите са еднакви само в частния случай, при който четирите странични честоти в матрицата са равни. Тогава задължително имаме a = d и b = c, т.е. матрицата е диагонално симетрична. Но това на практика се получава твърде рядко. Най-добре е, изследователят да има представа за това, коя от двете променливи може да се разглежда като причина, и коя – като следствие, и тогава да избере с коя разлика ще опише силата на връзката между променливите, съответно и силата на ефекта, който едната променлива оказва върху другата.

Друг недостатък на RD е, че еднакви стойности на разликата не съответстват на представите ни за еднакво силни връзки или ефекти. Например разлика от 10% или 0.1 може и да не се възприеме много сериозно, ако двете вероятности са 0.45 и 0.55. Но същата разлика може да придобие съвсем друго значение, ако вероятностите са 0.06 и 0.16. Особено клиницистите, а също и представителите на медиите, възприемат втората разлика много по-сериозно. По тази причина за описание на данните се използва и друга мярка, именно отношението на двете вероятности или рискове.

§ 4. Отношение на рискове, risk ratio, RR

Недостатъкът, че разликата от 10% може да информира за ефекти или връзки между променливите, които в действителност са различни по сила, се избягва чрез т.н. отношение между вероятности, или отношение между рискове, наричано и относителен риск (risk ratio, relative risk или RR). Вместо разликата между две вероятности p2 и p1, разглеждаме тяхното отношение RR = p2/p1. Ако приложим този метод към вероятностите 0.45 и 0.55, получаваме RR = 0.55/0.45 = 1.22; приложен към 0.06 и 0.16, получаваме RR = 0.16/0.06 = 2.66, което внушава значително по-силна връзка.

Ако имаме липса на връзка или ефект между променливите, двете вероятности са равни и RR = 1; това отношение поне теоретически може да се мени от нула до безкрайност. Отношението на рисковете се използва когато понятието риск в неговия ежедневен смисъл, като вероятност за опасност или нежелано събитие, има смисъл. Това е предимно при проспективни клинични и епидемиологични изследвания, но не и при ретроспективни. В § 2 по-горе бе изяснено защо рискът няма смисъл при ретроспективни изследвания.

Описанието на данни чрез отношение на рискове се използва широко и в медиите. То обаче, не е без недостатъци. Нека разгледаме примера в Табл.3. В § 3 вече пресметнахме, че вероятността да намерим шофьор между мъжете е p(да|M) = 0.9, а между жените е p(да|Ж) = 0.53. Отношението им е RR = 0.9/0.53 = 1.687, т.е. при мъжете вероятността да намерим шофьор е 1.687 пъти по-висока, отколкото при жените. Изглежда, че това число добре описва връзката между променливите „пол“ и „права“. Но на изследователя, студент или докторант, може да реши да дефинира „нежелано“ събитие, и то да бъде лице от дадена група да *не* може да шофира. Тогава думата „риск“ става оправдана. Рискът да попаднем на жена, която да не може да шофира, е p(не|Ж) = 35/75 = 0.47, рискът да попаднем на мъж, който не може да шофира, е p(не|M) = 10/100 = 0.1. Относителният риск пресмятаме като RR = 0.47/0.1 = 4.7. Тази стойност на RR внушава много по-силна връзка между променливите, отколкото RR = 1.687. С коя от двете да се опише силата на връзката?

Примерът илюстрира основния недостатък на индекса „отношение на рисковете“. Числото RR, с което се описват едни и същи данни, може силно зависи от това, как се дефинират числителят и знаменателят в отношението, и въобще от това, дали или как ще се дефинира понятието „риск“. Този недостатък се изтъква в по-сериозните текстове, които разглеждат описанието на връзката между дихотомни променливи (Fleiss & Berlin, 2009). Независимо от това, отношението RR между рисковете се използва и даже се препоръчва при описание на данни главно при клинични и епидемиологични изследвания (Sackett et al., 1996)

§ 5. φ-коефициент

Всеки читател, който е преминал основен курс по статистика, знае какво е корелационен коефициент, или индекс на Пирсън r. За подробности по пресмятането му, препоръчваме на читателя подробното разглеждане на понятието корелация в Калинов (2013). Индексът се използва масово, когато имаме група от N лица, всяко от които се характеризира с две променливи, измерени поне в интервална скала или скала на отношения. Това могат да са например ръст в сантиметри и тегло в килограми. Пресмятането на r може да се извърши лесно с помощта на програмата Excel. Просто трябва да си приготвим две колони от двойки данни, в едната от които да е нанесен ръста на лицата, а в другата – височината им. Пресмятането на r става с командата =pearson(данни за ръст; данни за височина).

При данни, които са получени от дихотомни променливи, може да се приложи подобна технология на пресмятане на връзка. За тази цел данните трябва първо да се кодират. Нека в примера на Табл. 2 кодираме жена с 1, мъж с 0, също и отговор „да“ с 1 и отговор „не“ с 0. Числата 1 и 0 се наричат „фиктивни“ променливи (dummy variables). В примера на Табл. 2 общият брой на лицата е N = 175. Приготвяме си две колони със 175 двойки данни, които съдържат нули и единици, и пресмятаме r по същия начин, както и при данните с ръст и тегло на лицата. Полученото число е равно на коефициента на Пирсън, но се обозначава с името φ–индекс (фи-коефициент).

Индексът φ може да се пресмята в случаи, при които не разполагаме със списък на суровите данни, а само с честоти в 2x2-матрица (Табл. 1). Тогава

$$\varphi = \frac{|a*d-b*c|}{\sqrt{(a+c)*(b+d)*(a+b)*(c+d)}}$$

Забелязваме, че в числителя на този израз се намира разликата в равенство (5). Когато тази разлика е равна на нула, правим извода, че няма връзка между променливите. Съответно и индексът φ е равен на нула. Така се получава в примера на Табл. 2, в нея a*d – b*c = 80*15 – 60*20 = 0. Това е удобство в сравнение с индекса RR, при който липсата на връзка се дава от стойността единица. Подобно и на r на Пирсън, φ не може да надхвърли стойността единица. Тъй като става въпрос за променливи в номинална скала, становище за посоката на връзката, положителна или отрицателна, не е смислено. По тази причина, удобно е φ да се дава с абсолютната му стойност, като вербално се пояснява в какво точно се изразява връзката между променливите.

На пръв поглед, φ представлява удобен индекс, с който да опишем силата на връзката, или силата на ефекта, който едната променлива оказва върху другата. Поне по принцип, φ може да се съпоставя с данни за r получени от данни в интервални скали. Тази възможност, обаче, е ограничена. Работата е там, че стойността на φ зависи от стойностите на страничните честоти. Ако четирите странични честоти са еднакви (и съответно матрицата е диагонално симетрична), то тогава при перфектна връзка между променливите φ може да достигне максималната стойност от единица. Не е така, ако страничните честоти са различни; тогава при една и съща по сила връзка, изчислена с другите индекси, стойността на φ силно се влияе от разликите между страничните честоти. Този факт илюстрираме с пример на две 2x2 матрици, при които всички други разглеждани индекси са еднакви (odds ratio, OR, разглеждаме след малко), но индексите φ се различават значително (Табл. 4). При едната от матриците (вдясно) страничните честоти са еднакви, по 500. При другата (вляво) обаче, страничните честоти се различават значително, което води до силно спадане на стойността на φ.

Таблица 4. Представени са две матрици, при които три от разглежданите в този текст индекси за сила на връзката са еднакви, но се различават по φ-индекса.

	М	Ж	
да	40	190	230
не	10	760	770
	50	950	1000

RD = (40/50) -(190/950) = 0.6
RR = (40/50)/(190/950) = 4
OR = (40*760)/(190*10) = 16
φ = 0.31

	М	Ж	
да	400	100	500
не	100	400	500
	500	500	1000

RD = (400/500) - (100/500) = 0.6
RR = (400/500)/(100/500) = 4
OR = (400*400)/(100*100) = 16
φ = 0.6

Тази особеност на φ-коефициента силно ограничава употребата му. Сравнения между силата на връзките, установени при две различни изследвания, може да се прави само ако разпределенията на четирите странични честоти са поне приблизително еднакви.

§ 6. Отношение на шансовете (odds ratio)

Първо нека разгледаме понятието odds. В този текст предлагаме превод на български „шанс“. Нека си представим, че играем на зар срещу партньор при следните условия. Печелим, ако при едно хвърляне се падне 6. Ако не се падне шестица, губим. Вероятността да спечелим е p = 1/6, вероятността да загубим е 5/6, или 1 - p. Вместо с две вероятности, играта може да се опише само с едно число, и това е тяхното отношение, p/(1 – p). С други думи, това е отношението на вероятността нещо да се случи към вероятността същото нещо да не се случи. За примера то е (1/6)/(5/6) = 1/5. Това отношение се нарича шанс, или odds. Казва се, че шансът да спечелим е 1 към 5. При честна игра, ако заложим левче на шестицата, и не се падне шестица, губим левчето. Но ако имаме късмет да се падне шестица, партньорът ни плаща 5 лева (като честно връща и левчето, което сме заложили). Така получаваме общо 6 лева, но печалбата ни е само 5 лева. Така шансът 1/5 ви информира какво губим и какво печелим при различните изходи от „игра на шестица“.

Нека си представим друга игра на зар, при която печелим, ако се падне някое от числата 1, 2, 3 или 4, и губим, ако се падне 5 или 6. Вероятността да спечелим е p = 4/6, а да загубим е 1 – p = 2/6. Шансът да спечелим е p/(1 – p) = (4/6)/(2/6) = 4/2 или 4 към 2. Абстрахираме се от сумите, които се залагат и изплащат. Очевидно е, че при втората игра шансът да спечелим е значително по-висок (4/2 = 2) в сравнение с шанса при предишната игра (1/5 = 0.2).

Шансът не е вероятност, това е число, което теоретически може да варира от нула до безкрайност. Числото информира за степента, в която даден изход е благоприятен; колкото е по-висок шансът, толкова е по-благоприятен изходът. Смислено е да определим колко пъти шансът да спечелим при втората игра е по-висок от шанса да спечелим при първата игра. За тази цел пресмятаме отношението на двата шанса, именно (4/2)/(1/5) = 2/0.2 = 10. Именно това е отношението на шансовете, или odds ratio, OR. То ни информира, че шансът да спечелим при втората игра е 10 пъти по-голям от шансът при първата игра.

Оказва се, че същият тип разсъждения са приложими в случая, когато трябва да опишем силата на връзката между дихотомни променливи. Нека отново разгледаме примера илюстриран на Табл. 1. Променливата „пол“ е с нива мъж и жена, а другата променлива е отговор „да“ или „не“ на зададения въпрос. Вероятността мъж да *е* отговорил с „да“ е

p(„да“|мъж) = a/(a + c)

Вероятността мъж да *не е* отговорил с „да“, т.е. да е отговорил с „не“, е

p(„не“/мъж) = c/(a + c)

Шансът мъж да е отговорил с „да“ е отношението на тези две вероятности,

p(„да“|мъж)/ p(„не“|мъж) = [a/(a + c)]/[c/(a + c)] = a/c

Така шансът мъж да отговори с „да“ се пресмята просто като отношението на честотите в клетките a и c. По същия начин пресмятаме шансът жена да отговори с „да“:

p(„да“|жена)/ p(„не“|жена) = [b/(b + d)]/[d/(b + d)] = b/d

Отношението на двата шанса, или odd ratio, OR, пресмятаме като

OR = (a/c)/(b/d) или
OR = (a*d)/(b*c) (7)

Виждаме силната прилика между равенства (6) и (7). При извеждането на равенство (6) определихме, че ако отношението (a*d)/(b*c) се получи равно на единица, това всъщност означава, че вероятността мъж да отговори с „да“ е равна на вероятността жена да отговори с „да“, и че няма връзка между двете променливи. Същото важи и за шансовете мъж да отговори с „да“ и жена да отговори с „да“. Ако отношението им OR = 1, тези шансове са равни, и връзка между променливите няма. Ако OR > 1, връзка между променливите има, тя се състои във факта, че шансът мъж да отговори с „да“ е по-висок. Ако OR < 1, шансът жена да отговори с „да“ е по-висок. Колкото повече OR се отклонява от единица, толкова по-силна е връзката между променливите. Така индексът OR е мярка за силата на връзката, и може да се използва за сравнения между данни от различни изследвания.

Предимствата на индекса OR са важни. В § 3 и § 4 показахме, че индексите „разлика между вероятностите“ RD и „отношение между рисковете” RR зависят от това, с кои вероятности се извършват пресмятанията. При пресмятането на OR това няма особено значение. От данните в Табл. 3 показахме, че „рискът“ да попаднем на лице без права за управление от жените е 4.7 пъти по-висок от рискът при мъжете, или RR = 4.7. Но вероятността мъж да шофира е само 1.687 пъти по-висока от вероятността жена да шофира (думата „риск“ не е подходяща за изхода „да може да шофира“). Това е объркващо. Нека сега пресметнем индекса OR за същите изходи.

Шансът мъж да шофира е a/b = 90/10 = 9.
Шансът жена да шофира е b/d = 40/35 = 1.142.
Отношението на двата шанса е OR = 9/1.142 = 7.87. Следователно шансът да намерим шофьор между мъжете е 7.87 пъти по-висок от шанса да намерим шофьор между жените. Нека пресметнем другия изход.
Шансът мъж да не шофира 10/90 = 0.111
Шансът жена да не шофира е 35/40 = 0.875, явно много по-голям от шанса при мъжете.
Отношението им е OR = 0.875/0.111 = 7.87, пак същото.

Може да дефинираме „шансовете“ по друг начин:
Шансът да попаднем на мъж измежду шофьорите е a/b = 90/40 = 2.25
Шансът да попаднем на мъж измежду не-шофьорите е c/d = 10/35 = 0.285
Отново OR = 2.25/0.285 = 7.875

Най-добре е да пресметнем OR просто като отношението (7) (наричано още и cross-product ratio)
OR = (a*d)/(b*c) = (90*35)/(40*10) = 7.875
С други думи, както и да смятаме всеки от шансовете, отношението им е винаги едно и също. Това е важно предимство на индекса OR. Проблем може да възникне, ако в една от четирите клетки на матрицата има твърде малък брой лица, или ако направо броят им е нула. В такъв случай отношението на шансовете също става нула, или става неопределено, ако нулата в клетката е в знаменателя на отношението. Въпросът се

решава с т.н. поправка за непрекъснатост, която се състои в това, към всяко от четирите числа в матрицата да се добави стойността 0.5. В литературата се препоръчва тази поправка с 0.5 да се прави винаги, независимо от това, дали има нула в матрицата или не. Това е най-простото решение, въпреки че се дискутират и други поправки (Haddock, Rindskopf &Shadish, 1998, Agresti, 1999).

При пресмятането на OR поставяхме по-високия шанс винаги в числителя на отношението им. Не е беда, ако разделим по-малкия шанс на по-големия. Ще получим реципрочната стойност 1/OR. При описания в текст или при презентация, трябва да се дава информация, как точно е съставено отношението на шансовете. Също така, при сравнение на данни от различни матрици, следва отношението на шансовете да се пресмята по един и същи начин.

Припомняме, че когато OR = 1, шансовете са еднакви и няма връзка между променливите. Стойността единица е „нулева точка“ при определяне на силата на връзката. Това не винаги е удобно. Изход от това положение е да се пресмята логаритъма на OR. По причини, които няма да дискутираме, се пресмята естествения логаритъм, при основа числото e = 2.718. В Excel това става с командата ln(число), но и някои джобни калкулатори имат същата функция. Например, имаме ln(7.875) = 2.064. Ако пресметнем ln(1/OR) = ln(1/7.875) = -2.064, т.е. същото число, но с отрицателен знак. Удобно е да използваме абсолютната стойност |ln(OR)| =|ln(1/OR)| като „абсолютна“ мярка за силата на връзката между променливите, без да се безпокоим относно кой шанс е бил в числителя и кой в знаменателя. Характерът на връзката следва да се описва с думи, и |ln(OR)| да показва силата ѝ. Логаритъм от единица е равен на нула, така че при |ln(OR)| = 0 връзка между променливите няма. Стойността нула представлява естествена нулева точка на индекса |ln(OR)|.

Впечатлението ни е, че в литературата се предпочита представянето на OR, а не на |ln(OR)|, като в числителя на равенство (7) се поставя по-високия шанс, така че винаги OR > 1. Но пресмятането на |ln(OR)| може да се окаже много полезно.

Съществува важна връзка между |ln(OR)| и индекса ES, или Cohen's d за оценка на сила на ефекта при разглеждане на две нормални разпределения от данни, измерени в интервална скала. Припомняме, че в най-простия му вид $ES = |m_1 - m_2|/s$, където m_1 и m_2 са двете средни стойности на разпределенията, а s е общото им стандартно отклонение. Установено е (виж Fleiss & Berlin, 2009), че е приблизително вярно равенството

$$ES = |\ln(OR)|*\pi/\sqrt{3} = |\ln(OR)|*3.14/1.73 = |\ln(OR)|*1.81 \quad (8)$$

съответно и

$$|\ln(OR)| = ES/1.81 \quad (9)$$

Следователно, с просто умножение на |ln(OR)| по 1.81 можем да сравняваме данни, описани с 2x2-матрица, с данни за ES, получени в интервална скала. Това е значително подобрява възможностите за количествени сравнения между резултати, получени от на пръв поглед съвсем различни изследвания.

Връзката между ES и |ln(OR)| може съществено да подпомогне интерпретацията на данни, описани с |ln(OR)|. На Cohen (1988) дължим класификацията на ES = 0.2, 0.5 и 0.8 като слаб, среден и силен ефекти. Като използваме равенство (9), с просто пресмятане можем да „преведем“ същата класификация като

слаб ефект - |ln(OR)| = 0.11
среден ефект - |ln(OR)| = 0.28
силен ефект - |ln(OR)| = 0.44

и след антилогаритмуване получаваме
слаб ефект - OR = 1.12
среден ефект - OR = 1.32
силен ефект - OR = 1.56

Въпреки че класификацията на Cohen е твърде ориентировъчна, тя може да подпомогне дипломанти и докторанти при оценката на силата на връзката, която се наблюдава в техните изследвания.

В литературата може да се проследи дискусия по въпроса, кой индекс е по-добър за описание на данни в 2x2-матрица, отношението на рисковете RR, или отношението на шансовете OR (Cumming, 2009, Sackett et al., 1996). Изтъква се факта, че RR е интуитивно по-разбираем индекс (Grimes & Schulz, 2008), особено от клиницисти или журналисти, които имат желанието да информират читателя за опасности, свързани с рисков начин на живот или рисково поведение. Разбирането на индекса може и да е проблем в областта на изследвания в медицината и епидемиологията. Мнението ни обаче е, че в областта на научните психологически изследвания колебания относно избора на индекс не трябва да има. Понятието „риск“ не е подходящо, или би звучало най-малкото странно, при описание на данни от експериментални изследвания. Що се отнася до медицинските изследвания, ще приведем забележката на Grrimes & Schulz (2008): “For most clinicians, odds ratios will remain… well, odd”. Изглежда, че проблемите на някои хора с разбирането на индекса OR не са свързани толкова със самия индекс, колкото с подготовката им.

Противниците на употребата на индекса OR признават едно негово основно предимство. Той може да се прилага при *всякакви* изследвания, проспективни и ретроспективни, и по-специално при т.н. case-control studies. Хубав пример за прилагането на OR може да се види в работата на Beautrais et al., (1996) Авторите са изследвали група от индивиди, които са предприемали сериозни (но неуспешни) опити за самоубийство. Регистрирано е предходното наличие на определени психични нарушения. Изследвана е и друга група от индивиди, които не са предприемали опит за самоубийство, като и при тях е регистрирано наличието на същите психични нарушения. Първата група са „случаите“ (cases), втората е контролната (control) група - типичен пример на case-control study. В рамките на размишленията може да се предположи, че наличието на психични нарушения е рисков фактор, който може да доведе до опит за самоубийство. Но, както отбелязахме в § 2, от такива данни риск, т.е.

вероятността лице с нарушения да направи опит за самоубийство, не може да се пресметне, и RR също не може да се пресметне. Така, чрез RR не може да се определи силата на връзката между дихотомните променливи „опит за самоубийство“ с нива „има“ и „няма“, и „психично нарушение“, с нива „има“ и „няма“. Отношението на шансовете OR обаче, може да се определи, и авторите именно по този начин демонстрират изключително силната и съществена връзка между двете променливи.

Шансовете винаги могат да се пресмятат и тяхното отношение OR да се сравнява с данни от други изследвания. Логаритмичната трансформация ln(OR) се използва при т.н. лог-линеен анализ. Той дава възможност за оценка на влиянието на различни странични променливи, които формират набор от няколко 2x2 матрици.

§ 7. Обобщение

Нека се върнем на мотивацията за написването на този текст. В увода цитирахме изисквания на редакторите на водещи списания, които силно препоръчват интерпретацията на данни да става въз основа на индекси за сила на ефекта, или на връзката между променливи. Именно тези индекси дават информация за съществеността на ефекта или връзката. Все повече специалисти се убеждават в това, че пресмятането на p-стойности и съответно на „значимости“ не носи информация за съществеността на получените резултати. Тук нямаме за задача да разгледаме всички доводи против използването на „значимостите“ при интерпретацията за съществеността на резултатите. Ще илюстрираме най-важния, според нас довод, именно че изводът за същественост зависи от обема на извадките.

При пресмятането на „значимостта“ на връзката между две дихотомни променливи обикновено се използва χ^2- критерия. Самата статистика се пресмята като $\chi^2 = \varphi^2 * N$ където N е общия брой на лицата в двете извадки. Ясно е, че при две изследвания с едни и същи стойности на φ-коефициента стойността на χ^2 ще е по-висока при изследването с по-голям брой лица, и съответно при него резултатът ще е „по-значим“. Но това съвсем не означава, че връзката между променливите при това изследване е по-съществена отколкото при другото.

Пресмятането на „ значимост“ не е безсмислено. То ни показва в какво степен може да вярваме, че ефектите, или зависимостите, които наблюдаваме в извадките, могат да се обобщят за цялата популация. Това не е маловажно. Но ако ефектът, получен от изследването на извадките е слаб и несъществен, големият брой лица и високата „значимост“, т.е. ниската стойност на p, не го правят по-съществен. Накратко, пресмятането на „значимост“ е смислено, но то не трябва да се използва за извод за същественост на връзката.

Някои читатели могат да кажат: „ами тогава защо да правим изследвания с голям брой лица, можем да използваме малък брой, и ако имаме късмет да получим силен ефект, показваме, че той е съществен“. Така е само на пръв поглед. Именно затова в изискванията на списанията за изследвания с добро качество фигурира и изискването за представяне на доверителен интервал. Въпросът за пресмятането на доверителните интервали го пропуснахме в текста по-горе, за да избегнем изтощителното разглеждане на формули. За разлика от индексите за сила на ефекта, при които броят на лицата в извадките не играе роля, размерът на доверителните интервали съществено се определя от обема на извадките. По-голям брой лица означава по-малък доверителен интервал, което пък от своя страна означава по-висока точност на определяне на силата на ефекта или на връзката в цялата популация. Високата точност прави изследването ценно, а не силният ефект, който се открива. Възможен е резултат, при който в извадката да се наблюдава отсъствие на връзка между променливите, например RD ≈ 0, но ако доверителният интервал на RD е малък, т.е. измерването е точно, то можем да вярваме, че и в популацията връзка няма. Ако доверителният интервал е голям, то твърдението, че в популацията връзка няма, не е обосновано. Връзка може и да има, но поради ниската точност на изследването, ние не можем да установим каква е и колко е тя. Пресмятането на доверителни интервали на индексите RD, RR и OR е представено в Допълнение В.

Читателят вероятно е забелязал предпочитанието към индекса OR, което недвусмислено изразихме при неговото описание. Данни, описани с този индекс, могат да се сравняват, когато са получени при всякакъв вид изследвания, проспективни и ретроспективни. Ако няма с какво да ги сравним, за охарактеризиране може да се използват реперите за слаб, среден и силен ефект. Препоръчваме този индекс на читателя.

При описание на данни от психологични изследвания, препоръчваме индексите RD и RR (или на аналогичното пресмятане на индексите с вероятности) да се използват с особена предпазливост, като се взимат пред вид техните недостатъци. Индексите са подходящи при обнародване на резултати в медиите, и въобще за аудитория от непрофесионалисти.

Индексът φ може да се използва при сравнения на данни, които отговарят на определени условия, т.е. данни, получени от 2x2 матрици с еднакви разпределения на страничните честоти. Дори и тогава, оценката на силата на връзката с φ може да е подвеждаща. Също така, пресмятането на доверителния интервал на φ е доста комплицирано.

Допълнение А: Правилото за независимост на две събития

Използваме дефиниция на независимост на две събития, чиито автор е останал неизвестен. Тя е следната: ако е изпълнено равенството

$$P(A|B) = P(A) \qquad (1A)$$

събитията А и В са независими. Дефиницията се чете по следния начин: ако вероятността $P(A|B)$ на събитието А при условие В, е равна на $P(A)$, т.е. на вероятността на А, това означава, че наличието или отсъствието на условието В не играе роля върху настъпването на А, или че А и В са независими.

Ако тези две вероятности не са равни, т.е. $P(A|B) \neq P(A)$, то А и В не са независими, между тях има връзка. Да разгледаме как от (1А) се получава правилото (1) в §1 на текста.

Нека дефинираме P(A) като вероятността, ако изберем случайно едно лице от групата, то да има права за управление. Тази вероятност се пресмята, като разделим броя на шофьорите (a + b) на общия брой на лицата в групата N, или тя е P(A) = (a + b)/N
Нека дефинираме условието B, и нека то да бъде лицето да е мъж. Вероятността, едно лице да има права за управление, *при условие, че то е мъж*, се пресмята като P(A|B) = a/(a + c). Това е вероятността, ако изберем случайно един от мъжете (условието B), да се окаже, че той има шофьорска книжка. По този начин, правилото (1A) придобива конкретния вид

$$a/(a + c) = (a + b)/N \qquad (2A)$$

Равенство (2A) може да се чете по следния начин: вероятността мъж да има шофьорска книжка е равна на вероятността кой да е от цялата група да има книжка. Ако равенството е изпълнено, то притежаването на шофьорска книжка няма връзка с това, дали лицето е мъж, или не.

Същото разсъждение може да се проведе, като пресметнем вероятността едно лице да има права за управление, *при условие, че е жена*. Тогава P(A|B) = b/(b + d). Следователно, ако

$$b/(b + d) = (a + b)/N \qquad (3A)$$

също можем да твърдим, че притежанието на шофьорска книжка няма връзка с това, дали лицето е жена или не е. Равенства (2A) и (3A) имат една и съща дясна страна. Те могат да се обединят като

$$a/(a + c) = b/(b + d)$$

което е и правилото (1) за независимост между двете променливи

Допълнение B. Доверителни интервали (ДИ) на индексите

При формулите за пресмятане са използвани символите и дефинициите на клетките в 2x2-матрицата на Табл.1.

B1. 95%-ен ДИ на RD, разлика между рискове или разлика между вероятности.

За подробното запознаване с въпроса за доверителния интервал на разлика между два относителни дяла препоръчваме Калинов (2013, стр. 275). Нека двете вероятности, или рискове, p_1 и p_2, са $p_1 = a/(a + c)$, $p_2 = b/(b + d)$ (Табл. 1).
95%-ен ДИ на разликата $|p_2 - p_1|$ се пресмята приблизително като
ДИ = $|p_2 - p_1| \pm 1.96*SE$
където SE е стандартната грешка, определена като

$$SE = \sqrt{P*Q*\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

където $P = (a + b)/N$, $Q = (1 - P)$ $n_1 = (a + c)$, $n_2 = (b + d)$

Числен пример с данните в Таблица 3:
p1 = 90/100 = 0.9
p2 = 40/75 = 0.533
p1-p2 = 0.367
P = (90 +40)/175 = 0.743
Q = (1 – 0.743) = 0.257
SE = $(0.743*0.257*(1/100+1/75))^{0.5}$ = 0.0668
Горната граница на ДИ е
0.367+1.96*0.0668 = 0.498
долната граница на ДИ е
0.367-1.96/0.0668 = 0.236
Вербалното твърдение е, че вероятността да се намери шофьор измежду мъжете е с 0.367 по-висока от вероятността да се намери шофьор измежду жените, като 95%-ният доверителен интервал на тази разлика е от 0.236 до 0.498.

B2. Доверителен интервал на относителния риск, RR

Нека двете вероятности, или рискове, p_1 и p_2, са p_1 =a/(a + c), p_2 = b/(b + d) (Табл. 1). ДИ на относителния риск RR = p_2/p_1 не може да се пресметне директно, но може да се пресметне след логаритмуването му, т.е. пресмята се първо ДИ на ln(RR). Тогава

ДИ = $\ln(p_2/p_1) \pm 1.96*SE$
където SE е стандартната грешка на ln(RR), определена (Fleiss & Berlin, 2009) като

$$SE = \sqrt{\frac{1-p_1}{n_1 * p_1} + \frac{1-p_2}{n_2 * p_2}}$$

За да се определят долната и горната граници на ДИ на самата стойност на RR следва границите на ln(RR) да се антилогаритмуват.

Числен пример с данните на Таблица 3:
p1 = 90/100 = 0.9
p2 = 40/75 = 0.533
RR = p1/p2 = 0.9/0.533 = 1.687
ln(RR) = ln(1.687) = 0.523
SE = $[(1-0.9)/(100*0.9)+(1-0.533)/(75*0.533)]^{0.5}$ = 0.113

Границите на 95%ния доверителен интервал на ln(RR) са
горна: 0.523 +1.96*0.113 = 0.744 и
долна: 0.523 – 1.96*0.113 = 0.403

След антилогаритмуване на тези граници получаваме границите на 95%-ния ДИ на самия RR:
горна: exp(0.744) = 2.105 и
долна: exp(0.403) = 1.496
Забележете, че след антилогаритмуването границите на ДИ не са симетрично разположени около стойността на RR = 1.687

Вербалното твърдение е, че вероятността да се намери шофьор измежду мъжете 1.687 пъти по-висока, отколкото измежду жените, като 95%-ният доверителен интервал на отношението е от 2.105 до 1.496.

В3. Доверителен интервал на отношението на шансовете, OR

Отново използваме данните в Таблица 3:
Шансът да открием шофьор между мъжете е a/c = 90/10 = 9
шансът да открием шофьор между жените е b/d = 40/35 = 1.423
Отношението им OR = (a/c)/(b/d) = 9/1.423 = 7.874
Същото OR се получава и ако пресметнем (a*c)/(b*d).
Доверителен интервал около тази стойност не може директно да се построи, но това е възможно около логаритъма на OR, ln(OR)= ln(7.874) = 2.063. Той е (Fleiss & Berlin, 2009) ln(OR) ± 1.96*SE, като

$$SE = \sqrt{1/a + 1/b + 1/c + 1/d}$$

За примера SE = $(1/90+1/40+1/10+1/35)^{0.5}$ = 0.405
горна граница на ln(OR): 2.063+1.96*0.405 = 2.858
долна граница на ln(OR): 2.063-1.96*0.405 = 1.267

За да се определят долната и горната граници на ДИ на самата стойност на OR, следва границите на ln(OR) да се антилогаритмуват. Те стават
горна граница: exp(2.858) = 17.44
долна граница: exp(1.267) = 3.55
Забележете, че след антилогаритмуването границите на ДИ не са симетрично разположени около стойността на OR = 7.974

Вербалното твърдение е, че шансът да се намери шофьор при мъжете е 7.874 пъти по-висок от шанса да се намери шофьор при жените, т.е. променливата „пол“ оказва много силен и съществен ефект върху променливата „права за кормуване“. 95%-ният доверителен интервал на отношението на шансовете е от 3.55 до 17.44.

Стандартната грешка SE, нужна за пресмятането на доверителния интервал на φ-индекса, се пресмята с изтощителна формула, която спестяваме в този текст. Читателят може да я намери в Fleiss & Berlin (2009), стр. 242.

ЛИТЕРАТУРА

Василев М (2014). Докладване на големина на ефекта и доверителни интервали: Преглед и начини за изчисляване. Psychological Thought, **7**, 37–54

Калинов К (2013). *Статистически методи в поведенческите и социалните науки*. Нов български университет, София (3-то издание)

Матеев С (2014). Същественост и значимост на резултати от научни изследвания в психологията. *Научен електронен архив на НБУ*, http://eprints.nbu.bg/2277/

Матеев С (2014а). Вероятност за превъзходство“ – лесно разбираем индекс за сила на ефекта при описание на данни. В: *Сборник научни доклади от VII-я Национален конгрес по психология, 2014* (ред. Джонев С, Димитров П & Матеева Н), Продуцентски център ЛМ ЕООД, София, стр. 337-346

Agresti A (1999). On logit confidence intervals for the odds ratio with small samples. Biomerics **55**, 597-602

Beautrais AL, Joyce PR, Mulder RT, David M, Fergusson DM, Deavoll BJ & Nightingale SK (1996). Prevalence and comorbidity of mental disorders in persons making serious suicide attempts: A case-control study. Am. J Psychiatry, **153**, 1009-1014

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum (2nd edition)

Cumming P (2009). The Relative Merits of Risk Ratios and Odds Ratios. Archives of Pediatric and Adolescent Medicine, **163**, 438-445

Fleiss JL & Berlin JA (2009). Effect sizes for dichotomous data. In: *Handbook of research synthesis and meta-analysis* (Edited by Cooper H, Hedges LV & Valentine JC), Russel Sage Foundation, New York (2nd edition), Chapter 13, 237-253

Grimes DA & Schulz KF (2008). Making sense of odds and odds ratios. Obstetrics & Gynecology, **111**, 423-426

Haddock CK, Rindskopf D & Shadish WR (1998). Using odds ratios as effect sizes for meta-analysis of dichotomous data: A primer on methods and issues. Psychological Methods, **3**, 339-353

Henson RK (2006). Effect-size measures and meta-analytic thinking in counseling psychology research. The Counseling Psychologist, **34**, 601-629

Kirk R (2003). The importance of effect magnitude. In: *Handbook of Research Methods in Experimental Psychology* (Edited by Davis, S.F.) Chapter 5, pp. 83 – 105, Blackwell Publishing Ltd

Kline RB (2004). *Beyond significance testing*. American Psychological Association, Washington

Rosnow RL & Rosenthal R (2003). Effect sizes for experimenting psychologists. Canadian Journal of Experimental Psychology, **57**, 221-237

Sackett DL, Deeks JJ & Altman DG (1996). Down with odds ratios! Evidence-Based Medicine, **1**, 164–6.