УДК: 314.48:519.235

Байесовский анализ показателей четырехпольных таблиц сопряженности

Обеснюк В. Ф.

Найдено новое выражение для эталонного распределения эпидемиологических оценок отношения рисков в четырехпольных методах исследования. При сопоставлении однородных групп распределение работоспособно в широких условиях при изучении как редких, так и частых событий. Инструментальная ценность формулы выше, прежде всего, в ситуациях с малым количеством наблюдаемых «случаев» или «контролей». С использованием нового распределения оценена, как очень высокая, достоверность наблюдения связи избыточной смертности работников ПО «Маяк» от рака легкого с радиационным облучением.

Ключевые слова: относительный риск, отношение шансов, интервал неопределенности, интервальное оценивание, эпидемиологический показатель.

Bayesian analysis of 2x2 contingency tables relative proportions

Obesnyuk V. F.

A new standard expression for distribution of epidemiological risk ratio estimates has been found out using two by two methods of study. Comparing homogeneous groups the distribution is operational within wide-ranging conditions studying both rare and frequent events. Support value of the distribution formula will increase primarily in cases with small amount of the cases or controls observed. Using the derived distribution excess mortality from lung cancer of the Mayak PA workers was associated with radiation exposure with high reliability of the observation.

Keywords: relative risk, odds ratio, uncertainty interval, interval estimation, epidemiological rate.

Введение

Выборочный метод при сравнительном изучении широко применяется в биологии, медицине, социологии и других науках для оценки рисков, оценки эффективности воздействия, оценке эффективности тестов. Результаты простейшего исследования по методу «case−control» [1−3], либо «case−base» [3, 5], записываются в так называемую четырехпольную таблицу сопряженности (contingency epidemiological table), которая позволяет выполнить оценку показателей отношения шансов (OR, odds ratio) и отношения правдоподобий (LR, likelihood ratio) при группировке по событиям; либо возможна оценка отношения шансов и относительного риска (RR, relative risk) при группировке по исследуемому признаку. Так как всякая выборка является случайной даже при работе с некоторыми однородными большими группами, эти показатели будут случайными, обладающими некоторой неопределенностью, которая дополнительно маскируется дискретным характером наблюдений (шум дискретизации). Однако вероятность наблюдения эффекта или относительного эффекта непрерывна, поэтому неопределенность величин OR или RR должна получить непрерывную оценку.

Материалы и методы

Разрешить противоречие между дискретностью наблюдений и непрерывностью показателей OR и RR можно лишь приближенно, выбирая частотный аппроксимационный метод [1−3, 6] и, соответственно, доверительное (confidence) оценивание, либо байесовский подход [7] с использованием некоторой доступной априорной информации о возможных распределениях выборочных и измеряемых величин (credible estimation). Заметим, что в данном случае частотный метод не имеет никаких преимуществ перед байесовским, так как тоже опирается на априорное предположение о существовании постоянных, но неизвестных популяционных частот в каждой из сравниваемых выборок; вместо них при построении областей неопределенности фактически используются центральные (точечные) оценки, не имеющие даже однозначного толкования для однократных выборок. Каждая из таких задач в конкретном исследовании всегда может быть решена численно, благодаря чему те или иные методики оценки нашли реализацию в ряде известных приближенных компьютерных алгоритмов, например, в открытой среде программирования «R» [8], пакете «LePAC» [9] и др. Как оказалось, байесовский подход также допускает строгое аналитическое представление непрерывной плотности распределения величин OR и RR, а значит, возможность получения сравнительно точных оценок, как для интервалов неопределенности и их центров, так и для связанных с этими распределениями вторичных характеристик (показатель согласованности «каппа Коуэна» и т.п.) при самых общих (неинформативных) априорных предположениях в рамках биномиального описания распределения дискретных событий в каждой независимой выборке.

Например, пусть при сравнении двух случайных выборок с объемами N (экспонированные) и L (интактные) наблюдалось M и K специфических «случаев», соответственно. Тогда при априорном (до выборки) предположении, что каждый индивидуум одной выборки имеет одинаковые шансы попасть в свою группу «случаев» этой же выборки, возможно строгое интегральное представление плотности вероятности наблюдения неизвестных показателей *OR* и *RR*, которое может быть выражено через известные специальные функции. Априорное предположение о наличии равных шансов соответствует одинаковой вероятности наблюдения целочисленных M среди M=0; 1; … N и K среди K=0; 1; … L, а также — равномерному распределению ожидаемых частот непрерывных параметров доли «случаев» в каждой из двух выборок до их формирования (группирование событий по методу «case-base»).

Таким предположениям соответствуют условные бета-распределения частот наблюдения специфических событий

$$\psi(\beta | M, N) = \beta^M (1-\beta)^{N-M} / B(M+1, N-M+1);$$

$$\varphi(\alpha | K, L) = \alpha^K (1-\alpha)^{L-K} / B(K+1, L-K+1)$$

в выборках, где

$$B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$$

— бета-функция Эйлера, определенная через гамма-функцию. Отношение *RR* абсолютных выборочных рисков само по себе риском не является, но в качестве отношения независимых случайных величин β и α имеет распределение, подчиняющееся законам преобразования распределений:

$$g(RR | M, N, K, L) = \int_0^{\min(1;\, 1/RR)} \alpha\, \psi(RR\, \alpha | M, N)\, \varphi(\alpha | K, L)\, d\alpha \,.$$

Таким образом, апостериорная плотность вероятности основного показателя сравнительного исследования — относительного риска — может быть вычислена точно (в рамках сделанных априорных предположений). В частности, можно показать, что

$$g(RR\,|\,M,N,K,L)=\begin{cases}\dfrac{B(M+K+2,\,L-K+1)\cdot RR^{M}}{B(M+1,\,N-M+1)\cdot B(K+1,\,L-K+1)}F(M-N,\,K+M+2;\,M+L+3;\,RR), & RR<1;\\[2ex] \dfrac{B(M+K+2,\,L+N-K-M+1)}{B(M+1,\,N-M+1)\cdot B(K+1,\,L-K+1)}, & RR=1;\\[2ex] \dfrac{B(M+K+2,\,N-M+1)}{B(M+1,\,N-M+1)\cdot B(K+1,\,L-K+1)\cdot RR^{K+2}}F\left(K-L,\,K+M+2;\,K+N+3;\,\dfrac{1}{RR}\right), & RR>1.\end{cases}$$

Здесь однократно наблюдаемые эмпирические значения *M, N, K, L* выступают в роли параметров распределения; *F(a, b; c; z)* — гипергеометрическая функция Гаусса. В области малых значений относительного риска плотность нарастает пропорционально $\sim RR^{M}$, а в области больших значений — убывает пропорционально $\sim 1/RR^{K+2}$. Иными словами, асимптотическое поведение распределения зависит в большей мере от наблюдаемого числа «случаев», чем от объемов выборок. Обнаруженный характер асимптотик противоречит ранее найденным приближениям [1−3, 6], имевшим приемлемую пригодность только для описания центральной части распределения.

Найденное распределение допускает точное вычисление математических ожиданий и дисперсий показателя относительного риска, а также его логарифма в предположении независимости выборок:

$$M\{RR\}=\frac{(M+1)\cdot(L+1)}{(N+2)\cdot K};\qquad D\{RR\}=\left(M\{RR\}\right)^2\left[\frac{(M+2)\cdot(N+2)\cdot K\cdot L}{(M+1)\cdot(N+3)\cdot(K-1)\cdot(L+1)}-1\right];$$

$$M\{\ln(RR)\}=\sum_{i=K+1}^{L+1}1/i-\sum_{j=M+1}^{N+1}1/j\;;\qquad D\{\ln(RR)\}=\sum_{i=K+1}^{L+1}1/i^2+\sum_{j=M+1}^{N+1}1/j^2\,.$$

Последние две величины могут быть использованы при записи логнормального приближения при описании распределения величины относительного риска в качестве аппроксимации вблизи своего центра.

По аналогии можно показать, что отношению шансов соответствует величина $(\beta\cdot(1-\alpha)/(\alpha\cdot(1-\beta))$. Для нее условное распределение плотности вероятности дается выражением

$$f(OR\,|\,M,N,K,L)=\frac{B(M+K+2,\,N-M+L-K+2)\cdot OR^{M}}{B(M+1,\,N-M+1)\cdot B(K+1,\,L-K+1)}F(N+2,\,K+M+2;\,N+L+4;\,1-OR),\;\; OR>0.$$

Однако оно шире распределения отношения рисков, и в области сравнительно частых специфических событий имеет хорошо различимое систематическое смещение относительно распределения RR. Таким образом, оценка OR, используемая в качестве приближения к искомой величине показателя RR во многих исследованиях, менее предпочтительна для применения. Замечено, что в ряде вычислительных пакетов программ гипергеометрическая функция, представленная в виде ряда, неверно табулируется в области отрицательного аргумента, поэтому для организации вычислений за основу было принято ее интегральное представление. Примеры наблюдаемой формы распределений $g(RR|\ M, N, K, L)$ и $f(OR|\ M, N, K, L)$ приведены на рисунке, из которого следует, что в центральной части они напоминают логнормальные распределения, но отклоняются от них при очень малых и очень больших значениях показателей. В качестве центра распределений более уместно использовать медиану или среднегеометрическое значение.

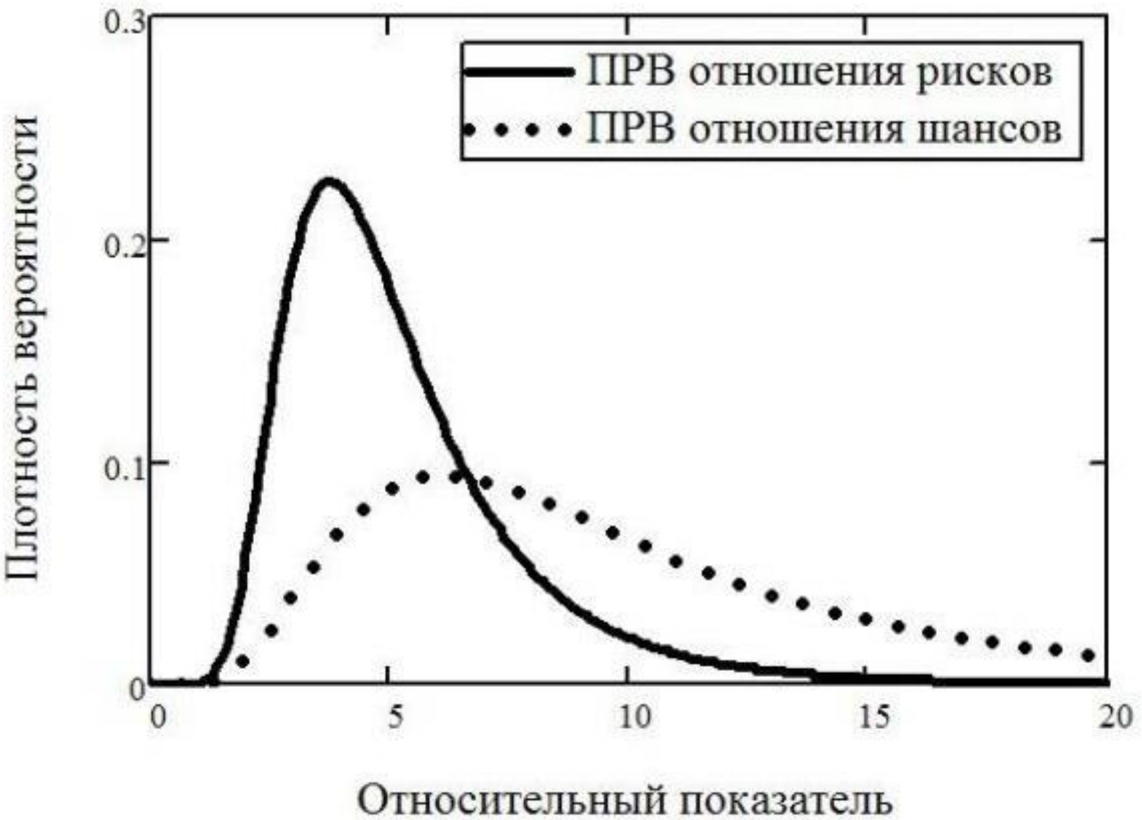

Рисунок 1 — Пример плотности распределения вероятности наблюдения относительных оценок RR и OR при сравнении двух случайных выборок. Интактные: 5 “случаев” из 50. Экспонированные: 13 “случаев” из 25

Результаты и обсуждение

Сопоставление байесовских интервальных оценок отношения рисков с частотными показало, что байесовский подход дает более узкие интервалы неопределенности для RR при малом количестве наблюдаемых случаев M и K, что представляется понятным в силу использования дополнительной информации об априорных распределениях, а также в силу большей математической строгости. Для исследований с «богатой» статистикой, когда N, L>100; M, K>5 и ($N-M$), ($L-K$) > 5, частотные и байесовские оценки практически совпадают. Сопоставление графиков распределения их отношения (с использованием непрерывной аппроксимации Анскомб [6]) для случая более редких наблюдений показано на рисунке 2 и рисунке 3.

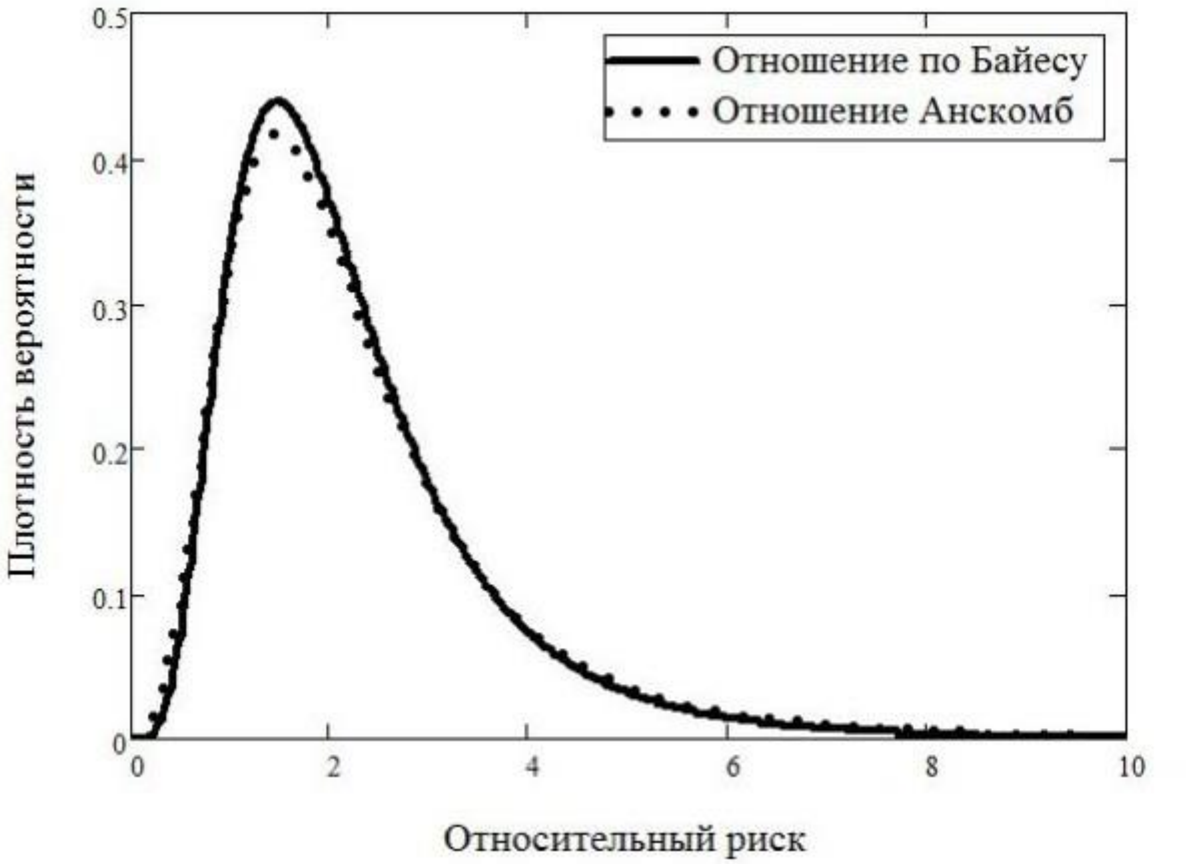

Рисунок 2 — Пример сопоставления распределений оценок относительного риска, полученных по байесовской и частотной методологии. Интактные: 5 “случаев” из 50. Экспонированные: 13 “случаев” из 25

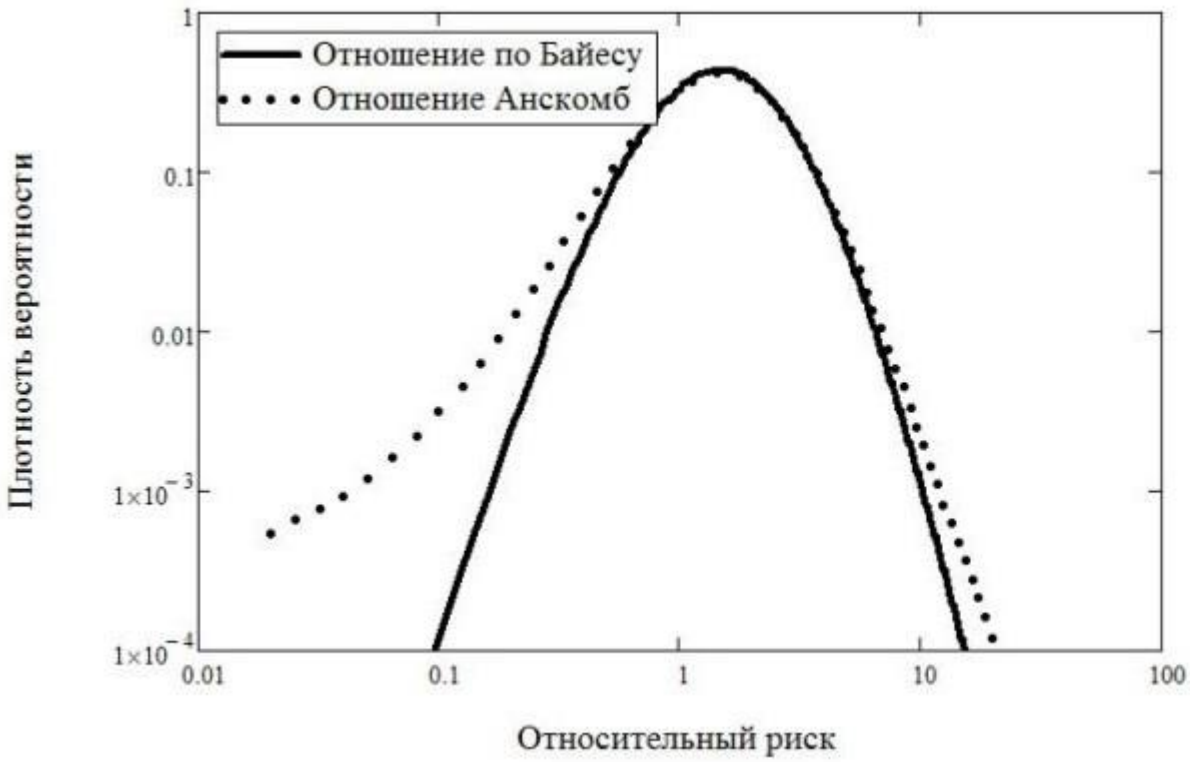

Рисунок 3 — То же, что и на рисунке 2, но в двойных логарифмических координатах.

Условные распределения $g(RR|\ M, N, K, L)$ и $f(OR|\ M, N, K, L)$ не могут быть асимптотически правильно описаны общепринятым логнормальным приближением. Тем не менее, вид функции $f(OR|\ M, N, K, L)$ частично оправдывает практику применения двухпараметрического логнормального приближения в качестве аппроксимации центральной части распределения Альбомом и Норрелом ([1], стр.87—89; оценка логит-границ Вульфа). Логнормальное приближение и найденные распределения относительного риска существенно различаются по поведению на «хвостах», что не позволяет производить интервальное оценивание со слишком большой величиной доверительной вероятности, пользуясь границами Вульфа. Кроме того, логнормальное приближение Вульфа дает визуальное отклонение в описании распределений по сравнению с найденным строгим представлением в исследованиях с низкой статистической мощностью.

Серьёзным преимуществом байесовского оценивания отношения рисков и возможности построения распределения по однократному наблюдению является возможность непосредственного описания альтернатив исследуемой гипотезы через распределение её показателей в сопоставлении с граничными величинами, свойственными нулевой гипотезе. Это открывает возможность прямого вычисления величины наблюдаемых ошибок II рода. Кроме того, изучение распределения, характеризующего нулевую гипотезу H_0 об отсутствии связи между маркерами и эффектами вместе с границами неопределенности показателя, оцененного для гипотезы H_1, позволяет также вычислить величину ошибки I рода с точностью, превышающей традиционное использование асимптотического критерия хи-квадрат с одной степенью свободы по единственному наблюдению, заведомо не подчиняющемуся закону хи-квадрат. Расчет плотности вероятности, характеризующей нулевую гипотезу, технически не сложнее, чем для гипотезы H_1. Для этого достаточно рассчитать и заполнить четырехпольную таблицу «ожидаемых» наблюдений, характерных для ситуации с отсутствием связи между изучаемыми маркерами и эффектами. Далее эти значения используются уже в качестве параметров «нулевого» байесовского распределения исследуемого показателя, по которому необходимо найти границы интервала неопределенности. В случае перекрытия обоих интервалов для гипотез H_0 и H_1 может возникать некоторая «серая зона неопределенности» изучаемого относительного показателя риска, для которой вероятность ошибок может превосходить использованный допустимый уровень значимости. Если же интервалы не перекрываются, «серая зона» отсутствует, и можно говорить об обоснованной достоверности изучаемой связи между маркером и эффектом. Ожидаемая таблица может быть рассчитана тем же способом, который обычно применяется для одновыборочного четырехпольного исследования в предположении статистической независимости наблюдений маркера и эффекта [2, 3]. Оказывается, результат расчета совпадает с таблицей, которая получилась бы в предположении, что связь между изучаемым маркером и эффектом практически полностью теряется при механическом смешивании двух сравниваемых выборок. Центр получающегося распределения для H_0 всегда лежит вблизи RR=1. Описанный способ оценки связи в двухвыборочном исследовании является консервативным в том смысле, что незначимость теста не будет означать отсутствия связи, в то время как значимость, напротив, будет являться указанием на её наличие.

Пример применения

В качестве примера применения байесовского подхода к интервальным оценкам основных показателей, а также к оценке их значимости, рассмотрим результаты исследования относительного пожизненного риска смерти от рака легкого в двух взаимоисключающих группах мужчин — работников химического комбината «МАЯК». Одна из групп характеризуется наличием документированных техногенных ненулевых эквивалентных доз на легкие от смешанного (альфа + гамма) облучения за 15 лет до момента наступления смерти. В другой группе облучение отсутствовало совсем (органные дозы меньше предела детектирования), либо оно произошло за меньший интервал времени до момента смерти работника (условно необлученного). В каждой из исследуемых групп часть работников умерла от рака легкого. Сводные результаты заимствованы из опубликованного описания когорты работников [4] и представлены в таблице 1. Требуется установить связь повышенного риска смерти от рака легкого с облучением ионизирующей радиацией, а также оценить достоверность выводов.

Таблица 1 — Сопряженность эффекта с облучением при наличии связи

		Рак легкого		
		Есть	Нет	∑
Маркер	Облученные	590	5718	6308
	Лагированна-янулевая доза	77	2721	2798

Этой таблице, построенной в соответствии с гипотезой о наличии связи между облучением и выходом рака легкого, соответствует таблица ожиданий (Таблица 2), полученная в предположении об отсутствии такой связи (гипотеза H_0).

Таблица 2 — Ожидаемая смертность от рака легкого при отсутствии связи

		Рак легкого		
		Есть	Нет	∑
Маркер	Облученные	462	5846	6308
	Лагированна-янулевая доза	205	2593	2798

Байесовские распределения показателя относительного риска для обеих гипотез приведены на рисунке 4. Им соответствуют следующие оценки показателя: в рамках гипотезы H_1 — RR=3,38 (95 % CI: 2,68…4,27); в рамках гипотезы H_0 — RR=1,00 (95 % CI: 0,85…1,17). За центральную оценку принята среднегеометрическая величина. Интервалы неопределенности не пересекаются, что указывает на значительно меньшую величину действительных ошибок I и II рода по сравнению с априорно установленным уровнем значимости 5 %. Это обстоятельство хорошо заметно из рисунке 4, что с огромным запасом указывает на высокую степень достоверности вывода о связи техногенного облучения с повышенным выходом рака легкого среди работников комбината. Заметим, что речь идет не только о высоких уровнях облучения, но и о дозах, близких к пределу детектирования, то есть о столь актуальных «малых дозах». Отсюда также ясно, что в облученной группе достоверно может быть оценена атрибутивная доля избыточных (техногенных) раков легкого: $AR=(RR-1)/RR=0{,}704$ (95 % CI: 0,627…0,766), что соответствует примерно 370…452 случаям. При этом «фоновая» специфическая смертность во всей когорте должна была бы наблюдаться на уровне ~250 событий. Важно указать, что линейная беспороговая теория, построенная на идее использования годовых показателей смертности, при средней эквивалентной дозе около 1 Зв [4] и уровне избыточной смертности от рака легкого около 20–40 %·Зв$^{-1}$ [10] приводит к ожиданию приблизительно 50—100 избыточных событий в когорте ПО «Маяк». Линейный участок зависимости доза—эффект действительно регистрируется в области свыше 0,3 Зв [4]. Это позволяет сделать вывод о том, что от 270 до 352 случаев смерти от рака легкого связано с облучением малыми дозами, где когортная кривая «доза—эффект» надлинейна. Иными словами, в когорте ПО «Маяк» наблюдалась значительная (в абсолютном исчислении) группа радиочувствительных лиц с повышенным уровнем риска на единицу дозы, которая погибла в первую очередь. Этот вывод хорошо согласуется с результатами ранее выполненного исследования, проведенного другим методом [4]. От общего объема когорты радиочувствительная группа составляет около 3,4 %. Такая величина оказывается сопоставима с долей лиц, умирающих от рака легкого без техногенного воздействия. В связи с наличием столь многочисленной радиочувствительной группы актуальность раннего выявления ее представителей в процессе профессионального отбора становится очень высокой, потому что такая группа фактически не охватывается действующими управленческими механизмами ограничения профессионального радиационного риска. Статистическая достоверность обоих исследований хорошо согласуется друг с другом.

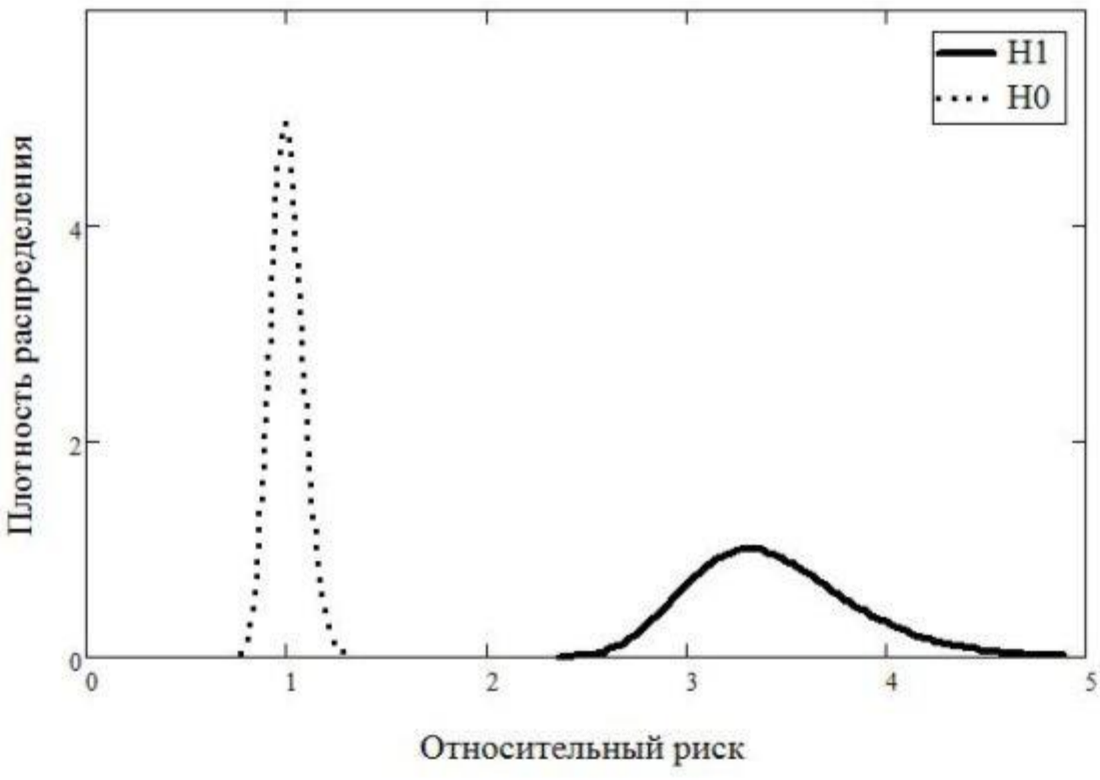

Рисунок 4 — Распределения оценок относительного риска для гипотез H_0 *и* H_1

Выводы

Найдено явное непрерывное выражение для распределения эпидемиологических оценок отношения рисков в четырехпольных методах исследования, которое можно считать точным при довольно общих априорных предположениях. Новое распределение может считаться эталонным при выполнении интервального оценивания. При сопоставлении однородных групп распределение работоспособно в широких условиях при изучении как редких, так и частых событий. Рекомендуемые интервальные оценки несколько отличаются от логит-интервалов Вульфа. Найденное распределение может быть использовано в современном программном обеспечении эпидемиологических, медико-статистических и биологических исследований. Инструментальная ценность формулы распределения будет выше, прежде всего, в ситуациях с малым количеством наблюдаемых «случаев» или «контролей». Корректное асимптотическое поведение найденного распределения позволяет производить прямые оценки величины ошибок I и II рода. Это позволило установить ошибочность традиционной процедуры оценивания величины p-value по критерию хи-квадрат с одной степенью свободы. С использованием нового распределения оценена, как очень высокая, достоверность наблюдения связи избыточной смертности работников ПО «Маяк» от рака легкого с радиационным облучением.

Литература

1. *Альбом, А., Норелл, С.* Введение в современную эпидемиологию // Таллинн: RNE, 1996. — 122 с.
2. *Гланц, С.* Медико-биологическая статистика // Пер. с англ. — М.: Практика, 1998. — 459 с.
3. Эпидемиологический словарь // Под ред. Дж. М. Ласта для Международной эпидемиологической ассоциации. — М.: ОИЗ, 2009. — 316 с.
4. *Обеснюк, В.Ф.* Прямые кумулятивные оценки дозового тренда пожизненного радиогенного риска смерти от рака легкого в когорте ПО «Маяк». // В сборнике "Источники и эффекты облучения работников ПО «Маяк» и населения, проживающего в зоне влияния предприятия. *Часть III* "/ ФМБА, ЮУрИБФ. — Челябинск: ЧДП, 2011. — С.98—117.
5. *Kupper, I.L., McMichael, A.J., Spirtas, R.* A Hybrid epidemiologic study design useful in estimating relative risk // Journal of the American Statistical Association. — 1975. — V.70. — P.524—528.
6. *Anscombe, F.J.* Transformations of Poisson, binomial and negative binomial data. // Biometrika, 1948, 35, P.246—254.
7. *Hey, J.D.* Data in doubt: An Introduction to Bayesian Statistical Inference for Economists // Oxford: Martin Robertson, 1983. — 316 p.
8. *Agresti, A.* URL: http://www.stat.ufl.edu/~aa/cda/R/bayes/index.html, дата обращения 07.02.2013.

9. *Lecoutre, B., Poitevineau, J.* LePAC version 2.0.41 (Program for the Analysis of Comparisons) URL: http://www.univ-rouen.fr/LMRS/Persopage/Lecoutre/PAC.htm, дата обращения 07.02.2013.

10.Effects of Ionizing Radiation. UNSCEAR 2006 Report., Vol. 1A // NY: United Nations Publication, 2008. — 383 P.