УДК 573 + 615.015 + 512.563

СОВМЕСТНОЕ ДЕЙСТВИЕ БИНАРНЫХ ФАКТОРОВ И РАВЕНСТВО РИСКОВ

Ю. В. Нагребецкая, А. И. Фаткуллина

Уральский Федеральный университет им. первого Президента России Б. Н. Ельцина, г. Екатеринбург, Россия

В работе исследуется широко применяемые в медико-биологических науках модели взаимодействия двух и трех независимых бинарных факторов в случае бинарного отклика. Основное внимание уделено формализации понятия «взаимодействие» и его связи с условием равенства разностей рисков. Показано, что это условие можно получить из приравнивания нулю всех членов взаимодействия в их представлении, применяемом в дисперсионном анализе. Получена и решена система равенства рисков для двух и трех факторов, а также показана ее инвариантность относительно действия группы симметрий эксперимента. Приведена интерпретация решения системы равенства рисков: в случае двух факторов эффект от их воздействия равен сумме эффектов от изолированного воздействия каждого фактора в отдельности; в случае трех факторов (а) эффект от воздействия всех этих факторов равен сумме эффектов от изолированного воздействия каждого фактора в отдельности; (б) эффект от воздействия любых двух факторов равен сумме эффектов от изолированного воздействия каждого фактора в отдельности при отсутствии воздействия третьего. Кроме того, проведено исследование решения равенства рисков для двух и трех факторов в рамках детерминистической модели, приводящей к булевой формализации этой модели. Доказано, что для значений булевой функции, представляющей данный отклик, система равенства рисков и для двух, и для трех факторов выполняется тогда и только тогда, когда эта булева функция существенно зависит не более чем от одной переменной.

Показано, что в рамках детерминистической модели понятие «совместное действие» бинарных факторов, введенное и исследованное в работах первого автора, полнее описывает взаимодействие этих факторов, чем нарушение условия равенства разностей рисков.

Ключевые слова: условие равенства разностей рисков; детерминистическая модель; бинарная теория достаточных причин; отклик; совместное действие бинарных факторов; булева функция; группа автоморфизмов; действие группы.

1. Введение

Проблема описания взаимодействия составляет одну из важнейших задач медико-биологических наук. Помимо теоретического интереса, эта задача имеет и очевидный экономический аспект. Например, если известно, что изолированное присутствие некоторых агентов приводит к увеличению частоты некоторого заболевания (для простоты будем считать, что одного и того же заболевания для всех агентов), то может оказаться, что совместное воздействие этих агентов имеет другую распространенность этого заболевания, чем та, которую исследователь ожидает на основании информации об изолированных эффектах. Таким образом, в зависимости от того, больше или меньше фактическая заболеваемость, чем ожидаемая, ресурсы, необходимые для купирования ситуации многофакторного воздействия, будут больше или меньше прогнозируемых на основании однофакторных эффектов.

Эта проблема имеет значительную специфику при рассмотрении факторов, имеющих небольшое число уровней, например, один-два уровня воздействия. Такая дискретная ситуация не дает возможности применить непрерывные аппроксимации даже в качестве приближения. Поэтому необходимо рассматривать действующие факторы так, как они есть (без аппроксимаций) и вводить характеристики их воздействия на основе этих точных данных.

Один из первых вариантов численного описания особенностей совместного действия двухуровневых факторов и такого же отклика был предложен в работе [1] на основе представления причинного механизма воздействия в рамках так называемой теории достаточных причин (в работе [1] для нее предложено название sufficient-component cause model). Апеллируя непосредственно к эпидемиологической ситуации, автор вводит коэффициент ICDR (Interaction Contrast of Disease Rates), который можно понимать как комбинацию разностей рисков действующих факторов при различных сочетаниях их уровней. Здесь важно отметить, что вид этой комбинации таков, что позволяет трактовать смысл коэффициента ICDR непосредственно. Иначе говоря, арифметическая структура этого коэффициента определяется его интерпретацией как коэффициента взаимодействия. Это делает затруднительным распространение коэффициента ICDR за пределы двухфакторного двухуровневого случая.

Ниже предложена формальная процедура получения коэффициента Купмана для двух двухуровневых факторов и двухуровневого отклика, основанная на идеях дисперсионного анализа. Она может быть легко распространена на большее число двухуровневых факторов, при этом, однако, вместо одного уравнения, выражающего отсутствие взаимодействия, таких уравнений будет несколько, так что они образуют систему уравнений.

2. Система равенства рисков для двух факторов

Основное наблюдение, которое позволяет рассматривать случай двухуровневых факторов как представляющий самостоятельный интерес, состоит в следующем. Рассмотрим два фактора (две независимые переменные) A, B, принимающих конечное число целых значений (уровней), а также результирующую (зависимую) переменную Y, вероятностная природа которой несущественна. Для каждой ячейки (i, j) обозначим теоретическое среднее значение отклика Y через η_{ij}. Имеющиеся данные позволяют провести анализ данных методами (двухфакторного) дисперсионного анализа [2–4] следующим образом. Пусть среднее значение отклика в ячейке (i, j) представлено в виде

$$\eta_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij},$$

где μ – *среднее значение* отклика во всей выборке,

α_i – *главный эффект* i-го уровня фактора A,

β_j – *главный эффект* j-го уровня фактора B,

γ_{ij} – *член взаимодействия* i-го уровня фактора A и j-го уровня фактора B.

Главные эффекты α_i и β_j, член взаимодействия γ_{ij} и среднее значение отклика μ можно представить в виде [4]:

$$\alpha_i = \eta_{i\bullet} - \mu,\ \beta_j = \eta_{\bullet j} - \mu,$$
$$\gamma_{ij} = \eta_{ij} - \alpha_i - \beta_j - \mu,\ \mu = \eta_{\bullet\bullet},$$

где точка означает усреднение значений η_{ij} по соответствующему индексу. Тогда

$$\gamma_{ij} = \eta_{ij} - \eta_{i\bullet} - \eta_{\bullet j} + \eta_{\bullet\bullet} \tag{1}$$

Далее мы можем рассмотреть полностью бинарную ситуацию, в которой факторы и отклик являются двухуровневыми переменными, уровни которых будут кодироваться числами 0 и 1. При этом приведенные выше соотношения между введенными выше величинами будут сохраняться. Действительно, нигде в приведенных выше определениях не требовалось, чтобы результирующая переменная имела какую-то определенную природу. Таким образом, если действующие факторы и отклик являются бинарными переменными, то взаимодействие между уровнями $A = i$, $B = j$, $i, j \in \{0, 1\}$ данных факторов в отклике также представляется равенством (1).

Важная особенность двухуровневой случайной переменной X, принимающей значение 0, 1 (бернуллиева случайная величина), состоит в том, что для нее выполняется равенство $E[X] = P(X = 1) = p$, где $E[...]$ – оператор математического ожидания.

В рассматриваемой полностью бинарной ситуации введем обозначения

$$p_{ij} = P(Y = 1 \mid A = i, B = j),\ i, j \in \mathbb{B},$$

где $\mathbb{B} = \{0, 1\}$.

Как обычно, в теоретическом анализе рассматриваются не выборочные характеристики, а их «истинные» значения, т. е. вместо усредненных величин берутся их математические ожидания, а вместо наблюдаемых частот появления события – вероятность их появления. Таким образом, условие отсутствия взаимодействия в данном случае имеет вид

$$\gamma_{ij} = p_{ij} - p_{i\bullet} - p_{\bullet j} + p_{\bullet\bullet} = 0,\ i, j \in \mathbb{B}$$

Это равенство хорошо известно в эпидемиологии [5, 6] как условие равенства разностей рисков, или условие отсутствия эффекта от воздействия двух данных факторов. Трактовка этого равенства основана на представлении его в виде

$$p_{ij} - p_{\bullet\bullet} = (p_{i\bullet} - p_{\bullet\bullet}) - (p_{\bullet j} - p_{\bullet\bullet})$$

с последующим толкованием каждой разности $p_{ij} - p_{\bullet\bullet}$, $p_{i\bullet} - p_{\bullet\bullet}$ и $p_{\bullet j} - p_{\bullet\bullet}$.

Решим систему равенства рисков для двух факторов, принимая следующие обозначения: $p_{i\bullet} = \frac{1}{2}(p_{i0} + p_{i1})$, $p_{\bullet j} = \frac{1}{2}(p_{0j} + p_{1j})$, $p_{\bullet\bullet} = \frac{1}{4}(p_{00} + p_{01} + p_{01} + p_{11})$

Теорема 1. Система из четырех уравнений (система равенства рисков)

$$p_{ij} - p_{i\bullet} - p_{\bullet j} + p_{\bullet\bullet} = 0,\ i, j \in \mathbb{B} \tag{2}$$

равносильна единственному уравнению

$$p_{00} - p_{01} - p_{10} + p_{11} = 0 \qquad (3)$$

Доказательство проводится непосредственными вычислениями.

Равенство (3) и представляет собой выражение коэффициента ICDR Купмана из работы [1].

Определение 1. Пусть значение уровня фактора *A* (фактора *B*) равно 1 при воздействии фактора *A* (фактора *B*) и равно 0 при отсутствии воздействия этого фактора. И пусть индексы i, $j \in \mathbb{B}$ не равны нулю одновременно. Величину $\Delta_{ij} = p_{ij} - p_{00}$ назовем *теоретическим эффектом* от воздействия фактора *A* (при отсутствии воздействия фактора *A*), если $i = 1$ ($i = 0$), и воздействия фактора *B* (при отсутствии воздействия фактора *B*), если $j = 1$ ($j = 0$). Разумеется, удобно предполагать, что $\Delta_{00} = 0$.

Пример 1. Δ_{10} – теоретический эффект от воздействия фактора *A* при отсутствии воздействия фактора *B*, Δ_{01} – теоретический эффект от воздействия фактора *B* при отсутствии воздействия фактора *A*, Δ_{11} – теоретический эффект от воздействия факторов *A* и *B* одновременно.

Везде ниже, говоря об эффектах воздействия или отстутствии воздействия факторов, мы будем иметь в виду теоретические эффекты.

Следствие 1. Система (2) *равенства рисков для двух факторов* равносильна уравнению

$$\Delta_{11} = \Delta_{01} + \Delta_{10} \qquad (4)$$

Равенство (4) можно интепретировать следующим образом: эффект от воздействия двух факторов равен сумме эффектов от изолированного воздействия каждого в отдельности при условии, что значение уровня каждого фактора равно 1 при воздействии этого фактора и равно 0 при отсутствии его воздействия.

3. Система равенства рисков для трех факторов

Пусть *A*, *B*, *C* – двухуровневые факторы. Тогда имеется 8 ячеек со значением отклика (случайной величины с двумя уровнями 0,1) при данном наборе факторов (0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1). Математическое ожидание (теоретическое среднее значение) отклика в этих ячейках обозначим через p_{ijk}, $i, j, k \in \mathbb{B}$. Поскольку отклик бинарный и принимает значение из $\mathbb{B}$, имеем

$$p_{ijk} = P(Y = 1 \mid A = i, B = j, C = k), \; i, j, k \in \mathbb{B}$$

Тогда возможно 12 взаимодействий двух факторов при всевозможных значениях уровней этих факторов при фиксированном значении третьего фактора и 8 взаимодействий трех факторов при всевозможных значениях уровней этих факторов (всего $12 + 8 = 20$ взаимодействий). Их характеризуют величины γ_{ij}^{AB}, γ_{ik}^{AC}, γ_{jk}^{BC} и γ_{ijk}^{ABC}, $i, j, k \in \mathbb{B}$ соответственно. Аналогично двухфакторному случаю можно получить следующие выражения для этих взаимодействий и равенство их нулю как выражение отсутствия всяких взаимодействий:

$$\begin{cases} \gamma_{ij}^{AB} = p_{ij\bullet} - p_{i\bullet\bullet} - p_{\bullet j\bullet} + p_{\bullet\bullet\bullet} = 0,\ i, j \in \mathbb{B} & (5.1) \\ \gamma_{ik}^{AC} = p_{i\bullet k} - p_{i\bullet\bullet} - p_{\bullet\bullet k} + p_{\bullet\bullet\bullet} = 0,\ i, k \in \mathbb{B} & (5.2) \\ \gamma_{jk}^{BC} = p_{\bullet jk} - p_{\bullet j\bullet} - p_{\bullet\bullet k} + p_{\bullet\bullet\bullet} = 0,\ j, k \in \mathbb{B} & (5.3) \\ \gamma_{ijk}^{ABC} = p_{ijk} - p_{ij\bullet} - p_{i\bullet k} - p_{\bullet jk} + p_{i\bullet\bullet} + p_{\bullet j\bullet} + p_{\bullet\bullet k} - p_{\bullet\bullet\bullet} = 0,\ i, j, k \in \mathbb{B}, & (5.4) \end{cases} \quad (5)$$

где, как и выше, точка означает усреднение по соответствующему индексу:

$$p_{ij\bullet} = \frac{1}{2}(p_{ij0} + p_{ij1}),\ p_{i\bullet\bullet} = \frac{1}{4}(p_{i00} + p_{i01} + p_{i01} + p_{i11}),\ p_{\bullet\bullet\bullet} = \frac{1}{8}\sum_{i,\,j,\,k\,\in\,\mathbb{B}} p_{ijk}$$

Лемма 1. Пусть $q_{ij} = p_{ij\bullet}$ для $i, j \in \mathbb{B}$, тогда $q_{i\bullet} = p_{i\bullet\bullet}$, $q_{\bullet j} = p_{\bullet j\bullet}$, $q_{\bullet\bullet} = p_{\bullet\bullet\bullet}$

Утверждения леммы 1 проверяются простыми вычислениями.

Лемма 2. Подсистема (5.4) системы (5) равносильна единственному уравнению

$$p_{000} - (p_{001} + p_{010} + p_{100}) + (p_{011} + p_{101} + p_{110}) - p_{111} = 0 \quad (6)$$

Доказательство проводится прямыми вычислениями.

Решим систему (5) равенства рисков для трех факторов.

Теорема 2. Система (5) из 20 уравнений равносильна системе (7) из 4 уравнений

$$\begin{cases} p_{000} - p_{010} - p_{100} + p_{110} = 0 & (7.1) \\ p_{000} - p_{001} - p_{100} + p_{101} = 0 & (7.2) \\ p_{000} - p_{001} - p_{010} + p_{011} = 0 & (7.3) \\ 2p_{000} - p_{001} - p_{010} - p_{100} + p_{111} = 0 & (7.4) \end{cases} \quad (7)$$

Доказательство. Сделав замену $q_{ij} = p_{ij\bullet}$, из теоремы 1 и леммы 1 получаем, что подсистема (5.1) системы (5) равносильна уравнению

$$p_{00\bullet} - p_{01\bullet} - p_{10\bullet} + p_{11\bullet} = 0 \quad (8)$$

Аналогично получаем уравнения

$$p_{0\bullet 0} - p_{0\bullet 1} - p_{1\bullet 0} + p_{1\bullet 1} = 0 \quad (9)$$

$$p_{\bullet 00} - p_{\bullet 01} - p_{\bullet 10} + p_{\bullet 11} = 0 \quad (10)$$

Таким образом, используя лемму 2, получаем, что система (5) равносильна системе из уравнений (6), (8), (9), (10).

Прибавим к левой и правой частям уравнения (6), соответственно, левую и правую части уравнения (8), умноженные на 2, и затем разделим каждую часть полученного уравнения на два, тогда получим уравнение (7.1). Аналогично получим уравнения (7.2) и (7.3). Подставив в уравнение (6) переменные p_{110}, p_{101}, p_{011}, выраженные из уравнений (7.1), (7.2), (7.3), получаем уравнение (7.4) и отсюда имеем равносильность системы уравнений (6), (7.1), (7.2), (7.3) системе (7).

Введем определение, аналогичное определению 1.

Определение 2. Пусть значение уровня фактора A (фактора B, фактора C) равно 1 при воздействии фактора A (фактора B, фактора C) и равно 0 при отсутствии воздействия этого фактора. И пусть индексы i, j, $k \in \mathbb{B}$ не равны нулю одновременно. Величину $\Delta_{ijk} = p_{ijk} - p_{000}$ назовем эффектом от воздействия фактора A (при отсутствии воздействия фактора A), если $i = 1$ ($i = 0$), воздействия фактора B (при отсутствии воздействия фактора B), если $j = 1$ ($j = 0$), воздействия фактора C (при отсутствии воздействия фактора C), если, соответственно, $k = 1$ ($k = 0$). Полагаем по определению, что $\Delta_{000} = 0$.

Пример 2. Δ_{101} – эффект от воздействия факторов A и C при отсутствии воздействия фактора B.

Из теоремы 2 и определения величин Δ_{ijk}, i, j, $k \in \mathbb{B}$, вытекает

Следствие 2. Система (5) равенства рисков для трех факторов равносильна системе

$$\begin{cases} \Delta_{110} = \Delta_{010} + \Delta_{100}, & (11.1) \\ \Delta_{101} = \Delta_{001} + \Delta_{100}, & (11.2) \\ \Delta_{011} = \Delta_{001} + \Delta_{010}, & (11.3) \\ \Delta_{111} = \Delta_{001} + \Delta_{010} + \Delta_{100} & (11.4) \end{cases} \tag{11}$$

Интерпретация этого результата состоит в следующем: (а) эффект от воздействия трех факторов равен сумме эффектов от изолированного воздействия каждого в отдельности; (б) эффект от воздействия любых двух факторов равен сумме эффектов от изолированного воздействия каждого в отдельности при отсутствии воздействия третьего. При этом предполагается, что значение уровня каждого фактора равно 1 при воздействии этого фактора и равно 0 при отсутствии этого воздействия.

4. Детерминистическая модель для двух и трех факторов

Рассмотрим детерминистическую модель [7] для двух факторов, где $p_{ij} \in \mathbb{B}$ для любых i, $j \in \mathbb{B}$. Тогда величину $p_{ij} \in \mathbb{B}$ можно рассматривать как значение булевой функции $f \in \mathbb{B}(x_1, x_2)$ от значений уровней факторов $x_1 = i$, $x_2 = j$: $p_{ij} = f(i, j)$. Здесь $\mathbb{B}(x_1, x_2)$ – множество всех булевых функций от переменных x_1, x_2. Булева функция f тогда представляет отклик, а переменные x_1, x_2 – факторы. Для дальнейшего исследования удобно рассматривать булевы функции, представленные в виде ДНФ (дизъюнктивной нормальной формы).

Теорема 3. Пусть для некоторой функции $f \in \mathbb{B}(x_1, x_2)$ имеет место равенство $p_{ij} = f(i, j)$, i, $j \in \mathbb{B}$, и для величин p_{ij} справедлива система (2). Тогда $f \in \{\mathbf{0}, \mathbf{1}, x_1, x_2, \bar{x}_1, \bar{x}_2\}$.

Доказательство. Пусть f – ненулевая булева функция из $\mathbb{B}(x_1, x_2)$. По теореме 1 система (2) равносильна уравнению (3) и уравнению (4). Пусть $p_{11} = f(1, 1) = 1$. Если $p_{00} = f(0, 0) = 1$, то из уравнений (3) или (4) легко следует, что $p_{01} = p_{10} = f(0, 1) = f(1, 0) = 1$, т. е. $f = \mathbf{1}$. Если $f(0, 0) = 0$, то из этих же уравнений $f(1, 0) = 1$, $f(0, 1) = 0$ или $f(1, 0) = 0$, $f(0, 1) = 1$, т. е. $f = x_1$ или $f = x_2$ соответственно. В случае $p_{11} = f(1, 1) = 0$ аналогично получаем, что $f = \bar{x}_1$, или $f = \bar{x}_2$.

Рассмотрим детерминистическую модель для трех факторов, где $p_{ijk} \in \mathbb{B}$ для любых i, j, $k \in \mathbb{B}$. Аналогично случаю двух факторов величину $p_{ijk} \in \mathbb{B}$ можно рассматривать как значение булевой функции $f \in \mathbb{B}(x_1, x_2, x_2)$ от значений уровней факторов $x_1 = i$, $x_2 = j$, $x_3 = k$: $p_{ijk} = f(i, j, k)$. Здесь $\mathbb{B}(x_1, x_2, x_3)$ – множество всех булевых функций от переменных x_1, x_2, x_3. Для трех факторов справедливо утверждение, аналогичное теореме 3.

Теорема 4. Пусть для некоторой булевой функции $f \in \mathbb{B}(x_1, x_2, x_3)$ имеет место равенство $p_{ijk} = f(i, j, k)$, $i, j, k \in \mathbb{B}$, и для величин p_{ijk} справедлива система (5). Тогда $f \in \{\mathbf{0}, \mathbf{1}, x_1, x_2, x_3, \overline{x}_1, \overline{x}_2, \overline{x}_3\}$.

Доказательство. Заметим, что $\Delta_{ijk} = f(i, j, k) - f(0, 0, 0)$. Пусть f – ненулевая булева функция из $\mathbb{B}(x_1, x_2, x_3)$, тогда снова сначала рассмотрим ситуацию, когда $f(1, 1, 1) = 1$. Возможно два случая: $f(0, 0, 0) = 1$ и $f(0, 0, 0) = 0$. Рассмотрим первый случай, тогда $\Delta_{111} = 0$ и $\Delta_{ijk} \le 0$ для любых $i, j, k \in \mathbb{B}$. Следовательно, из уравнения (11.4) имеем $\Delta_{100} = \Delta_{010} = \Delta_{001} = 0$. Тогда из уравнений (11.1)–(11.3) получаем равенства $\Delta_{110} = \Delta_{101} = \Delta_{011} = 0$, а это и значит, что $f = \mathbf{1}$. Рассмотрим второй случай, тогда $\Delta_{111} = 1$ и $\Delta_{ijk} \ge 0$ для любых $i, j, k \in \mathbb{B}$. Снова применяем уравнение (11.4): ровно одна из величин $\Delta_{100}, \Delta_{010}, \Delta_{001}$ равна 1, пусть без ограничения общности это будет Δ_{100}, т. е. $\Delta_{100} = 1$, $\Delta_{010} = 0$, $\Delta_{001} = 0$. Применяем уравнения (11.1)–(11.3) и получаем: $\Delta_{110} = \Delta_{101} = 1$, $\Delta_{011} = 0$. Таким образом, $f(1, j, k) = 1$, $f(0, j, k) = 0$ для всех $j, k \in \mathbb{B}$, а это значит, что $f = x_1$. Таким образом, при условии $f(1, 1, 1) = 1$ имеем $f \in \{\mathbf{1}, x_1, x_2, x_3\}$. При условии $f(1, 1, 1) = 0$ совершенно аналогично можно получить, что $f \in \{\overline{x}_1, \overline{x}_2, \overline{x}_3\}$.

Следствие 3. В случае детерминистической модели для двух (трех) факторов система (2) (система (5)) равенства рисков выполняется для тех и только тех булевых функций, которые существенно зависят не более чем от одной переменной.

Пример 3. При $n = 2$ система (2) равенства рисков выполняется только для функций

$$f \in \{\mathbf{0}, \mathbf{1}, x_1, \overline{x}_1, x_2, \overline{x}_2\}$$

и не выполняется для всех остальных функций $f \in \{x_1 \vee x_2, \overline{x}_1 \vee x_2, x_1 \vee \overline{x}_2, \overline{x}_1 \vee \overline{x}_2, x_1x_2, \overline{x}_1x_2, x_1\overline{x}_2, \overline{x}_1\overline{x}_2, x_1x_2 \vee \overline{x}_1\overline{x}_2, \overline{x}_1x_2 \vee x_1\overline{x}_2\}$.

Пример 4. При $n = 3$ система равенства рисков (5) выполняется только для функций $f \in \{\mathbf{0}, \mathbf{1}, x_1, \overline{x}_1, x_2, \overline{x}_2, x_3, \overline{x}_3\}$ и не выполняется, например, для функций

$$f \in \{x_1 \vee x_2, x_1 \vee x_2 \vee x_3, x_1x_2, x_1x_2x_3, x_1\overline{x}_2 \vee \overline{x}_1x_2, x_1x_2 \vee x_3\}$$

5. Совместное действие бинарных факторов в бинарном отклике

Следствие 3 означает, что если считать, что отсутствие взаимодействия данных факторов в булевой функции от двух (трех) переменных выражается системой равенства рисков (2) (системой (5)), то любая булева функция, существенно зависящая более чем от одной переменной, представляет взаимодействие факторов.

В работах [8–11] была развита теория совместного действия бинарных факторов в бинарном отклике в рамках концепции бинарной теории достаточных причин, применяемой в токсикологии, эпидемиологии и доказательной медицине. Значение отклика, равное 1, в этих приложениях означает обычно заболевание или летальный исход. В частности, в этих работах было дано более тонкое определение взаимодействия бинарных факторов (переменных) в бинарном отклике (булевой функции). Приведем это определение.

Определение 3 [8–10]. Будем говорить, что в булевой функции $f \in \mathbb{B}(x_1, x_2)$ ($f \in \mathbb{B}(x_1, x_2, x_3)$) имеет место *совместное действие* двух (трех) переменных $x = x_1, x_2$ ($x = x_1, x_2, x_3$), если существует такой набор $\alpha = (\alpha_1, \alpha_2) \in \mathbb{B}^2$ ($\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{B}^3$), что конъюнкция $x^\alpha = x_1^{\alpha_1}x_2^{\alpha_2}$ ($x^\alpha = x_1^{\alpha_1}x_2^{\alpha_2}x_3^{\alpha_3}$) является простым импликантом [12–14] булевой функции *f*. Будем говорить, что в этом случае совместное действие в булевой функции достигается для $x = \alpha$.

Здесь для $\gamma \in \mathbb{B}$ символ x^γ обозначает *x*, если $\gamma = 1$, и $\bar{x}$, если $\gamma = 0$.

Пример 5. При $n = 2$ совместное действие переменных x_1, x_2 есть в булевых функциях

$$f \in \{x_1x_2,\ \bar{x}_1x_2,\ x_1\bar{x}_2,\ \bar{x}_1\bar{x}_2,\ x_1x_2 \vee \bar{x}_1\bar{x}_2,\ \bar{x}_1x_2 \vee x_1\bar{x}_2\}$$

и нет во всех остальных булевых функциях

$$f \in \{\mathbf{0},\ \mathbf{1},\ x_1,\ \bar{x}_1,\ x_2,\ \bar{x}_2,\ x_1 \vee x_2,\ \bar{x}_1 \vee x_2,\ x_1 \vee \bar{x}_2,\ \bar{x}_1 \vee \bar{x}_2\}$$

Пример 6. При $n = 3$ совместное действие переменных x_1, x_2, x_3 есть, например, в булевых функциях $f \in \{x_1x_2x_3,\ \bar{x}_1\bar{x}_2\bar{x}_3 \vee x_1x_2,\ \bar{x}_1\bar{x}_2\bar{x}_3 \vee x_1x_2x_3\}$ и нет его, например, в булевых функциях

$$f \in \{\mathbf{0},\ \mathbf{1},\ x_1,\ \bar{x}_1,\ x_1x_2,\ x_1 \vee x_3,\ \bar{x}_1 \vee x_2 \vee \bar{x}_3,\ x_1x_3 \vee x_2x_3\}$$

Из следствия 3 и определения 3 непосредственно вытекает

Следствие 4. Если в булевой функции $f \in \mathbb{B}(x_1, x_2)$ ($f \in \mathbb{B}(x_1, x_2, x_3)$) есть совместное действие переменных x_1, x_2 (x_1, x_2, x_3), то для *f* не выполняется система (2) (система (5)) равенства рисков, где $p_{ij} = f(i, j)$, $i, j \in \mathbb{B}$ ($p_{ijk} = f(i, j, k)$, $i, j, k \in \mathbb{B}$).

Обратное неверно, а именно: 1) для $n = 2$ в булевых функциях вида $f = x_1^{\alpha_1} \vee x_2^{\alpha_2}$, где $\alpha_1, \alpha_2 \in \mathbb{B}$, по следствию 3 не выполняется система (2) равенства рисков, но по определению 3 нет совместного действия переменных x_1, x_2; 2) для $n = 3$ в булевых функциях вида $f = x_1^{\alpha_1} \vee x_2^{\alpha_2} \vee x_3^{\alpha_3}$, $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{B}$, аналогично не выполняется система (5) равенства рисков, но по определению 3 нет совместного действия переменных x_1, x_2, x_3. Таким образом, в рамках детерминистической модели определение 3 совместного действия бинарных факторов более точно характеризует взаимодействие бинарных факторов, чем традиционное определение через невыполнение системы равенства рисков.

Рассмотрим примеры откликов (булевых функций) для двух факторов (двух переменных) и взаимодействие этих факторов в них.

Пример 7. Известно, что риск развития рака легких среди курящих значительно повышается в присутствии амфиболового асбеста [9, 15]. В этом случае бинарным фактором x_1 можно считать курение, точнее, $x_1 = 1$ для курящего человека и $x_1 = 0$ для некурящего, бинарным фактором x_2 – присутствие амфиболового асбеста ($x_2 = 1$ – амфиболовый асбест присутствует, $x_2 = 0$ – отсутствует), а булевой функцией, характеризующей отклик, функцию $f = x_1x_2$ – развитие рака. По определению 3 в *f* есть совместное действие факторов x_1, x_2 при уровнях значений факторов $x_1 = x_2 = 1$, а значит, по следствию 3 не выполняется система (2) равенства рисков, где $p_{ij} = f(i, j)$, $i, j \in \mathbb{B}$.

Пример 8. При одном из обследований несовершеннолетних пациентов выявилась ситуация, когда у детей, проживавших в местности, где произошел выброс соединений и свинца, и кадмия, результаты анализов были практически такими же, как и у детей, проживавших в местности, где не было выброса ни того, ни другого. В то же время, у детей, проживавших в местности, где были выбросы только соединений свинца или только соединений кадмия, результаты анализов были плохими. Интересно, что сами по себе соединения свинца и кадмия не нейтрализуют друг друга. В данном случае бинарным фактором x_1 можно считать присутствие выбросов соединений свинца, а фактором x_2 – присутствие соединений кадмия. Булевой функцией, характеризующей отклик, является функция $f = \bar{x}_1 x_2 \vee x_1 \bar{x}_2$. В f по определению 3 есть совместное действие факторов x_1, x_2 при уровнях значений факторов $x_1 = 0$, $x_2 = 1$ (проживание в местности, с выбросами только соединений кадмия) или при $x_1 = 1$, $x_2 = 0$ (проживание в местности, с выбросами только соединений свинца), и не выполняется система (2) равенства рисков.

Пример 9. Известно, что цианиды в промышленных отходах токсичны для аквафауны. Но если в этих отходах присутствуют соединения никеля, которые сами по себе тоже является токсичным, то ионы никеля могут образовывать с цианид-ионами нетоксичные соединения [8]. Здесь, как и в предыдущем примере, булева функция, характеризующая отравление аквафауны, зависящая от факторов x_1 – присутствия цианидов и x_2 – присутствия никеля, равна $f = \bar{x}_1 x_2 \vee x_1 \bar{x}_2$. И так же, как и в предыдущем примере, мы имеем совместное действие факторов x_1, x_2 при уровнях значений факторов $x_1 = 0$, $x_2 = 1$ (присутствие только никеля) или при $x_1 = 1$, $x_2 = 0$ (присутствие только цианидов) и невыполнение системы (2) равенства рисков.

Булева функция следующего примера может возникнуть при описании любых двух несовместимых лекарственных препаратов.

Пример 10. При лечении заболеваний щитовидной железы могут назначать высокие дозы препарата Йодомарина, действующим веществом которого является йодид калия. Если при этом для лечения артериальной гипертензии назначаются калийсберегающие диуретики, то может произойти гиперкалиемия, что значительно ухудшит состояние организма. Таким образом, булева функция, означающая в данном случае ослабление общего состояния здоровья, для пациентов с заболеванием щитовидной железы и артериальной гипертензией равна $f = \bar{x}_1 \bar{x}_2 \vee x_1 x_2$, где фактор x_1 – прием высоких доз препарата Йодомарина (Й), x_2 – прием калийсберегающих диуретиков (КД). В функции f есть совместное действие факторов x_1, x_2 при уровнях значений факторов $x_1 = 0$, $x_2 = 0$ (отсутствие приема высоких доз Й и приема КД) или при $x_1 = 1$, $x_2 = 1$ (одновременный прием двух препаратов), и, конечно, не выполняется система (2) равенства рисков.

Пример 11. Если рассматривать факторы x_1, x_2 как курение и злоупотребление пациентами спиртных напитков соответственно, а откликом считать булеву функцию f, означающую развитие у курящих пациентов или пациентов, злоупотребляющих спиртными напитками, сердечно-сосудистых заболеваний, то $f = x_1 \vee x_2$, поскольку это заболевание возникает при присутствии хотя бы одного из этих двух факторов. В этой булевой функции нет совместного действия переменных x_1, x_2 по определению 3, хотя по следствию 4 система (2) равенства рисков не выполняется.

Теперь рассмотрим примеры откликов в случае трех факторов и взаимодействие этих факторов в них.

Пример 12. Пусть фактор x_1 – низкая усваиваемость железа, фактор x_2 – аллергия на препараты железа, фактор x_3 – низкая свертываемость крови, а f – булева функция, означающая присутствие у пациентов анемии. Тогда $f = x_1 x_2 x_3$. Действительно, это заболевание проявляется только при присутствии всех этих трех факторов. В этой булевой функции есть совместное действия переменных x_1, x_2, x_3 по определению 3, и по следствию 4 не выполняется система равенства рисков (5).

В работах [9, 11] введено и изучено понятие совместного действия двух бинарных факторов в отклике, зависящем от трех факторов.

Определение 4. В булевой функции (отклике) $f \in \mathbb{B}(x_1, x_2, x_3)$ имеется совместное действие двух булевых переменных (факторов) $x_I = (x_i)_{i \in I}$ для двухэлементного упорядоченного подмножества I множества $\{1, 2, 3\}$, если в этой функции (отклике) есть совместное действие этих двух переменных при фиксированном значении третьей переменной $x_i = \beta$, $i \notin I$, $\beta \in \mathbb{B}$.

В этих же работах показано, что если в булевой функции есть совместное действие трех переменных, то в ней есть и совместное действие любых двух переменных.

Пример 13. Во всех булевых функциях из примера 6, где есть совместное действие трех переменных, есть и совместное действие любых двух переменных. Кроме того, в следующих булевых функциях из примера 6 по определению 4 есть совместное действие двух переменных, но по определению 3 нет совместного действия трех переменных: в $f = x_1x_2$ есть совместное действие переменных x_1, x_2 для значений уровней $x_1 = 1$ и $x_2 = 1$ при фиксированном значении уровня $x_3 = 0$ или $x_3 = 1$; в $f = x_1x_3 \vee x_2x_3$ есть совместное действие переменных x_1, x_3 для значений уровней $x_1 = 1$ и $x_3 = 1$ при фиксированном значении уровня $x_2 = 0$ или совместное действие переменных x_2, x_3 для значений уровней $x_2 = 1$ и $x_3 = 1$ при фиксированном значении уровня $x_1 = 0$.

Пример 14. Если рассматривать зависимость отклика f – тяжелого течения болезни COVID-19 от факторов x_1 – сахарного диабета, x_2 – возраста выше 65 лет и x_3 – вакцинации от COVID-19, то легко понять, что $f = x_1\bar{x}_3 \vee x_2\bar{x}_3$. В этой булевой функции нет совместного действия трех переменных x_1, x_2, x_3 по определению 3, но по определению 4 есть совместное действие переменных x_1, x_3 для значений $x_1 = 1$ (присутствие сахарного диабета) и $x_3 = 0$ (отсутствие прививки) при фиксированном значении переменной $x_3 = 0$ (возраст не выше 65), а также есть совместное действие переменных x_2, x_3 для значений $x_2 = 1$ (возраст выше 65) и $x_3 = 0$ (отсутствие прививки) при фиксированном значении переменной $x_1 = 0$ (отсутствии сахарного диабета). Кроме того, по следствию 4 не выполняется система (5) равенства рисков.

Введение и исследование понятия совместного действия двух факторов в отклике, зависящем от трех факторов, – это еще одно обстоятельство, подчеркивающее, что определение совместного действия бинарных факторов в бинарном отклике, развитое в работах [8–11], в рамках детерминистической модели более точно описывает взаимодействие факторов, чем просто невыполнение системы равенства рисков, соответствующей данному отклику.

6. Инвариантность системы равенства рисков относительно группы симметрий эксперимента

При исследовании совместного действия двух (трех) бинарных факторов в отклике неважно, как кодируются уровни факторов. Так возникает группа симметрий эксперимента, являющаяся группой G_2 (G_3) автоморфизмов булева квадрата $\mathbb{B}^2$ (булева куба $\mathbb{B}^3$) как графа, две вершины которого соединены ребром тогда и только тогда, когда расстояние Хэмминга между ними равно 1 [8, 16, 17]. Автоморфизмы булева квадрата $\mathbb{B}^2$ (булева куба $\mathbb{B}^3$) естественным образом продолжаются на булеву алгебру $\mathbb{B}(x_1, x_2)$ ($\mathbb{B}(x_1, x_2, x_3)$). Действие группы G_2 (G_3) на булевой алгебре $\mathbb{B}(x_1, x_2)$ ($\mathbb{B}(x_1, x_2, x_3)$) заключается в следующих последовательно выполненных операциях: любой перестановке переменных в каждой функции и навешивании отрицаний на некоторые переменные. Группа $\mathbb{B}(x_1, x_2)$ ($\mathbb{B}(x_1, x_2, x_3)$) изоморфна группе симметрий квадрата (куба) [17, 18]. Действие группы G_2 (G_3) на алгебре $\mathbb{B}(x_1, x_2)$ ($\mathbb{B}(x_1, x_2, x_3)$) порождает естественное разбиение этой алгебры на орбиты действия этой группы – типы совместного действия факторов [8, 16–18]. Орбиты булевой функции f относительно этого действия будем обозначать через $\langle f \rangle$. Совместное действие двух (трех) факторов инвариантно относительно действия группы G_2 (G_3) [8, 9, 18].

Группа G_2 (G_3), действуя на булевом квадрате $\mathbb{B}^2$ (на булевом кубе $\mathbb{B}^2$), действует и на множестве переменных $P_2 = \{p_{ij} \mid i, j \in \mathbb{B}\}$ ($P_3 = \{p_{ijk} \mid i, j, k \in \mathbb{B}\}$). Поэтому группу G_2 (G_3) можно считать подгруппой группы $Sym(P_2)$ ($Sym(P_3)$) перестановок множества переменных P_2 (P_3). Хорошо известно представление группы $Sym(P_2)$ ($Sym(P_3)$) перестановочными матрицами [19]. Таким образом, группа $Sym(P_2)$ ($Sym(P_3)$) изоморфным образом вкладывается в группу $GL(2^2, \mathbb{Q})$ ($GL(2^3, \mathbb{Q})$) всех невырожденных матриц четвертого (восьмого) порядка над полем $\mathbb{Q}$ рациональных чисел или, что то же самое, в группу всех обратимых линейных преобразований векторного пространства $\mathbb{Q}^{2^2}$ ($\mathbb{Q}^{2^3}$) над полем $\mathbb{Q}$ [20]. Кроме того, поскольку векторное пространство $\mathbb{Q}^{2^2}$ ($\mathbb{Q}^{2^3}$) изоморфно векторному пространству $\mathbb{Q}[P_2]$ $\mathbb{Q}[P_3]$) всех многочленов с коэффициентами из $\mathbb{Q}$ от переменных из G_2 (G_3), группа G_2 (G_3) действует также и на $\mathbb{Q}[P_2]$ ($\mathbb{Q}[P_3]$).

Теорема 5. Система (2) равенства рисков (система (5) равенства рисков) инвариантна относительно действия группы G_2 (G_3).

Доказательство. Легко понять, что система (2) инвариантна относительно действия группы G_2. Аналогично каждая из подсистем (5.1), (5.2), (5.3), (5.4) инвариантна относительно действия группы G_3, а значит, и система (5) инвариантна относительно действия этой группы.

Таким образом, о выполнении или невыполнении системы равенства рисков можно говорить не только для данной булевой функции f, но для всего класса $\langle f \rangle$.

Рассмотрим обобщение определений 1, 2.

Определение 5. Пусть значение уровня фактора A (фактора B) равно $\alpha_1 \in \mathbb{B}$ ($\alpha_2 \in \mathbb{B}$) при воздействии фактора A (фактора B) и равно $\bar{\alpha}_1$ ($\bar{\alpha}_2$) при отсутствии воздействия этого фактора. Пусть индексы $i, j \in \mathbb{B}$ не равны нулю одновременно. Величину $\Delta_{\alpha,ij} = p_{i\oplus\alpha_1, j\oplus\alpha_2} - p_{\alpha_1,\alpha_2}$, где $\alpha = (\alpha_1, \alpha_2) \in \mathbb{B}^2$ и $\oplus$ – сложение по модулю 2, назовем эффектом от воздействия фактора A (при отсутствии воздействия фактора A), если $i = 1$ ($i = 0$), и воздействия фактора B (при отсутствии воздействия фактора B), если $j = 1$ ($j = 0$). Полагаем по определению, что $\Delta_{\alpha,00}$.

Обобщением следствия 1 является утверждение

Следствие 5. Для любого набора $\alpha = (\alpha_1, \alpha_2) \in \mathbb{B}^2$ система (2) равенства рисков равносильна уравнению

$$\Delta_{\alpha,11} = \Delta_{\alpha,01} + \Delta_{\alpha,10}. \tag{12}$$

Доказательство. Очевидно, многочлен $\Delta_{\alpha,ij} \in \mathbb{Q}[P_2]$ является образом многочлена $\Delta_{ij} \in \mathbb{Q}[P_2]$ при действии сдвига $t_\alpha(x) = x \oplus \alpha$ на множестве $\mathbb{Q}[P_2]$. Учитывая, что $t_\alpha(x) \in G_2$, по следствию 1 и теореме 5 получаем требуемое.

Пример 15. Пусть значения уровней факторов A, B равны $\alpha_1 = 1$ и $\alpha_2 = 0$, соответственно, при воздействии этих факторов. Тогда по следствию 5 система равенства рисков (2) равносильна равенству (12):

$$p_{1\oplus1,1\oplus0} - p_{10} = (p_{0\oplus1,1\oplus0} - p_{10}) + (p_{1\oplus1,0\oplus0} - p_{10}),$$

т. е. равенству $p_{01} - p_{10} = (p_{11} - p_{10}) + (p_{00} - p_{10})$.

Следствие 5 можно интерпретировать следующим образом: эффект от воздействия двух факторов равен сумме эффектов от изолированного воздействия каждого в отдельности независимо от значений уровней факторов, при котором происходит воздействие этих факторов.

Аналогично можно ввести величины $\Delta_{\alpha,ijk} = p_{i\oplus\alpha_1, j\oplus\alpha_2, k\oplus\alpha_3} - p_{\alpha_1,\alpha_2,\alpha_3}$ для $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{B}^3$, если индексы i, j, $k \in \mathbb{B}$ не равны нулю одновременно, принимая, что $\Delta_{\alpha,000} = 0$, и доказать аналог следствия 5.

Следствие 6. Для любого набора $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{B}^3$ система (5) равенства рисков равносильна системе

$$\begin{cases} \Delta_{\alpha,110} = \Delta_{\alpha,010} + \Delta_{\alpha,100} \\ \Delta_{\alpha,101} = \Delta_{\alpha,001} + \Delta_{\alpha,100} \\ \Delta_{\alpha,011} = \Delta_{\alpha,001} + \Delta_{\alpha,010} \\ \Delta_{\alpha,111} = \Delta_{\alpha,001} + \Delta_{\alpha,010} + \Delta_{\alpha,100}. \end{cases}$$

Интерпретировать этот результат можно следующим образом: (а) эффект от воздействия трех факторов равен сумме эффектов от изолированного воздействия каждого в отдельности; (б) эффект от воздействия любых двух факторов равен сумме эффектов от изолированного воздействия каждого в отдельности при отсутствии воздействия третьего. И этот результат не зависит от значений уровней факторов, при котором происходит воздействие этих факторов.

7. Заключение

В настоящей работе:

1. получена и решена система равенства рисков для двух и трех бинарных факторов в случае бинарного отклика;
2. показано, что система равенства рисков для двух и трех факторов инвариантна относительно действия группы симметрий экспер带мента;
3. получена интерпретация решения системы равенства рисков для двух факторов: эффект от воздействия двух факторов равен сумме эффектов от изолированного воздействия каждого в отдельности независимо от значений уровней факторов, при котором происходит воздействие этих факторов;
4. получена интерпретация решения системы равенства рисков для трех факторов: (а) эффект от воздействия трех факторов равен сумме эффектов от изолированного воздействия каждого в отдельности; (б) эффект от воздействия

любых двух факторов равен сумме эффектов от изолированного воздействия каждого в отдельности при отсутствии воздействия третьего; этот результат не зависит от значений уровней факторов, при котором происходит воздействие этих факторов;

5. показано, что в рамках детерминистической модели и для двух, и для трех факторов система равенства рисков выполняется только для откликов, существенно зависящих не более чем от одного фактора;
6. показано, что в рамках детерминистической модели определение совместного действия двух или трех факторов в бинарном отклике, данное в работах [8–11], более полно описывает взаимодействие этих факторов, чем невыполнение системы равенства рисков.

8. Список литературы

1. *Koopman*, *J. S.* Interaction between discrete causes / J. S. Koopman // Am J Epidem. – 1981. – Vol. 113, № 6. – С. 716–724. – DOI: 10.1093/oxfordjournals.aje.a113153.
2. *Фишер*, *Р. А.* Статистические методы для исследователей / Р. А. Фишер. – М. : Госстатиздат, 1958. – 268 с.
3. *Кендалл*, *М.* Многомерный статистический анализ и временные ряды / М. Кендалл, А. Стьюарт. – М. : Наука, 1976. – 737 с.
4. *Шеффе*, *Г.* Дисперсионный анализ / Г. Шеффе. – М.: Наука, 1980. – 511 с.
5. *Rothman*, *K. J.* Modern epidemiology: Third edition / K. J. Rothman, S. Greenland, T. L. Lash // Modern Epidemiology. – 2011. – P. 1–758. – EDN YDXWPS.
6. *Szklo*, *M.* Epidemiology: Beyond the Basics / M. Szklo, F. J. Nieto. – 2nd ed. Boston, MA: Jones and Bartlee Publishers, 2007. – 578 с. – ISBN-128411659X.
7. *Miettinen*, *O. S.* Causal and preventive interdependence. Elementary principles / O. S. Miettinen // Scand J. Work Environ Health. – 1982. – Vol. 8, № 3. – С. 159–168. – DOI: 10.5271/sjweh.2479.
8. *Nagrebetskaya*, *J. V.* Joint action of binary factors in the sufficient causes theory and its classification / J. V. Nagrebetskaya, V. G. Panov // International J. of Innovative Technology and Exploring Engineering. – 2019. – Vol. 9, No. 1. – P. 2146–2153. – DOI 10. 35940/ ijitee.A4702.119119. – EDN LQILAE.
9. *Нагребецкая*, *Ю. В.* Булева модель частичного взаимодействия бинарных факторов / Ю. В. Нагребецкая, В. Г. Панов // Вестник Омского университета. – 2021. – Т. 26, № 3. – С. 10–19. – DOI 10.24147/1812-3996.2021.26(3).10-19. – EDN XVLXOK.
10. *Нагребецкая*, *Ю. В.* Степень взаимодействия бинарных факторов в теории достаточных причин / Ю. В. Нагребецкая, В. Г. Панов // Системный анализ в медицине : Материалы XIII Международной научной конференции, Благовещенск, 26–27 сентября 2019 г. – Благовещенск: Дальневосточный научный центр физиологии и патологии дыхания, 2019. – С. 31–34. – EDN ETTKWN.
11. *Нагребецкая, Ю. В.* Обобщение понятия взаимодействия n факторов в теории достаточных причин и его свойства / Ю. В. Нагребецкая, В. Г. Панов // Системный анализ в медицине : Материалы XIII Международной научной конференции, Благовещенск, 26–27 сентября 2019 г. – Благовещенск : Дальневосточный научный центр физиологии и патологии дыхания, 2019. – С. 35–38. – EDN SMROHS.
12. *Яблонский*, *С. В.* Введение в дискретную математику : учеб. пособие для студентов вузов, обучающихся по специальности «Прикладная математика» / С. В. Яблон-

ский. – 4 изд., стер. – М. : Высшая школа, 2003. – 384 с. – (Высшая математика). – ISBN 5-06-004681-8. – EDN QJLZCB.

13. *Лидл*, *Р*. Прикладная абстрактная алгебра: учеб. пособие / Р. Лидл, Г. Пильц; пер. с англ. – Екатеринбург : Изд-во Урал. ун-та, 1996. – 744 с.

14. *Crama*, *Y*. Boolean functions. Theory, Algorithms, and Applications / Y. Crama, P. L. Hammer // Cambridge: Cambridge University Press, 2011. – 711 c.

15. *Boffetta*, *P*. Occupational exposure to asbestos and man-made vitreous fibres and risk of lung cancer: A multicentre case-control study in Europe / P. Boffetta, R. Carel, A. C. Olsson [et al.] // Occupational and Environmental Medicine. – 2007. – Vol. 64, No. 8. – P. 502–508. – DOI 10.1136/oem.2006.027748. – EDN LKNXVV.

16. *Panov*, *V. G*. Boolean algebras and classification of interactions in sufficient-component cause model / V. G. Panov, J. V. Nagrebetskaya // International Journal of Pure and Applied Mathematics. – 2015. – Vol. 98, No. 2. – P. 239–259. – DOI 10.12732/ijpam.v98i2.7. – EDN UENIDL.

17. *Panov*, *V. G*. Classification of combined action of binary factors and Coxeter groups / V. G. Panov, J. V. Nagrebetskaya // J. of Discrete Mathematical Sciences and Cryptography. – 2018. – Vol. 21, No. 3. – P. 661–677. – DOI 10.1080/09720529.2016.1222733. – EDN XXXMFV.

18. *Панов*, *В. Г*. Алгебраическая трактовка двухфакторной бинарной теории достаточных причин / В. Г. Панов, Ю. В. Нагребецкая // Труды СПИИРАН. – 2013. – № 3(26). – С. 277–296. – EDN QIXKOD.

19. *Супруненко*, *Д. А*. Группы матриц / Д. А. Супруненко. – М. : Наука, 1972. – 352 с.

20. *Винберг*, *Э. Б*. Курс алгебры / Э. Б. Винберг. – М. : МЦНМО, 2011. – 592 с. ISBN 978-5-94057-685-3. – EDN SDSFZZ.

Сведения об авторах:

Нагребецкая Юлия Вацлавовна, к. ф.-м. н., доцент, Институт естественных наук и математики, департамент математики, механики и компьютерных наук УрФУ, ул. Тургенева, 4, г. Екатеринбург, Россия. Эл. почта: I.V.Nagrebetckaia@urfu.ru.

Фаткуллина Аэлита Ильфатовна, студентка Института естественных наук и математики, департамент математики, механики и компьютерных наук УрФУ.

JOINT ACTION OF BINARY FACTORS AND RISKS' EQUALITY

J. V. Nagrebeskaya, A. I. Fatkullina

Ural Federal University named after the first President of Russia B. N. Yeltsin, Ekaterinburg, Russia

In this paper we study the models of interaction of two and three independent binary factors in the case of binary response, which are widely used in medical and biological sciences. The main attention is paid to the formalization of the concept of "interaction" and its connection with the risks' difference equality. It is shown that this equality can be obtained from equating all interaction terms to zero in their representation used in the analysis of variance. The system of risks' differences equations for two and three factors is obtained and solved, and its invariance with respect to the action of a symmetry group of experiment is shown. The interpretation of the solution for the system of risks' differences equations is presented: in the case of two factors the effect of their impact is equal to the sum of effects of isolated action of each factor separately; in the case of three factors (a) the effect of all these factors is equal to the sum of effects of isolated action of each factor separately; (b) the effect of any two factors is equal to the sum of effects of isolated action of each factor separately in the absence of the third one. In addition, a study of the solution of equality of risks for two and three factors within the framework of the deterministic model, leading to a Boolean formalization of this model, was carried out. It is proved that for the values of the Boolean function representing a given response, the system of risks' differences equations for both two and three factors is satisfied if and only if this Boolean function significantly depends on no more than one variable.

It is shown that within the framework of the deterministic model, the concept of "joint action" of binary factors, introduced and studied in the works of the first author, more accurately describes the interaction of these factors than violation of risks' differences equations.

Key words: risks' differences equations; deterministic model; binary theory of sufficient causes; response; joint action of binary factors; Boolean function; automorphism group; group action.

References

1. *Koopman*, *J. S.* Interaction between discrete causes / J. S. Koopman // Am J Epidem. – 1981. – Vol. 113, № 6. – C. 716–724. – DOI: 10.1093/oxfordjournals.aje.a113153.
2. *Fisher*, R. A. Statistical methods for research workers / R. A. Fisher. – Moskow : Gosstatizdat. – 1958. – 268 p.
3. *Kendall*, *M. G.* Design and analysis, and time-series / M. G. Kendall, A. Stuart. – Moskow : Nauka. – 1976. – 737 p.
4. *Scheffe*, *G.* Analysis of variance / G. Scheffe. – Moskow : Nauka, 1980. – 511 p.

5. *Rothman*, *K. J.* Modern epidemiology: Third edition / K. J. Rothman, S. Greenland, T. L. Lash // Modern Epidemiology: Third Edition, 2011. – P. 1–758. – EDN YDXWPS.
6. *Szklo*, *M.* Epidemiology: Beyond the Basics / M. Szklo, F. J. Nieto. – 2nd ed. Boston, MA: Jones and Bartlee Publishers, 2007. – 578 c. – ISBN-128411659X.
7. *Miettinen*, *O. S.* Causal and preventive interdependence. Elementary principles / O. S. Miettinen // Scand J. Work Environ Health. – 1982. – Vol. 8, № 3. – C. 159–168. – DOI: 10.5271/sjweh.2479.
8. *Nagrebetskaya*, *J. V.* Joint action of binary factors in the sufficient causes theory and its classification / J. V. Nagrebetskaya, V. G. Panov // International J. of Innovative Technology and Exploring Engineering. – 2019. – Vol. 9, No. 1. – P. 2146–2153. – DOI 10. 35940/ijitee.A4702.119119. – EDN LQILAE.
9. *Nagrebetskaya*, *J. V.* Boolean model of partial interaction of binary factors / J. V. Nagrebeskaya, V. G. Panov // Guardian of the Omsk University. – 2021. – Vol. 26, № 3. – P. 10–19. – DOI 10.24147/1812-3996.2021.26(3).10–19.
10. *Nagrebetskaya*, *J. V.* Interaction degree of binary factors in the sufficient causes theory / J. V. Nagrebetskaya, V. G. Panov // Materials of the XIII International Scientific Conference «System analysis in medicine» (SAM 2019) September 19–20, 2019. Blagoveshchensk. – P. 31–34. – DOI: 10.12737/collection_5d8335e34b6a76.02467823.
11. *Nagrebetskaya*, *J. V.* Обобщение понятия взаимодействия n факторов в теории достаточных причин и его свойства // Materials of the XIII International Scientific Conference «System analysis in medicine» (SAM 2019) September 19–20, 2019. Blagoveshchensk. – P. 35–38. – DOI: 10.12737/collection_5d8335e34b6a76.02467823.
12. *Yablonskiy*, *S. V.* Introduction to Discrete Mathematics: Textbook for university students studying in the specialty «Applied Mathematics» / S. V. Yablonskiy. – 4th ed. – Moskow: High School, 2003. – 384 p. – (High Mathematics). – ISBN 5-06-004681-8. – EDN QJLZCB.
13. *Lidl*, *R.* Applied Abstrct Algebra: Textbook / R. Lidl, G. Pilz: Translation from English. – Ekaterinburg: Publisher Ural State University, 1996.– 744 p.
14. *Crama*, *Y.* Boolean functions. Theory, Algorithms, and Applications / Y. Crama, P. L. Hammer // Cambridge: Cambridge University Press, 2011. – 711 c.
15. *Boffetta*, *P.* Occupational exposure to asbestos and man-made vitreous fibres and risk of lung cancer: A multicentre case-control study in Europe / P. Boffetta, R. Carel, A. C. Olsson [et al.] // Occupational and Environmental Medicine. – 2007. – Vol. 64, No. 8. – P. 502–508. – DOI 10.1136/oem.2006.027748. – EDN LKNXVV.
16. *Panov*, *V. G.* Boolean algebras and classification of interactions in sufficient-component cause model / V. G. Panov, J. V. Nagrebetskaya // International Journal of Pure and Applied Mathematics. – 2015. – Vol. 98, No. 2. – P. 239–259. – DOI 10.12732/ijpam. v98i2.7. – EDN UENIDL.
17. *Panov*, *V. G.* Classification of combined action of binary factors and Coxeter groups / V. G. Panov, J. V. Nagrebetskaya // J. of Discrete Mathematical Sciences and Cryptography. – 2018. – Vol. 21, No. 3. – P. 661–677. – DOI 10.1080/09720529.2016.1222733. – EDN XXXMFV.
18. *Panov*, *V. G.* An Algebraic Interpretation of the Two-Factor Binary Theory of Sufficient Causes / V. G. Панов, J. V. Nagrebetskaya // Proceedings of SPIIRAN. – 2013, № 3(26). – C. 277–296. – EDN QIXKOD.
19. *Suprunenko*, *D. A.* Groups of Matrices / D. A. Suprunenko. – Moskow : Nauka, 1972. – 352 p.
20. *Vinberg*, *E. B.* Algebra course / E. B. Vinberg. – Moskow: MCCMO, 2011. – 592 p. – ISBN 978-5-94057-685-3. – EDN SDSFZZ.