

УДК 519.213

ВОССТАНОВЛЕНИЕ ПОЛИМОДАЛЬНЫХ ПЛОТНОСТЕЙ ВЕРОЯТНОСТИ ПО ЭКСПЕРИМЕНТАЛЬНЫМ ДАННЫМ В СТРУКТУРАХ СО СТОХАСТИЧЕСКИМИ СВОЙСТВАМИ

© 2014 г.

В.Б. Куликов

Нижегородский государственный технический университет им. Р.Е. Алексеева

vb.kulikov@yandex.ru

Поступила в редакцию 12.11.2013

Рассматриваются результаты восстановления функций плотности распределения вероятности основных показателей в иммунологии. Иммунологическая система человека описывается в стохастических терминах. В качестве идентифицируемой плотности распределения принимается приближенное решение интегрального уравнения Фредгольма первого рода. Регуляризация указанной обратной задачи выполняется на семействе гладких (тригонометрических) функций по выборкам клинических анализов ограниченного объема. Предложенный подход позволяет успешно восстанавливать полимодальные плотности распределения и носит универсальный характер в сфере естествознания.

Ключевые слова: многомодальные законы распределения вероятностей, случайная величина, выборки ограниченного объема, иммунология, обратная задача, метод регуляризации по А.Н. Тихонову.

Введение

В публикации представлен подход к восстановлению полимодальных плотностей распределения вероятностей значительного ряда показателей иммунной системы человека, рассматриваемой в качестве стохастической динамической среды. Исследуются иммунологические характеристики пациентов, получавших курсы антибактериальной терапии.

В естествознании, технике, медицине имеется множество структур, требующих при их исследовании «тонкой» идентификации сложных законов стохастического поведения.

В иммунологии, в частности, обнаруживаются многомодальные распределения характеристик клеток крови и лимфы при медикаментозном и полевым воздействию [1]. Изучение таких структур требует обоснованных методов выявления законов распределения описывающих их случайных величин (СВ) и процессов. В биологии, медицине распространено использование числовых точечных параметров СВ – выборочного среднего и дисперсии выборки, а также гистограмм. Часто объем экспериментального материала из-за сложности и высокой стоимости (цена анализаторов для проточной цитофлуориметрии может достигать десяти и более миллионов рублей) недостаточен для обоснованного применения существующих методик проверки статистических гипотез, а необходимый объем выборки зависит от закона распределения СВ.

Кроме указанных способов оценки законов распределения стохастических показателей используется метод Парзена–Розенблатта [2]. В этом методе вся информация о случайной величине на этапе грубого оценивания содержится в оценке выборочной дисперсии. Данное обстоятельство определяет некорректность метода Парзена–Розенблатта, так как одно и то же значение дисперсии может быть получено как для одномодальных, так и для многомодальных распределений.

Важным обстоятельством является то, что в живой природе относительно много показателей имеют в норме широкий интервал допустимых значений. В этом отражается их зависимость от географического, временного факторов, генетических особенностей организмов. Так называемая «диффузность» – рассеяние изучаемых характеристик по диапазону значений – проявляется фактически при анализе большинства представленных выборок в биологии.

Поэтому законы распределения указанных и других показателей в иммунологии имеют специфические особенности в виде значительных уровней дисперсии, сложных законов распределения – многомодальных, негауссовых, негладкого типа [3]. Такого рода характеристики свидетельствуют о нелинейных эффектах в процессах образования/апоптоза клеток крови и лимфы, в частности, пролиферации Т-лимфоцитов [4]. При этом характерно, что изучаемые выборки часто имеют малый объем (от нескольких десятков отсчетов до ста).

Отмеченные особенности требуют соответствующих методов исследования. Основу подхода автора статьи составили методы восстановления плотностей распределения, заложенные школой академика А.Н. Тихонова для обратных задач математической физики [5]. Обратные задачи, как правило, имеют некорректность в постановке: множественность «решения» и его чувствительность к погрешности исходных данных. Решение такого рода задач основано на методах регуляризации.

В работе данный метод решения некорректно поставленных задач был применен в области практической медицины. Ряд результатов указанной школы был реализован в виде программного обеспечения для приближенного решения интегрального уравнения Фредгольма I рода. Подынтегральная функция плотности вероятности является искомой величиной. Правая часть уравнения соответствовала эмпирической функции распределения для каждого показателя, например, уровня лейкоцитов, В-лимфоцитов, иммуноглобулинов, фагоцитарных чисел. Всего учитывалось тридцать три иммунных показателя.

Использовались ограничения на решения – непрерывность законов распределения изучаемых иммунологических показателей, их сосредоточенность на некотором отрезке (по диапазону изменения), гладкость формы плотности распределения.

С учетом последнего фактора восстановление плотностей распределения всех иммунных показателей велось в классе тригонометрических функций с ограничением количества членов разложения N в зависимости от объема L наблюдаемых данных минимизацией гарантированного риска. Применение указанного подхода к обширному материалу иммунологических показателей позволило построить эмпирические законы распределения, классифицировать весь объем данных и свести его к структурированной и строгой системе.

Идентификация законов распределения

В таблице 1 для примера представлен фрагмент классифицированных данных, полученных в результате восстановления эмпирической плотности распределения для пациентов-мужчин. Рисунок 1 демонстрирует типичные графики восстановленных плотностей распределения вероятности по классам (результат интенсивной антибактериальной терапии). Рассмотрим их подробнее.

Как видно из представленных плотностей распределения иммунологических показателей,

только распределение № 9 имеет сходство с наиболее важным в теории вероятностей распределением – нормальным или близким к нему при определенных параметрах обобщенным распределением Рэлея. Последнее есть обобщение на случай, когда средние a_1 и a_2 не равны нулю, распределения на плоскости модуля вектора с координатами в виде независимых гауссовых случайных величин с одинаковыми дисперсиями. Указанное распределение имеет аналитическое выражение для плотности вероятности и находит широкое применение в статистической радиотехнике, радиофизике, механике сплошных сред, спектроскопии.

Обобщенное распределение Рэлея записывается таким образом:

$$p(x) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2 + a^2}{2\sigma^2}\right) I_0\left(\frac{xa}{\sigma^2}\right) \text{ при } x > 0, \\ p(x) = 0 \text{ при } x < 0,$$

где a – полярный радиус центра нормального распределения, $I_0\left(\frac{xa}{\sigma^2}\right)$ – функция Бесселя мнимого аргумента, σ^2 – дисперсия нормального компонента вектора.

Распределение № 4 по своему виду сходно с логарифмически нормальным, зеркально отображенным (относительно оси ординат) распределением. Данный закон в исходном виде хорошо описывает поведение микро- и мезочастиц в биологии, медицине, статистической физике пластично деформируемых сплошных сред (например, поликристаллических соединений металлов) [6]. Логарифмически нормальное распределение в теории вероятностей также представляется в аналитической форме.

Иммунологические показатели, распределенные по классу № 3, представляют, по видимому, предельную кривую плотности вероятности для класса № 4 при концентрации значений случайной величины в области нижней границы рассеяния.

Немногочисленный класс № 5 в первом приближении может быть описан любым односторонним распределением: треугольным (Симпсона), односторонним Гаусса (со значительной дисперсией), неравнобочным трапецеидальным распределением.

Шестой класс, за исключением области самых верхних значений, можно отнести к разряду распределений четвертого класса. Есть также основание для хорошей аппроксимации его формульной зависимостью вида

$$p(x) = \frac{a\lambda}{\pi} \left\{ \frac{1}{\lambda^2 + (x_0 + x)^2} + \frac{1}{\lambda^2 + (x_0 - x)^2} \right\},$$

Таблица 1

Классификация законов распределения иммунологических показателей, восстановленных методом решения обратных задач (пациенты – мужчины)

| 1 | Классификационные признаки и характеристики | | | | | | |
|---|---|---------------------|-------------------------|---|---|---------------------|--|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Иммунологические показатели (мед. норма) | Объем выборки, L | Класс распределения | Степень полимодальности | Аппроксимация нормальным распределением | Число членов разложения для $p(x) \sim N$ | Центр распределения | Вероятность попадания в интервал нормы |
| Лейкоциты, млн/л (4–9) | 71 | класс 7 | 1 | да | 4 | 8.70 | 0.65 |
| Лимфоциты, % (19–37) | 71 | класс 1 | 2 | нет | 4 | 28.4 | 0.69 |
| Лимфоциты, млн/л (0.7–3.8) | 71 | класс 7 | 1 | да | 4 | 2.43 | 0.91 |
| Нейтрофилы п/я, % (1–5) | 70 | класс 1 | 2 | нет | 4 | 2.23 | 0.59 |
| Нейтрофилы с/я, % (47–72) | 71 | класс 4 | 1 | нет | 3 | 59.4 | 0.69 |
| Эозинофилы, % (1–5) | 70 | класс 5 | 1 | нет | 2 | 2.46 | 0.42 |
| Моноциты, % (2–10) | 71 | класс 3 | 1 | нет | 2 | 7.45 | 0.45 |
| T-лимфоциты, % (40–90) | 40 | класс 2 | 3 | нет | 5 | 36.7 | 0.24 |
| T-лимфоциты, млн/л (0.5–3.0) | 38 | класс 7 | 1 | нет | 4 | 0.93 | 0.86 |
| B-лимфоциты, % (2–30) | 40 | класс 4 | 1 | нет | 2 | 15.0 | 0.97 |
| B-лимфоциты, млн/л (0.03–0.9) | 38 | класс 1 | 2 | нет | 5 | 0.40 | 0.97 |
| Нулевые клетки, % (2–35) | 40 | класс 4 | 1 | нет | 2 | 48.1 | 0.06 |

где a, λ – параметры распределения, x_0 – абсцисса наибольшего значения функции $p(x)$, т.е. мода. В теории спектрального оценивания стационарных случайных процессов указанной функцией определяется спектральная плотность для процессов с корреляционной функцией вида

$$k(\tau) = a \exp\{-\lambda|\tau|\} \cos(x_0 \tau).$$

Распределение, отнесенное к восьмому классу, встретилось один раз («индекс нагрузки» у женщин). Значения этого иммунологического показателя концентрируются в узкой области значений, а кривая распределения приближается по форме к дельта-функции. Можно предположить, что это нетипичная, квазидискретная форма распределения для иммунной системы – своего рода артефакт. Показатель «индекс нагрузки» – решение негладкого типа. Характерно, что два десятка членов не достигли аппроксимации на удовлетворительном уровне. Исследование возможностей аппроксимации

для такого рода явлений в иммунологии и других биологических системах, а также в технике, технологических процессах и управлении представляет собой актуальную задачу.

Классы распределений 1, 2 и 7 в некотором смысле примыкают друг к другу. Два первых относятся к семейству кругловершинных многомодальных, причем первый асимметричен и его плотность вероятности имеет две различных моды, а второй – полимодален с числом мод три и более. Моды высших порядков существенно ниже по своему уровню относительно главной моды. Следует отметить, что в этом классе обнаружено почти симметричное трехмодальное распределение (T-лимфоциты, % – мужчины). В первом приближении плотности вероятностей для классов 1, 2 можно трактовать как суперпозицию гауссовых, соответствующим образом масштабированных и разнесенных по области изменения изучаемых показателей.

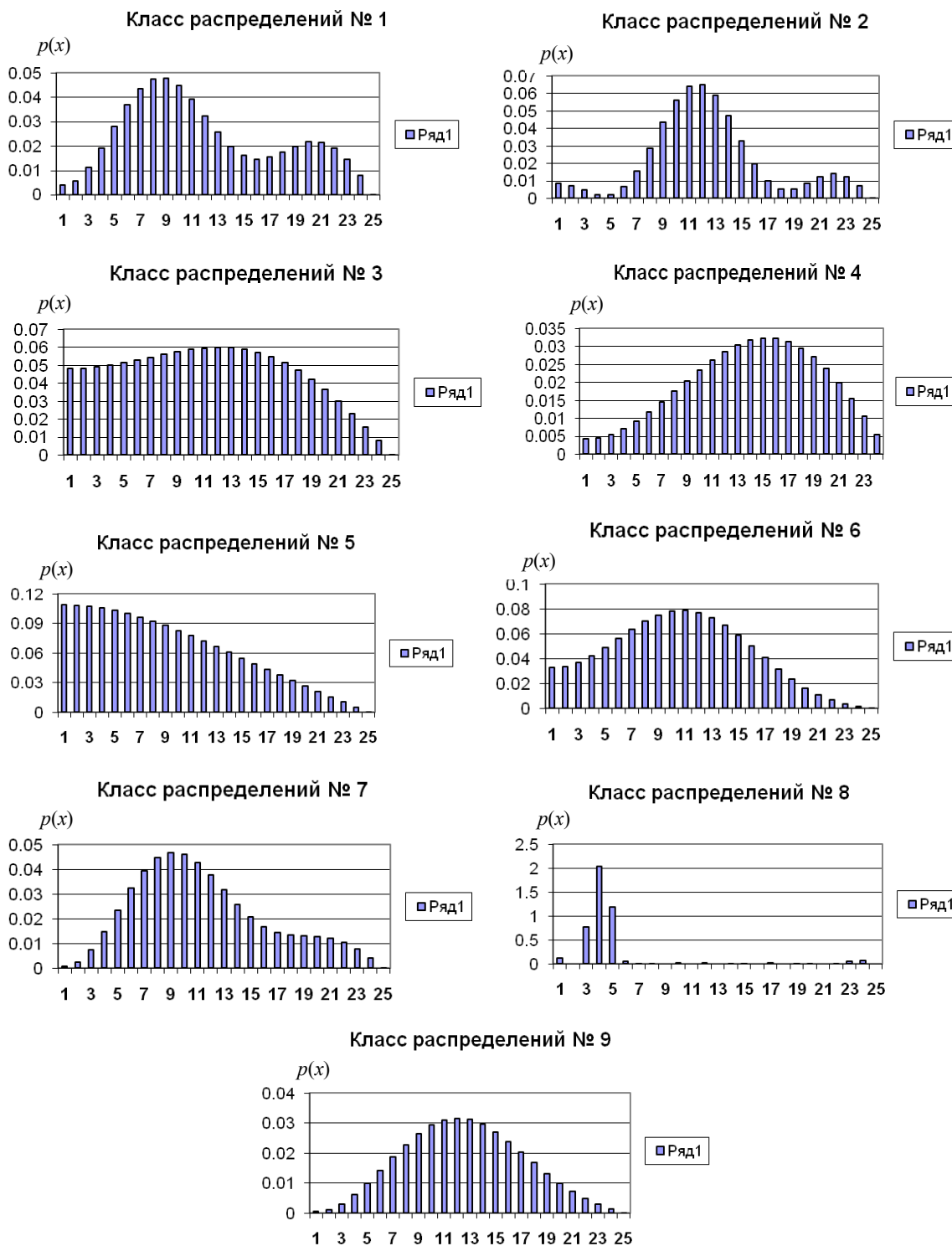


Рис. 1. Иммунологические показатели: примеры классов распределений, идентифицированных по выборкам малого объема методом регуляризации

Одномодальное распределение класса 7 может интерпретироваться как не проявленная по второй, меньшей моде, форма двухмодального распределения класса 1.

Принцип «сложности» $N=N(L)$ оценки плотности распределения получил наглядное выражение в количестве требуемых для решения гармоник –

минимум (3–7) для колоколообразных функций, в том числе содержащих несколько локальных мод; максимум (20–25) – для компактно локализованных (малые уровни дисперсии).

Полученные численные значения восстановленных функций в дискретных точках (необходимого объема) используются для вычисления

моментов любого порядка, а также энтропийных характеристик случайной величины, представляющей иммунологический параметр.

Для изучения функциональных состояний иммунной системы представляет интерес обнаружение многомодальных распределений у целого ряда показателей. В этом смысле статистическое (стохастическое по своей природе) «поведение» части иммунных тел после интенсивной антибактериальной терапии можно сравнить, в частности, с многоорбитальным (по энергиям) распределением возбужденных электронов в теории лазерных эффектов, двухфотонно-возбуждаемых нелинейных процессов излучения люминесцентных спектров в генетических структурах [7, 8]. Форма же распределения может ассоциироваться с волновыми функциями или распределением интенсивности для фраунгоферовой дифракционной картины, а также с другими фундаментальными физическими закономерностями.

В практическом плане сравнение известных оценок плотности вероятности иммунологических, гормональных и других показателей, исследование их трансформаций в состоянии здоровья и при терапевтических воздействиях позволит вести мониторинг методов лечения, а также выявить глубинные связи изучаемых явлений в клеточной биологии, микробиологии, клинической медицине с универсальными законами абиотического мира.

Кроме рассмотренного подхода исходный объем лабораторно-клинических данных – матрица размером примерно 80×30 – «пациенты–иммунные показатели» подвергался корреляционному анализу, подтвердившему многотаксонный характер полей корреляции и многомодальность целого ряда зависимостей.

Интерпретация данных на основе двоичной матрицы норм иммунологических показателей

Результаты предыдущего раздела позволяют сформулировать прикладной метод исследования иммунологических данных. Метод заключается в следующем.

1. Полагаем границы нормы по каждому показателю согласно официально признанным. При этом большинство показателей имеет весьма широкую норму, с отношением максимальных значений нормы к минимальным от 1.5 до 5 и выше.

2. Таблицу иммунологических показателей в исходном виде (нормализация результатов анализов) преобразуем в таблицу в двоичной форме, т.е. особым образом сжимаем исходную ин-

формацию. При этом в каждой ячейке двоичной матрицы записывается единица, если значение результата исходной ячейки лежит в пределах нормы, и нуль – в противоположном случае. Иначе говоря, нулевые значения элементам двоичной матрицы обеспечивают те отсчеты показателя, которые или ниже, или выше нормы. Фактический фрагмент двоичной матрицы для десяти пациентов и 12 показателей демонстрирует таблица 2.

3. Эти кодированные данные отличаются от исходных лабораторных тем, что максимально емко, т.е. только по двум градациям, представляют последние для дальнейшей обработки. Как показывается ниже, предлагаемый подход обладает значительной наглядностью и простотой при анализе в плоскости «пациент–показатель», естественным образом формирует алгоритмы и методы обработки иммунологических данных. Подобного рода бинарные матрицы применяются в теории графов.

4. С учетом изложенного выше, на основе двоичной иммунологической матрицы введем соответствующие функции. Лучше всего это демонстрируется в среде электронных таблиц Microsoft Office – Excel. Первая интегральная функция (рис. 2а) показывает распределение (%) количества пациентов, имеющих норму по m показателям ($m = 1, 2, 3, \dots$). В нашем случае число выбранных основных показателей равно 12. Например, когда пациенты – мужчины, гистограмма этой функции показывает, что пациенты с «числом норм» 6, 7, 8 и 9 составляют 88.1% – это соответственно 10, 11, 9 и 7 человек из 42.

Вторая интегральная функция (рис. 2б) характеризует число пациентов (%), имеющих норму по m и более ($m = 1, 2, 3, \dots$) показателям. Так, число мужчин с нормой по восьми и более показателям равно 40.5%. Обе функции формируются соответствующим подсчетом «единиц» по строкам двоичной иммунологической матрицы. Термин «интегральность» понимается в том смысле, что суммирование «единиц» по каждому пациенту ведется относительно всех показателей в целом, без выделения каких-то из них.

Определение числа норм по отдельным показателям, т.е. по столбцам, дает дифференциальную функцию иммунологической матрицы (рис. 2в). Данное распределение записано в предпоследней строке табл. 2. Эта функция характеризует результат по каждому отдельному показателю после лечения антибиотиками. Так, показатели 4, 5 и 7 у мужчин (нулевые клетки (%), фагоцитоз латекса (%), индекс нагрузки

Таблица 2

Фрагмент двойной матрицы иммунологических показателей и результаты обработки

| A | B | C | D | E | F | G | H | I | J | K | L |
|-------------|-------------|-------------|-------------|-------------|-------------|------------|-------------|-------------|-------------|-------------|-------------|
| 1 | 1 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 0 | 1 | 1/0 | 1/0 | 0 | 0 | 0 | 0 | 0 | 1/0 | 1/0 | 1/0 |
| 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1/0 | 1 | 1 |
| 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 1/0 | 1 | 1 |
| 1 | 1 | 1 | 0 | 0 | 1/0 | 0 | 0 | 1 | 1/0 | 1 | 1 |
| 1 | 1/0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 1/0 | 1/0 | 1 |
| 1 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 1/0 | 1 | 1 |
| 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 1 |
| 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 1 |
| 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1/0 | 1 | 1/0 | 1 | 1 |
| 88.1 | 83.3 | 95.2 | 11.9 | 14.3 | 54.8 | 4.8 | 33.3 | 57.1 | 81.0 | 97.6 | 95.2 |
| 78.6 | 76.2 | 85.7 | 0.0 | 14.3 | 42.9 | 0.0 | 14.3 | 42.9 | 7.1 | 83.3 | 64.3 |

A – лимфоциты, млн/л
 B – Т-лимфоциты, млн/л
 C – В-лимфоциты, млн/л
 D – нулевые клетки, %
 E – фагоцитоз латекса, %
 F – фагоцитарное число
 G – индекс нагрузки, ед.
 H – иммуноглобулин А, г/л
 I – иммуноглобулин М, г/л
 J – иммуноглобулин G, г/л
 K – аутоантитела к коллагену, титр
 L – аутоантитела к коже, титр

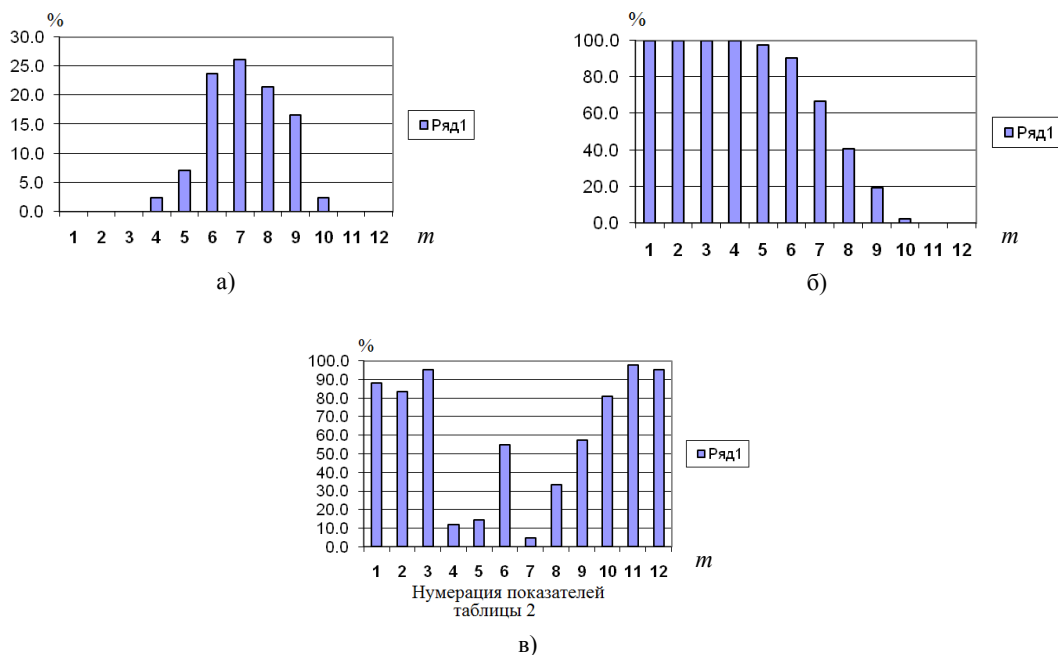


Рис. 2. Интегральные (а, б) и дифференциальная (в) функции (по показателям) двойной иммунологической матрицы по официальным нормам: а – число пациентов, имеющих норму по m ($m = 1, \dots, 12$) показателям, %; б – число пациентов, имеющих норму по m и более ($m = 1, \dots, 12$) показателям, %; в – число пациентов, имеющих норму по каждому из m показателей, %

(единиц) имеют уровень норм ниже 15% (11.9%, 14.3%, 4.8% соответственно). Уровень же других не опускается ниже 54.8%.

Введенные таким образом функции содержат классификационные аспекты для иммуно-

логических показателей и обладают важнейшим свойством – наглядно характеризуют глубину иммунодефицита после проведенной терапии, а также устойчивость результата относительно вариации границ официальной нормы.

Результаты исследований фактического иммунологического материала в рамках вероятностного подхода, введенные понятия двоичной иммунологической матрицы и производных от нее функций определяют дополнительные возможности анализа данных.

А именно, учитывая значительную величину поля нормы практически по всем показателям, необходимо определенным образом ввести ограничение на используемую норму и оценить полученные результаты по более жесткому критерию. Предлагаемый принцип позволяет ввести понятие устойчивости статистической характеристики иммунологического показателя и разработать алгоритм такой оценки.

Для этого вычисляем середину каждой официальной нормы как полусумму нижней и верхней границ; из таблицы 1 определяем выборочное среднее значение каждого показателя; по направлению его сдвига от официальной середины изменяем, например, на 20% соответствующую границу вовнутрь поля нормы. Таким образом, получается новая норма показателя, у которой, по сравнению с официальной, изменена одна из границ. При этом может увеличиться нижняя или уменьшиться верхняя граница исходной нормы.

Рассмотрим этот алгоритм на примере трех показателей для мужчин: 1) лимфоциты, млн/л; 2) Т-лимфоциты, млн/л; 3) иммуноглобулин G, г/л. Выборочное среднее (табл. 1) первого показателя 2.43 млн/л смещено вверх относительно середины нормы, равной 2.25 млн/л; второго – 0.93 млн/л – смещено вниз относительно середины нормы 1.75 млн/л; выборочное среднее третьего 11.1 г/л смещено вверх относительно середины нормы 6.085 г/л. Соответственно на 20% изменяются границы норм: верхняя граница первого показателя уменьшается до 3.04 млн/л, нижняя граница второго показателя увеличивается до 0.60 млн/л, верхняя граница третьего показателя уменьшается до 9.60 г/л. При этом другая граница каждого показателя остается без изменения, а интервал нормы становится меньше.

В качестве обоснования значения 20% коэффициента «ужесточения» вполне правомерно использовать классическое правило для оценки доверительного интервала выборочного среднего случайной величины.

Для этого рассчитаны двусторонние доверительные интервалы для каждого иммунологического показателя при вероятностях P_d , равных 0.90 и 0.95. При это число усредняемых отсчетов везде более 30. Другими словами, среднее значение любого показателя распределено по нормальному закону. Анализ этих вычислений

показывает, что значения доверительных интервалов лежат в пределах 15–30% от выборочного среднего. Отсюда 20% – средний интервал неопределенности выборочного среднего и значение смещения соответствующей границы нормы. Возможны также и другие варианты сжатия поля нормы.

После расчета новых норм по всем показателям заново определим двоичную иммунологическую матрицу, записывая «единицы» по попаданию исходного значения показателя уже в новые пределы. По сравнению с первичной, эта двоичная матрица становится более разреженной, т.е. содержит больше нулей (в таблице 2 новые данные указаны через наклонную черту). Так, для прежнего примера (пациенты-мужчины) доля «единиц» в матрице уменьшилась с 60% до 42%; а для женщин – с 59% до 43% по аналогичным совокупным показателям. При этом, возвращаясь к рассмотрению показателей у мужчин, видим, что вновь сформированные интегральные и дифференциальные функции претерпели значительные изменения (рис. 3). Максимум функции (а) сместился на $m = 5$ и теперь доминируют пациенты с числом норм по 4, 5 и 6 показателям, и их доля составляет 76.2%. Доля пациентов с числом норм 7, 8, 9 составляет всего 14.3%. Эти результаты отражаются также на быстром спаде интегральной функции (б).

«Дифференциальная функция» – число норм по отдельному показателю – также несет значительные «потери» – об этом наглядно свидетельствует гистограмма функции и процентные значения в нижней строке табл. 2. Такие показатели, как нулевые клетки, %, индекс нагрузки, ед., выпали из жесткой нормы полностью; фагоцитоз латекса, %, остался на прежнем уровне; остальные показатели снизили уровень норм от 10 до 30%, но самый значительный спад наблюдается у иммуноглобулина G, г/л (J-показатель).

Число пациентов с нормой по данному показателю сократилось с 81.0% до 7.1%. Объяснение данному результату дают табл. 1 и вид плотности вероятностного распределения. Этот показатель, в среднем оставаясь в пределах исходной нормы, сосредоточен вблизи верхней границы и имеет малое рассеяние. Поэтому в поле широкой нормы для имеющейся выборки результат по иммуноглобулину G вполне положительный. Но, учитывая допустимую погрешность выборочного результата, необходимо иметь запас достоверности на величину неопределенности (в среднем 20%) и посмотреть на выборку с более жестких позиций, т.е. по норме (9.60–0.18 г/л).

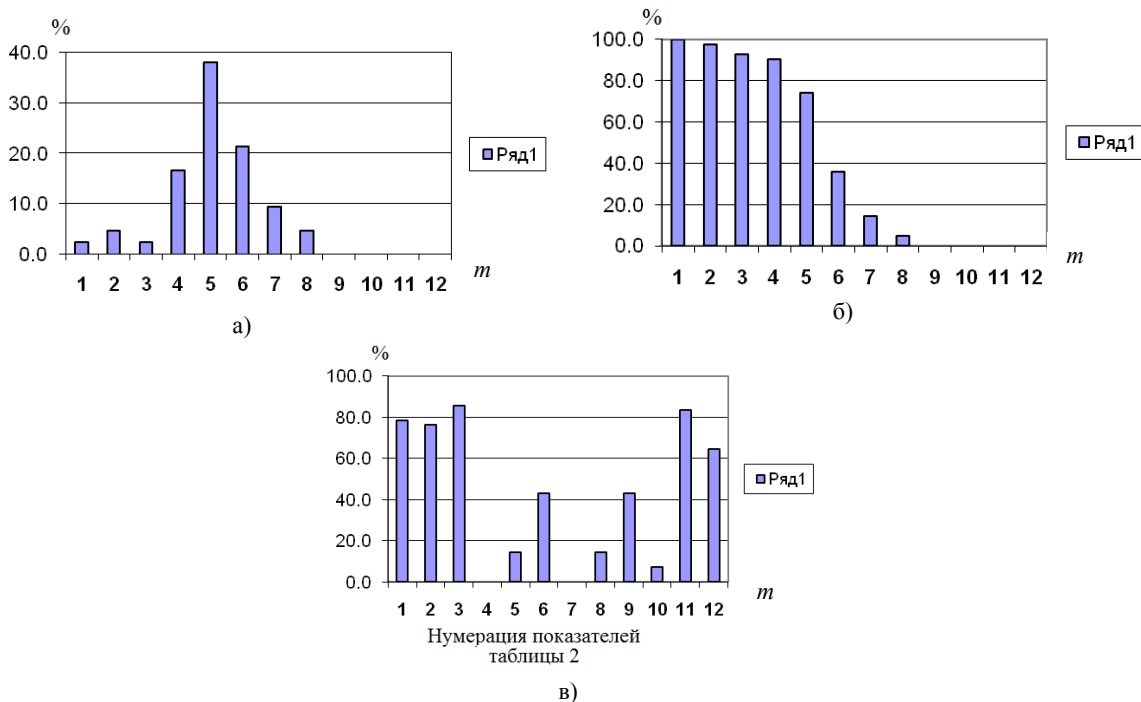


Рис. 3. Интегральные (а, б) и дифференциальная (по показателям) функция (в) двоичной иммунологической матрицы по «сжатым» нормам: а – число пациентов, имеющих норму по m ($m = 1, \dots, 12$) показателям %; б – число пациентов, имеющих норму по m и более ($m = 1, \dots, 12$) показателям, %; в – число пациентов, имеющих норму по каждому из m показателей, %

Если и в этом случае величина дифференциальной функции получается значительной, то делается вывод об устойчивости благоприятного результата по данному иммунологическому показателю. Очевидно, что иммуноглобулин G к таким (см. вышеприведенную динамику) отнести нельзя.

Напротив, ряд из пяти других показателей, а именно: 1) лимфоциты, млн/л; 2) Т-лимфоциты, млн/л; 3) В-лимфоциты, млн/л; 4) аутоантитела к коллагену, титр; 5) аутоантитела к коже, титр, – образует группу показателей, устойчивых относительно вариации границ нормы. После проведенного «испытания» первые три имеют уровень норм не ниже 75%, четвертый – 83%, пятый – 64%. Можно охарактеризовать их как «надежно благополучные».

Показатели следующего ряда: 1) нулевые клетки, %; 2) индекс нагрузки, ед.; 3) фагоцитоз латекса, %; 4) иммуноглобулин А, г/л; 5) иммуноглобулин G, г/л, – можно отнести к показателям с подавленной нормой. Следующие показатели – фагоцитарное число, ед., и иммуноглобулин М, г/л, – достаточно устойчивы в классе со средним уровнем норм порядка 50%. Другими словами, информационная энтропия их близка к максимальной величине.

Рассмотрение указанных иммунологических показателей у женщин подтверждает обнаруженную закономерность. Учитывая независимость этой выборки от мужской, получаем до-

полнительные аргументы в обоснование выбранного подхода к классификации иммунологических данных.

Таким образом, предлагаемый подход содержит в себе дополнительную степень свободы при изучении характеристик иммунной системы по ограниченным выборкам. При этом в условиях, когда, зачастую, нет возможности увеличить объем выборки, он позволяет более обоснованно и надежно интерпретировать результаты терапии и состояние пациентов. На основе функций, получаемых от двоичных матриц, производится не только анализ, но и классификация иммунологических показателей по указанному принципу. На основе двоичной матрицы «норм» автором развивается метод оценивания состояния иммунной системы, который предполагается изложить в следующей публикации.

Кроме того, рассмотренный принцип конструктивен в том смысле, что предлагает расчетный алгоритм оценки. В целом, полученные выводы важны не только в исследовательском плане, но и в практической сфере для преодоления «тирании многомерности» экспериментального материала.

Заключение

Выполненная идентификация эмпирических двух- и полимодальных распределений по выборкам малого объема изложенным способом позволяет считать его перспективным в области

медицины, биологии и других сферах естествознания. При этом исследование на модельных примерах показывает, что более простой, альтернативный метод Парзена–Розенблатта по эффективности и разрешающей способности значительно уступает применяемому подходу.

Дальнейшее исследование возможностей математических методов для реконструкции распределений, описываемых негладкими плотностями (фрактальная особенность структур с нелинейной динамикой), актуально. В частности, для законов распределения, не имеющих производных (или разрывных), перспективной представляется аппроксимация системами функций Хаара или Уолша при корректировке условий алгоритма регуляризации исходной задачи.

Предложенная интерпретация данных на основе двоичной матрицы «норм» дает дополнительную степень верификации клинического материала.

Список литературы

1. Куликов В.Б. Многомодальные законы распределения случайных величин в сложных стохастических системах и их идентификация // Матер. XIX междунар. научн.-практ. конф. «Информационные системы и технологии» ИСТ-2013, Н. Новгород, 16–18 апреля 2013 г. С. 246–247.
2. Поршев С.В., Копосов А.С. Использование аппроксимации Розенблатта–Парзена для восстановления непрерывной случайной величины с ограниченным одномодальным законом распределения // Научный журнал КубГАУ. 2013. № 92(08). С. 1–14.
3. Зорин Д.Ю., Иванченко М.В. Динамика иммунного ответа в многоклональных популяциях Т-клеток: роль иммунопротеасомы // Вестник Нижегородского университета им. Н.И. Лобачевского. 2012. № 5(2). С. 97–101.
4. Бочаров Г.А., Лузянина Т.Б., Розе Дирк. Математические технологии анализа пролиферации Т-лимфоцитов по данным проточной цитофлуориметрии // Российский иммунологический журнал. 2009. Т. 3(12). № 1. С. 13–22.
5. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука, 1986. 288 с.
6. Аратский Д.Б., Леонтьев Е.А., Морозов О.А. и др. Информационно-оптимальные методы в физике и обработке экспериментальных данных. Монография. Н. Новгород: Издательство Нижегородского университета, 1992. 146 с.
7. Агалцов А.М., Гаряев П.П., Горелик В.С. и др. Двухфотонно-возбуждаемая люминесценция в генетических структурах // Квантовая электроника. 1996. Т. 23. № 2. С. 181–184.
8. Kawabe Y., Wang L., Nakamura T., Ogata N. Thin-film lasers based on dye-deoxyribonucleic acid-lipid complexes // Applied Physics Letters. 2002. V. 81. Issue 8. August 19. P. 1372–1374.

RESTORATION OF MULTIMODAL PROBABILITY DENSITY FUNCTIONS FROM EXPERIMENTAL DATA IN STRUCTURES WITH STOCHASTIC PROPERTIES

V.B. Kulikov

The restoration of probability density functions for main indicators in immunology is considered. The human immune system is described in terms of a stochastic process. An approximate solution of the Fredholm integral equation of the first kind is taken as an identifiable probability distribution. The regularization of the mentioned inverse problem is applied to the family of smooth (trigonometric) functions from volume-limited samples of clinical analyses. The proposed approach allows successful restoration of multimodal probability density functions and is universal in the field of natural sciences.

Keywords: multimodal laws of probability distribution, random variable, volume-limited samples, immunology, inverse problem, Tikhonov regularization method.

References

1. Kulikov V.B. Mnogomodal'nye zakony raspredelenija sluchajnyh velichin v slozhnyh stohasticheskikh sistemah i ih identifikacija // Mater. XIX mezhdunar. nauchn.-prakt. konf. «Informacionnye sistemy i tehnologii» IST-2013, N. Novgorod, 16–18 aprelja 2013 g. S. 246–247.
2. Porshev S.V., Kuposov A.S. Ispol'zovanie approksimacii Rozenblatta–Parzena dlja vosstanovlenija nepreryvnoj sluchajnoj velichiny s ogranichenym odnomodal'nyim zakonom raspredelenija // Nauchnyj zhurnal KubGAU. 2013. № 92(08). S. 1–14.
3. Zorin D.Ju., Ivanchenko M.V. Dinamika immunnogo otveta v mnogoklonal'nyh populacijah T-kletok: rol' immunoproteasomy // Vestnik Nizhegorodskogo universiteta im. N.I. Lobachevskogo. 2012. № 5(2). S. 97–101.
4. Bocharov G.A., Luzjanina T.B., Roze Dirk. Matematicheskie tehnologii analiza proliferacii T-limfocitov po dannym protochnoj citofluorimetrii // Rossijskij immunologicheskij zhurnal. 2009. T. 3(12). № 1. S. 13–22.
5. Tihonov A.N., Arsenin V.Ja. Metody reshenija nekorrektnyh zadach. M.: Nauka, 1986. 288 s.
6. Aratskij D.B., Leont'ev E.A., Morozov O.A. i dr. Informacionno-optimal'nye metody v fizike i obrabotke jeksperimental'nyh dannyh. Monografija. N. Novgorod: Izdatel'stvo Nizhegorodskogo universiteta, 1992. 146 s.
7. Agal'cov A.M., Garjaev P.P., Gorelik V.S. i dr. Dvuhfotonno-vozbuzhdaemaja ljuminescencija v geneticheskikh strukturah // Kvantovaja jelektronika. 1996. T. 23. № 2. S. 181–184.
8. Kawabe Y., Wang L., Nakamura T., Ogata N. Thin-film lasers based on dye-deoxyribonucleic acid-lipid complexes // Applied Physics Letters. 2002. V. 81. Issue 8. August 19. P. 1372–1374.