РАДИОТЕХНИКА И ЭЛЕКТРОНИКА, 2024, том 69, № 1, с. 76–87

ТЕОРИЯ И МЕТОДЫ ОБРАБОТКИ СИГНАЛОВ

УДК 621.391.2

НОВЫЕ МЕТОДЫ ПРИНЯТИЯ СТАТИСТИЧЕСКИХ РЕШЕНИЙ В УСЛОВИЯХ ОГРАНИЧЕННОГО ОБЪЕМА НАБЛЮДЕНИЙ И ПРИ АПРИОРНОЙ ПАРАМЕТРИЧЕСКОЙ НЕОПРЕДЕЛЕННОСТИ

 Ф. А. Мкртчян

Фрязинский филиал Института радиотехники и электроники им. В.А. Котельникова РАН, пл. Введенского, 1, Фрязино Московской обл., 141190 Российская Федерация

E-mail: ferd47@mail.ru

Поступила в редакцию 20.03.2023 г.
После доработки 20.03.2023 г.
Принята к печати 31.03.2023 г.

Разработан новый обобщенный адаптивный алгоритм обучения принятию статистических решений для экспоненциальных семейств распределений при априорной параметрической неопределенности в условиях выборок малого объема. Приведено обобщенное решающее правило, полученное методом оценки неизвестных параметров распределений, а также решающее правило, удовлетворяющее необходимым условиям оптимальности: постоянству средней вероятности ошибки первого рода и несмещенности. Рассмотрены конкретные решающие процедуры для частных распределений, полученных от обобщенного алгоритма. Приведены численные примеры. Показана эффективность разработанной оптимальной процедуры для выборок малого объема.

Ключевые слова: априорная параметрическая неопределлнность, обобщенный адаптивный алгоритм, экспоненциальное семейство распределений, выборки малого объема

DOI: 10.31857/S0033849424010064, **EDN:** KZVMIA

ВВЕДЕНИЕ

Первостепенное значение при мониторинге окружающей среды приобретают организация массового сбора информации об изучаемой системе, оперативность ее обработки и достоверная интерпретация данных наблюдений на основе аналитических и численных математических моделей.

Современный этап развития экспериментальных радиофизических и оптических методов исследования окружающей среды характеризуется переходом от пассивного сбора информации об изучаемом объекте к постановке целенаправленных экспериментов. С практической точки зрения важным является синтез комплексной системы сбора и обработки информации об окружающей среде, объединяющей дистанционные и контактные измерения, составляющие основу систем геоинформационного мониторинга.

Основной смысл концепции дистанционного мониторинга состоит в соединении в систему средств сбора данных, методов их обработки, математических моделей природных объектов, компьютерных средств реализации алгоритмов и моделей с широким спектром сервисного обеспечения при визуализации результатов мониторинга [1–3].

Эффективное решение этих задач невозможно без широкого внедрения в практику исследований автоматизированных систем сбора, хранения и обработки данных на базе современных компьютеров с применением технологии открытых систем.

В последнее время разрабатываются методы и алгоритмы компьютерного анализа двумерных изображений земной поверхности. Ведется работа по построению моделей формирования этих двумерных полей и решаются задачи классификации явлений, анализа изображений на изучаемом пространстве. Уже созданные методы и алгоритмы обладают способностью преодолевать такие трудности, как отрывочность и нестационарность информации, наличие малых статистически неоднородных выборок [1–3].

В настоящее время основной тенденцией в построении крупных проблемно-ориентированных информационных систем является использование распределенных баз данных и знаний, использование компьютеров различных классов и производителей, использование локальных и глобальных сетей. При этом возникают сложности в использовании информации баз данных и знаний, реализованных на различных системах управления базами данных (СУБД) и программного обеспечения, разработанного на разных платформах [1–3].

Преодоление указанных сложностей основано на применении технологии открытых информационных систем, использующей стандартные интерфейсы между всеми программно-аппаратными компонентами среды. Важнейшим этапом является построение профиля – набора согласованных стандартов для данной области применения.

Развитие систем дистанционного мониторинга окружающей среды требует решения ряда задач организации потоков данных измерений. Среди этих задач одной из важных является задача принятия статистического решения о наличии на обследуемой части земной поверхности того или иного явления. Одной из особенностей условий сбора информации для такого решения является невозможность получения статистических выборок больших объемов. Поэтому необходимы разработка и исследование оптимальных алгоритмов принятия статистических решений для выборок малого объема при информационных ограничениях.

Для случая, когда число наблюдений достаточно велико, задача решается методом оценки параметров вероятностных распределений. Этот метод эффективен при неограниченном росте объемов выборок, на основе которых производится оценка параметров. При ограниченных объемах выборок полученное методом оценки параметров решающее правило не удовлетворяетнеобходимымусловиямоптимальности: постоянству средней вероятности ошибки первого рода и несмещенности.

Цель данной работы – разработка нового обобщенного адаптивного алгоритма обучения принятию статистических решений для экспоненциальных семейств распределений при априорной параметрической неопределенности в условиях выборок малого объема.

1. ПОСТАНОВКА ЗАДАЧИ

Пусть имеются три случайные величины $\xi_i = (i = 1, 2, 3)$ с плотностями распределения вероятностей $f_{\xi_i}\left(x^{(i)} / \omega_i\right)$ $(i = 1, 2, 3)$, известные с точностью до значения параметров $\omega_i \in \Omega(i = 1, 2, 3)$, при этом для значения параметра ω_3 имеются две альтернативы: 1) $H_1 : \omega_3 = \omega_1$ и 2) $H_2 : \omega_3 = \omega_2$.

Задача состоит в построении решающего правила, которое позволяло бы по n_i ($i = 1, 2, 3$) независимым наблюдениям случайных величин

$$\xi_1 : x^{(i)*} = \left(x_1^i, x_2^i, \dots, x_{n_i}^i\right), (i = 1, 2, 3), \tag{1}$$

указать, какая из альтернатив, H_1 или H_2, выбирается в качестве решения.

Решающее правило может быть задано с помощью бинарной функции $\varphi\left(x^{(1)*}, x^{(2)*}, x^{(3)*}\right)$, причем если $\varphi\left(x^{(1)*}, x^{(2)*}, x^{(3)*}\right) = 0$, то принимается гипотеза H_1, если же $\varphi\left(x^{(1)*}, x^{(2)*}, x^{(3)*}\right) = 1$, то принимается гипотеза H_2. В поставленной выше задаче выборочные значения $x^{(1)*} = \left(x_1^i, x_2^i, \dots, x_n^i\right)$, $(i = 1, 2)$, случайных величин $\xi_i = (i = 1, 2)$ являются обучающими. При фиксированных значениях параметров ω_1 и ω_2 решающее правило указанного типа со статистической точки зрения характеризуется двумя величинами [1–5]:

1) $\alpha\left(\varphi, \omega_1, \omega_2, x^{(1)*}, x^{(2)*}\right)$ – вероятностью ошибки первого рода, т.е. принятия гипотезы H_2, в случае когда имеет место гипотеза H_1;

2) $\beta\left(\varphi, \omega_1, \omega_2, x^{(1)*}, x^{(2)*}\right)$ – вероятностью совершения ошибки второго рода, т.е. принятия гипотезы H_1, в случае когда имеет место гипотеза H_2.

Аналитически α и β будут выражаться следующим образом:

$$\alpha\left(\varphi, \omega_1, \omega_2, x^{(1)*}, x^{(2)*}\right) =$$
$$= \int \varphi\left(x^{(1)*}, x^{(2)*}, x^{(3)*}\right) f_{\xi_3}\left(x^{(3)*} / \omega_1\right) dx^{(3)*},$$

$$\beta\left(\varphi,\ \omega_1,\ \omega_2,\ x^{(1)*},\ x^{(2)*}\right) =$$
$$= \int\left[1-\varphi\left(x^{(1)*},\ x^{(2)*},\ x^{(3)*}\right)\right] f_{\xi_3}\left(x^{(3)*}/\ \omega_2\right) dx^{(3)*}.$$

Поскольку α и β являются случайными величинами, то для характеристики качества применяемого решающего правила используют математические ожидания $\bar{\alpha}(\varphi,\ \omega_1,\ \omega_2)$ и $\bar{\beta}(\varphi,\ \omega_1,\ \omega_2)$ [1–5]:

$$\begin{gathered}\bar{\alpha} = \iiint \varphi\left(x^{(1)*},\ x^{(2)*},\ x^{(3)*}\right) f_{\xi_1}\left(x^{(1)*}/\ \omega_1\right) f_{\xi_2} \times \\ \times\left(x^{(2)*}/\ \omega_2\right) f_{\xi_3}\left(x^{(3)*}/\ \omega_1\right) dx^{(1)*} dx^{(2)*} dx^{(3)*} \\ \bar{\beta} = \iiint\left[1-\varphi\left(x^{(1)*}, x^{(2)*}, x^{(3)*}\right)\right] f_{\xi_1}\left(x^{(1)*}/\ \omega_1\right) f_{\xi_2} \times \\ \times\left(x^{(2)*}/\ \omega_2\right) f_{\xi_3}\left(x^{(3)*}/\ \omega_2\right) dx^{(1)*} dx^{(2)*} dx^{(3)*}.\end{gathered} \tag{2}$$

В некоторых случаях из соображений чисто аналитического характера ограничиваются рассмотрением класса решающих правил, для которых равенство

$$\bar{\alpha}(\varphi,\ \omega_1,\ \omega_2) = \alpha_0 \tag{3}$$

выполняется при всех возможных значениях параметров ω_1 и ω_2. Этот класс называется классом подобных решающих правил уровня α_0.

Если существует решающее правило φ^*, такое что неравенство

$$\bar{\beta}\left(\varphi^*,\ \omega_1,\ \omega_2\right) \le \bar{\beta}(\varphi,\ \omega_1,\ \omega_2)$$

имеет место при любых возможных значениях параметров ω_1, ω_2 и любых решающих правил из класса (в предположении, что как φ^*, так и φ принадлежат к рассматриваемому классу), то такое решающее правило φ^* называется равномерно наиболее мощным решающим правилом. Впервые П.П. Данковым [4] была найдена оптимальная процедура обучения ЭВМ принятию статистических решений для нормально распределенных случайных величин с неизвестными средними и известными дисперсиями [4, 5]. Поскольку равномерно наиболее мощное решающее правило существует только для весьма ограниченного класса задач с параметрической априорной неопределенностью, то на практике при построении тех или иных решающих правил принимают одно из указанных выше (3) требований. Достаточно часто ограничиваются требованием выполнения двух обязательных условий оптимальности: а) постоянство средней вероятности ошибки первого рода $\bar{\alpha}$ (условие подобия (3)), б) несмещенность:

$$1-\bar{\beta} \ge \bar{\alpha}, \tag{4}$$

т.е. средняя вероятность правильного обнаружения не должна быть меньше средней вероятности ложных тревог.

Как будет показано далее, решающие процедуры, основанные на теории оценок параметров, не удовлетворяют указанным двум обязательным условиям оптимальности. Поэтому возникает задача улучшения процедуры различения классов сигналов по ограниченному ряду наблюдений выборок каждого класса. Тем не менее ниже предложена методика улучшения процедуры, основанная на теории оценки параметров. В результате для экспоненциальных семейств распределений получены процедуры, удовлетворяющие двум обязательным условиям оптимальности.

2. НОВЫЕ ОПТИМАЛЬНЫЕ ПРОЦЕДУРЫ ПРИНЯТИЯ СТАТИСТИЧЕСКИХ РЕШЕНИЙ ДЛЯ ЭКСПОНЕНЦИАЛЬНЫХ СЕМЕЙСТВ РАСПРЕДЕЛЕНИЙ

Предположим, что случайные величины из (1) имеют следующие плотности распределения вероятностей

$$f_{\xi_1}\left(x^{(i)}\ /\ \omega_i\right) = u\left(\omega_i\right)\omega\left(x^i\right)\exp\left[-g\left(\omega_i\right)\nu\left(x^i\right)\right],$$
$$i = 1,2,3, \tag{5}$$

где u, ω, g, ν - действительные монотонные функции [1–5].

Можно показать, что совместное условное распределение наблюдаемых последовательностей $x^{(i)*}\left(x_1^{(i)},\ldots,x_{n_i}^{(i)}\right)$ $(i = 1,\ 2,\ 3)$ случайных величин $\xi_i\left(i = 1,2,3\right)$ при условии, что заданы значения статистик

$$x_i = \sum_{j=1}^{n_i} \nu\left(x_j^{(i)}\right),\ \ i = 1,\ 2,\ 3, \tag{6}$$

не зависит от параметров ω_i ($i = 1,\ 2,\ 3$). Следовательно, система трех случайных величин $\left(x_1,\ x_2,\ x_3\right)$ является достаточной статистикой для параметров ω_1, ω_2, ω_3 и решающая функция зависит только от значений этих статистик. В качестве конкретного вида для ν и ω рассмотрим степенные функции $\nu\left(t\right) = t^m$ и $\omega\left(t\right) = t^k$, часто используемые в радиофизических экспериментах для аппроксимации данных измерений. В дальнейшем нам необходимо иметь вид распределения статистик (6). Для нахождения распределения суммы

$$x_3 = \sum_{j=1}^{n_3} \nu\left(x_j^{(3)}\right)$$

сначала найдем распределение случайной функции $\nu\left(x_j^{(3)}\right) = \left(x_j^{(3)}\right)^m$, а затем с помощью метода математической индукции определим и распределим суммы. Поскольку случайная величина $x_j^{(3)}$ имеет распределение

$$f_{\xi_3}\left(x_j^{(3)} / \omega_3\right) = u(\omega_3)\left(x_j^3\right)^k \exp\left[-g(\omega_3)\left(x_j^3\right)^m\right], \quad (7)$$

то распределение функции $\nu\left(x_j^{(3)}\right) = \left(x_j^{(3)}\right)^m$ можно найти посредством следующего выражения [5]:

$$q\left[\left(x_j^{(3)}\right)^m\right] = f\left[v^{-1}\left(x_j^3\right)\right]\left|v'^{-1}\left(x_j^3\right)\right|,$$

где $v^{-1}\left(x_j^{(3)}\right)$ – функция, обратная v.

В результате получаем

$$q\left[\left(x_j^{(3)}\right)^m\right] = \frac{u(\omega_3)}{m}\left[\left(x_j^{(3)}\right)^m\right]^{(k-m+1)/m} \times$$
$$\times \exp\left[-g(\omega_3)\left(x_j^{(3)}\right)^m\right].$$

Исходя из выражения для распределения композиции случайных величин [5] и используя математическую индукцию, можно показать, что распределение суммы

$$x_3 = \sum_{j=1}^{n_3} \nu\left(x_j^{(3)}\right) = \sum_{j=1}^{n_3} \nu\left(x_j^{(3)}\right)^m$$

описывается функцией

$$\nu(x_3 / \omega_3) = \frac{u^{n_3}(\omega_3)}{m^{n_3}} \exp\left[-g(\omega_3)x_3\right] \times$$
$$\times x_3^{(n_3 k - m + n_3)/m} \frac{\left[\Gamma\left(\frac{k+1}{m}\right)\right]^{n_3}}{\Gamma\left(\frac{n_3(k+1)}{m}\right)}. \quad (8)$$

Для

$$x_1 = \sum_{j=1}^{n_1}\left(x_j^{(1)}\right)^m \text{ и } x_2 = \sum_{j=1}^{n_1}\left(x_j^{(2)}\right)^m$$

распределения $\nu(x_1 / \omega_1)$, $\nu(x_2 / \omega_2)$ имеют вид аналогичный (8). Совместная плотность распределения вероятностей независимых случайных величин x_1, x_2, x_3 запишется в форме

$$h(x_1, x_2, x_3 / \omega_1, \omega_2, \omega_3) = \nu(x_1 / \omega_1)\nu(x_2 / \omega_2)\nu(x_3 / \omega_3) =$$
$$= \frac{u(\omega_1)^{n_1} u(\omega_2)^{n_2} u(\omega_3)^{n_3} x_1^{n_1'-1} x_2^{n_2'-1} x_3^{n_3'-1}\left[\Gamma\left(\frac{k+1}{m}\right)\right]^{n_1+n_2+n_3} \mu}{m^{n_1+n_2+n_3}\Gamma(n_1')\Gamma(n_2')\Gamma(n_3')},$$

при $x_1, x_2, x_3 > 0$, где

$$n_1' = \frac{n_1(k+1)}{m}; \; n_2' = \frac{n_1(k+1)}{m}; \; n_3' = \frac{n_3(k+1)}{m};$$
$$\mu = \exp\left[-g(\omega_1)x_1 + g(\omega_2)x_2 + g(\omega_3)x_3\right].$$

3. РЕШАЮЩЕЕ ПРАВИЛО, ОСНОВАННОЕ НА ОЦЕНКЕ ПАРАМЕТРОВ РАСПРЕДЕЛЕНИЙ

Классический метод решения задачи различения двух гипотез основывается на достаточно развитой теории точечных оценок параметров вероятностных распределений. Понятие "оценка" совпадает с понятием "статистика", т.е. с понятием произвольной (измеримой) функции наблюдаемых значений случайной величины. Для рассматриваемого распределения (5) с неизвестными параметрами ω_1, ω_2 и ω_3 по методу максимума правдоподобия [3–5] можно получить следующие выражения для оценок этих параметров:

$$L\left(x_1^{(i)}, \ldots, x_{n_i}^{(i)}\right) = \prod_{j=1}^{n_i} f_{\xi_3}\left(x_j^{(i)} / \omega_i\right) u(\omega_i)^{n_i}$$
$$\prod_{j=1}^{n_i} \omega\left(x_j^{(i)}\right) \exp\left[-g(\omega_i)\sum_{j=1}^{n_i} \nu\left(x_j^{(i)}\right)\right], \quad i = 1, 2, 3.$$

Из условия $\partial \ln L / \partial \omega_1 = 0$ получим

$$\frac{u'(\omega_i)}{u'(\omega_i)g(\omega_i)} = \frac{\sum_{j=1}^{n_i} \nu\left(x_j^{(i)}\right)}{n_i}, \; i = 1,2,3.$$

Задавая конкретные виды функции $u(\omega) = a\omega^L$ и $g(\omega) = b\omega^M$, где a, b, M, L – действительные числа, получим

$$\omega_i^M = \frac{n_i L}{bM \sum_{j=1}^{n_i} \nu\left(x_j^{(i)}\right)}, \; i = 1,2,3. \tag{9}$$

Такая оценка параметра θ является достаточной, несмещенной, эффективной, а последовательность этих оценок является состоятельной [1, 2, 5], и, следовательно, оценку нельзя улучшить. Однако ниже будет показано, что решающее правило, получаемое с помощью таких оценок, не является наилучшим и, более того, не удовлетворяет основным требованиям, предъявляемым к решающим правилам, если только объемы выборок не увеличиваются неограниченно. Применение метода оценки параметров вероятностных распределений для получения решающего правила выполняется по следующей схеме.

По результатам наблюдений $x^{(1)*} = \left(x_1^{(1)},...,x_{n_1}^{(1)}\right)$ и $x^{(2)*} = \left(x_1^{(2)},...,x_{n_2}^{(2)}\right)$ случайных величин ξ_1 и ξ_2 производится оценка параметров соответствующих распределений. В нашем случае, как следует из (9), для параметров ω_1 и ω_2 имеем соответственно следующие оценки:

$$\theta_1 = \left(\frac{n_1 L}{bMx_1}\right)^{1/M}, \; \theta_2 = \left(\frac{n_2 L}{bMx_2}\right)^{1/M}. \tag{10}$$

Далее все делается так, как если бы параметры распределений при альтернативах H_1 и H_2 действительно имели значение θ_1 и θ_2 соответственно.

Способ построения решающего правила для значений $x^{(3)*} = \left(x_1^{(3)},...,x_{n_3}^{(3)}\right)$ случайной величины ξ_3, когда их распределение как при гипотезе H_1, так и при гипотезе H_2 заданы, основан на фундаментальной лемме Неймана–Пирсона. Строится отношения правдоподобия

$$\Lambda\left(x^{(3)*} / \theta_1,\theta_2\right) = \frac{f_{\xi_3}\left(x_1^{(3)},...,x_{n_3}^{(3)} / \theta_2\right)}{f_{\xi_3}\left(x_1^{(3)},...,x_{n_3}^{(3)} / \theta_1\right)}$$

и выбирается порог $c(\theta_1,\theta_2)$, определяющий область принятия гипотезы H_2, т.е. $\Lambda\left(x^{(3)*}/\theta_1,\theta_2\right) > c(\theta_1,\theta_2)$.

Порог $C(\theta_1,\theta_2)$ выбирается таким образом, чтобы $\alpha = \alpha_0$:

$$\Lambda\left(x^{(3)*} / \theta_1,\theta_2\right) = \left[\frac{u(\theta_2)}{u(\theta_1)}\right]^{n^3} \times$$
$$\times \exp\left\{-\left[g(\theta_2) - g(\theta_1)\right]\sum_{j=1}^{n_3} \nu\left(x_j^3\right)\right\} > c, \tag{11}$$

а

$$\ln \Lambda\left(x^{(3)*} / \theta_1,\theta_2\right) = n_3 \ln\left[\frac{u(\theta_2)}{u(\theta_1)}\right] -$$
$$-\left[g(\theta_2) - g(\theta_1)\right]\sum_{j=1}^{n_3} \nu\left(x_j^{(3)}\right) > \ln c.$$

При конкретных видах функции

$$\nu(z_i) = z_i^m; \; u(\theta_1) = a\theta_1^L;$$
$$u(\theta_2) = a\theta_2^L; g(\theta_1) = b\theta_1^M; g(\theta_2) = b\theta_2^M \tag{12}$$

имеем

$$\ln \Lambda = n_3 \ln\left(\frac{\theta_2}{\theta_1}\right)^L -$$
$$- b\left(\theta_2^M - \theta_1^M\right)\sum_{j=1}^{n_3} \nu\left(x_j^{(3)}\right)^m > \ln c. \tag{13}$$

Далее следует рассматривать два случая.

Первый случай: $\theta_1 < \theta_2$. Из (13) получим

$$\sum_{j=1}^{n_3}\left(x_j^{(3)}\right)^m > \ln\left[\frac{1}{c}\left(\frac{\theta_2}{\theta_1}\right)^{n_3 L}\right] \times$$
$$\times b^{-1}\left(\theta_2^M - \theta_1^M\right)^{-1} = c(\theta_1,\theta_2). \tag{14}$$

Используя формулу (8) и подставляя конкретный вид функции $u(\theta_1)$ и $g(\theta_1)$ из (12) при альтернативе H_1 находим

$$\nu(x_3 / \theta_1) = \frac{\left(a\theta_1^L\right)^{n_3}}{m^{n_3}} \times$$
$$\times \exp\left[-b\theta_1^M x_3\right]_{x_3}^{\frac{n_3 k - m + n_3}{m}} \frac{\left[\Gamma\left(\frac{k+1}{m}\right)\right]^{n_3}}{\left[\Gamma\left(\frac{n_3(k+1)}{m}\right)\right]}. \tag{15}$$

Тогда получим

$$\alpha_0 = \int_0^{c'} \nu\left(x_3 / \theta_1\right) dx_3 = G_{n_3,m,\kappa}\left(\theta_1^M bc\right), \quad (16)$$

где

$$G_{n_3,m,\kappa}(r) = \frac{\left[\Gamma\left(\frac{k+1}{m}\right)^{n_3} a^{n_3} b^{-\frac{n_3(k+1)}{m}}\right]}{\Gamma\left(\frac{n_3(k+1)}{m}\right) m^{n_3}} \times \quad (17)$$
$$\times \int_0^r x_3^{\frac{n_3(k+1)}{m}-1} \exp(-x_3)\, dx_3.$$

Из (16) следует, что $c = G^{-1}{}_{n_3,m,\kappa}(\alpha_0)\theta_1^{-M} b^{-1}$, где $r = G^{-1}{}_{n_3,m,\kappa}(l)$ есть функция, обратная функции $l = G_{n_3,m,\kappa}(r)$. Область принятия гипотезы H_2 определяется неравенством $x_3 < G^{-1}{}_{n_3,m,\kappa}(\alpha_0)\theta_1^{-M} b^{-1}$. Используя выражение для θ_1 и θ_2 из (15) в пространстве (x_1, x_2, x_3) область принятия альтернативы можно определить неравенствами

$$\frac{x_1}{n_1} > \frac{x_2}{n_2};\ n_1 L x_3 < M x_1 G^{-1}_{n_3,m,\kappa}(\alpha_0).$$

Вводя обозначения

$$s = \frac{x_1}{x_3};\ t = \frac{x_2}{x_3}, \quad (18)$$

получаем, что в рассматриваемом случае область принятия гипотезы H_2 в пространстве (s,t) определится неравенствами

$$t < s\frac{n_2}{n_1};\ s > \frac{n_1 L}{M G^{-1}_{n_3,m,\kappa}(\alpha_0)}. \quad (19)$$

Второй случай: $\theta_1 > \theta_2$. Из (13) получаем

$$\sum_{j=1}^{n_3} \left(x_j^{(3)}\right)^m > \ln\left[c\left(\frac{\theta_2}{\theta_1}\right)^{nL}\right] \times \quad (20)$$
$$\times b^{-1}\left(\theta_1^M - \theta_2^M\right)^{-1} = c(\theta_1, \theta_2).$$

Из неравенства (20), рассуждая аналогично случаю $\theta_1 \le \theta_2$, получаем, что в пространстве (s,t) область принятия гипотезы H_2 определится неравенствами

$$t > s\frac{n_2}{n_1},\ s > \frac{n_1 L}{M G^{-1}_{n_3,m,\kappa}(1-\alpha_0)}. \quad (21)$$

Окончательно находим, что область принятия гипотезы H_2 есть объединение областей, определяемых соотношениями (19) и (21) (рис. 1).

Отметим, что решающее правило не зависит от принятой системы измерения. Для определения вероятностей ошибок первого и второго рода вычислим плотность распределения вероятностей статистик s и t. Перейдем от статистики (x_1, x_2, x_3) с плотностью распределения вероятностей (1) к статистике (s,t,w), определяемой соотношениями $x_1 = sx_3$, $x_2 = tx_3$, $x_3 = w$. Якобиан этого преобразования равен

$$\mathrm{I} = \begin{vmatrix} \frac{\partial x_1}{\partial s} & \frac{\partial x_1}{\partial t} & \frac{\partial x_1}{\partial w} \\ \frac{\partial x_2}{\partial s} & \frac{\partial x_2}{\partial t} & \frac{\partial x_2}{\partial w} \\ \frac{\partial x_3}{\partial s} & \frac{\partial x_3}{\partial t} & \frac{\partial x_3}{\partial w} \end{vmatrix} = w^2.$$

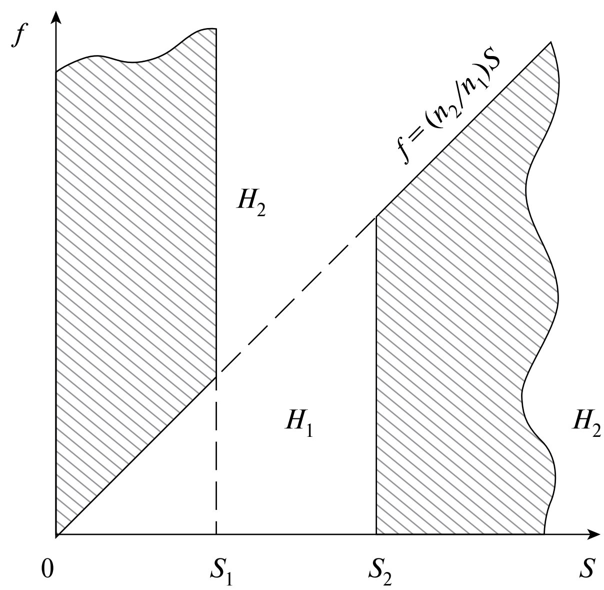

Рис. 1. Области принятия гипотез для классического решающего правила.

Отсюда плотность распределения вероятностей статистики (s,t,w) принимает вид

$$H(s,t,w) = h(x_1,x_2,x_3 / \omega_1,\omega_2,\omega_3)w^2 = \\ = \frac{u(\omega_1)^{n_1} u(\omega_2)^{n_2} u(\omega_3)^{n_3} \Psi S^{n_1'-1} t^{n_2'-1} w^{n_1'+n_2'+n_3'-1} \left[\Gamma(N)\right]^n}{m_n \Gamma(n_1')\Gamma(n_2')\Gamma(n_3')}, \tag{22}$$

где

$$n = n_1 + n_2 + n_3,$$

$$\Psi = \exp\left\{-W\left[g(\omega_1)S + g(\omega_2)t + g(\omega_3)\right]\right\},$$

$$N = \frac{\kappa+1}{m},$$

$$n_1' = n_1 N,\ n_2' = n_2 N,\ n_3' = n_3 N.$$

Распределение статистики (s,t) получим интегрированием функции (22) по переменной w на интервале $(0,\infty)$. В результате находим

$$f(s,t) = \frac{u(\omega_1)^{n_1} u(\omega_2)^{n_2} u(\omega_3)^{n_3} B s^{n_1'-1} t^{n_2'-1}}{m^n \left[g(\omega_1)s + g(\omega_2)t + g(\omega_3)\right]^{n_1'+n_2'+n_3'}},$$

где

$$B = \frac{\Gamma(n_1'+n_2'+n_3')\left[\Gamma\left(\frac{\kappa+1}{m}\right)\right]^n}{\Gamma(n_1')\Gamma(n_2')\Gamma(n_3')}.$$

Для конкретного вида функций u и g из (12) имеем

$$f(s,t) = \frac{a^n \omega_1^{n_1 L} \omega_2^{n_2 L} \omega_3^{n_3 L} B s^{n_1'-1} t^{n_2'-1}}{m^n \left[b\left(\omega_1^M s + \omega_2^M t + \omega_3^M\right)\right]^{n'}}, \tag{23}$$

где $n' = n_1' + n_2' + n_3'$. Обозначая $\lambda = \omega_2 / \omega_1$ и задавая условие $L = \frac{(\kappa+1)M}{m}$, преобразуем $f(s,t)$ при гипотезе H_1 к виду

$$f_1(s,t) = \frac{a^n \lambda^{n_2 L} B s^{n_1'-1} t^{n_1'-1}}{m^n \left[b\left(s + 1 + \lambda^M t\right)\right]^{n'}},$$

а при гипотезе H_2—

$$f_2(s,t) = \frac{a^n \lambda^{(n_1+n_2)L} B s^{n_1'-1} t^{n_2'-1}}{m^n \left[b\left(s + \lambda^M (t+1)\right)\right]^{n'}}. \tag{24}$$

Из рис. 1 видно, что вероятности ошибок первого и второго рода можно определить следующим образом:

$$\alpha(\lambda) = \left[P_1(s_1) - Q_1(s_1)\right] + \left[Q_1(\infty) - Q_1(s_2)\right],$$

$$\beta(\lambda) = 1 - \left[P_2(s_1) - Q_2(s_1)\right] + \left[Q_2(\infty) - Q_2(s_2)\right],$$

где

$$P_j(y) = \int_0^y \int_0^\infty f_j(s,t)\,ds dt,$$

$$Q_j(y) = \int_0^y \int_0^{s\frac{n_2}{n_1}} f_j(s,t)\,ds dt,\ (j = 1,2). \tag{25}$$

4. РЕШАЮЩЕЕ ПРАВИЛО, УДОВЛЕТВОРЯЮЩЕЕ НЕОБХОДИМЫМ УСЛОВИЯМ ОПТИМАЛЬНОСТИ

Для случайных величин, принадлежащих к экспоненциальному семейству распределений с априорной параметрической неопределенностью можно построить несмещенное решающее правило, которое является подобным заданному уровню α_0. Другими словами, решающее правило будет удовлетворять двум обязательным условиям оптимальности, (3) и (4). Искомое решающее правило выберем в классе решающих правил, представленных на рис. 2.

Для оптимизации свободными остаются только два параметра s_1 и s_2.

Требования в процессе оптимизации следующие:

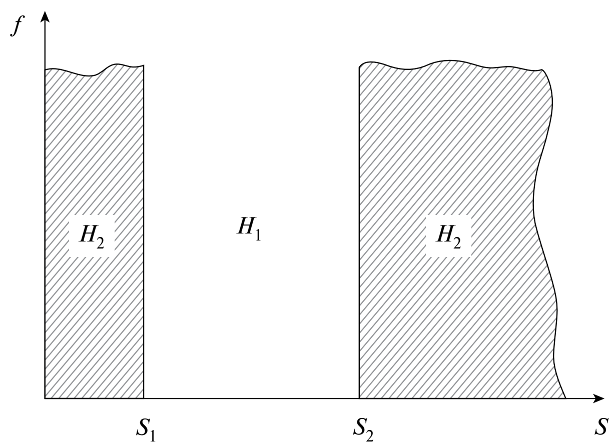

Рис. 2. Области принятия гипотез для оптимального решающего правила; область принятия гипотезы H_2 заштрихована.

1) $\max \alpha(s_1, s_2, \lambda) = \alpha_0$, $0 < \lambda < \infty$, (26)

2) $\min D(s_1, s_2, \lambda)$, $0 < \lambda < \infty$, должен принимать для оптимальной пары (s_1, s_2) максимальное значение.

Здесь α_0 — заданное число $(0 < \alpha_0 < 1)$ и $D(s_1, s_2, \lambda) = 1 - \beta(s_1, s_2, \lambda)$. Выражение для $\alpha(s_1, s_2, \lambda)$ и $D(s_1, s_2, \lambda)$ получаются с помощью P_1 и P_2 из (25):

$$\alpha(s_1, s_2, \lambda) = = A_2 \left\{ \begin{array}{c} 1 + B_1 \sum\limits_{j=0}^{n_1'-1} c(j) \times \\ \times \left[\dfrac{1}{(s_2+1)^{n_3'+j}} - \dfrac{1}{(s_1+1)^{n_3'+j}} \right] \end{array} \right\} = \alpha_0, \tag{27}$$

где введены обозначения

$$A_2 = \frac{a^n \left[\Gamma\left(\dfrac{\kappa+1}{m} \right) \right]^n}{m^n b^{n'}}; \quad B_1 = \frac{\Gamma(n_1' + n_3')}{\Gamma(n_3')}; \tag{28}$$

$$D(s_1, s_2, \lambda) = A_2 \left\{ \begin{array}{c} 1 + B_1 \sum\limits_{j=0}^{n_1'-1} c(j) \times \\ \times \left[\lambda^{M_j + n_3 L} \left(s_2 + \lambda^M \right)^{-n_3' - j} - \right. \\ \left. - \lambda^{M_j + n_3 L} \left(s_1 + \lambda^M \right)^{-n_3' - j} \right] \end{array} \right\}.$$

Задачу (26) решим методом неопределенных множителей Лагранжа. Для этого составим функционал

$$F = D(s_1, s_2, \lambda) + \varpi\left[\alpha(s_1, s_2, \lambda) - \alpha_0 \right] =$$
$$= A_2 \left\{ 1 + B_1 \sum_{j=0}^{n_1'-1} c(j) \left[\frac{1}{(s_2 + \lambda^m)^{n_3'+j}} - \frac{1}{(s_1 + \lambda^m)^{n_3'+j}} \right] \lambda^{M_j + n_3} \right\} + \tag{29}$$
$$+ \mu \left\{ A_2 \left[1 + B_1 \sum_{j=0}^{n_1'-1} c(j) \left[\frac{1}{(s_2+1)^{n_3'+j}} - \frac{1}{(s_1+1)^{n_3'+j}} \right] \right] - \alpha_0 \right\},$$

где μ — неопределенный множитель Лагранжа, определяемый из условия (27).
Составляем уравнения

$$\frac{\partial F}{\partial \lambda} = A_2 B_1 \sum_{j=0}^{n_1'-1} c'(j) \times$$
$$\times \left[\frac{s_2}{(s_2 + \lambda^M)^{n_3'+j+1}} - \frac{s_1}{(s_1 + \lambda^M)^{n_3'+j+1}} \right] \lambda^{M_j + n_3 L^{-1}} = 0,$$
$$\frac{\partial F}{\partial s_1} = A_2 B_1 \sum_{j=0}^{n_1'-1} c'(j) \times$$
$$\times \left[\frac{\lambda^{M_j + n_3 L}}{(s_2 + \lambda^M)^{n_3'+j+1}} - \frac{\mu}{(s_1+1)^{n_3'+j+1}} \right] = 0,$$
$$\frac{\partial F}{\partial s_2} = A_2 B_1 \sum_{j=0}^{n_1'-1} c'(j) \times$$
$$\times \left[\frac{\lambda^{M_j + n_3 L}}{(s_2 + \lambda^M)^{n_3'+j+1}} - \frac{\mu}{(s_2+1)^{n_3'+j+1}} \right] = 0,$$

где $c'(j) = \dfrac{(-1)^j}{(n_1' - 1 - j)!\, j!}$.

После некоторых преобразований получим следующую систему для определения s_1 и s_2:

$$\left\{ \begin{array}{l} A_2 \left\{ 1 + B_1 \sum\limits_{j=0}^{n_1'-1} c(j) \left[\dfrac{1}{(s_2+1)^{n_3'+j}} - \dfrac{1}{(s_1+1)^{n_3'+j}} \right] = \alpha_0 \right. \\ \left(\dfrac{s_1}{s_2} \right)^{n_1'} = \left(\dfrac{s_1+1}{s_2+1} \right)^{n_3'+n_1'} \end{array} \right. \tag{30}$$

где $c(j) = \dfrac{(-1)^j}{(n_1' - 1 - j)!\, j (n_3' + j)}$.

Систему (30) в общем случае можно решить численными методами на компьютере. Изложенная здесь методика позволяет получить решающие правила для частных распределений, которые входят в экспоненциальное семейство распределений. Для этого выражения конкретных функций *u* и *g* из (12) подставим в (7) и получим

$$f_{\xi_3}\left(x_j^{(3)} / \omega_3\right) = a\omega_3^L\left(x_j^{(3)}\right)^K \exp\left(-b\omega_3^M\left(x_j^{(3)}\right)^m\right). \quad (31)$$

Перечислим конкретные виды распределений (31), представленные в табл. 1:

$$\omega\exp(-\omega x)$$

– экспоненциальное распределение;

$$\omega x\exp\left(-0.5\omega x^2\right)$$

– распределение Релея;

$$\omega^2 x\exp(-\omega x)\frac{\omega^{m'}}{\Gamma(m')}x^{m'-1}\cdot e^{-\omega x}$$

– гамма-распределение;

$$m'\omega^{-1}x^{m'-1}\exp\left(-\omega^{-1}x^{m'}\right)$$

– распределение Вейблла (m' – известный действительный параметр);

$$\omega^{-1/2}(2\pi)^{-\frac{1}{2}}\exp\left(-0.5x^2\omega^{-1}\right)$$

– нормальное распределение с нулевым средним значением;

$$\frac{2}{(m'-1)'} - (m'/\omega)^{m'}x^{2m'-1}\exp\left(-m'x^2\omega^{-2}\right)$$

– распределение Накагами (m' - целочисленный параметр).

Рассмотрим частные случаи решающего правила, получаемого методом оценки параметров. В табл. 2 приведены примеры процедур, основанных на методе оценки неизвестных параметров, рассмотрены допустимые вероятности ошибок первого рода при $\alpha_0 = 0.5, 0.1, 0.2$, объемы обучающих и контрольных выборок от 1 до 20.

Приведены также значения порогов s_1 и s_2 и максимальные вероятности ошибок первого рода.

Как видно из табл. 2, максимальная вероятность ошибки первого рода для многих параметров в 25...8 раз превышает допустимую величину.

Зависимости $\alpha(\lambda)$ и $\beta(\lambda)$ для примера $n_1 = n_2 = n_3$, $\alpha_0 = 0,2$ приведены на рис. 3.

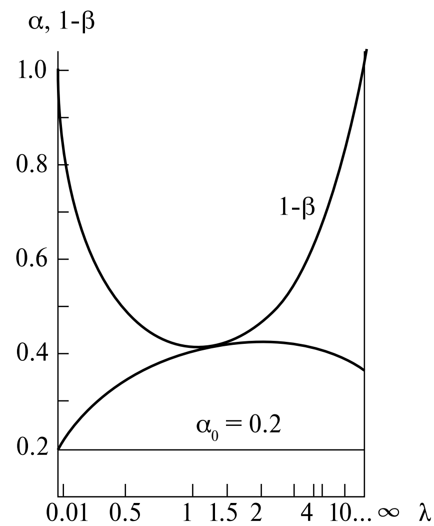

Рис. 3. Графики вероятностей ошибок первого и второго рода для классической решающей процедуры.

Таблица 1. Типы распределения, входящие в экспоненциальное семейство

Тип распределения $f_\xi(x/\omega)$	Набор параметров семейства распределений (31)					
	L	M	K	m	a	b
Экспоненциальное распределение	1	1	0	1	1	1
Распределение Релея	1	1	1	2	1	1/2
Гамма-распределение	m	1	1	1	$\frac{1}{\Gamma(m')}$	1
Распределение Вейблла	–1	–1	–1	$m'-1$	m'	1
Нормальное распределение с нулевым средним значением	–1	–2	0	2	$\frac{1}{\sqrt{2\pi}}$	1/2
Распределение Накагами	$-m'$	–1	$2m'-1$	2	$\frac{2m'^{m'}}{\Gamma(m')}$	m'

Таблица 2. Примеры классических решающих правил

$n_1 = n_2 = n_3$	$\alpha_0 = 0.05$			$\alpha_0 = 0.1$			$\alpha_0 = 0.2$		
	s_1	s_2	$\max\limits_{\lambda} \alpha(\lambda)$	s_1	s_2	$\max\limits_{\lambda} \alpha(\lambda)$	s_1	s_2	$\max\limits_{\lambda} \alpha(\lambda)$
1	0.33	14.3	0.257	0.43	10	0.312	0.62	4.48	0.42
2	0.44	6.25	0.233	0.53	4	0.303	0.66	2.43	0.412
3	0.49	3.75	0.224	0.58	2.72	0.304	0.7	2	0.409
4	0.53	3.08	0.218	0.61	2.35	0.296	0.73	0.74	0.416
5	0.56	2.63	0.24	0.64	2.08	0.318	0.74	1.61	0.424
6	0.6	2.4	0.23	0.66	1.93	0.312	0.78	1.54	0.386
7	0.61	2.2	0.223	0.68	1.82	0.31	0.79	1.49	0.37
8	0.63	2.1	0.22	0.7	1.74	0.3	0.8	1.45	0.365
9	0.64	2	0.215	0.73	1.67	0.29	0.81	1.43	0,36
10	0,65	1,9	0,2	0,76	1,56	0,28	0,82	1,4	0.35
15	0.7	1.66	0.16	0.78	1.47	0.22	0.83	1.3	0.32
20	0.73	1.55	0.11	0.79	1.4	0.17	0.85	1.24	0.3

Рассмотрим частные случаи для решающего правила, удовлетворяющего условиям оптимальности (табл. 3).

Рассмотрены допустимые вероятности ошибок первого рода для $\alpha_0 = 0.5$, 0.1, 0.2, объемы обучающих и контрольных выборок от 1 до 20, а также приведены значения порогов s_1 и s_2.

Нетрудно показать, что в рассмотренных процедурах минимальное значение $D = 1 - \beta$ достигается при $\lambda = 1$. На рис. 4 приведены графики функции $\alpha(\lambda)$ и $D = 1 - \beta(\lambda)$ для случая $n_1 = n_2 = n_3 = 5$, $\alpha_0 = 0.2$.

Отметим, что при $\lambda > 1$ на оси абсцисс откладывается величина $2 - 1/\lambda$. Ясно, что найденное решающее правило является подобным уровню α_0 и несмещенным.

Для остальных примеров эти графики аналогичные.

В качестве меры эффективности оптимальной процедуры по отношению к процедуре, получаемой методом оценки параметров, рассмотрим величину $\eta_{\alpha_0} = 1 - \dfrac{\alpha_0}{\alpha_{\max}}$, где α_0 – допустимая вероятность ошибки первого рода для обеих рассмотренных процедур, а $\alpha_{\max}$ – это фактическая вероятность ошибки первого рода при процедуре, получаемой методом оценки неизвестных параметров.

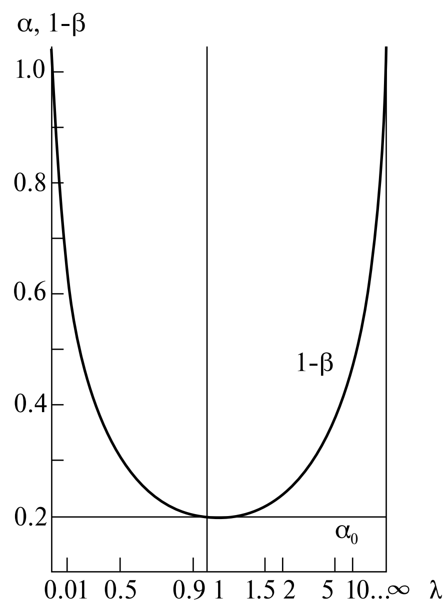

Рис. 4. Графики вероятностей ошибок первого и второго рода для оптимальной решающей процедуры.

Таблица 3. Примеры оптимальных решающих процедур

$n_1 = n_2 = n_3$	$\alpha_0 = 0.05$		$\alpha_0 = 0.1$		$\alpha_0 = 0.2$	
	s_1	s_2	s_1	s_2	s_1	s_2
1	0.33	33.33	0.5	20	0.11	9.09
2	0.1	10	0.15	6.66	0.24	4.16
3	0.16	6.25	0.23	4.34	0.32	3.12
4	0.21	4.76	0.28	3.57	0.38	2.63
5	0.26	3.84	0.33	3.03	0.43	2.32
6	0.29	3.44	0.37	2.7	0.46	2.174
7	0.32	3.125	0.4	2.5	0,49	2.04
8	0,35	2.857	0.42	2.38	0.52	1.926
9	0.38	2.63	0.45	2.22	0.54	1.8519
10	0.42	2.38	0.48	2.08	0.56	1.7857
15	0.51	1.96	0.54	1.85	0.68	1.47
20	0.56	1.8	0.62	1.61	0.75	1.33

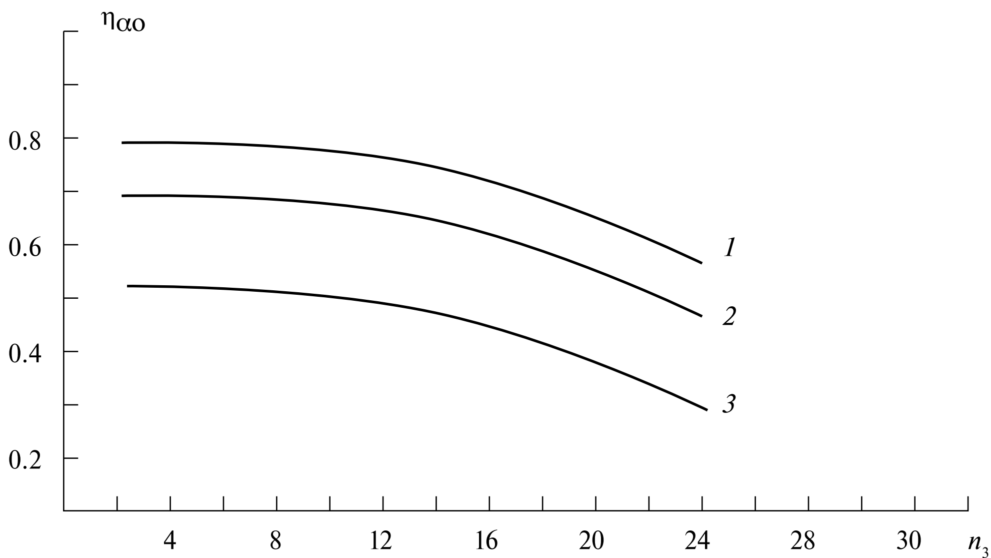

Рис. 5. Соотношение решающих процедур для α_0 = 0.05 (*1*), 0.1 (*2*) и 0.2 (*3*).

На рис. 5 показаны зависимости η_{α_0} от длины выборки n_3 при разных α_0.

С ростом n_3, η_{α_0} уменьшается, при $n_3 \to \infty$ обе процедуры становятся эквивалентными.

Как видно из рис. 5, при увеличении n_3 уменьшается η_{α_0}, а при $n_3 \to \infty$, $\alpha_{max} \to \alpha_0$, $\eta_{\alpha_0} \to 0$ обе процедуры становятся эквивалентными.

ЗАКЛЮЧЕНИЕ

Таким образом, в данной работе получены следующие результаты.

Предложен новый подход к решению задачи обучения принятию статистических решений по выборкам ограниченного объема для экспоненциальных семейств распределений при априорной параметрической неопределенности.

Разработано решающее правило, удовлетворяющее необходимым условиям оптимальности: постоянству средней вероятности ошибки первого рода и несмещенности. Рассматриваются конкретные решающие процедуры для частных распределений, полученных от обобщенного алгоритма.

Показана эффективность разработанной оптимальной процедуры для выборок малого объема.

ФИНАНСИРОВАНИЕ РАБОТЫ

Работа выполнена в рамках госзадания Института радиотехники и электроники им. В.А. Котельникова РАН.

СПИСОК ЛИТЕРАТУРЫ

1. *Арманд Н.А., Крапивин В.Ф., Мкртчян Ф.А.* Методы обработки данных радиофизического исследования окружающей среды. М.: Наука, 1987.
2. *Мкртчян Ф.А.* Оптимальное различение сигналов и проблемы мониторинга. М.: Наука, 1982.
3. *Mkrtchyan F.A., Varotsos C.A.* // Water, Air, & Soil Pollution. 2018. V. 229. № 8. Article No. 273.
4. *Данков П.П.* // РЭ. 1965. Т. 10. № 10. С. 1774.
5. *Леман Э.* Проверка статистических гипотез. М.: Наука, 1979.

NEW METHODS FOR STATISTICAL DECISION MAKING IN CONDITIONS OF A LIMITED VOLUME OF OBSERVATIONS AND WITH A PRIORIY PARAMETRIC UNCERTAINTY

F. A. Mkrtchyan

Fryazino Branch Kotelnikov Institute of Radio Engineering and Electronics, Russian Academy of Sciences, Fryazino, Moscow region, 141190 Russia

E-mail: ferd47@mail.ru

A new generalized adaptive algorithm for learning to make statistical decisions for exponential families of distributions with a priori parametric uncertainty in conditions of small samples has been developed. A generalized decision rule is presented, obtained by estimating unknown parameters of distributions, as well as a decision rule that satisfies the necessary optimality conditions: constancy of the average probability of a type I error and unbiasedness. Specific decision procedures for partial distributions obtained from a generalized algorithm are considered. Numerical examples are given. The effectiveness of the developed optimal procedure for small samples is shown.

Keywords: a priori parametric uncertainty, general adaptive algorithm, exponential family of distributions, small samples