

УДК 681.518.54

О. Ю. КРОПАЧЕК**ИССЛЕДОВАНИЕ СВОЙСТВ КВАДРАТИЧНОЙ ДИСКРИМИНАНТНОЙ ФУНКЦИИ ПРИ ДИАГНОСТИРОВАНИИ ДИНАМИЧЕСКИХ ПРОЦЕССОВ**

Розглянута ймовірнісна модель квадратичного дискримінуючого перетворення вхідних випадкових векторних вимірювальних сигналів в інформаційній системі технічної і медичної діагностики, коли коефіцієнти функції перетворення випадкові. Показано, що така модель – це суперпозиція умовних (за класами станів, що діагностуються) центрального та нецентрального F-розподілів Фішера, число ступенів яких залежить від об'ємів навчальних вибірок (ОНВ), як за компонентами вхідного вектору, так і за класами станів. Запропоновано метод оцінювання достовірності діагностики, як функції не лише розмірності вхідного вектору вимірюваних інформативних ознак, але і функції значень ОНВ.

Ключові слова: достовірність, вірогідність, F-розподіл, вирішальна функція.

Рассмотрена вероятностная модель квадратичного дискриминирующего преобразования входных случайных векторных измерительных сигналов в информационной системе технической и медицинской диагностики, когда коэффициенты функции преобразования случайны. Показано, что такая модель – это суперпозиция условных (по классам диагностируемых состояний) центрального и нецентрального F-распределений Фишера, число степеней которых зависит от объемов обучающих выборок (ООВ), как по компонентам входного вектора, так и по классам состояний. Предложен метод оценивания достоверности диагностики, как функции не только размерности входного вектора измеряемых информативных признаков, но и функции значений ООВ.

Ключевые слова: достоверность, вероятность, F-распределения, решающая функция.

We consider the probability model quadratic discriminating converting input random vector measurement signals in the information system of technical and medical diagnostics, where the transformation function coefficients are random. It is shown that this model – it is a superposition of the conditional (on classes of diagnosed conditions) central and non-central Fisher's F-distribution, the number of degrees which depends on the volume of training samples, both components of the input vector, and the classes of states. A method for estimating the reliability of the diagnosis, as a function of not only the dimension of the input vector measured informative features, but also functions volume of training samples values.

Keywords: accuracy, probability, F-distribution, decision function.

Постановка проблеми. Повышение эффективности работы любых информационных систем контроля, диагностики, идентификации невозможно без учета априорной информации, связанной с исходной неопределенностью состояний контролируемых или диагностируемых объектов. Кроме того, уменьшение такой неопределенности связано с повышением точности измерительных преобразований и уменьшения остаточной неопределенности решений, принимаемых в ходе контроля или функциональной диагностики.

Известные вероятностные подходы в теории информации и информационной теории измерений позволяют рассчитывать и анализировать количество ожидаемой информации для моделей измерительных преобразований и задач кодирования случайных измерительных сигналов. Теория же оценивания достоверности полученного количества информации в задачах контроля и диагностики практически не разработана. Это ограничивает возможности теоретического анализа и совершенствования информационных моделей оптимального синтеза информационно-измерительных систем контроля и диагностики объектов со случайными свойствами.

Анализ литературы. Проблема снижения рисков диагностических решений в условиях неустранимой априорной неопределенности свойств объектов как технической, так и медицинской диагностики, всегда являлась предметом вероятностно-статистического анализа процедур преобразования первичной измерительной информации во вторичные логические решения. При этом, степень и глубина изучения проблемы

связывалась со сложностью математической модели процедуры такого дискриминантного преобразования [1–4]. Лучше всего изучена проблема для простых (линейных) процедур, реализованных в виде параметрических линейных дискриминантных функций (ДФ) [1, 5–7]. Квадратичные функции, учитывающие априорную информацию большего (в разы) объема исследованы – недостаточно. [8]. По крайней мере, отсутствуют статистически обоснованные модели функциональной связи между объемами обучающих выборок, используемых для синтеза коэффициентов ДФ и интервальными оценками достоверности получаемых решений.

Цель статьи. Изложение методики оценивания достоверности диагностики, когда для принятия решений используется квадратичная ДФ со случайными коэффициентами, дисперсии которых зависят от объема многомерных и многократных измерений на этапе обучения информационной системы диагностики.

Модель измерительно-логического преобразования. Проблема априорной неопределенности. Такая модель включает в себя:

- математическую модель решающей (дискриминантной) функции $g(\bar{X})$, зависящей от вектора $\bar{X} = (X_1, \dots, X_n)$ входных измерительных сигналов X_1, \dots, X_n ;

- логическую модель принятия одного из двух решений (ω_0 – если состояние объекта диагностики θ_0 , ω_1 – если состояние θ_1).

Выбор решений производят в соответствии с правилом [1–4]:

$$\begin{cases} \omega_0, & \text{если } g(\bar{X}) > 0; \\ \omega_1, & \text{если } g(\bar{X}) \leq 0. \end{cases} \quad (1)$$

На рис. 1. представлена структурная схема, иллюстрирующая измерительно-логическое преобразование случайного измерительного вектора \bar{X} в логические решения ω_0 или ω_1 .

Логические правила (1) фактически реализуют альтернативный статистический тест с двумя (основной H_0 и альтернативной H_1) гипотезами:

H_0 – состояние исследуемого объекта есть θ_0 ,

H_1 – состояние объекта есть θ_1 .

Случайность результатов измерений x_1, \dots, x_n

компонент X_1, \dots, X_n вектора \bar{X} обуславливает появление рисков первого (α) и второго (β) рода [1–4].

$$\begin{cases} \alpha = P[\omega_1 | \theta_0] \\ \beta = P[\omega_0 | \theta_1] \end{cases} \quad (2)$$

Риски (2) минимальны, если полностью определены постоянные коэффициенты решающей функции $g(\bar{X})$. Для случая квадратичной функции $g(\bar{X})$, ее коэффициентами являются вектора условных (по состояниям θ_0 и θ_1) средних $\mu_{(0)}$, $\mu_{(1)}$ и дисперсионные единичные матрицы $D_{(0)}$, $D_{(1)}$, как комплексные числовые характеристики независимых нормально распределенных компонент вектора \bar{X} .

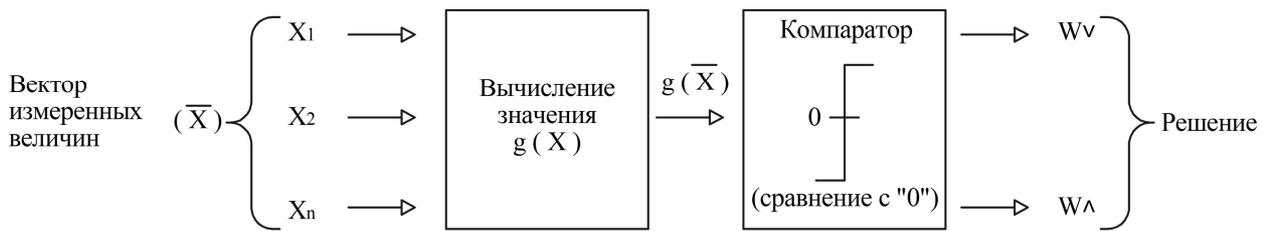


Рис. 1 – Структурная схема измерительно-логического преобразования

Выражение для квадратичной решающей функции (в векторной форме) имеет вид [1, 7]:

$$g(\bar{X}) = (\bar{X} - \mu_{(1)})' D_{(1)} (\bar{X} - \mu_{(1)}) - (\bar{X} - \mu_{(0)})' D_{(0)} (\bar{X} - \mu_{(0)}) + \ln |D_{(1)}| - \ln |D_{(0)}| \quad (3)$$

Если коэффициенты функции (3) оцениваются, в ходе обучения системы диагностики, по конечным образцовым выборкам вектора \bar{X}

$$\left\{ \bar{X}_1^{(0)}, \dots, \bar{X}_{N1}^{(0)} \right\}, \left\{ \bar{X}_1^{(1)}, \dots, \bar{X}_{N2}^{(1)} \right\} \quad (4)$$

для состояний θ_0 и θ_1 , то оценки таких коэффициентов – это случайные величины. В этом случае дискриминантная функция (3) становится случайной оценкой и зависит не только от случайного вектора \bar{X} , но и от случайных коэффициентов $\mu_{(0)}^*$, $\mu_{(1)}^*$, $D_{(0)}^*$, $D_{(1)}^*$, что приводит к увеличению рисков α и β диагностики.

Следует признать, и это проблема, что рассчитать реальные, увеличенные значения рисков диагностики, очень сложно. Это создает неопределенность в оценке реальных диагностических свойств проектируемых информационных систем диагностики, причем безотносительно к задачам (техническая или медицинская диагностика).

Процедура уменьшения параметрической неопределенности. Оценка $g^*(\bar{X})$ функции $g(\bar{X})$ (3), должна рассматриваться как условная случайная величина $\delta^{(k)}$, меняющая свои вероятностные свойства (условную плотность $f(\delta/\theta_k)$) в зависимости от значения индекса k :

$$f(g^*(\bar{X})/\theta_k) = f(\delta/\theta_k) = \begin{cases} f(\delta/\theta_0), & k = 0; \\ f(\delta/\theta_1), & k = 1. \end{cases} \quad (5)$$

Различие между плотностями $f(g(\bar{X})/\theta_k)$ и $f(g^*(\bar{X})/\theta_k)$ всегда порождает отрицательное смещение при оценке среднего риска диагностики (оценка занижена) [1], и смещение это растет с уменьшением ООВ. Фактически, это приводит к завышению достоверности решений для проектируемых информационно-измерительных систем (ИИС) контроля и диагностики, поскольку обеспечить представительные ООВ по состояниям θ_0 и θ_1 никогда не удастся.

Предлагается процедура коррекции распределений $f(g^*(\bar{X})/\theta_k)$, $k = 0, 1$, для уменьшения смещения оценок среднего риска диагностики. Такая процедура включает следующую последовательность этапов статистического анализа.

1. Оценка значений кумулянтов (семиинвариантов) случайных величин $\delta^{(0)}$ и $\delta^{(1)}$ (для $k = 0$ и $k = 1$). Порядок r кумулянтов должен быть достаточно большой ($r \geq 4$), чтобы имелась возможность восстановления плотностей $f(\delta/\theta_0)$ и $f(\delta/\theta_1)$, хотя бы в виде распределений Пирсона [9].

2. Восстановление (оценка) плотностей распределения $f(\delta/\theta_0)$ и $f(\delta/\theta_1)$ рядами Эджворта [9], когда $r \geq 4$.

3. Интегрирование полученных модельных распределений (в виде рядов Эджворта) для оценивания рисков первого (α) и второго (β) рода и расчета среднего риска (\bar{R}) диагностики.

4. Использование рассчитанных рисков для получения минимально смещенной оценки достоверности решений проектируемой ИИС технической и медицинской диагностики.

Кумулянтный анализ квадратичной ДФ. Представим векторную модель (3) квадратичной ДФ

как статистику $\delta^{(K)}$ от скалярных случайных величин $x_i^{(K)}, m_i^{(0)}, m_i^{(1)}, S_i^{(0)2}, S_i^{(1)2}$ ($\kappa = 0, 1$).

$$\delta^{(K)} = \sum_{i=1}^n \left(\frac{x_i^{(K)} - m_i^{(1)}}{S_i^{(1)}} \right)^2 - \sum_{i=1}^n \left(\frac{x_i^{(K)} - m_i^{(0)}}{S_i^{(0)}} \right)^2 + \sum_{i=1}^n \left(\frac{S_i^{(1)}}{S_i^{(0)}} \right)^2, \quad (6)$$

где $m_i^{(0)}, m_i^{(1)}$ – оценки элементов векторов средних $\mu_{(0)}, \mu_{(1)}$, полученные по образцовым выборкам (4);

$S_i^{(0)}, S_i^{(1)}$ – оценки элементов дисперсионных матриц $D_{(0)}, D_{(1)}$, полученные по выборкам (4);

$x_i^{(K)}$ – реализуемая i -го значения контрольной выборки $\{x_1^{(K)}, \dots, x_n^{(K)}\}$ с заданным номером состояния:

$$K = \begin{cases} 0, & \text{если } \theta \in \theta_0; \\ 1, & \text{если } \theta \in \theta_1. \end{cases}$$

Пусть объемы образцовых выборок по состояниям θ_0 и θ_1 конечны и одинаковы ($N_1 = N_2 = N$). Статистику $\delta^{(1)}$ можно рассматривать как линейную комбинацию трех случайных величин $\varepsilon^{(K)}, \eta^{(K)}$ и ω :

$$\delta^{(K)} = \varepsilon^{(K)} - \eta^{(K)} + \omega, \quad (7)$$

где случайные величины $\varepsilon^{(K)}$ и $\eta^{(K)}$ имеют взаимно различные (центральные или нецентральные) F -распределения дисперсионного отношения Фишера, зависящие от значения K (табл. 1) [10].

Таблица 1 – Варианты распределений статистик $\varepsilon^{(K)}$ и $\eta^{(K)}$

K	F-распределения Фишера	
	Статистика $\varepsilon^{(K)}$	Статистика $\eta^{(K)}$
0	нецентральное	центральное
1	центральное	нецентральное

Статистика ω от вектора контрольных реализаций не зависит и может рассматриваться как параметр сдвига условных плотностей $f(\delta^{(0)}), f(\delta^{(1)})$ статистики $\delta^{(K)}$ ($K = \overline{0, 1}$), обеспечивающего равенство этих плотностей в точке $\delta^{(0)} = \delta^{(1)} = 0$. Кроме того, эта статистика не влияет на высшие, начиная со второго порядка, кумулянты (семиинварианты) [11, 12] статистики $\delta^{(K)}$. Влияние отсутствует и на любые, начиная со второго, начальные и центральные моменты распределений $f(\delta^{(0)}), f(\delta^{(1)})$.

Случайные статистики $\varepsilon^{(K)}$ и $\eta^{(K)}$ удобно рассматривать как суммы из n случайных величин $\varepsilon_i^{(K)}$ и $\eta_i^{(K)}$

$$\begin{cases} \varepsilon^{(K)} = \sum_{i=1}^n \varepsilon_i^{(K)}; \\ \eta^{(K)} = \sum_{i=1}^n \eta_i^{(K)}. \end{cases} \quad (8)$$

Каждая из случайных величин $\varepsilon_i^{(K)}$ и $\eta_i^{(K)}$ имеет либо центральное $F_{1, N-1}$ либо нецентральное $F_{1, N-1, V_i^{(K)}, \lambda_i^{(K)}}$ распределения Фишера с числом степеней свободы l и $(N-1)$, а также дисперсией $V^{(K)}$ и параметром нецентральности $\lambda^{(K)}$:

$$\begin{cases} V_i^{(0)} = \left(\frac{S_i^{(0)}}{S_i^{(1)}} \right)^2 \frac{N}{(N+1)} + (N+1)^{-1}; \\ V_i^{(1)} = \left(\frac{S_i^{(1)}}{S_i^{(0)}} \right)^2 \frac{N}{(N+1)} + (N+1)^{-1}; \\ \lambda_i^{(0)} = \left(\frac{m_i^{(0)} - m_i^{(1)}}{S_i^{(1)}} \right)^2 \frac{N}{(N+1)}; \\ \lambda_i^{(1)} = \left(\frac{m_i^{(1)} - m_i^{(0)}}{S_i^{(0)}} \right)^2 \frac{N}{(N+1)}. \end{cases} \quad (9)$$

Поскольку случайные величины $\varepsilon_i^{(K)}$ и $\eta_i^{(K)}$ образуют двумерную случайную величину $\delta_i^{(K)}$, то кумулянты этой величины $\kappa_r^{(K)}$ ($r=1, 2, \dots$), имея порядок r , зависят не только от кумулянтов $\kappa_{(\varepsilon)r}^{(K)}$ и $\kappa_{(\eta)r}^{(K)}$ случайных величин $\varepsilon_i^{(K)}$ и $\eta_i^{(K)}$, но и от совместных кумулянтов $\kappa_{(\varepsilon)(\eta)}^{(K)}$, отображающих корреляционные связи между величинами $\varepsilon_i^{(K)}$ и $\eta_i^{(K)}$ [11]. Введем согласно [11] обозначения кумулянтов двумерной случайной величины с двумя нижними индексами d и l ($\kappa_{dl}^{(K)}$), где первый индекс относится к порядку центрального момента величины $\varepsilon_i^{(K)}$, а второй – величины $\eta_i^{(K)}$. Собственные кумулянты отдельных величин $\varepsilon_i^{(K)}$ и $\eta_i^{(K)}$ обозначим как:

$$\begin{cases} \kappa_{(\varepsilon)r}^{(K)} = \kappa_{r0}^{(K)}; \\ \kappa_{(\eta)r}^{(K)} = \kappa_{0r}^{(K)}, \end{cases} \quad (10)$$

где r – порядок кумулянтов отдельно величин $\varepsilon_i^{(K)}$ и $\eta_i^{(K)}$.

Первый кумулянт ($r=1$) случайной величины $\delta^{(K)}$, с учетом последнего слагаемого выражения (5), имеет вид:

$$\kappa_1^{(K)} = \sum_{i=1}^n (\kappa_{10i}^{(K)} - \kappa_{01i}^{(K)}) + \sum_{i=1}^n \ln \left(\frac{S_i^{(1)}}{S_i^{(0)}} \right)^2. \quad (11)$$

Согласно табл. 1 значение индекса $\kappa_{10i}^{(0)} = \left(\frac{N-1}{N-3} \right) (V_i^{(0)} + \lambda_i^{(0)})$, (0 или 1) будет определять вид уравнения для кумулянтов $\kappa_{10i}^{(K)}$ и $\kappa_{01i}^{(K)}$ одномерных распределений величин $\varepsilon_i^{(K)}$ и $\eta_i^{(K)}$ [9–12]:

$$\left\{ \begin{aligned} \kappa_{10i}^{(0)} &= \left(\frac{N-1}{N-3} \right) (V_i^{(0)} + \lambda_i^{(0)}); \\ \kappa_{10i}^{(1)} &= \left(\frac{N-1}{N-3} \right); \\ \kappa_{01i}^{(0)} &= \left(\frac{N-1}{N-3} \right); \\ \kappa_{01i}^{(1)} &= \left(\frac{N-1}{N-3} \right) (V_i^{(1)} + \lambda_i^{(1)}). \end{aligned} \right. \quad (12)$$

С учетом (12) общее выражение (11) распадется на два, отдельно для $K=0$ и отдельно для $K=1$ (что характерно для всех $\kappa_r^{(K)}$ с нечетными индексами r)

$$\left\{ \begin{aligned} \kappa_1^{(0)} &= \left(\frac{N-1}{N-3} \right) \sum_{i=1}^n (V_i^{(0)} + \lambda_i^{(0)} - 1) + \sum_{i=1}^n \ln \left(\frac{S_i^{(1)}}{S_i^{(0)}} \right); \\ \kappa_1^{(1)} &= \left(\frac{N-1}{N-3} \right) \sum_{i=1}^n (1 - V_i^{(1)} + \lambda_i^{(1)}) + \sum_{i=1}^n \ln \left(\frac{S_i^{(1)}}{S_i^{(0)}} \right). \end{aligned} \right. \quad (13)$$

Кумулянты порядка $r \geq 2$ величины $\delta^{(K)}$ могут быть найдены по известному [10] уравнению:

$$\kappa_r^{(K)} = \sum_{i=1}^n \sum_{l=0}^r \left[\frac{r!(-1)^l}{l!(r-l)!} \cdot \kappa_{(r-l)l} \right], \quad (14)$$

где $(r-l)$ и l – соответственно первый и второй индексы ($d=r-l$ и l) совместного кумулянта $\kappa_{(r-l)l}$ величин $\varepsilon_i^{(K)}$ и $\eta_i^{(K)}$.

Можно показать [9, 10], что совместные кумулянты $\kappa_{dl}^{(K)}$ ($d \neq 0, l \neq 0$) являются линейными функциями параметров $V_i^{(K)}, \lambda_i^{(K)}$ уравнения (9) и начальных моментов $M[\dot{x}^2], M[\dot{x}^3], \dots, M[\dot{x}^{2(d+l)}]$ центрированной и нормированной нормально распределенной случайной величины ($\dot{x} \approx NORM(0, 1)$). Учтем, что моменты величины \dot{x} с нечетной степенью равны нулю, а моменты величины \dot{x} с четными степенями порядка $b \leq 2(d+l)$ равны

$$M[\dot{x}^b] = 1 \cdot (1+2) \cdot \dots \cdot (1+b-2). \quad (15)$$

Для порядка $r=2$ будем иметь для i -й компоненты вектора \bar{X} :

$$\left\{ \begin{aligned} \kappa_{11i}^{(K)} &= 3V_i^{(K)} + \lambda_i^{(K)2} - \left(\frac{N-1}{N-3} \right)^2 (V_i^{(K)} + \lambda_i^{(K)}); \\ \kappa_{2i}^{(K)} &= \frac{2(N-1)^2}{(N-3)(N-5)} \cdot \left\{ \left(\frac{N-2}{N-3} \right) + \right. \\ &\quad \left. + \left[V_i^{(K)} + 2\lambda_i^{(K)}V_i^{(K)} + \left(\frac{V_i^{(K)} + \lambda_i^{(K)}}{N-3} \right)^2 \right] \right\} - \\ &\quad - 2\kappa_{11i}^{(K)}. \end{aligned} \right. \quad (16)$$

Из уравнений (12), с учетом выражений (13), (14), следует, что

$$\begin{aligned} \chi_2^{(K)} &= \frac{2(N-1)^2}{(N-3)(N-5)} \cdot \sum_{i=1}^n \left\{ \left(\frac{N-2}{N-3} \right) + \right. \\ &\quad \left. + V_i^{(K)} + 2\lambda_i^{(K)}V_i^{(K)} + \left(\frac{V_i^{(K)} + \lambda_i^{(K)}}{N-3} \right)^2 \right\} - \\ &\quad - 2 \sum_{i=1}^n \left[3V_i^{(K)} + \lambda_i^{(K)} - \left(\frac{N-1}{N-3} \right)^2 (V_i^{(K)} + \lambda_i^{(K)}) \right]. \end{aligned} \quad (17)$$

Из уравнения (17) следует, что кумулянты $\chi_r^{(K)}$ четных порядков ($r=2, 4, 6, \dots$) не зависят от значения индекса K .

Кумулянты $\chi_r^{(K)}$ третьего и более высоких порядков могут быть рассчитаны по известным, вероятностно обоснованным правилам.

Правило 1. По уравнению (14) определяется общая зависимость $\chi_r^{(K)}$ от одномерных ($\chi_{r0}^{(K)}, \chi_{0r}^{(K)}$) и совместных ($\chi_{dl}, d \neq 0, l \neq 0$) кумулянтов. Например, для $r=3$ имеем:

$$\chi_3^{(K)} = \sum_{i=1}^n [\chi_{30i}^{(K)} - 3\chi_{21i}^{(K)} + 3\chi_{12i}^{(K)} - \chi_{03i}^{(K)}]. \quad (18)$$

Правило 2. Для одномерных кумулянтов используют известные выражения для кумулянтов центрального и нецентрального F -распределений [9–11], с учетом значений K из табл. 1. Например, для $r=3$ будем иметь:

а) для центрального F -распределения

$$\chi_{30i}^{(K)} = \chi_{03i}^{(K)} = \frac{8(N-1)^3(N^2-3N+2)}{(N-3)^3(N-5)(N-7)}, \quad (19)$$

б) для нецентрального F -распределения

$$\begin{aligned} \chi_{30i}^{(K)} = \chi_{03i}^{(K)} &= \frac{8(N-1)^3}{(N-3)(N-5)(N-7)} \times \\ &\times \left[V_i^{(K)3} + 3\lambda_i^{(K)}V_i^{(K)} + \frac{3(V_i^{(K)2} + 2\lambda_i^{(K)}V_i^{(K)})}{(N-3)} \right]. \end{aligned} \quad (20)$$

Правило 3. Используя метод кумулянтных скобок [11], определяют выражения для расчета совместных кумулянтов, как функций начальных моментов

$M_{dl} \{ \varepsilon_i^{(K)d} \cdot \eta_i^{(K)l} \}$. Например, для $r=3$ имеем:

$$\begin{aligned} \chi_{21i}^{(K)} &= M \{ \varepsilon_i^{(K)2} \cdot \eta_i^{(K)} \} - 2\chi_{10i}^{(K)} M \{ \varepsilon_i^{(K)} \cdot \eta_i^{(K)} \} - \\ &- (\chi_{20i}^{(K)} + \chi_{10i}^{(K)2}) \cdot \chi_{01i}^{(K)} + 2\chi_{10i}^{(K)2} \cdot \chi_{01i}^{(K)}, \end{aligned} \quad (21)$$

$$\begin{aligned} \chi_{12i}^{(K)} &= M \{ \varepsilon_i^{(K)} \cdot \eta_i^{(K)2} \} - 2\chi_{01i}^{(K)} M \{ \varepsilon_i^{(K)} \cdot \eta_i^{(K)} \} - \\ &- (\chi_{02i}^{(K)} + \chi_{01i}^{(K)2}) \cdot \chi_{10i}^{(K)} + 2\chi_{10i}^{(K)} \cdot \chi_{01i}^{(K)2}. \end{aligned} \quad (22)$$

Совместные начальные моменты для выражений (21), (22) имеют вид:

$$\left\{ \begin{aligned} M \{ \varepsilon_i^{(K)} \cdot \eta_i^{(K)} \} &= 3V_i^{(K)2} + \lambda_i^{(K)2}; \\ M \{ \varepsilon_i^{(1)2} \cdot \eta_i^{(1)} \} &= 15V_i^{(1)2} + 3\lambda_i^{(1)2}; \\ M \{ \varepsilon_i^{(0)2} \cdot \eta_i^{(0)} \} &= 15V_i^{(0)4} + 16V_i^{(0)2}\lambda_i^{(0)2} + \lambda_i^{(0)4}; \\ M \{ \varepsilon_i^{(1)} \cdot \eta_i^{(1)2} \} &= 15V_i^{(1)4} + 16V_i^{(1)2}\lambda_i^{(1)2} + \lambda_i^{(1)4}; \\ M \{ \varepsilon_i^{(0)} \cdot \eta_i^{(0)2} \} &= 15V_i^{(0)2} + 3\lambda_i^{(0)2} \end{aligned} \right. \quad (23)$$

Используя совместные начальные моменты (23) и одномерные $\chi_{10_i}^{(k)}$, $\chi_{01_i}^{(k)}$ (выражения (12)) кумулянты, вычисляют совместные кумулянты (21). Затем, применяя уравнения (19), (20) определяют кумулянты $\chi_{30_i}^{(k)}$ и $\chi_{03_i}^{(k)}$. Наконец, полученные результаты подставляют в уравнение (18) для определения отдельно, $\chi_3^{(0)}$ и $\chi_3^{(1)}$. Полученные уравнения достаточно громоздки (но вероятно абсолютно адекватны) и в данной статье не приводятся.

Приведенные правила поэтапного расчета кумулянтов $\chi_r^{(k)}$ могут быть запрограммированы для достаточно больших порядков r . Громоздкость уравнений – это плата за вероятностную и статистическую корректность результатов, особенно, при небольших объемах обучающих выборок N .

Восстановление плотности распределения по кумулянтам ДФ. Прямое восстановление плотности квадратичной ДФ рядами Эджворта или Грама-Шарлье, построенных по конечному числу кумулянтов, неприемлемо [11], поскольку такие восстановленные плотности распределения могут иметь отрицательные значения. Более того, они хорошо представляют только модели симметричных распределений, например, близких закону Гаусса.

В исследуемой задаче использовался комбинационный подход к восстановлению плотности $f(\delta/\theta_k)$ на базе трехпараметрического гамма-распределения [13] с коррекцией формы последнего усеченным рядом Эджворта.

Базовое гамма-распределение имеет стандартный вид:

$$\tilde{f}(\delta/\theta_k) = \frac{(\delta - v_k)^{a_k}}{\theta_k^{a_k+1} \Gamma(a_k+1)} e^{-\frac{\delta - v_k}{\theta_k}}, \quad \delta \in [v, \infty), \quad (24)$$

где $\Gamma(a_k+1)$ – гамма-функция [9].

Параметры a_k , v_k , θ_k являются функциями первых трех кумулянтов $\chi_r^{(k)}$ ($r = 1, 2, 3$):

$$\begin{cases} a_k = 4 \left(\frac{\chi_2^{(k)^3}}{\chi_2^{(k)^2}} \right) - 1; \\ \theta_k = \frac{\chi_3^{(k)}}{2\chi_2^{(k)}}; \\ v_k = \chi_1^{(k)} - 2 \left(\frac{\chi_2^{(k)^2}}{\chi_3^{(k)}} \right). \end{cases} \quad (25)$$

Пронормируем случайную величину δ :

$$\overset{\circ}{\delta} = \frac{\delta - \chi_1^{(k)}}{\sqrt{\chi_2^{(k)}}}. \quad (26)$$

Представим нормированные плотности $f\left(\overset{\circ}{\delta}/\theta_k\right)$ и $\tilde{f}\left(\overset{\circ}{\delta}/\theta_k\right)$ рядами Эджворта [11] и вычтем, формально, второй ряд из первого с переносом

$\tilde{f}\left(\overset{\circ}{\delta}/\theta_k\right)$ в правую часть. Тогда плотность $f\left(\overset{\circ}{\delta}/\theta_k\right)$ будет функцией нормированного гамма-распределения (24) и высших, начиная с четвертого кумулянтных коэффициентов

$$\gamma_r^{(k)} = \frac{\chi_r^{(k)}}{\left(\sqrt{\chi_2^{(k)}}\right)^r}, \quad r \geq 4. \quad (27)$$

Кроме этого, плотность $f\left(\overset{\circ}{\delta}/\theta_k\right)$ будет определяться условными, по значению индекса k , полиномами Эрмита $H_r^{(k)}$ [14] порядков $r = 4, 5, \dots$ соответствующих вариантам $k = 0$ и $k = 1$:

$$\begin{aligned} f\left(\overset{\circ}{\delta}/\theta_k\right) &= \tilde{f}\left(\overset{\circ}{\delta}/\theta_k\right) + (2\pi)^{-1/2} \exp\left(-\frac{\overset{\circ}{\delta}^2}{2}\right) \times \\ &\times \left[\frac{H_4^{(k)}\left(\gamma_4^{(k)} - \frac{6}{(a_k+1)}\right) + \frac{H_5^{(k)}}{120}}{24\left(\gamma_4^{(k)} - \frac{6}{(a_k+1)}\right) + \frac{H_5^{(k)}}{120}} \cdot \right. \\ &\left. \cdot \left[\gamma_5^{(k)} - \frac{24}{(a_k+1)^{3/2}} \right] + \frac{H_7^{(k)}\left(\gamma_4^{(k)} - \frac{6}{(a_k+1)}\right) + \dots}{144\left(\gamma_4^{(k)} - \frac{6}{(a_k+1)}\right) + \dots} \right] \quad (28) \end{aligned}$$

Полученное модельное распределение, на базе трехпараметрической плотности (24), практически лишено недостатков, присущих усеченным рядам Эджворта, даже если порядок высших кумулянтов $\chi_r^{(k)}$ ограничен значением $r = 7$.

Практическое использование квадратичной ДФ для неинвазивной оценки диагностической эффективности биомедицинских признаков. Полученные математические модели оценивания рисков диагностических решений, при ограничениях на объем N первичной измерительной информации, был использован в задаче диагностики сердечных повреждений I и II ступеней [15] (состояния θ_0 и θ_1 , соответственно), когда характерная клиническая картина сердечной травмы отсутствует. Применяемые, в этом случае, неинвазивные методы диагностического обследования обусловлены длительностью времени ожидания, ограниченного фактором нестабильности гемодинамики.

Представляется интересным проанализировать диагностическую эффективность количественных и качественных биомедицинских признаков, используемых, фактически, на этапе сортировки пациентов с потенциально возможными сердечными повреждениями.

В табл. 2 и табл. 3 представлены результаты оценивания среднего риска \bar{R} диагностики отдельно по каждому количественному (табл. 2) и качественному признаку (табл. 3 – измерения проведены в бальной системе экспертного оценивания). Все результаты первичных измерений предварительно были пронормированы (приведены к интервалу $[0, 5]$), что устранило размерность признаков, но не отразилось на их вероятностно-статистических свойствах.

Экспериментальной базой проведенных исследований служило государственное предприятие «Институт общей и неотложной хирургии

им. В.Т. Зайцева национальной Академии медицинских наук Украины».

Таблица 2 – Оценки средних диагностических рисков \bar{R} для количественных признаков

Признак	Оценки средних и дисперсии				Объем выборки N	Средний риск \bar{R}
	$\mu_{(0)}^*$	$\mu_{(1)}^*$	$D_{(0)}^*$	$D_{(1)}^*$		
Объем кровопотери	1,421	2,608	0,689	1,482	68	0,316
Лейкоциты	2,081	2,968	0,259	0,174	39	0,318
Эритроциты	2,766	2,190	1,073	0,364	24	0,381
АД-давление	4,294	3,941	1,431	1,928	79	0,429
Пульс	3,263	3,488	1,611	2,370	81	0,454
АД-нижнее	4,542	4,251	1,566	2,983	80	0,455

Таблица 3 – Оценки средних диагностических рисков \bar{R} для качественных признаков

Признак	Оценки средних и дисперсии				Объем выборки N	Средний риск \bar{R}
	$\mu_{(0)}^*$	$\mu_{(1)}^*$	$D_{(0)}^*$	$D_{(1)}^*$		
Кожа и слизистые	1,350	1,758	0,156	0,378	70	0,392
Степень шока	1,461	2,39	0,777	2,008	38	0,401
Язык	1,133	1,432	0,121	0,252	54	0,407
Состояние	1,611	2,168	0,419	0,834	78	0,413
Сознание	1,137	1,452	0,122	0,563	82	0,449

Из табл. 2 видно, что наиболее информативными для неинвазивной диагностики сердечных травм являются «Объем кровопотери», «Лейкоциты», «Эритроциты». Все признаки в табл. 2 и табл. 3 расставлены, для удобства анализа, в порядке убывания \bar{R} .

Из табл. 3 видно, что диагностическая ценность качественных признаков ниже, чем признаков количественных. Более того эта ценность качественных признаков практически одинакова (\bar{R} изменяется незначительно).

Следует отметить, что максимальный порядок используемых для оценки \bar{R} кумулянтов квадратичной ДФ, равняется пяти ($r \leq 5$).

Выводы. 1. Разработана вероятностная модель статистической коррекции оцениваемых диагностических рисков при использовании квадратичной дискриминантной функции с неопределенными коэффициентами.

2. Разработана методика восстановления качества диагностических свойств квадратичной ДФ при ограничении обучающей измерительной информации для проектируемой ИИС параметрической диагностики.

3. Предложенные математические модели кумулянтного анализа квадратичной ДФ и комбинационной модели восстановления ее плотности распределения вероятностей показали их практическую эффективность (при ранжировании по минимуму среднего риска широко используемых нормативных биомедицинских диагностических признаков).

Список литературы

1. Раудис Ш. Ограниченность выборки в задачах классификации / Ш. Раудис // Статистические проблемы управления. – Вильнюс. – 1976. – Вып. 18. – С. 1–185.
2. Мигуценко Р. П. Исследование влияния ограниченности априорной информации на вид и размер достоверности диагностики / Р. П. Мигуценко // Вестник БГТУ им. В. Г. Шухова. – Белгород : БГТУ им. В. Г. Шухова. – 2014. – № 6. – С. 201–204.
3. Montgomery D. C. Introduction to Statistical Quality Control / D. C. Montgomery. – New York : John Wiley & Sons, 2001. – 796 p.
4. Малайчук В. П. Інформаційно-вимірвальні технології неруйнівного контролю : навч. посібник / В. П. Малайчук, О. В. Мозговой, О. М. Петренко. – Дніпропетровськ : РВВ ДНУ, 2001. – 240 с.
5. Уткин Л. В. Модель классификации на основе неполной информации о признаках в виде их средних значений / Л. В. Уткин, Ю. А. Жук, И. А. Селиховкин // Искусственный интеллект и принятие решений. – 2012. – № 3. – С. 71–81.
6. Щапов П. Ф. Синтез информационной модели процедуры альтернативной функциональной диагностики / П. Ф. Щапов, Р. П. Мигуценко // Приборы и методы измерений. – Минск. – 2014. – Вып. 2. – С. 94–100.
7. Щапов П. Ф. Теоретичні та практичні засади систем контролю та діагностування складних промислових об'єктів : монографія / П. Ф. Щапов, Р. П. Мигуценко, О. Ю. Крончак. – Х. : НТУ «ХПИ», 2015. – 260 с.
8. Щапов П. Ф. Повышение достоверности контроля и диагностики объектов в условиях неопределенности : монография / П. Ф. Щапов, О. Г. Аврунин. – Х. : ХНАДУ, 2011. – 191 с.
9. Королюк В. С. Справочник по теории вероятностей и математической статистике / В. С. Королюк, Н. И. Портенко / ред. В. С. Королюк. – К. : Наукова думка, 1978. – 584 с.
10. Шефе Г. Дисперсионный анализ / Г. Шефе / пер. с англ. Б. А. Севостьянов. – 2-е изд. – М. : Наука, 1980. – 512 с.
11. Малахов А. Н. Кумулянтный анализ случайных негауссовых процессов и их преобразований / А. Н. Малахов. – М. : Сов. радио, 1978. – 376 с.
12. Мигуценко Р. П. Контроль состояния динамических объектов с помощью однопараметровых тестовых статистик / Р. П. Мигуценко, О. Ю. Крончак // Вестник Казахской

- академии транспорта и коммуникаций им. Гынышпаева. – Алматы. – 2014. – № 2 (87). – С. 23–28.
13. Шемякин А. Е. Новый подход к построению объективных априорных распределений: информация Хеллингера / А. Е. Шемякин // Прикладная экометрика. – 2012. – № 4 (28). – С. 124–136.
 14. Смирнов Н. В. Курс теории вероятностей и математической статистики для технических приложений / Н. В. Смирнов, И. В. Дунин-Барковский. – М.: Наука, 1969. – 512 с.
 15. Бойко В. В. Хірургія серцевих ушкоджень (особливості сучасної доктрини) / В. В. Бойко, П. М. Замятин, І. В. Полівенок [та ін.]. – Харків: «Промінь», 2014. – 156 с.
- References (transliterated)**
1. Raudis, Sh. "Ogranichennost' vyborki v zadachakh klassifikatsii." *Statisticheskie problemy upravleniya*. Vol. 18. Vilnius, 1976. 1–185. Print.
 2. Migushchenko, R. P. "Issledovanie vliyaniya ogranichennosti apriornoj informatsii na vid i razmer dostovernosti diagnostiki." *Vestnik BGTU im. V. G. Shukhova*. No. 6. Belgorod: BGTU im. V. G. Shukhova, 2014. 201–204. Print.
 3. Montgomery, D. C. *Introduction to Statistical Quality Control*. New York: John Wiley & Sons, 2001. Print.
 4. Malaychuk, V. P., O. V. Mozhovoy and O. M. Petrenko. *Informatsiynno-vymiryuvalni tekhnolohiyi neruchynivnoho kontrolyu*. Dnipropetrovsk: RVV DNU, 2001. Print.
 5. Utkin, L. V., Yu. A. Zhuk and I. A. Selikhovkin. "Model' klassifikatsii na osnove nepolnoy informatsii o priznakakh v vide ikh srednikh znacheniy." *Iskusstvennyy intellekt i prinyatie resheniy*. No. 3. 2012. 71–81. Print.
 6. Shchapov, P. F., and R. P. Migushchenko. "Sintez informatsionnoy modeli protsedury al'ternativnoy funktsional'noy diagnostiki." *Pribory i metody izmereniy*. Vol. 2. Minsk, 2014. 94–100. Print.
 7. Shchapov, P. F., R. P. Myhushchenko and O. Yu. Kropachek. *Teoretychni ta praktychni zasady system kontrolyu ta diahnostuvannya skladnykh promyslovykh ob'yektiv*. Kharkiv: NTU «HPI», 2015. Print.
 8. Shchapov, P. F. and O. G. Avrunin. *Povyshenie dostovernosti kontrolya i diagnostiki ob'ektov v usloviyakh neopredelennosti*. Kharkov: KhNADU, 2011. Print.
 9. Korolyuk, V. S., and N. I. Portenko. *Spravochnik po teorii veroyatnostey i matematicheskoy statistike*. Ed. V. S. Korolyuk. Kiev: Naukova dumka, 1978. Print.
 10. Shefe, G. *Dispersionnyy analiz*. 2nd ed. Moscow: Nauka, 1980. Print.
 11. Malakhov, A. N. *Kumulyantnyy analiz sluchaynykh negaussovykh protsessov i ikh preobrazovaniy*. Moscow: Sov. radio, 1978. Print.
 12. Migushchenko, R. P., and O. Yu. Kropachek. "Kontrol' sostoyaniya dinamicheskikh ob'ektov s pomoshch'yu odnparametrovykh testovykh statistic." *Vestnik Kazakhskoy akademii transporta i kommunikatsiy im. Tynyshpaeva*. No. 2 (87). Almaty, 2014. 23–28. Print.
 13. Shemyakin, A. E. "Novyy podkhod k postroeniyu ob'ektivnykh apriornykh raspredeleniy: informatsiya Khellingera." *Prikladnaya ekometrika*. No. 4 (28). 2012. 124–136. Print.
 14. Smirnov, N. V., and I. V. Dunin-Barkovskiy. *Kurs teorii veroyatnostey i matematicheskoy statistiki dlya tekhnicheskikh prilozheniy*. Moscow: Nauka, 1969. Print.
 15. Boyko, V. V., et al. *Khirurgiya sertsevykh ushkodzen' (osoblyvosti suchasnoy doktryny)*. Kharkiv: Promin', 2014. Print.

Поступила (received) 31.10.2016

Бібліографічні описи / Библиографические описания / Bibliographic descriptions

Исследование свойств квадратичной дискриминантной функции при диагностировании динамических процессов / О. Ю. Кропачек // Вісник НТУ «ХПІ». Серія: Гідравлічні машини та гідроагрегати. – Х. : НТУ «ХПІ», 2016. – № 41 (1213). – С. 25–31. – Бібліогр.: 15 назв. – ISSN 2411-3441.

Исследование свойств квадратичной дискриминантной функции при диагностировании динамических процессов / О. Ю. Кропачек // Вісник НТУ «ХПІ». Серія: Гідравлічні машини та гідроагрегати. – Х. : НТУ «ХПІ», 2016. – № 41 (1213). – С. 25–31. – Библиогр.: 15 назв. – ISSN 2411-3441.

Studying the properties quadratic discriminant function in the diagnosis of dynamic processes / O. Yu. Kropachek // Bulletin of NTU "KhPI". Series: Hydraulic machines and hydrounits. – Kharkov : NTU "KhPI", 2016. – No. 41 (1213). – P. 25–31. – Bibliogr.: 15. – ISSN 2411-3441.

Відомості про авторів / Сведения об авторах / About the Authors

Кропачек Ольга Юрьевна – кандидат технічних наук, доцент, Національний технічний університет «Харківський політехнічний інститут», доцент кафедри «Теоретичні основи електротехніки»; тел.: (057) 707-69-61; e-mail: kropachek@ukr.net.

Кропачек Ольга Юрьевна – кандидат технических наук, доцент, Национальный технический университет «Харьковский политехнический институт», доцент кафедры «Теоретические основы электротехники»; тел.: (057) 707-69-61; e-mail: kropachek@ukr.net.

Kropachek Olha Yuriivna – Candidate of Technical Sciences (Ph. D.), Docent, National Technical University "Kharkiv Polytechnic Institute", Associate Professor at the Department of Theoretical Foundations of Electrical Engineering; tel.: (057) 707-69-61; e-mail: kropachek@ukr.net.