Научная статья
УДК 51-7
DOI 10.26297/2312-9409.2024.3.8

О методах нахождения вероятностей ошибок 1-го и 2-го рода при бинарной классификации

Николай Андреевич Белоусов[1✉], Денис Сергеевич Махов[2], Михаил Юрьевич Захаров[3]

[1,2]Краснодарское высшее военное орденов Жукова и Октябрьской Революции Краснознаменное училище имени генерала армии С.М. Штеменко, Краснодар, Россия
[3]Кубанский государственный университет, Краснодар, Россия
[1]loki911112@mail.ru✉

Аннотация. Предлагается исследование различных подходов к нахождению и анализу ошибок 1-го и 2-го рода и их вероятностей при проверке статистических гипотез. Проанализированы структурная и математическая составляющие указанной проблемы системной оценки достоверности статистических гипотез.

Ключевые слова: системный подход, системный анализ, статистические гипотезы, критическая область, ошибки 1-го и 2-го рода, уровень значимости, мощность критерия, случайное событие, вероятность, условная вероятность, плотность распределения случайной величины, мода случайной величины

Original article

About methods for finding probabilities of errors of the first and second kind in binary classification

Nikolay A. Belousov[1✉], Denis S. Mahov[2], Michail Y. Zaharov[3]
[1,2]Krasnodar Higher Military Orders of Zhukov and the October Revolution Red Banner School named after General of the Army S.M. Shtemenko, Krasnodar, Russia
[3]Kuban State University, Krasnodar, Russia
[1]loki911112@mail.ru ✉

Abstract. The study of various approaches to finding and analyzing errors of the first and second kind and their probabilities when testing statistical hypotheses is proposed. The structural and mathematical components of the specified problem of the systematic assessment of the reliability of statistical hypotheses are analyzed.

Keywords: a systematic approach, system analysis, statistical hypotheses, critical area, errors of the first and second kind, significance level, criterion power, random event, probability, conditional probability, distribution density of a random variable, mode of a random variable

Учёт, нахождение и анализ ошибок 1-го и 2-го рода, их вероятностей являются важными составляющими статистической обработки данных различных процессов, важным элементом системного подхода к исследованию последних. В данной работе делается попытка системного анализа структурной и математической составляющих указанной проблемы.

Известно по меньшей мере о двух подходах (математических методах) нахождения вероятностей ошибок 1-го и 2-го рода [1, 2]. Эти подходы, как будет описано ниже, достаточно сильно отличаются друг от друга и образуют своеобразную структуру возможных путей решения проблемы. Назовем эти подходы (методы) условно “прикладной” [1] и “теоретический” [2].

Актуальность предлагаемой работы обусловлена тем, что она помогает корректному моделированию и оптимизации вероятностей ошибок 1-го и 2-го рода, что в свою очередь позволяет (оптимально) улучшать достоверность бинарной классификации в различных процессах.

Напомним, что ошибка 1-го рода – отвергнуть правильную (основную) гипотезу, а ошибка 2-го рода – принять неправильную (основную) гипотезу [3].

Рассмотрим “прикладной” подход (метод) нахождения вероятностей ошибок 1-го и 2-го рода [1]. Основной гипотезой будем считать: Г = {объект действительно годный}, альтернативной гипотезой: Б = {объект действительно бракованный}. Введем также события:$Г_{конт}=¿${объект признан годным (по результатам контроля)}, $Б_{конт}=¿$ {объект признан бракованным (по результатам контроля)}. В этом случае вероятностям ошибок 1-го и 2-го будут, соответственно, условные вероятности $P(Б_{конт}|Г)$ и $P¿$). В обозначениях [1]

$$P(Б_{конт}|Г)=P_{Г-Б}=\int_{-\infty}^{C} f(x)\int_{C-x}^{+\infty} \varphi(¿\xi)\, d\xi dx, ¿ \quad (1)$$

где $f(x)$ – плотность вероятности распределения контролируемой величины x (для действительно годного объекта $x \leq C$);

φ(ξ) – плотность вероятности распределения погрешности оценки значения контролируемой величины x;

$$P(Г_{конт}|Б)=P_{Б-Г}=\int_{C}^{+\infty} f(x)\int_{-\infty}^{C-x} \varphi(¿\xi)\, d\xi dx. ¿ \quad (2)$$

Нужно отметить, что правая часть (1) верна и взята из [1] без изменений. В отличие от правой части (2), в которой в [1] допущена ошибка в верхнем пределе внутреннего интеграла
(*x*- *C*, а правильно: *C* -*x*).

Также очевидно, что при указанном подходе

$$P(Г)=\int_{-\infty}^{C} f(x)dx; P(Б)=\int_{C}^{+\infty} f(x)dx. \qquad (3)$$

Как видно из (1) и (2) плотность $f(x)$ вероятности распределения контролируемой величины x – одна и та же для действительно годного и для действительно бракованного объектов (в связи с этим уместно предположение о том, что $f(x)$ может быть плотностью двухмодального распределения, а соответствующие моды – условные математические ожидания контролируемой величины x для годных и для бракованных изделий).

Другими словами, плотность $f(x)$ построена с учетом как действительно годных, так и действительно бракованных объектов. И границей (в смысле контролируемой величины x) между этими совокупностями объектов является значение $x = C$ (границей во многом формальной), см. формулы (3).

Ошибки же 1-го и 2-го рода, как следует из (1) и (2), возникают при рассматриваемом подходе за счет соответствующих погрешностей оценки (измерения) контролируемой величины x. Что представляется недостаточно полным объяснением причин указанных ошибок и определенным недостатком соответствующих математических моделей (1) и (2). В качестве положительных моментов можно отметить относительную простоту моделей (1), (2) и (3).

В [2] приведен следующий “теоретический” подход (метод) нахождения вероятностей ошибок 1-го и 2-го рода.

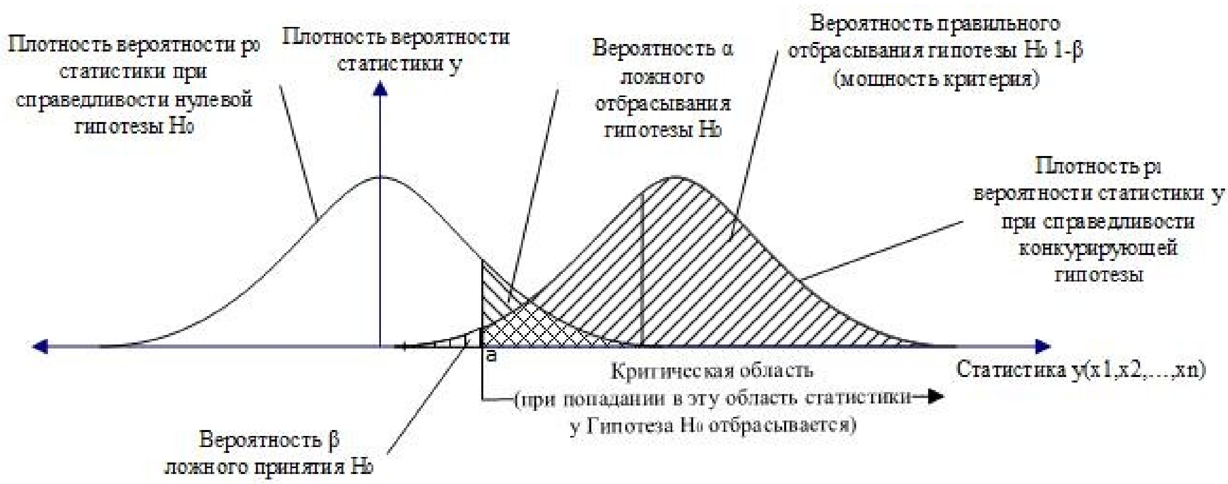

Рис. 1. Плотности вероятности статистики.
Figure 1. Probability densities of statistics.

Обозначим плотности вероятности статистики y при справедливости нулевой гипотезы H_0через $p_0(y)$, а при справедливости конкурирующей гипотезыH_1 – через $p_1(y)$ (см. рис. 1).

Тогда, очевидно, вероятности ошибок 1-го и 2-го рода – соответственно, α и β, вычисляются следующим образом:

$$\alpha = \int_{a}^{+\infty} p_0(y)\,dy, \qquad (4))$$

$$\beta = \int_{-\infty}^{a} p_1(y)\,dy, \qquad (5)$$

соответственно вероятность 1-α правильного принятия гипотезы H_0(где α – уровень значимости) и вероятность 1-β правильного отбрасывания H_0 (мощность критерия) вычисляются по формулам

$$1-\alpha = 1-\int_{a}^{+\infty} p_0(y)\,dy = \int_{-\infty}^{a} p_0(y)\,dy; \qquad (6)$$

$$1-\beta=1-\int_{-\infty}^{a} p_1(y)\,dy=\int_{a}^{+\infty} p_1(y)\,dy\,. \tag{7}$$

Как видно из вышеизложенного, при рассматриваемом подходе (методе) совокупности “действительно годных” и “действительно бракованных” объектов будут отличаться объективно – соответствующие плотности вероятности статистик у них будут различные ($p_0(y) и\, p_1(y)$¿. Представляется уместным предположение о том, что $p_0(y) и\, p_1(y)$ реализации условной плотности вероятности статистики y при различных условиях – гипотезах H_0 и H_1 соответственно. Таким образом, ошибки 1-го
и 2-го рода при данном подходе обусловлены объективными причинами. Математические модели (4) и (5) ((6) и (7)) представляются вполне адекватными и обоснованными. В качестве рекомендаций для развития указанного подхода можно назвать разработку способа нахождения вероятностей (справедливости) гипотез H_0 и H_1:$p(H_0), p(H_1)$, а также способа построения плотности вероятности $p_1(y)$ при известной плотности $p_0(y)$ (либо наоборот – построение $p_0(y)$ при известной $p_1(y)$).

Как показал вышеприведённый анализ двух подходов (методов) нахождения вероятностей ошибок 1-го и 2-го рода ([1], [2]), эти методы достаточно сильно отличаются друг от друга и образуют структуру возможных путей оценки достоверности проверки статистических гипотез. Отличия “прикладного” и “теоретического” подходов обусловлены наличием в первом случае одной – общей для обеих гипотез – плотности вероятности распределения статистики, а во втором случае – своей плотности распределения статистики для каждой из гипотез H_0 и H_1. Также в “прикладном” подходе используется такая прикладная характеристика, как плотность вероятности распределения погрешности оценки значения статистики.

По итогам рассмотрения обоих подходов можно сделать следующие выводы. “Прикладной” подход имеет более простые, хотя и неполные математические модели и возможно лучше адаптирован к практическому применению за счет явного учета погрешности оценки статистики. “Теоретический” подход использует вполне адекватные и обоснованные математические модели, применение которых несколько затруднено в силу необходимости определения некоторых параметров.

В заключение необходимо отметить целесообразность дальнейшего исследования “теоретического” подхода [2], в частности, в направлении анализа вероятностей, связанных с ошибками 1-го и 2-го рода. Актуальность подобных исследований обусловлена необходимостью оптимального улучшения достоверности бинарной классификации.

Список источников

1. ГОСТ Р 8.731-2010. Системы допускового контроля. Основные положения (2010) // Государственная система обеспечения единства измерений. М.: Стандартинформ.

2. *Корн Г., Корн Т.* Справочник по математике для научных работников и инженеров. 2-е изд. М.: Изд-во «Наука», 1973. 832 с.

3. *Гмурман В.Е.* Теория вероятностей и математическая статистика: учебник для вузов. 12-е изд. М.: Изд-во "Юрайт", 2023. 479 с.

References

1. *GOST* R 8.731-2010. *Sistemy dopuskovogo kontrolya. Osnovnye polozheniya* (2010) // *Gosudarstvennaya systema obespecheniya edinstva izmerenii* (GOST R 8.731-2010. Admission control systems. Basic provisions (2010) // State system for ensuring the uniformity of measurements), Moscow.

2. Korn G., Korn T.*, Spravochnik po matematike dlya nauchnyh rabotnikov i ingenerov* (Handbook of Mathematics for researchers and engineers), vol. 2, Moscow, 1973, p. 832.

3. Gmurman V.E., *Teoria veroyatnostei I matematicheskaya statistika: uchebnik dlya vuzov* (Probability theory and mathematical statistics: textbook for universities), vol. 12, Moskow, 2023, p. 479.