

УДК 662.831.32
MSC2010 05C40+05C70

© М. А. Гузев^{1,2}, Г. Ш. Цициашвили^{1,2}

Алгоритм прогнозирования критических событий по обобщенному показателю

В настоящей заметке дано обоснование алгоритма прогнозирования редких критических событий.

Ключевые слова: *критические события, обучающая выборка, частота событий.*

Введение

В работе [1] рассмотрена задача прогнозирования критического события в горной шахте по значению обобщенного показателя, вычисленного на основании данных акустического мониторинга [2]. С точки зрения обработки данных специфика этой задачи состоит в том, что критические события являются редкими. Настоящая задача была решена путем выбора способа прогнозирования критического события по результатам анализа частот прогнозирования наличия критического события и его отсутствия в обучающей выборке.

Однако в [1] не было представлено полного математического обоснования алгоритма построения этого правила. В настоящей заметке восполнен этот информационный пробел.

1. Основные результаты

Пусть задана обучающая выборка, состоящая из двумерных векторов (x_k, y_k) , $k = 1, \dots, n$, где $x_k > 0$ — значение обобщенного показателя, y_k — логическая переменная, характеризующая наличие критического события. Если $y_k = \times$, то событие присутствует в момент времени k , в противном случае $y_k = \bullet$ (рис. 1).

Для предсказания критического события сформулируем следующее правило: если $x_k \geq c$, где $c > 0$, то в момент k происходит критическое событие, в противном

¹ Дальневосточный федеральный университет, 690950, г. Владивосток, ул. Суханова, 8.

² Институт прикладной математики ДВО РАН, 690041, Владивосток, ул. Радио, 7.

Электронная почта: guzev@iam.dvo.ru (М. А. Гузев), guram@iam.dvo.ru (Г. Ш. Цициашвили).



Рис. 1. Представление обучающей выборки на прямой набором символов \times , \bullet .

случае прогнозируется отсутствие критического события. Пусть в обучающей выборке содержится n_{\times} моментов наступления критических событий и $n_{\bullet} = n - n_{\times}$ моментов отсутствия критических событий. Обозначим $n_{\times}(c)$ количество правильных прогнозов критического события и $n_{\bullet}(c)$ — количество правильных прогнозов отсутствия критического события при заданном значении c (рис. 2, 3).



Рис. 2. Определение числа $n_{\times}(c)$: точки \times правее c .

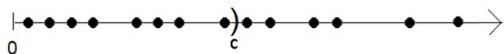


Рис. 3. Определение числа $n_{\bullet}(c)$: точки \bullet левее c .

Введем частоты $\rho_{\times}(c) = n_{\times}(c)/n_{\times}$, $\rho_{\bullet}(c) = n_{\bullet}(c)/n_{\bullet}$ прогноза критического события и прогноза отсутствия критического события. В [1] возникла задача корректного определения значения c , задающего границу для правил прогноза. Для этого можно потребовать, чтобы выполнялось соотношение $\rho_{\times}(c) \geq a$, где величина a определяется экспертным методом и характеризует допустимую частоту ошибок прогнозирования. При выполнении этого соотношения поставим задачу максимизации $\rho_{\bullet}(c)$ и обозначим через c_* решение этой оптимизационной задачи:

$$\max(\rho_{\bullet}(c) : \rho_{\times}(c) \geq a) \Rightarrow c = c_*. \tag{1}$$

Рассмотрим вспомогательную оптимизационную задачу:

$$\max(c : \rho_{\times}(c) \geq a) \Rightarrow c = c^*. \tag{2}$$

Теорема 1. Задачи (1), (2) имеют единственные решения c_* , c^* , и справедливо равенство $c_* = c^*$.

Доказательство. Из определения функции $\rho_{\times}(c)$, $c \geq 0$, следует, что она является ступенчатой, невозрастающей и непрерывной слева (рис. 4), а функция $\rho_{\bullet}(c)$, $c \geq 0$, является ступенчатой, неубывающей и непрерывной справа (рис. 5).

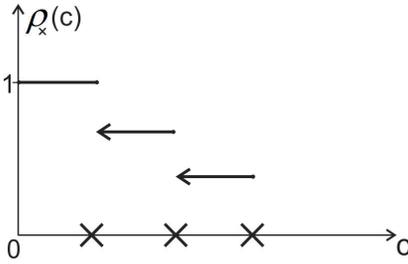


Рис. 4. Вид функции $\rho_{\times}(c)$.

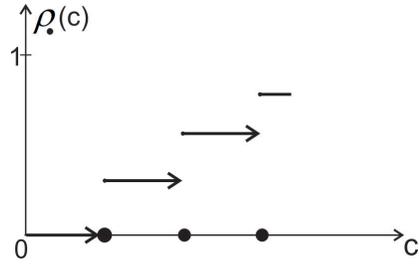


Рис. 5. Вид функции $\rho_{\bullet}(c)$.

Из свойств функции $\rho_{\times}(c)$ следует, что множество $\{c : \rho_{\times}(c) \geq a\} = [0, c^*]$ и значит точка c^* определяется единственным способом. В свою очередь,

$$\max(\rho_{\bullet}(c) : \rho_{\times}(c) \geq a) = \max(\rho_{\bullet}(c) : c \in [0, c^*]). \tag{3}$$

Так как функция $\rho_{\bullet}(c)$, $c \geq 0$, является неубывающей и непрерывной справа, то ее максимум на отрезке $[0, c^*]$ достигается в единственной точке c_* . При этом возможны два варианта расположения c_* по отношению к точкам \bullet . Для первого из них на рис. 6 точка c_* совпадает с точкой \bullet , для второго на рис. 7 точка c_* лежит или между соседними точками \bullet , или между нулем и ближайшей к нему точкой \bullet .

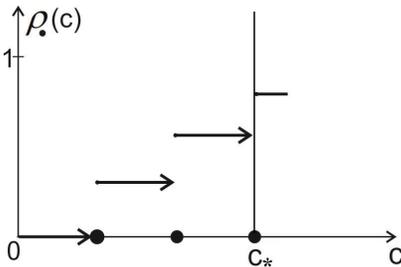


Рис. 6. Первый вариант расположения c_* .

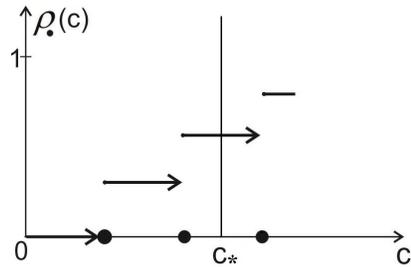


Рис. 7. Второй вариант расположения c_* .

□

Замечание 1. При вычислении c_* требуется знание всей выборки, состоящей из n элементов, а при вычислении c^* требуется знание только ее части, состоящей из n_{\times} элементов. Вследствие того, что критические события являются редкими, справедливо соотношение $n_{\times} \ll n$.

Замечание 2. Функция $\rho_{\times}(c)$ характеризует степень безопасности работы в шахте, а функция $\rho_{\bullet}(c)$ характеризует отсутствие ложных тревог, что влияет на экономические результаты деятельности предприятия.

Список литературы

- [1] M. A. Guzev, I. Yu. Rasskazov, G. Sh. Tsitsiashvili, “Algorithm of potentially burst-hazard zones dynamics representation in massif of rocks by results of seismic-acoustic monitoring”, *Procedia Engineering*, **191**, (2017), 36–42.
- [2] И. Ю. Рассказов, *Контроль и управление горным давлением на рудниках Дальневосточного региона*, Горная книга, Москва, 2008.

Поступила в редакцию
13 февраля 2019 г.

Исследование выполнено при частичной финансовой поддержке РФФИ (проект № 17-07-00177).

Guzev M. A., Tsitsiashvili G. Sh. An algorithm for predicting the critical events in the generalized indicator. *Far Eastern Mathematical Journal*. 2019. V. 19. No 1. P. 20–23.

ABSTRACT

In this note, the substantiation of the algorithm for predicting rare critical events is given.

Key words: *critical events, training sample, frequency of event.*