А. А. Грушо, Н. А. Грушо, Е. Е. Тимонина (Москва, МГУ, РГГУ). **Существование состоятельных последовательностей статистических критериев в дискретных задачах.**

Пусть задано конечное множество $X = \{x_1, \ldots, x_m\}$. Через X^∞ обозначим множество бесконечных последовательностей, где каждый элемент последовательности принадлежит X. Пусть $\mathcal{A}$ — σ-алгебра, порожденная цилиндрическими множествами. Пусть P_0 — вероятностная мера на измеримом пространстве $(X^\infty, \mathcal{A})$. Кроме того, пусть задано семейство вероятностных мер P_θ, $\theta \in \Theta$, на том же измеримом пространстве.

Рассмотрим проекции введенных вероятностных мер на первые n координат случайных последовательностей из X^∞ и обозначим их соответственно $P_{0,n}$ и $P_{\theta,n}$. Относительно данных мер рассмотрим для каждого n задачу проверки статистической гипотезы $H_{0,n} : P_{0,n}$ против сложной альтернативы $H_{1,n} : \{P_{\theta,n},\, \theta \in \Theta\}$. Критерий уровня значимости α_n описывается критическим множеством S_n, $P_{0,n}(S_n) \leqq \alpha_n$ и мощностью критерия $W_n(\theta) = P_{\theta,n}(S_n)$. Последовательность статистических критериев с критическими множествами S_n называется состоятельной, если $P_{0,n}(S_n) \to 0$ и $W_n(\theta) \to 1$ для каждого $\theta \in \Theta$.

Статистический критерий можно описать функцией принадлежности полученного в наблюдении случайного элемента критическому множеству критерия. В дискретных задачах можно говорить о сложности вычисления функции принадлежности некоторого элемента заданному множеству. Таким образом, можно говорить о сложности вычисления статистического критерия. В рассмотренной выше модели мы имеем последовательность статистических критериев для каждого n при $n \to \infty$. Отсюда следует возможность асимптотической оценки сложности последовательности критериев.

Соотношение между статистическими свойствами критерия и сложностью его вычисления редко исследовалась в дискретной математике. Естественно рассматривать свойство состоятельности последовательности критериев и асимптотическую сложность этой последовательности критериев. Например, в работе [1] было показано, что задача выявления выделенного множества вершин в случайной графе имеет состоятельное решение, когда мощность выделенного множества вершин растет со скоростью $C \ln n$, где n — число вершин графа. Однако состоятельная последовательность критериев асимптотически имеет высокую сложность. Если рассматривать просто вычислимые критерии, то соответствующая граница выявляемости выделенного множества становится равной $C\sqrt{n}$.

Вопрос о том, может ли любое снижение сложности привести к потере состоятельности, остался открытым. Кроме того, можно ли считать, что снижение сложности последовательности критериев и сохранении свойства состоятельности достигается за счет отбрасывания некоторого множества альтернатив? В работе, представленной данным сообщением, в рамках приведенной в начале статьи модели дается положительное решение этих двух задач.

Назовем одноточечным распределением вырожденное распределение вероятностей, сосредоточенное в одной точке пространства X^∞. Пусть P_0 — это одноточечное распределение, сосредоточенное в точке $\omega = (x_1, x_1, \ldots, x_1, \ldots)$, все координаты которой равны x_1. В качестве альтернатив рассмотрим множество одноточечных распределений $\{P_\theta\}$, $\theta \in X^\infty$, $\theta \neq \omega$. В качестве основной операции для оценки сложности возьмем сравнение очередного знака наблюденной последовательности с x_1. Для каждого n проекция $P_{0,n}$ является вырожденным распределением на X^n, сосредоточенном в точке ω_n, где ω_n — вектор длины n, состоящий из координат x_1. Каждая из альтернатив $P_{\theta,n}$ также является вырожденным распределением, сосредоточенным в точке θ_n. В качестве критического множества при проверке гипотезы $H_{0,n}$ против сложной альтернативы $H_{1,n}$ положим $S_n = X^n \backslash \{\omega_n\}$. Для этого критерия $P_{0,n}(S_n) = 0$ и $P_{\theta,n}(S_n) = 1$ для любого θ, для которого проекция $\theta_n \neq \omega_n$.

Сложность вычисления данного критерия не превосходит n для всех точек пространства X^n.

Таким образом определенная последовательность критериев является состоятельной. В самом деле,

$$\lim_{n\to\infty} P_{0,n}(S_n) = 0$$

и для каждого $\theta \in X^\infty$, $\theta \neq \omega$, существует n такое, что $\theta_n \neq \omega_n$. Тогда, начиная с этого n, мощность $W_n(\theta) = 1$.

Сложность вычисления любого другого критерия может быть снижена только за счет того, что мы не будем сравнивать каждую координату наблюденного вектора из X^n с единственным допустимым значением x_1. Тогда если сложность последовательности статистических критериев равна $o(n)$, то это значит, что при вычислении функции принадлежности критическому множеству не просматривается значительная часть координат. Будем считать, что последовательность множеств не просмотренных координат монотонно не убывает с ростом n. Тогда последовательность критических множеств D_n для такой последовательности критериев будет обладать свойством

$$(X^n \setminus D_n) \times X^\infty \supset \{\omega,\, \theta\},$$

т. е. для всех n, начиная с некоторого, в цилиндрические множества, порожденные критическими множествами критериев, не будет входить хотя бы одна последовательность $\theta \neq \omega$. Значит для этой последовательности θ мощность критерия не стремится к 1.

Таким образом мы доказали, что снижение сложности критерия приводит к потере свойства состоятельности для соответствующей последовательности критериев. Вместе с тем, если в последнем примере отбросить из альтернатив все те последовательности, которые никогда не встречаются в последовательности цилиндрических множеств, порожденных критическими множествами критериев, то полученная последовательность критериев снова будет состоятельной против нового класса альтернатив.

Работа поддержана грантами РФФИ, проект 07-01-00484, проект 07-07-00236.

СПИСОК ЛИТЕРАТУРЫ

1. *Грушо А. А.* Некоторые статистические задачи на графах. — Матем. заметки, 1984, т. 36, в. 2.