Доклады Академии наук СССР
1973. Том 209, № 1

УДК 519.281

МАТЕМАТИКА

Н. П. САЛИХОВ

АСИМПТОТИЧЕСКИЕ СВОЙСТВА ВЕРОЯТНОСТЕЙ ОШИБОК КРИТЕРИЕВ ДЛЯ РАЗЛИЧЕНИЯ НЕСКОЛЬКИХ МУЛЬТИНОМИАЛЬНЫХ СХЕМ ИСПЫТАНИЙ

(Представлено академиком Ю. В. Линником 22 V 1972)

Поведение вероятностей ошибок критериев, основанных на фиксированном объеме выборки, при стремлении объема выборки к бесконечности, изучено достаточно хорошо в случае различения двух статистических гипотез. В настоящей работе рассмотрен случай нескольких гипотез специального вида. Пусть гипотеза H_t, $t\in 1,\ldots,m$, состоит в том, что последовательность $b_1, b_2, \ldots$ является последовательностью независимых испытаний по мультиномиальной схеме с исходами $a_1,\ldots,a_r$; вероятности появления которых при каждом испытании равны соответственно $p_{t1},\ldots,p_{tr}$. Предполагается, что $m\geqslant 2$, $r\geqslant 2$, $p_{t1}>0,\ldots,p_{tr}>0$, $\sum\limits_{i=1}^{r} p_{ti}=1$ для $t\in 1,\ldots,m$, $\sum\limits_{i=1}^{r}|p_{ti}-p_{si}|>0$ для $t\neq s$.

Испытания $b_1,\ldots,b_n$ будем характеризовать точкой $\mathbf{u}=(x_1/n,\ldots \ldots, x_r/n)$, в которой x_k — количество появлений исхода a_k среди $b_1,\ldots \ldots, b_n$. Пусть Ω_n — пространство возможных значений $\mathbf{u}$. Гипотеза H_t определяет на Ω_n вероятностную меру $\mu_t^n(\mathbf{u})=n!(x_1!\ldots x_r!)^{-1}p_1^{x_1}\ldots \ldots p_{tr}^{x_r}$.

Пусть A_n^0 — множество всех разбиений $e^n=(e_0^n, e_1^n,\ldots,e_m^n)$ пространства Ω_n на $m+1$ взаимно непересекающихся множеств e_i^n. Каждое разбиение e^n определяет следующий нерандомизированный критерий для различения альтернативных гипотез H_t, основанный на испытаниях $b_1,\ldots,b_n$. Если характеризующая испытания $b_1,\ldots,b_n$ точка $\mathbf{u}\in e_t^n$, то принимается гипотеза H_t при $t\in 1,\ldots,m$ и не принимается ни одна из гипотез при $t=0$. Разбиение e^n будем характеризовать вектором потерь $(\alpha_1^n(e_1^n),\ldots,\alpha_m^n(e_m^n))$, в котором $\alpha_t^n(e_t^n)=1-\mu_t^n(e_t^n)$ — вероятность ошибки, совершаемой, когда не принимается истинная гипотеза H_t.

Пусть A^0 — множество всех последовательностей $\{e^n\}=(e^1, e^2,\ldots)$, в которых $e^n\in A_n^0$ при каждом $n\geqslant 1$. В дальнейшем будем рассматривать некоторые непустые множества $N\subseteq A^0$ и $J\subseteq 1,\ldots,m$. Для любой последовательности $\{e^n\}\in A^0$ представляет интерес рассмотреть поведение при $n\to\infty$ величины $P_n(e^n, J)=\max\alpha_t^n(e_t^n)$ по $t\in J$.

Определение 1. *J-Мерой асимптотической эффективности* для последовательности $\{e^n\}\in A^0$ назовем величину $M(\{e^n\}, J)=-\overline{\lim}\,T_n(e^n, J)$, $n\to\infty$, где $T_n(e^n,J)=n^{-1}\ln P_n(e^n,J)$, если $P_n(e^n,J)>0$, $T_n(e^n,J)=-\infty$, если $P_n(e^n,J)=0$.

Определение 2. Последовательность $\{g^n\}\in N$ назовем *J-оптимальной* последовательностью множества $N\subseteq A^0$, если $M(\{g^n\}, J)= -\lim T_n(g^n, J)=M(N, J)$, $n\to\infty$, где $M(N, J)=\sup M(\{e^n\}, J)$ по $\{e^n\}\in N$.

Последовательность $\{g^n\}$ имеет следующие свойства:

1) Если $\{e^n\}\in N$, $M(\{e^n\}, J)<M(N, J)$, то $\lim P_n(g^n, J)(P_n(e^n, J))^{-1}=0$, $n\to\infty$.

2) $P_n(g^n, J) = \exp[-nM(N, J) + o(n)]$ и не существует последовательности $\{e^n\} \in N$ такой, что $P_n(e^n, J) = \exp[-nM + o(n)]$, где $M > M(N, J)$.

В связи с этим представляет интерес задача о нахождении J-оптимальных последовательностей и величин $M(N, J)$ для заданных N и J.

1. Множества N. Пусть $A_n = \{e^n\colon e^n \in A_n^0,\ e_0^n = \phi\}$, а C_n (соответственно B_n) — подкласс всех допустимых (байесовских) разбиений класса A_n [3]. Подкласс B_n представляет собой сумму множеств

$$B_n(\lambda) = \left\{e^n\colon e^n \in A_n, \quad \sum_{i=1}^{m} \lambda_i \alpha_i^n(e_i^n) = \min \sum_{i=1}^{m} \lambda_i \alpha_i^n(g_i^n) \text{ по } g^n \in A_n\right\} \quad \text{по}$$

всем векторам $\lambda = (\lambda_1, \ldots, \lambda_m)$ априорных вероятностей λ_i [3]. Заметим, что $B_n(\lambda) = B_n(c\lambda)$ при любом $c > 0$.

В дальнейшем в качестве множества N рассмотрим множество A^0 и аналогично определяемые множества A, B, C, а также следующие множества, зависящие от некоторого $I \subseteq 1, \ldots, m$:

$$P(I) = \{\{e^n\}\colon \{e^n\} \in A^0,\ \lim n^{-1} \ln \mu_t^n(e_t^n) = 0,\ n \to \infty, \quad \text{для } t \in I\},$$

$$P_1(I) = \{\{e^n\}\colon \{e^n\} \in A^0,\ \lim \mu_t^n(e_t^n) = 1,\ n \to \infty, \quad \text{для } t \in I\},$$

$$Q(I) = \{\{e^n\}\colon \{e^n\} \in A^0,\ \lim n^{-1} \ln \alpha_t^n(e_t^n) = 0,\ n \to \infty, \quad \text{для } t \in I\},$$

$$Q_1(I) = \{\{e^n\}\colon \{e^n\} \in A^0,\ \lim \alpha_t^n(e_t^n) = 1,\ n \to \infty, \quad \text{для } t \in I\}.$$

Справедливы включения $B \subseteq C \subseteq A \subseteq A^0$, $P_1(I) \subseteq P(I) \subseteq \bigcup_{I=1}^{m} P(I) \subseteq A^0$, $Q_1(I) \subseteq Q(I) \subseteq \bigcup_{I=1}^{m} Q(I) \subseteq A^0$.

Последовательности

$$\{e^n\} \in \left(A^0 \setminus \bigcup_{I=1}^{m} P(I)\right) \cap \left(A^0 \setminus \bigcup_{I=1}^{m} Q(I)\right)$$

не имеют большого практического значения, так как для любой такой последовательности при каждом $t \in 1, \ldots, m$ найдутся сколь угодно большие n такие, что $\mu_t^n(e_t^n) > 1 - \exp(-n\varepsilon_t)$ при $\varepsilon_t > 0$, и одновременно найдутся сколь угодно большие n такие, что $\mu_t^n(e_t^n) < \exp(-n\delta_t)$ при $\delta_t > 0$.

2. Построение J-оптимальных последовательностей и вычисление величин $M(N, J)$. J-Оптимальные последовательности введенных выше множеств N можно построить, используя только разбиения из множеств $B_n(\lambda_n)$ при специальном выборе последовательности $\{\lambda_n\}$ векторов $\lambda_n = (\lambda_{1n}, \ldots, \lambda_{mn})$. Эти разбиения можно построить с помощью обобщенной леммы Неймана — Пирсона [3].

Величины $M(N, J)$ можно вычислить, используя только расстояния Кульбака — Лейблера и Чернова. Расстоянием Кульбака — Лейблера от точки $\mathbf{p}_t = (p_{t1}, \ldots, p_{tr})$ до точки $\mathbf{p}_s = (p_{s1}, \ldots, p_{sr})$ назовем величину $I(\mathbf{p}_t, \mathbf{p}_s) = \sum_{i=1}^{r} p_{ti} \ln\left(\frac{p_{ti}}{p_{si}}\right)$. Рассмотрим отрезок кривой

$$u_i(x) = p_{si}^{1-x} p_{ti}^{x} \left(\sum_{\nu=1}^{r} p_{s\nu}^{1-x} p_{t\nu}^{x}\right)^{-1}, \quad i \in 1, \ldots, r, \quad x \in [0, 1],$$

соединяющий точки $\mathbf{p}_t$ и $\mathbf{p}_s$. На этом отрезке существует единственная точка $\mathbf{u}(z)$, где z — единственный положительный корень уравнения $\sum_{i=1}^{r} p_{si}^{1-x} p_{ti}^{x} \ln\left(\frac{p_{ti}}{p_{si}}\right) = 0$, такая, что

$$I(\mathbf{u}(z), \mathbf{p}_t) = I(\mathbf{u}(z), \mathbf{p}_s) = -\ln \sum_{i=1}^{r} p_{si}^{1-z} p_{ti}^{z}.$$

Таблица 1

Множество N	J-Оптимальная последовательность множества N	$M(N, J)$
A^0, B, C, A; $P(I)$, $P_1(I)$, $I \subseteq J$; $Q(I)$, $Q_1(I)$, $I \cap J = \phi$	$\{f^n\}: f^n \in B_n(\lambda_n)$, $\lambda_{tn} = \begin{cases} 1 \text{ для } t \in J, \ n \geqslant 1, \\ 0 \text{ для } t \in 1, \ldots m \setminus J, \ n \geqslant 1, \end{cases}$	∞, если $\lvert J \rvert = 1$; M_J, если $\lvert J \rvert > 1$
$P(I)$, $P_1(I)$, $I \setminus J \neq \phi$	$\{g^n\}: g^n \in B_n(\lambda_n)$, $\lambda_{tn} = \begin{cases} 1 & \text{для } t \in I \setminus J, \ n \geqslant 1, \\ \exp\,[n(M_{IJ} - \varepsilon_n)] & \text{для } t \in J, \ n \geqslant 1, \\ 0 & \text{для } t \in 1, \ldots, m \setminus (I \cup J), \ n \geqslant 1 \end{cases}$	M_{IJ}
$Q(I)$, $Q_1(I)$, $I \cap J \neq \phi$	$\{e^n\}$ — любая последовательность множества N	0

Последнюю величину обозначим через $I_c(\mathbf{p}_t, \mathbf{p}_s)$ и назовем расстоянием Чернова ([7], [8]) между точками $\mathbf{p}_t$, $\mathbf{p}_s$. (На возможность геометрического представления расстояния Чернова указал автору Н. Н. Ченцов.)

Основные результаты приведены в табл. 1.

В табл. 1 $\{\varepsilon_n\}$ — любая монотонно убывающая последовательность положительных чисел, удовлетворяющая условиям $\lim \varepsilon_n = 0$, $\lim n\varepsilon_n^2 = \infty$, $n \to \infty$, а величины M_J, M_{IJ} вычисляются по формулам

$$M_J = \min I_c(\mathbf{p}_t, \mathbf{p}_s) \text{ по } t \in J, \ s \in J, \ t > s,$$

$$M_{IJ} = \begin{cases} \delta(I, J), & \text{если } \lvert J \rvert = 1, \\ \min(\delta(I, J), M_J, & \text{если } \lvert J \rvert > 1, \end{cases}$$

где $\delta(I, J) = \min I(\mathbf{p}_t, \mathbf{p}_s)$ по $t \in I \setminus J$, $s \in J$.

В некоторых случаях можно указать и другие J-оптимальные последовательности, что видно из следующих теорем.

Теорема 1. *Пусть $B_n \subseteq F_n \subseteq A_n^0$ при $n \geqslant 1$.*

Тогда при $\lvert J \rvert > 1$ для величин $P_n(\bar{g}^n, J) = \min P_n(e^n, J)$ по $e^n \in F_n$ справедливо соотношение

$$\lim n^{-1} \ln P_n(\bar{g}^n, J) = -M_J, \ n \to \infty.$$

Таким образом, например, при $F_n = B_n$ последовательность $\{\bar{g}^n\}$ является J-оптимальной последовательностью множества B.

Теорема 2. *Если J состоит из одного элемента $s \in 1, \ldots, m$, а $I = 1, \ldots, m \setminus J$, и множество $N(I) = P(I) \cap Q(I) \cap C \neq \phi$, то любая его последовательность является J-оптимальной последовательностью множества $P(I)$.*

Для случая $m = 2$ из результатов работ ([6], [10], [11]) можно получить утверждения, близкие к некоторым утверждениям настоящего пункта. При $m \geqslant 2$ для доказательства утверждений данного пункта используются результаты работ ([4], [5], [7], [9]).

3. Связь с результатами Чернова и Вайды. Следующие теоремы можно доказать с помощью результатов пункта 2.

Теорема 3. *Пусть $B_n \subseteq F_n \subseteq A_n^0$ для $n \geqslant 1$; $a_{tn} > 0$ для $n \geqslant 1$, $t \in J$; $\lim n^{-1} \ln a_{tn} = 0$, $n \to \infty$, для $t \in J$.*

Тогда при $\lvert J \rvert > 1$ для величин $g(F_n, J) = \min \sum_{t \in J} a_{tn} \alpha_t^n (e_t^n)$ по $e^n \in F_n$ справедливо соотношение

$$\lim n^{-1} \ln g(F_n, J) = -M_J, \quad n \to \infty.$$

Для случая, когда $m=2$, $J=1, \ldots, m$, $a_{tn}=c_t>0$ для $t \in J$, $n \geqslant 1$, $F_n = B_n$ для $n \geqslant 1$ утверждение теоремы следует из результатов работ [7, 8].

Теорема 4. *Пусть* $|J|>1$,

$$h_n(J) = 1 - |J|^{-1} \sum_{\mathbf{u} \in \Omega_n} \left(\sum_{t \in J} (\mu_t^n(\mathbf{u}))^2 \Big/ \sum_{t \in J} \mu_t^n(\mathbf{u}) \right),$$

$$H_n(J) = -|J|^{-1} \sum_{\mathbf{u} \in \Omega_n} \sum_{t \in J} \mu_t^n(\mathbf{u}) \ln \left(\mu_t^n(\mathbf{u}) \Big/ \sum_{t \in J} \mu_t^n(\mathbf{u}) \right).$$

Тогда

$$\underline{\lim}\, n^{-1} \ln H_n(J) \geqslant \lim n^{-1} \ln h_n(J) = -M_J, \quad n \to \infty.$$

Это утверждение доказывается с помощью оценок Вайды [1]:

$$1 - \sqrt{1 - h_n(J)} \leqslant |J|^{-1} \sum_{t \in J} \alpha_t^n (f_t^n) \leqslant h_n(J) \leqslant (\ln 4)^{-1} H_n(J),$$

где $\{f^n\}$ — последовательность из табл. 1. Относительно величины $H_n(J)$ из результатов работ [2, 12] следует более точное утверждение $\lim n^{-1} \ln H_n(J) = -M_J$, $n \to \infty$.

Автор выражает благодарность Н. Н. Ченцову и Г. П. Климову за обсуждение результатов работы.

Поступило
19 V 1972

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] И. Вайда, Проблемы передачи информации, 4, 1, 9 (1968). [2] И. Вайда, Тр. I Всесоюзн. симпозиума по статистическим проблемам в технической кибернетике. Идентификация и аппаратура для статистических исследований, 1970, стр. 137. [3] С. Р. Рао, Линейные статистические методы и их применения, М., 1968. [4] И. Н. Санов, Математич. сборн., 42 (84), 1, 11 (1957). [5] Б. С. Флейшман, Конструктивные методы оптимального кодирования для каналов с шумами, М., 1963. [6] Э. М. Хазен, Методы оптимальных статистических решений и задачи оптимального управления, М., 1968. [7] H. Chernoff, Ann. Math. Stat., 23, 4, 493 (1952). [8] H. Chernoff, Ann. Math. Stat., 27, 1, 1 (1956). [9] W. Hoeffding, Ann. Math. Stat., 36, 2, 369 (1965). [10] O. Krafft, D. Plachky, Ann. Math. Stat., 41, 5, 1646 (1970). [11] C. R. Rao, J. Roy. Statist. Soc., B, 24, 1, 46 (1962). [12] I. Vajda, Kybernetika, 6, 4, 306 (1970).