УДК 519.24 : 621.391.2

Я. Ю. НИКИТИН, Р. П. ФИЛИМОНОВ

(Ленинград)

АСИМПТОТИЧЕСКАЯ ОТНОСИТЕЛЬНАЯ ЭФФЕКТИВНОСТЬ НЕКОТОРЫХ НЕПАРАМЕТРИЧЕСКИХ ПРАВИЛ ОБНАРУЖЕНИЯ В СХЕМЕ ДВУХКАНАЛЬНОЙ ОБРАБОТКИ

1. **Введение.** Создание устройств автоматической обработки изображений, в частности систем распознавания объектов когерентно-оптическими методами [1, 2], тесно связано с применением правил обнаружения сигналов, работающих в условиях априорной неопределенности вероятностных характеристик изображений и помех. Наиболее распространенные способы преодоления априорной неопределенности — использование параметрических инвариантных либо непараметрических правил обнаружения. Указанные правила не являются оптимальными, поэтому практическое значение приобретают методы сравнительной оценки их качества, позволяющие обосновать эмпирические процедуры выбора.

В статье [3] такие оценки получены на основе понятия асимптотической относительной эффективности (АОЭ) по Питмену для ряда параметрических инвариантных к уровню помех правил обнаружения. Наряду с инвариантными, при решении задач распознавания естественно испытать и различные непараметрические правила. АОЭ по Питмену некоторых наиболее употребительных из них вычислены в настоящей работе. При этом в отличие от [3] используется техника так называемых «приближенных наклонов» по Бахадуру, предложенная самим Бахадуром [4] и недавно развитая Виэндом [5].

2. **Постановка задачи и описание правил обнаружения.** Рассматривается двухвыборочная задача некогерентного обнаружения сигнала в шумах неизвестного уровня, описанная в [3] и заключающаяся в том, что на выходе двухканальной схемы наблюдения регистрируются две независимые выборки объема n: $x_1, x_2, \ldots, x_n$ и $y_1, y_2, \ldots, y_n$. В отсутствие сигнала обе выборки извлечены из распределения Рэлея с плотностью

$$f_0(x)=\begin{cases}\dfrac{x}{\sigma^2}e^{-\frac{x^2}{2\sigma^2}}, & x\geqslant 0;\\ 0, & x<0.\end{cases} \tag{1}$$

В присутствии сигнала одна выборка остается чистым шумом, а другая содержит смесь сигнала и шума и распределена по закону Райса с плотностью

$$f_\theta(x)=\begin{cases}\dfrac{x}{\sigma^2}e^{-\frac{x^2+\theta^2}{2\sigma^2}}I_0\left(\dfrac{x\theta}{\sigma^2}\right), & x\geqslant 0;\\ 0, & x<0.\end{cases} \tag{2}$$

Требуется проверить гипотезу $H_0:\theta=0$ об отсутствии сигнала против альтернативы $H_1:\theta>0$ о наличии сигнала при неизвестном мешающем параметре σ^2.

Обозначим через $F_n(x)$ и $G_n(x)$ эмпирические функции распределения, построенные соответственно по первой и второй выборкам, и пусть $H_n(x)=(1/2)(F_n(x)+G_n(x))$. Рассмотрим следующие двухвыборочные непараметрические правила обнаружения, основанные на статистиках:

$$D_{n,n} = \sqrt{\frac{n}{2}} \sup_x |F_n(x) - G_n(x)|; \tag{3}$$

$$T_{n,n} = \sqrt{\frac{n}{2}} \left[\int_0^\infty (F_n(x) - G_n(x))^2 \, dH_n(x) \right]^{1/2}; \tag{4}$$

$$U_{n,n} = \sqrt{\frac{n}{2}} \left[\int_0^\infty \left(F_n(x) - G_n(x) - \int_0^\infty (F_n(y) - G_n(y)) \, dH_n(y) \right)^2 dH_n(x) \right]^{1/2}; \tag{5}$$

$$W_{n,n} = \sqrt{2n} \left(\int_0^\infty F_n(x) \, dG_n(x) - \frac{1}{2} \right). \tag{6}$$

Первые две из этих статистик являются двухвыборочными вариантами классических статистик Колмогорова и ω^2 Крамера — Мизеса. Статистика $U_{n,n}$ — вариант так называемой статистики Ватсона [6]. Наконец, статистика $W_{n,n}$ — хорошо известная статистика Вилкоксона в форме Манна — Уитни [7]. Подчеркнем, что для дальнейшего удобно рассматривать статистики $T_{n,n}$ и $U_{n,n}$, а не их квадраты, как это обычно принято.

3. **Приближенные наклоны по Бахадуру.** В работе [4] Бахадур предложил в качестве меры асимптотической эффективности правила обнаружения, основанного на последовательности тестовых статистик $\{t_n\}$, рассматривать «приближенный наклон» этой последовательности, определяемый следующим образом.

Пусть $F(x)$ — предельная функция распределения $\{t_n\}$ при основной гипотезе H_0:

$$\lim_{n\to\infty} P(t_n < x \mid H_0) = F(x), \tag{7}$$

причем найдется постоянная a, $0<a<\infty$, такая, что

$$\ln[1-F(x)] = -(ax^2/2)[1+o(1)], \quad x\to\infty. \tag{8}$$

Предположим, кроме того, что существует такая функция $b(\theta)$, $0< <b(\theta)<\infty$, что при любом $\varepsilon>0$

$$\lim_{n\to\infty} P\left(\left| \frac{t_n}{\sqrt{n}} - b(\theta) \right| > \varepsilon \mid H_1 \right) = 0, \tag{9}$$

где θ — параметр, определяющий альтернативу H_1. Тогда приближенный наклон $c^*(\theta)$ последовательности тестовых статистик $\{t_n\}$ определяется равенством

$$c^*(\theta) = a[b(\theta)]^2. \tag{10}$$

Используя (10), вычислим приближенные наклоны статистик (3)—(6). Для вычисления коэффициента a, согласно определению (7), необходимо располагать явным видом предельных распределений соответствующих статистик при нулевой гипотезе. Из [8] следует, что предельное распределение величины $D_{n,n}$, точно такое же, как и у обычной статистики Колмогорова, т. е.

$$\lim_{n\to\infty} P(D_{n,n} < x) = K(x) = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2x^2}. \tag{11}$$

Предельные распределения статистик $T^2_{n,n}$ и $U^2_{n,n}$ также совпадают с их одновыборочными вариантами, что доказано соответственно в [9] и [6]. Поэтому справедливы соотношения:

$$\lim_{n\to\infty} P\left(T^2_{n,n} < x\right) = 1 - \frac{1}{\pi}\sum_{k=1}^{\infty}(-1)^{k-1}\int_{(2k-1)\pi}^{2k\pi}\exp\left(-\frac{\lambda^2 x^2}{2}\right)\sqrt{-\lambda\sin\lambda}\,d\lambda; \tag{12}$$

$$\lim_{n\to\infty} P\left(U^2_{n,n} < x\right) = K(\pi x) = \sum_{k=-\infty}^{\infty}(-1)^k e^{-2k^2\pi^2x^2}. \tag{13}$$

Что касается статистики $W_{n,n}$, то хорошо известно, что ее распределение асимптотически нормально (см., например, [10], гл. 31), поэтому можно записать

$$\lim_{n\to\infty} P\left(W_{n,n} < x\right) = \Phi\left(x\sqrt{3}\right), \tag{14}$$

где $\Phi(x)$ — функция распределения стандартного нормального закона.

Используя явный вид предельных распределений (11)—(14), с помощью относительно элементарных вычислений найдем соответственно коэффициенты a рассматриваемых статистик:

$$a_D = 4,\ a_T = \pi^2,\ a_U = 4\pi^2,\ a_W = 3. \tag{15}$$

Явный вид выражений для $b(\theta)$ элементарно получается из (9) на основании теоремы Гливенко (см., например, [11], гл. 6). Переходя в заключение от плотностей (1) и (2) к соответствующим им функциям распределения $F_0(x)$ и $F_\theta(x)$, полагая $H_\theta(x) = 1/2(F_0(x) + F_\theta(x))$ и учитывая (10) и (15), окончательно получаем

$$c^*_D(\theta) = \sup_x |F_\theta(x) - F_0(x)|^2; \tag{16}$$

$$c^*_T(\theta) = \frac{\pi^2}{4}\int_0^\infty (F_\theta(x) - F_0(x))^2\, dH_\theta(x); \tag{17}$$

$$c^*_U(\theta) = \pi^2\left[\int_0^\infty (F_\theta(x) - F_0(x))^2\, dH_\theta(x) - \left(\int_0^\infty (F_\theta(x) - F_0(x))\, dH_\theta(x)\right)^2\right]; \tag{18}$$

$$c^*_W(\theta) = 3\left[\int_0^\infty F_0(x)\, dF_\theta(x) - \frac{1}{2}\right]^2. \tag{19}$$

Подобные представления приближенных наклонов для случая одновыборочных статистик имеются в [5].

Ограничиваясь рассмотрением наиболее важного в приложениях случая слабых сигналов, найдем главные части асимптотики приближенных наклонов (16)—(19) при $\theta \to 0$. С этой целью разложим функции H_θ и F_θ в ряды по θ^2 в окрестности нуля и, удерживая первые два члена разложений, подставим результаты в выражения (16)—(19). В результате получим, что локально (при $\theta \to 0$) приближенные наклоны изучаемых правил обнаружения имеют вид:

$$c^*_D(\theta) \sim \frac{1}{4e^2}\frac{\theta^4}{\sigma^4}; \tag{20}$$

$$c^*_T(\theta) \sim \frac{\pi^2}{216}\frac{\theta^4}{\sigma^4}; \tag{21}$$

$$c_U^*(\theta) \sim \frac{5\pi^2}{1728}\frac{\theta^4}{\sigma^4}; \tag{22}$$

$$c_W^*(\theta) \sim \frac{3}{64}\frac{\theta^4}{\sigma^4}. \tag{23}$$

Данные результаты позволяют вычислять так называемую локальную приближенную асимптотическую эффективность по Бахадуру [4] одного правила обнаружения по отношению к другому, для чего достаточно найти предел отношения приближенных наклонов соответствующих правил при $\theta \to 0$.

4. **Асимптотическая относительная эффективность по Питмену.** В настоящее время в большинстве приложений математической статистики, и в частности в теории обнаружения сигналов, доминирующее положение занимает сравнение различных правил обнаружения на основе определения АОЭ по Питмену (см., например, [10], гл. 25 и [12], гл. 1, 2). Это обусловлено отчасти тем, что питменовская эффективность есть число, а это для многих приложений представляется особенно ценным.

Для нахождения АОЭ по Питмену с помощью приближенных наклонов следует принять во внимание следующее утверждение. Для широкого класса статистик АОЭ по Питмену совпадает с локальной приближенной эффективностью по Бахадуру, методика вычисления которой была изложена в п. 3. Условие принадлежности рассматриваемых статистик к указанному классу следует для статистики Вилкоксона из [4], а для трех остальных статистик — из [5]. Строго говоря, в [5] была рассмотрена только одновыборочная задача, однако переход к двухвыборочному случаю не требует, по существу, использования никаких новых идей. Обозначив через $E_{t,t'}$ АОЭ по Питмену правила обнаружения t по отношению к t', получим с учетом (20)—(23)

$$E_{D,W} = 16/(3e^2) = 0{,}722; \quad E_{U,W} = (5\pi^2)/81 = 0{,}609;$$
$$E_{T,W} = (8\pi^2)/81 = 0{,}974 \text{ и т. д.}$$

Сравним теперь рассмотренные выше непараметрические правила обнаружения с параметрическими инвариантными, изучавшимися в [3].

Прямому сравнению обоих классов правил обнаружения препятствует то обстоятельство, что в [3] использовалось классическое определение АОЭ [10], и поэтому меры эффективности параметрических правил обнаружения лежат в другой шкале измерений. Для установления соответствия между обеими шкалами вычислим меру эффективности статистики Вилкоксона $W_{n,n}$ тем же методом, что и в [3]. Используя результаты § 31. 59 из [10], нетрудно получить, что мера эффективности этой статистики эквивалентна при $n \to \infty$ величине $3n/8\sigma^4$. Обозначив через $\Pi_{n,n}$ правило обнаружения Прокофьева

$$\Pi_{n,n} = \sqrt{n}\left(\sum_{i=1}^{n} x_i^2 \Big/ \sum_{i=1}^{n} y_i^2\right),$$

через $\Lambda_{n,n}$ — статистику логарифмического контраста

$$\Lambda_{n,n} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \ln\frac{x_i}{y_i}$$

и через $K_{n,n}$ — статистику простого контраста с параметром $\alpha = 2$

$$K_{n,n} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \sqrt{\frac{x_i}{y_i}}$$

и использовав значения мер эффективности этих правил, найденные в [3], получим следующие значения АОЭ статистики Вилкоксона по отношению к предыдущим:

$$E_{W,\pi}=3/4;\ E_{W,\Lambda}=\pi^2/8;\ E_{W,K}=(6(4-\pi))/\pi.$$

Пользуясь полученным результатом и очевидной формулой

$$E_{t,t'}=E_{t,t''}E_{t'',t'},$$

нетрудно вычислить АОЭ по Питмену для любых пар правил обнаружения из рассмотренных нами семи правил. В качестве примера ограничимся значениями АОЭ рассмотренных правил по отношению к локально-оптимальному правилу Прокофьева:

$$E_{W,\pi}=3/4=0{,}750;\quad E_{T,\pi}=(2\pi^2)/27=0{,}731;$$
$$E_{\Lambda,\pi}=6/\pi^2=0{,}608;\quad E_{D,\pi}=4/e^2=0{,}541;$$
$$E_{U,\pi}=(5\pi^2)/108=0{,}457;\quad E_{K,\pi}=\pi/(8(4-\pi))=0{,}457.$$

Полученные результаты вводят следующее упорядочение среди рассмотренных правил обнаружения: наиболее высокой эффективностью по Питмену обладает статистика Прокофьева, а за ней следуют статистики $W_{n,n}$, $T_{n,n}$, $\Lambda_{n,n}$, $D_{n,n}$, $K_{n,n}$ и $U_{n,n}$.

ВЫВОДЫ

В условиях рассмотренной задачи локально-оптимальное правило Прокофьева заметно превосходит по своей асимптотической эффективности конкурирующие с ним правила, основанные на статистиках группы контраста и употребительных непараметрических статистиках. В то же время сравнительный анализ правил обнаружения позволяет предположить, что при отсутствии сведений о структуре оптимального правила для решения задач распознавания эффективными могут оказаться правила, основанные на ранговых и знаковых статистиках, а также смешанные знаково-ранговые алгоритмы [12]. Еще один пример приложения результатов данной работы — задача последетекторного обнаружения слабых сигналов (см., например, [13, 14]), где рассмотренная модель сигналов и помех является достаточно распространенной.

Анализируя полученные результаты, следует, однако, иметь в виду, что при отклонениях фактических распределений от рассмотренной теоретической модели, например, вследствие так называемого «загрязнения распределений» упорядочение правил обнаружения может оказаться существенно иным. Иллюстрацией к изложенному может служить известный результат Ходжеса и Лемана (см. [10], § 31.62), согласно которому АОЭ $E_{W,t}$-статистики Вилкоксона по отношению к t-статистике Стьюдента в случае проверки гипотезы о сдвиге в нормальных выборках равна $3/\pi=0{,}95$. Однако при переходе к распределениям, отличным от нормального, значения $E_{W,t}$ могут быть сколь угодно велики (и даже обращаться в бесконечность), но никогда не могут быть ниже значения 0,864. Ситуации подобного рода могут оказаться в ряде случаев дополнительным аргументом в пользу применения в задачах рассматриваемого типа не только оптимальных инвариантных, но и неоптимальных непараметрических правил обнаружения.

Другой вывод состоит в том, что техника вычисления АОЭ по Питмену, основанная на понятии приближенного наклона, часто оказывается проще и удобнее традиционной техники вычисления этой характеристики. Кроме того, для ряда статистик, основанных на эмпирических функциях распределения, это пока единственный метод вычисления питменовской эффективности.

ЛИТЕРАТУРА

1. **Карпова О. М., Нежевенко Е. С., Уманцев Г. Д.** Распознавание изображений известной формы на фотоснимках.— «Автометрия», 1975, № 3, с. 68—72.
2. **Козлов О. А., Нежевенко Е. С., Потатуркин О. И.** Распознавание изображений в когерентно-оптических системах с применением контурных эталонов.— «Автометрия», 1976, № 6, с. 36—44.
3. **Никитин Я. Ю., Филимонов Р. П., Шубина Е. П.** Расчет асимптотической относительной эффективности некоторых инвариантных правил обнаружения в схеме двухканальной обработки.— «Автометрия», 1978, № 2, с. 51—57.
4. **Bahadur R. R.** Stochastic comparison of tests.— "Ann. Math. Stat.", 1960, vol. 31, p. 276—295.
5. **Wieand H. S.** A condition under which the Pitman and Bahadur approaches to efficiency coincide.— "Ann. of Stat.", 1976, vol. 4, p. 1003—1011.
6. **Watson G. S.** Goodness-of-fit tests on a circle. P. I, II.— "Biometrika", 1961, vol. 48, p. 109—114; 1962, vol. 49, p. 57—63.
7. **Mann H. B., Whitney D. R.** On a test whether one of the two random variables is stochastically larger than the other.— "Ann. Math. Stat.", 1947, vol. 18, p. 50—60.
8. **Смирнов Н. В.** Оценка расхождения между эмпирическими кривыми распределения в двух независимых выборках.— «Бюл. МГУ, сер. А», 1939, т. 2, с. 3—14.
9. **Rosenblatt M.** Limit theorems associated with variants of the von Mises statistic.— "Ann. Math. Stat.", 1952, vol. 23, p. 617—623.
10. **Кендалл М., Стьюарт А.** Статистические выводы и связи. М., «Наука», 1973.
11. **Рао С. Р.** Линейные статистические методы и их применение. М., «Наука», 1968.
12. **Левин Б. Р.** Теоретические основы статистической радиотехники, кн. 3. М., «Сов. радио», 1976.
13. **Левин Б. Р.** Теоретические основы статистической радиотехники, кн. 2. М., «Сов. радио», 1975.
14. **Прокофьев В. Н.** Инвариантное правило некогерентного обнаружения сигнала на фоне шумов неизвестного уровня.— «Радиотехника и электроника», 1973, т. XVIII, вып. 3, с. 547—553.

Поступила в редакцию 17 июля 1978 г.;
окончательный вариант — 30 ноября 1978 г.

УДК 519.27 : 621.391.2

Л. М. КЕНИН

(Воронеж)

АЛГОРИТМ СВЕРХЭФФЕКТИВНОЙ ОЦЕНКИ МАЛЫХ ЗНАЧЕНИЙ ПАРАМЕТРА РАСПРЕДЕЛЕНИЯ

Проблема оценки параметров сигнала на фоне помех и получения эффективных алгоритмов обработки особенно остро стоит в случае малых отношений сигнал/помеха (С/П) и малых объемов выборки.

Существующие методы оценки, например метод максимального правдоподобия, в этих условиях не являются эффективными. При малых отношениях С/П и малых объемах выборки может произойти смещение оценки и дисперсия оценки может отличаться от нижней границы дисперсии оценки, определяемой неравенством Рао — Крамера, как в случае оценки параметров распределения случайной величины [1, 2], так и для параметров случайного процесса [3].

Известны сверхэффективные оценки [1, 4], имеющие дисперсию в ограниченном диапазоне значений измеряемого параметра ниже, чем это следует из неравенства Рао — Крамера; такие оценки возможны при разрывных функциях распределения случайной величины или при разрывных зависимостях оценки от данных наблюдения [4]. Однако на практике эффективные и сверхэффективные оценки важно иметь для наиболее распространенных ситуаций.