11. **Писаренко В. Ф.** Спектральная оценка максимальной энтропии и ее использование для определения частот гармоник.— В кн.: Вычислительная сейсмология. М.: Наука, 1975, № 8, с. 83—109.
12. **Кей С. М., Марпл С. Л.** Современные методы спектрального анализа. Обзор.— ТИИЭР, 1981, т. 69, № 11, с. 5—51.
13. **Ulrych T. J., Clayton R. W.** Time series modelling and maximum entropy.— Physics of the Earth and Planetary Interiors, 1976, vol. 12, N 2/3, p. 189—199.
14. **Маркел Дж. Д., Грэй А. Х.** Линейное предсказание речи.— М.: Связь, 1981, с. 305.
15. **Клаербоут Дж. Ф.** Теоретические основы обработки геофизической информации с приложением к разведке нефти.— М.: Недра, 1982, с. 304.
16. **Marple L.** A new autoregressive spectrum analysis algorithm.— IEEE Trans. on Acoustics, Speech and Signal Proces., 1980, vol. ASSP-28, N 4, p. 441—454.
17. **Зеленков А. В.** О выборе алгоритма вычисления комплексного кепстра.— Радиотехника и электроника, 1981, т. 24, № 10, с. 2095—2109.
18. **Митяшев Б. Н.** Определение временного положения импульсов при наличии помех.— М.: Сов. радио, 1962, с. 199.
19. **Зеленков А. В.** Комплекс программ СОМСЕР вычисления кепстра мощности, фазового кепстра, комплексного кепстра и обратной свертки.— Инф. бюл. ВНТИЦ «Алгоритмы и программы», 1981, № 1-2, П004617.

Поступила в редакцию 10 января 1983 г.

УДК 519.21

А. Н. ЕФИМОВ, Е. В. КРИВОРУКОВ

(Москва — Харьков)

МЕТРОЛОГИЧЕСКИЕ МОДЕЛИ И АЛГОРИТМЫ КОСВЕННОГО ОЦЕНИВАНИЯ ПРИ МАЛОМ ЧИСЛЕ НАБЛЮДЕНИЙ

В системах обработки данных зачастую функциональная связь наблюдаемого аргумента x с измеряемой величиной $y=f(x)$ априори известна. При этом можно выделить две различные модели наблюдения.

В первой из них наблюдения x_i $(i=\overline{1,\ N})$ — выборочные значения случайного по своей природе аргумента x с моментами m, σ^2, а измеряемым параметром является определенная характеристика величины y, чаще всего математическое ожидание $M(y)=M[f(x)]$. Во второй модели значения аргумента x_i представляют собой сумму неизвестной величины m, несущей полезную информацию, и помехи η, как правило, центрированной, т. е. $x_i=m+\eta_i$. При этом измеряют величину $y_0=f(m)$.

Очевидно, что измеряемые величины в моделях 1 и 2 различны, так как $M[f(x)]\neq f(m)$ при нелинейных $f(x)$. В метрологии указанные модели приводят к косвенным измерениям 1- и 2-го типов [1].

Кажется естественным, что должны различаться и оценочные функции $\varphi_1(x_1, \ldots, x_N)$ для $M(y)$ и $\varphi_2(x_1, \ldots, x_N)$ для y_0. Действительно, если нам ничего неизвестно о вероятностной структуре аргумента x, то для оценивания y_0 возможен лишь алгоритм (назовем его «а») $\bar{x}=\sum x_i/N$, $\hat{y}_0=f(\bar{x})$, а для оценки $M(y)$ — инверсная последовательность операций (обозначим ее «б») $y_i=f(x_i)$, $\bar{y}=\sum y_i/N$.

Однако, как будет показано ниже, при определенных условиях оба алгоритма могут применяться и для оценивания одной и той же величины.

Авторами в работе [2] был введен в рассмотрение новый фактор — объем выборки. В настоящей работе, существенно дополняя [2], исследуется влияние объема выборки наблюдений на правило выбора более точного алгоритма. Демонстрируется, как одна и та же модель наблюдения может (при уменьшении объема выборки) потребовать изменения алгоритма оценивания, причем более точным окажется тот, который приводит к несостоятельным псевдооценкам.

и $\Delta_{II}(N)$ разность среднеквадратических алгоритмов. Рассмотрим две модели наблюдений соответственно:

$$\Delta_I(N) = M\left[\hat{y} - M(y)\right]^2 - M\left[\bar{y} - M(y)\right]^2 = [M(y) - M(\hat{y})]^2 + \sigma^2_{\hat{y}} - \sigma^2_y/N, \tag{1}$$

где $M(y)$, σ^2_y, $M(\hat{y})$, $\sigma^2_{\hat{y}}$ — математическое ожидание и дисперсия величин $y = f(x)$ и $\hat{y} = f(\bar{x})$;

$$\Delta_{II}(N) = M[\hat{y} - f(m)]^2 - M[\bar{y} - f(m)]^2 = \Delta_I(N) - 2[M(y) - f(m)] \times \times [M(y) - M(\hat{y})]. \tag{2}$$

Знак $\Delta_i(N)$, $i = 1, 2$, свидетельствует о преимуществе того или иного метода оценивания: при $\Delta_i(N) < 0$ точнее $\hat{y}$, и наоборот.

Для 1-й модели оценка «а» является смещенной несостоятельной псевдооценкой, а оценка $\bar{y}$ (выборочное среднее) состоятельная и не имеет смещения. Поэтому при больших N предпочтительнее оценка $\bar{y}$, и выигрыш при ее применении составляет $\Delta_I(N) \xrightarrow[N\to\infty]{} [M(y) - f(m)]^2$.

Для 2-й модели наблюдения, наоборот, оценка «а» асимптотически несмещенная и состоятельная, а $\bar{y}$ — смещенная псевдооценка, и при больших N, естественно, $\hat{y}$ предпочтительнее. Проигрыш при использовании $\bar{y}$ составляет $\Delta_{II}(N) \xrightarrow[N\to\infty]{} -[M(y) - f(m)]^2$, т. е. при $N \to \infty$ $\Delta_I(N) = -\Delta_{II}(N)$. Именно так и рекомендует применять алгоритмы «а» и «б» теория измерений.

Однако, когда число N невелико, асимптотические свойства оценок $\bar{y}$ при 1-й модели наблюдения и $\hat{y}$ при 2-й модели уже не имеют значения и может оказаться целесообразным перейти к псевдооценкам (т. е. от $\bar{y}$ к $\hat{y}$, и наоборот), если их среднеквадратическая ошибка будет меньше, чем у конкурирующей оценки.

Факт, состоящий в большей точности псевдооценки, назовем «эффектом малой выборки» и укажем условия, при которых он может иметь место.

Первая модель наблюдения. В работе [2] было получено выражение $\Delta_I(N)$ для 1-й модели наблюдения:

$$\Delta_I(N) = \left(\tilde{\mu}_3 - \mu_3/N\right)[f'(m) f''(m)]^2 + \left(\tilde{\mu}_4 - \mu_4/N\right) f'(m) f'''(m)/3 + \left(\tilde{\mu}_4 - \mu_4/N + (N-1)\mu_2^2/N\right)[f''(m)]^2/4 + \sum_{i=5}^{\infty}\Big\{2\left(\tilde{\mu}_i - \mu_i/N\right) f^{(i-1)}(m) \times \times f'(m)/(i-1)! + \sum_{k=2}^{i-2} f^{(i-k)}(m) f^{(k)}(m)\left[\tilde{\mu}_i - \mu_i/N + (N-1)\mu_{i-k}\mu_k/N - - \tilde{\mu}_{i-k}\mu_k - \mu_{i-k}\tilde{\mu}_k\right]/(i-k)!k!\Big\}. \tag{3}$$

Здесь μ_i, $\tilde{\mu}_i$ — центральные моменты величин x и $\bar{x}$, $f^{(i)}(m)$ — i-е производные $y = f(x)$ в точке m. Анализ выражения (3) показал, что при небольших объемах выборок смещенная псевдооценка $\hat{y}$ оказывается предпочтительней выборочного среднего $\bar{y}$, если $y = f(x)$ принадлежит классу абсолютно и вполне монотонных функций [3], а класс распределений $F(x)$ достаточно широк.

Теперь попытаемся найти сочетания $F(x)$ и $y=f(x)$, для которых выборочное среднее $\bar{y}$ всегда предпочтительней при оценивании $M(y)$. Для этого используем тот же метод анализа знаков слагаемых выражения $\Delta_{\mathrm{I}}(N)$, что был применен в [2]. Потребовав симметричности $F(x)$, приходим к следующему набору условий, при котором все слагаемые (3) неотрицательны: $\Delta_{\mathrm{I}}(N)\geqslant 0$, т. е. $\bar{y}$ точнее $\widehat{y}$:

$$\operatorname{sgn}[f'(m)]\neq\operatorname{sgn}[f'''(m)],\ \operatorname{sgn}[f'(m)]\neq\operatorname{sgn}[f^{(i-1)}(m)],\ \operatorname{sgn}[f^{(i-k)}(m)]\neq \neq\operatorname{sgn}[f^{(k)}(m)],\quad i=2h. \tag{4}$$

Условия (4) не могут выполняться одновременно, однако выявляют характерную тенденцию в поведении тех функций $y=f(x)$, для которых $\Delta_{\mathrm{I}}(N)\geqslant 0$, а именно: последовательность производных одинаковой четности образует знакопеременный ряд. В класс подобных функций входят, в частности, некоторые тригонометрические функции.

Пример 1. Пусть $y=\cos x$ — функция, попадающая в описанный выше класс, и x — нормальная случайная величина с неизвестными m и σ^2. Выясним поведение $\Delta_{\mathrm{I}}(N)$:

$$\Delta_{\mathrm{I}}(N)=\left[N\left(1-e^{-2\sigma^2/N}\right)+e^{-2\sigma^2}-1\right]\Big/2N+\cos^2 m\left[N\left(e^{-\sigma^2}-2e^{-(N+1)\sigma^2/2N}\right)+ + e^{-\sigma^2}-e^{-2\sigma^2}\right]\Big/N. \tag{5}$$

(Моменты для (5) определены согласно [4, формула (3), с. 923].)

Нетрудно показать, что $\Delta_{\mathrm{I}}(N)>0$ при $N>1$ независимо от значений m и σ^2, т. е. для $y=\cos x$ выборочное среднее $\bar{y}=\sum\cos x_i/N$ всегда точнее псевдооценки $\widehat{y}=\cos\left(\sum x_i/N\right)$ при оценивании $M(\cos x)$.

Вторая модель наблюдения. Перейдем теперь ко 2-й модели наблюдения. Напомним, что здесь по наблюдениям $x_i=m+\eta_i$ необходимо оценить величину $y_0=f(m)$.

Для анализа $\Delta_{\mathrm{II}}(N)$ применим ту же методику, что и в [2] при исследовании поведения $\Delta_{\mathrm{I}}(N)$. Предъявляя к $F(x)$ и $y=f(x)$ вновь те же требования и подставляя в $\Delta_{\mathrm{II}}(N)$ выражения соответствующих моментов, получаем

$$\begin{aligned}\Delta_{\mathrm{II}}(N)&=\left(\tilde{\mu}_3-\mu_3/N\right)[f'(m)f''(m)]^2+\left(\tilde{\mu}_4-\mu_4/N\right)f'(m)f'''(m)/3+\\&+\left(\tilde{\mu}_4-\mu_4/N-(N-1)\mu_2^2/N\right)[f''(m)]^2/4+\sum_{i=5}^{\infty}\Big\{2\left(\tilde{\mu}_i-\mu_i/N\right)\times\\&\times f^{(i-1)}(m)f'(m)/(i-1)!+\sum_{k=2}^{i-2}f^{(i-k)}(m)f^{(k)}(m)\left[\tilde{\mu}_i-\mu_i/N-\right.\\&\left.-(N-1)\mu_{i-k}\mu_k/N\right]/(i-k)!(k)!\Big\}.\end{aligned} \tag{6}$$

В данном соотношении третье слагаемое всегда отрицательно. В силу этого есть смысл искать условия, при которых $\Delta_{\mathrm{II}}(N)\leqslant 0$, т. е. когда оценка $\widehat{y}$ точнее псевдооценки $\bar{y}$.

Класс симметричных распределений. В этом случае $\mu_i=0$, $i= =2h+1$, так что выражение (6) принимает вид

$$\begin{aligned}\Delta_{\mathrm{II}}(N)&=\left(\tilde{\mu}_4-\mu_4/N\right)f'(m)f'''(m)/3+\left(\tilde{\mu}_4-\mu_4/N-(N-1)\mu_2^2/N\right)[f''(m)]^2/4+\\&+\sum_{\substack{i=2h\\ n=3,4,\dots}}^{\infty}\Big\{2\left(\tilde{\mu}_i-\mu_i/N\right)f^{(i-1)}(m)f'(m)/(i-1)!+\sum_{k=2}^{i-2}f^{(i-k)}(m)f^{(k)}(m)\times\\&\times\left[\tilde{\mu}_i-\mu_i/N-(N-1)\mu_{i-k}\mu_k/N\right]/(i-k)!\,k!\Big\}.\end{aligned} \tag{7}$$

Достаточным условием для того, чтобы все слагаемые (7) были отрицательными и $\Delta_{\mathrm{II}}(N)<0$ при $N>1$, является равенство знаков производных $y = f(x)$ одинаковой четности. Подобные функции уже встречались в [2]; они составляют подкласс абсолютно и вполне монотонных функций, для которых был обнаружен эффект малых выборок при исследовании $\Delta_{I}(N)$ в [2].

Класс несимметричных распределений $F(x)$ с $\mu_i > 0$, $i = 2h + 1$. Для $\Delta_{II}(N) < 0$ при любых N достаточно, чтобы все производные $y = f(x)$ имели одинаковые знаки. Класс подобных функций также фигурировал в [2] — это абсолютно монотонные функции.

Наблюдается любопытная инверсия: при одних и тех же сочетаниях $F(x)$ и $y = f(x)$ для 1-й модели наблюдения при малых N оказывается предпочтительней псевдооценка, для второй же модели наблюдения всегда точнее состоятельная оценка.

Пример 2. Пусть по зашумленным наблюдениям $x_i = m + \eta_i$ (η — центрированная нормально распределенная помеха с параметром σ^2) необходимо оценить значение $y_0 = e^m$. Функция $y = e^x$ принадлежит классу вполне монотонных функций, поэтому оценка $\widehat{y} = e^{\bar{x}}$ должна быть всегда предпочтительней псевдооценки $\bar{y} = \sum e^{x_i}/N$. Проверим это, исследовав поведение

$$\Delta_{II}(N) = e^{2m}\left[e^{2\sigma^2/N} - 2e^{\sigma^2/2N} + 1 - e^{2\sigma^2/N} + e^{\sigma^2}/N - \left(1 - e^{\sigma^2/2}\right)^2\right] =$$
$$= -e^{2m}z\left(z^{4N-1} + (N-1)z^{2N-1} - 2Nz^{N-1} - Nz^3 + 2N\right)/N.$$

Нетрудно убедиться, что выражение в скобках всегда положительно, так как оно имеет один кратный корень $z = 1$ и положительно при $z \to \infty$. Следовательно, $\Delta_{II}(N)$ отрицательно при $N > 1$, т. е. $\widehat{y}$ всегда точнее $\bar{y}$.

Эффект малых выборок при определенных сочетаниях $F(x)$ *и* $y = f(x)$. Потребуем симметричности $F(x)$. При этом $\Delta_{II}(N)$ имеет вид (7) и можно заметить, что с ростом N положительные слагаемые убывают, стремясь к нулю. Далее видно, что число положительных слагаемых в (7) особенно велико в случае, когда производные $y = f(x)$ одинаковой четности образуют знакопеременный ряд. Следовательно, обнаружить эффект малых выборок при 2-й модели наблюдения наиболее вероятно в случае, когда производные преобразования $f(x)$ одинаковой четности имеют чередующиеся знаки и N мало, т. е. именно тогда, когда при 1-й модели наблюдения предпочтительней оказывается состоятельная оценка. Преобразования указанного типа фигурировали выше, к ним относятся $\cos x$, $\sin x$. Так, при $y = \cos x$, $x_i = m + \eta_i$ (модель примера 2), исследуя

$$\Delta_{II}(N) = \cos^2 m\left[e^{-2\sigma^2/N} - e^{-2\sigma^2}/N - (N-1)e^{-\sigma^2}/N + 2\left(e^{-\sigma^2} - e^{-\sigma^2/2N}\right)\right] +$$
$$+ \left(N - 1 - Ne^{-2\sigma^2/N} + e^{-2\sigma^2}\right)/2N,$$

можно указать интервалы $z \in [0; 0{,}15]$ и $z \in [0{,}85; 1]$, $z = e^{-\sigma^2/4}$, $N = 2$, при которых $\Delta_{II}(2) > 0$, т. е. предпочтительнее псевдооценка.

ЛИТЕРАТУРА

1. **Кудряшова Ж. Ф., Рабинович С. Г.** Методы обработки результатов наблюдений при косвенных измерениях.— В кн.: Тр. метрологических ин-тов СССР. Л.: Энергия, 1975, вып. 172(232), с. 3—58.
2. **Ефимов А. Н., Криворуков Е. В.** Исследование эффективности оценок и «псевдооценок» среднего при малом числе наблюдений.— Автометрия, 1983, № 1, с. 11—18.
3. **Феллер В.** Введение в теорию вероятностей и ее приложения.— М.: Мир, 1967, т. 2.
4. **Градштейн И. С., Рыжик И. М.** Таблицы интегралов, сумм, рядов и произведений.— М.: Физматгиз, 1971.

Поступила в редакцию 25 января 1983 г.;
окончательный вариант — 9 февраля 1984 г.