УДК 62-50

В.Н. КРАСНИКОВ, А.Б. ЛЕЩЕНКО

Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Украина

РЕШЕНИЕ ЗАДАЧИ ОЦЕНКИ ПАРАМЕТРОВ ПО ЭКСПЕРИМЕНТАЛЬНЫМ ДАННЫМ

Рассмотрена задача оценивания интересующих исследователя параметров непосредственно по выборочным данным для случая произвольной симметричной функции плотности распределения вероятностей ошибок измерений.

оценивание параметров, опытные данные, закон распределения случайных величин, плотность распределения, метод максимального правдоподобия

Числовые информационные потоки любой природы, а тем более потоки измерительной информации, сопровождающие многие виды человеческой деятельности, требуют оптимальной по точности стратегии не только проведения, но и обработки измерений. Особенно важно это в случаях, когда по большому числу экспериментальных данных определяется ограниченное количество параметров, характеризующих некоторое реальное явление. Примерами могут служить как определения элементов орбиты космического аппарата по данным большого числа наблюдений, так и сложнейшие явления в экономике и социологии. При решении задач рассматриваемого типа исследуемые явления заменяется его математической моделью, зависящей от m параметров $\overline{q_1, q_m}$ состояния этой модели. Целью задачи является определение значений параметров состояния используемой модели по результатам измерения некоторых других величин $\overline{d_1, d_n}$, связанных с параметрами состояния зависимостями вида

$$\widetilde{d_i} = d_i\left(q_1, q_2, ..., q_m\right) + \xi_i,\ \ i = \overline{1, n}\,.$$

Эти равенства неточны, так как они содержат суммарные погрешности измерения и модели – ξ_i. Отметим, что наиболее широко используются вероятностные модели погрешностей, при которых задаются некоторые вероятностные характеристики суммарных погрешностей ξ_i. Так, например, в работе [1] была поставлена и решена задача представления законов распределения случайных величин в виде рядов по системам ортонормированных многочленов, коэффициенты которых определяются через числовые характеристики самих эмпирических данных, что позволяет их использовать в виде подходящих моделей распределения.

В работе [2] была поставлена и успешно решена задача отыскания алгоритма оценивания интересующих исследователя параметров непосредственно по выборочным данным для случая симметричной функции плотности распределения вероятностей погрешностей измерений.

Как развитие метода решения задач оценки неизвестных параметров, основанного на принципе соответствия функции распределения полученным эмпирическим данным [2], в данной работе представлен еще один интересный для практического использования подход.

Пусть имеется выборка регистрируемых данных, по которой строится эмпирическая функция распределения

$$y_n = F_э\left(x_n\right),\ \ n = \overline{1, N}\,.$$

Эмпирическая функция распределения $F_э\left(x_n\right)$ (или диаграмма накопленных частот) практически строится так. На оси абсцисс указывают значения наблюдений x_n. Значения по оси ординат равно нулю левее точки $x_{\min}$ а в точке $x_{\min}$ и далее во всех других точках x_n диаграмма имеет скачок (ступеньку), равный $1/N$. Если существует несколько совпадающих значений x_n, то в этом месте на диаграмме происходит скачок, равный λ/N, где λ – число совпадающих точек. Понятно, что для величин $x > x_{\max}$, значение диаграммы накопленных частот равно единице. Отметим, что при $N \to \infty$ $F_э\left(x_n\right) \to F\left(x\right)$.

Эмпирическому закону распределения поставим в соответствие теоретический закон

$$F\left(x\right) = \int\limits_{-\infty}^{x} f\left(t\right)dt,$$

где
$$f\left(t\right) = \frac{1}{\sqrt{2\pi S_2}} e^{-\frac{(t-S_1)^2}{2S_2}} \left[1 + A_1\left(t - S_1\right)^2 + \right.$$
$$\left. + \ldots + A_m\left(t - S_1\right)^{2m}\right], \qquad (1)$$

а S_1 и S_2 – эмпирические средние и дисперсия [2]; так, чтобы сумма квадратов была минимальная

$$\sum_{n=1}^{N} [F(x_n) - y_n]^2 \Rightarrow \min .$$

Здесь за y_n принимаем середину ступеньки интегрального закона $F_Э(x_n)$, построенного из вариационного ряда значений x_n.

Таким образом, применяя для эмпирического закона распределения метод наименьших квадратов, найдем неизвестные коэффициенты A_i для функции плотности распределения, и, применяя метод максимального правдоподобия, получим алгоритм оценивания искомых неизвестных параметров.

Рассмотрим функции

$$\Phi_O(x) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi S_2}} e^{-\frac{(\tau-S_1)^2}{2S_2}} d\tau,$$

$$\Phi_k(x) = \int\limits_{-\infty}^{x} \frac{1}{\sqrt{2\pi S_2}} e^{-\frac{(\tau-S_1)^2}{2S_2}} (\tau - S_1)^{2k} d\tau,$$

через которые удобно выразить функцию плотности распределения $f(t)$ искомого теоретического закона $F(x)$.

Интегрирование по частям $\Phi_k(x)$ приводит к уравнению

$$\Phi_k(x) = \frac{1}{\sqrt{2\pi S_2}} \int\limits_{-\infty}^{x} e^{-\frac{(\tau-S_1)^2}{2S_2}} (\tau - S_1)^{2k} d\tau =$$

$$= \begin{vmatrix} \tau - S_1 = t \\ d\tau = dt \end{vmatrix} = \frac{1}{\sqrt{2\pi S_2}} \int\limits_{-\infty}^{x-S_1} t^{2k} e^{-\frac{t^2}{2S_2}} dt =$$

$$= \begin{vmatrix} \dfrac{1}{2S_2} = p \\ 2p = \dfrac{1}{S_2} \end{vmatrix} = \frac{1}{\sqrt{2\pi S_2}} \int\limits_{-\infty}^{x-S_1} t^{2k} e^{-pt^2} dt =$$

$$= \begin{vmatrix} u = t^{2k-1}, du = (2k-1)t^{2k-2}dt \\ te^{-pt^2} dt = dv, V = -\dfrac{1}{2p} e^{-pt^2} \end{vmatrix} =$$

$$= \frac{1}{\sqrt{2\pi S_2}} \left[-\frac{t^{2k-1}}{2p} e^{-pt^2} \right] \Bigg|_{-\infty}^{x-S_1} +$$

$$+ \frac{2k-1}{2p} \int\limits_{-\infty}^{x-S_1} t^{2k-2} e^{-pt^2} dt$$

или в итоге

$$\Phi_k(x) = -\frac{(x - S_1)^{2k-1} e^{-p(x-S_1)^2}}{2p\sqrt{2\pi S_2}} +$$
$$+ \frac{2k+1}{2p} \Phi_{k-1}(x).$$

Таким образом, получена рекуррентная формула, позволяющая записать $\Phi_k(x)$ через $\Phi_O(x)$:

$$\Phi_k(x) = -\frac{(x - S_1)^{2k-1} e^{-p(x-S_1)^2}}{2p\sqrt{2\pi S_2}} + \frac{2k-1}{2p} \times$$
$$\times \left[-\frac{(x - S_1)^{2k-3} e^{-p(x-S_1)^2}}{2p\sqrt{2\pi S_2}} + \right.$$
$$\left. + \frac{2k-3}{2p} \Phi_{k-2}(x) \right] =$$

$$= \frac{e^{-p(x-S_1)^2}}{2p\sqrt{2\pi S_2}}\left[(x-S_1)^{2k-1} + \frac{2k-1}{2p}(x-S_1)^{2k-3} + \right.$$
$$+ \frac{(2k-1)(2k-3)}{(2p)^2}(x-S_1)^{2k-5} + \ldots$$
$$\left. + \frac{(2k-1)(2k-3)\ldots 3\cdot 1}{(2p)^{k-1}}(x-S_1)\right] +$$
$$+ \frac{(2k-1)!!}{(2p)^k}\Phi_O(x) =$$
$$= (2k-1)!!\left[S_2^k \Phi_O(x) - \sqrt{\frac{S_2}{2\pi}} - e^{-\frac{(x-S_1)^2}{2S_2}} \times \right.$$
$$\left. \times \sum_{r=0}^{k-1} \frac{S_2^r (x-S_1)^{2k-2r-1}}{(2k-2r-1)!!}\right].$$

Поскольку теоретический закон распределения задаем в виде

$$F(x_n) = \Phi_O(x) + \sum_{k=1}^{m} A_k \Phi_k(x),$$

то применяя для аппроксимации эмпирического закона $F_Э(x_n)$ метод наименьших квадратов минимизируем функцию

$$\Psi(A_k) = \sum_{n=1}^{N}\left[F(x_n) - y_n\right]^2 =$$
$$= \sum_{n=1}^{N}\left[\sum_{k=1}^{m} A_k \Phi_k(x_n) + \Phi_O(x_n) - y_n\right]^2 \Rightarrow \min.$$

Из условия $\partial\Psi / \partial A_\gamma = 0$, получаем систему линейных уравнений для определения неизвестных коэффициентов A_k:

$$\sum_{k=1}^{m} A_k \left[\sum_{n=1}^{N} \Phi_k(x_n)\Phi_\gamma(x_n)\right] =$$
$$= \sum_{n=1}^{N}\left[y_n - \Phi_O(x_n)\right]\Phi_\gamma(x_n), \gamma = \overline{1, m}. \quad (2)$$

Эта система линейных уравнений (2) имеет решение, поскольку определитель матрицы коэффициентов при неизвестных A_k является определителем Грама.

Найденные в результате решения этой системы уравнений коэффициенты A_k можно теперь подставить в выражение для плотности вероятностей (1) и, используя метод максимального правдоподобия, как это было показано нами в [2], придем к системе уравнений относительно оцениваемых параметров.

Заключение

Достоинство описанных подходов к вопросу оценивания параметров состоит в том, что искомые параметры отыскиваются из системы уравнений, полученных с учетом эмпирических значений среднего и дисперсии и оговоренного [2] ограничения на симметрию плотности распределения.

Литература

1. Красников В.Н., Лещенко А.Б. Моделирование законов распределения случайных величин по эмпирическим данным // Радіоелектронні і комп'ютерні системи. – 2005. – № 4 (12). – С. 108-111.

2. Красников В.Н., Лещенко А.Б. Решение задачи нахождения алгоритма оценивания параметров по эмпирическим данным для произвольной симметрической функции распределения случайных величин // Радіоелектронні і комп'ютерні системи. – 2006. – № 2 (14). – С. 101-103.

Поступила в редакцию 16.11.2006

Рецензент: д-р физ.-мат. наук, проф. А.Г. Николаев, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков.