V. D. Konakov, A non-parametric estimate of a probability density function, *Teor. Veroyatnost. i Primenen.*, 1972, Volume 17, Issue 2, 377–379

ЛИТЕРАТУРА

[1] Н. В. Смирнов, О критерии симметрии закона распределения случайной величины, ДАН СССР, 56, 1 (1947), 13—16.

[2] J. Hemelrijk, A family of parameterfree tests for symmetry, Proc. Kon. Ned. Akad. (section of sciences), 53 (1950), 945—955; 1186—1198.

[3] Ю. Н. Тюрин, Об оценивании функции распределения, Теория вероят. и ее примен., XV, 3 (1970), 567—568.

[4] И. И. Гихман, А. В. Скороход, Введение в теорию случайных процессов, М., Изд-во «Наука», 1965.

[5] M. Rosenblatt, Limit theorems associated with variants of the von Mises statistic, Ann. Math. Statist., 23, 4 (1952), 617—623.

[6] Д. М. Чибисов, Некоторые теоремы о предельном поведении эмпирической функции распределения, Труды Матем. ин-та им. В. А. Стеклова АН СССР, L XXI, (1964), 104—112.

[7] А. И. Орлов, Оценки скорости сходимости к пределу распределений некоторых статистик, Теория вероят. и ее примен., XVI, 3 (1971), 583—584.

[8] Butler Calvin, A test for symmetry using the sample distribution function, Ann. Math. Statist., 40, 6 (1969), 2209—2210.

[9] J. L. Doob, Heuristic approach to the Kolmogorov — Smirnov theorems, Ann. Math. Statist., 20 (1949), 393—403.

[10] Koul Hira Hal, Asymptotic normality of random rank statistics Ann. Math. Statist., 41, 6 (1970), 2144—2149.

ON TESTING THE SYMMETRY OF DISTRIBUTION

A. I. ORLOV (MOSCOW)

(*Summary*)

Let $F(x)$ be an unknown distribution function. The hypothesis to be tested is the symmetry of $F(x)$ at $x = 0$: $F(-x) = 1 - F(x)$. A test of the ω^2-test type is constructed, and its asymptotic properties are investigated.

НЕПАРАМЕТРИЧЕСКАЯ ОЦЕНКА ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ

В. Д. КОНАКОВ

Ряд методов статистической классификации основан на вычислении отношения правдоподобия. Например, в простейшем случае классификации по двум совокупностям с наличием обучающих последовательностей строят по этим последовательностям оценки соответствующих плотностей распределения, а затем очередное наблюдение Z относят к первой или второй совокупности, в зависимости от того, больше или меньше 1 отношение упомянутых оценок в точке Z. Отсюда понятна необходимость иметь оценки неизвестных плотностей распределения, обладающие достаточно хорошими свойствами, такими как несмещенность, состоятельность и т. п. Построению подобных оценок посвящен ряд работ [1] — [4]. В работах [2] — [4] рассмотрен случай одномерных наблюдений. Обобщение результатов [2] и [3] на многомерный случай дано в работе [1].

В статье [1] найден класс оценок плотности распределения, которые являются асимптотически несмещенными, состоятельными и асимптотически нормальными. Ниже будет приведен класс оценок, обладающих аналогичными свойствами, не входящий в класс оценок [1].

Рассмотрим N взаимно независимых p-мерных случайных векторов $X_i = (X_{i1}, \ldots, X_{ip})$, $i = 1, 2, \ldots, N$, имеющих одинаковую функцию распределения $F(\mathbf{x})$, $\mathbf{x} = (x_1, \ldots, x_p) \in R_p$. Предполагается, что F абсолютно непрерывна и что соответствующая характеристическая функция принадлежит L_1.

Пусть $\mathbf{B}_N = (B_{N1}, \ldots, B_{Np})$ — последовательность векторов, удовлетворяющих условиям:

$$\lim_{N\to\infty} \min(B_{N1}, \ldots, B_{Np}) = +\infty, \qquad \lim_{N\to\infty} \frac{1}{N} \prod_{k=1}^{p} B_{Nk} = 0. \tag{1}$$

Теорема. *Пусть* $\mathbf{x}_0$ *— точка непрерывности плотности распределения* $f(\mathbf{x})$ *векторов* $\mathbf{X}_1, \mathbf{X}_2, \ldots$. *Тогда, если характеристическая функция этого распределения абсолютно интегрируема, то функция*

$$f_N(\mathbf{x}_0) = \frac{1}{\pi^p N} \sum_{l=1}^{N} \prod_{j=1}^{p} \frac{\sin B_{Nj}(X_{lj} - x_{0j})}{X_{lj} - x_{0j}} \tag{2}$$

является асимптотически несмещенной, состоятельной и асимптотически нормальной оценкой для $f(\mathbf{x}_0)$.

Доказательство. Пусть $\varphi(\mathbf{t})$ и $\varphi_N(\mathbf{t})$ — характеристические функции истинного $F(\mathbf{x})$ и эмпирического $F_N(\mathbf{x})$ распределений соответственно, $\mathbf{t} = (t_1, \ldots, t_p)$. Очевидно, $\mathrm{E}\,\varphi_N = \varphi$.

Так как $\varphi \in L_1$, то имеет место формула обращения ([5], стр. 271):

$$f(\mathbf{x}_0) = (2\pi)^{-p} \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \exp\{-i(\mathbf{t}, \mathbf{x}_0)\}\, \varphi(\mathbf{t})\, d\mathbf{v}, \qquad d\mathbf{v} = dt_1 \ldots dt_p.$$

Для того чтобы получить отсюда оценку для $f(\mathbf{x})$, заменим φ на φ_N, а чтобы интеграл был конечен, рассмотрим его в конечных пределах $(-B_{NK}, B_{NK})$, $k = 1, 2, \ldots, p$, где B_{NK} — числа, удовлетворяющие условию (1):

$$f_N(\mathbf{x}_0) = (2\pi)^{-p} \int_{-B_{N1}}^{B_{N1}} \ldots \int_{-B_{Np}}^{B_{Np}} \exp\{-i(\mathbf{t}, \mathbf{x}_0)\}\, \varphi_N(\mathbf{t})\, d\mathbf{v}. \tag{3}$$

Подставив вместо φ_N ее выражение в виде суммы и учтя, что интеграл от синуса в симметричных пределах равен нулю, после несложных преобразований можно убедиться, что искомая оценка выражается формулой (2).

Асимптотическая несмещенность такой оценки является следствием формулы обращения, представления (3) и условий (1). Более того, можно показать, что для любой $\varphi(\mathbf{t}) \in L_1$ последовательность $\mathbf{B}_N$ всегда можно выбрать таким образом, что имеет место факт сильной асимптотической несмещенности нашей оценки, т. е.

$$(\mathrm{E} f_N(\mathbf{x}_0) - f(\mathbf{x}_0))\, (\mathrm{D} f_N(\mathbf{x}_0))^{-1/2} \to 0.$$

Конструктивный выбор $\mathbf{B}_N$ зависит от свойств гладкости функции плотности $f(\mathbf{x})$.

Для доказательства состоятельности достаточно убедиться, что $\mathrm{D}\, f_N(\mathbf{x}_0) \to 0$. В силу независимости и одинаковой распределенности векторов $\mathbf{X}_1, \ldots, \mathbf{X}_N$ достаточно показать, что математическое ожидание квадрата первого слагаемого суммы (2) есть $o(N)$. Итак, пусть

$$V_N(\mathbf{x}_0) = \prod_{j=1}^{p} \frac{\sin B_{Nj}(X_{1j} - x_{0j})}{X_{1j} - x_{0j}}.$$

Полагая $y_i = B_{Ni}(x_i - x_{0i})$, $i = 1, \ldots, p$, получим:

$$\mathrm{E} V_N^2(\mathbf{x}_0) = \left(\prod_{i=1}^{p} B_{Ni}\right) \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \left(\prod_{j=1}^{p} \frac{\sin^2 y_j}{y_j^2}\right) f(\mathbf{x})\, d\mathbf{v}, \quad d\mathbf{v} = dy_1 \ldots dy_p.$$

Произведение в скобках под знаком интеграла допускает мажоранту, удовлетворяющую всем требованиям p-мерного окна ([1], стр. 46). В качестве такой мажоранты можно взять произведение $\psi(y_1)\ldots\psi(y_p)$, где $\psi(y)=1$ при $|y|\leqslant\pi/2$ и $\psi(y)=y^{-2}$ при $|y|>\pi/2$. В силу леммы 2 работы Мёрфи ([1], стр. 46)

$$\lim_{N\to\infty}\left(\prod_{i=1}^{p}B_{Ni}\right)^{-1}\mathrm{E}V_N^2(\mathbf{x}_0)=\pi^p f(\mathbf{x}_0).$$

Согласно условиям (1) величина $\prod_{i=1}^{p}B_{Ni}$ есть $o(N)$, поэтому $\mathrm{E}V_N^2(\mathbf{x}_0)=o(N)$ при $N\to\infty$.

Для доказательства асимптотической нормальности оценки $f_N(\mathbf{x}_0)$ достаточно проверить условие Ляпунова для некоторого $\sigma>0$:

$$N^{-\sigma}(\mathbf{D}V_N)^{-1-\sigma}\mathrm{E}\,|V_N-\mathrm{E}V_N|^{2(1+\sigma)}\to 0\ (N\to\infty). \tag{4}$$

Условие (4) проверим, например, для $\sigma=1$. Аналогично приведенному выше доказательству факта: $\mathrm{E}V_N^2=o(N)$, нетрудно убедиться в справедливости соотношений:

$$\mathrm{E}V_N^k(\mathbf{x}_0)\sim\left(\prod_{i=1}^{p}B_{Ni}^{k-1}\right)f(\mathbf{x}_0)\int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty}\left(\prod_{j=1}^{p}\frac{\sin y_j}{y_j}\right)^k d\mathbf{y}\,(k=2,3,4),$$

откуда немедленно следует условие (4). Теорема доказана.

Автор благодарен Л. Н. Большеву и С. А. Айвазяну за внимание к работе и ценные замечания при подготовке статьи к печати.

Поступила в редакцию 29.6.71

ЛИТЕРАТУРА

[1] V. K. Murthy, Nonparametric estimation of multivariate densities with applications, Multivariate Analysis, vol. 1, Academic Press, New York — London, 1966, 43—55.

[2] V. K. Murthy, Estimation of probability density, Ann. Math. Statist. **36**, 3 (1965), 1027—1031.

[3] V. K. Murthy, Estimation of jumps, reliability and hazard rate, Ann. Math. Statist., **36**, 3 (1965), 1032—1040.

[4] E. Parzen, On estimation of a probability density function and mode, Ann. Math. Statist., **33**, 3 (1962), 1065—1076.

[5] С. Бохнер, Лекции об интегралах Фурье, М., Физматгиз, 1962.

A NON-PARAMETRIC ESTIMATE OF A PROBABILITY DENSITY FUNCTION

V. D. KONAKOV (MOSCOW)

(*Summary*)

Non-parametric estimate (2) of multivariate density is proposed. The estimate is obtained by the inverse Fourier transformation of the empirical distribution characteristic function. Our estimate is similar to that of Murthy [1], but in our case the p-dimentional window is not absolutely integrable. We prove that our estimate is consistent, asymptotically unbiased and asymptotically normal.