Доклады Академии наук СССР
1970. Том 191, № 4

УДК 519.21

МАТЕМАТИКА

Г. П. КЛИМОВ

О ФИДУЦИАЛЬНОМ ПОДХОДЕ В СТАТИСТИКЕ

(Представлено академиком А. Н. Тихоновым 4 IX 1969)

Пусть $\mathscr{P} = \{P_\theta,\ \theta \in \Omega\}$ — семейство распределений, заданных на одном и том же выборочном пространстве (X, B) и абсолютно непрерывных по отношению к фиксированной мере μ. Плотность P_θ относительно μ обозначим $p(x|\theta)$. Рассмотрим класс $\mathscr{P}_\nu$ (априорных) распределений P, заданных на параметрическом пространстве (Ω, F) и абсолютно непрерывных относительно меры ν. Предположим, $0 < p(x) = \int p(x|\theta)\,d\nu(\theta) \leqslant \infty$ почти для всех (по μ) x.

Определение. Пусть $\{\Omega_n\}_{n\geqslant 1}$ — произвольная монотонно возрастающая последовательность подмножеств из Ω такая, что $\bigcup \Omega_n = \Omega$, $\nu(\Omega_n) < \infty$. Обозначим через $\{P_n\}_{n\geqslant 1}$ последовательность распределений из $\mathscr{P}_\nu$ таких, что распределение P_n сосредоточено на Ω_n и имеет максимальную энтропию $-\int p_n(\theta)\ln p_n(\theta)\,d\nu(\theta)$ среди таких же распределений. Здесь $p_n(\theta)$ — плотность распределения P_n по мере ν. Пусть $p_n(\theta|x)$ — плотность апостериорного распределения (по мере ν), вычисленная по формуле Байеса, отправляясь от априорной плотности $p_n(\theta)$. Тогда распределение P_x на (Ω, F), определенное почти для всех x и имеющее плотность по ν, равную $p(\theta|x) = \lim p_n(\theta|x)$, назовем фидуциальным распределением (соответствующим наблюдению x). Плотность $p(\theta|x)$ вычисляется однозначно и равна $p(x|\theta)/p(x)$. Тем самым фидуциальное распределение (фид. р.) P_x определенно однозначно с точностью до меры ν. Сейчас будет приведена статистическая модель, для которой требования инвариантности статистических выводов однозначно выделяют меру ν, а значит и фид. р. P_x.

Пусть G есть группа измеримых преобразований X на X. Сделаем следующие предположения.

П.1. $\theta_1 = \theta_2 \Leftrightarrow P_{\theta_1}(E) = P_{\theta_2}(E)$ для всякого $E \in B$.

П.2. Семейство $\mathscr{P}$ замкнуто по отношению к преобразованиям из G. Так, если случайная величина (сл. в.) x имеет распределение P_θ и $g \in G$, то существует $\theta^* \in \Omega$ такое, что сл. в. gx имеет распределение P_{θ^*}. В силу П.1 такой элемент θ^* единственный. Тем самым каждому $g \in G$ соответствует отображение g^* множества Ω в себя. Множество таких отображений g^* обозначим через G^*.

П.3. Всякий элемент $g^* \in G^*$ есть отображение Ω на себя. Теперь множество G^* превращается в группу.

П.4. Группы G и G^* изоморфны.

П.5. Группа G преобразований множества X строго транзитивна, т. е. для любых точек $x, y \in X$ существует единственное преобразование $g \in G$, переводящее x в y: $y = gx$.

П.6. Группа G^* преобразований множества Ω строго транзитивна.

Элементы группы G теперь могут быть использованы для описания точек $x \in X$ и точек $\theta \in \Omega$. Для этого выберем произвольные точки x_0 и θ_0 из X и Ω соответственно. Эти точки будут играть роль эталонных точек (масштабных элементов), по отношению к которым будут «измеряться» остальные точки из X и Ω. Так как для каждого $x \in X$ существует единственное преобразование $g_x \in G$, переводящее x_0 в $x = g_x x_0$, то тем самым элементы $x \in X$ параметризуются элементами $g_x \in G$. Аналогично, так как для каждого $\theta \in \Omega$ существует единственное преобразование $g_\theta^* \in G^*$, переводящее θ_0 в $\theta = g_\theta^* \theta_0$, то элементы $\theta \in \Omega$ описываются элементами $g_\theta^* \in G^*$. В силу же изоморфизма групп G и G^* элементу $g_\theta^* \in G^*$ соответствует элемент $g_0 \in G$. Мы указали взаимно однозначное соответствие между элементами множеств X, Ω, G, G^*. В частности, это соответствие порождает σ-алгебру F подмножеств множества Ω, исходя из σ-алгебры B подмножеств множества X. Кроме того, пользуясь таким соответствием, будем одной и той же буквой x обозначать как элемент из X, так и соответствующий ему элемент g_x из G. То же самое сделаем и для элементов из Ω. Например, под $\theta^{-1}x$ понимается $g_\theta^{-1} g_x \in G$.

Предположим, что выполнены следующие два принципа: прицип инвариантности фид. р. и принцип инвариантности энтропии фид. р. относительно выбора «масштабных элементов» x_0 и θ_0. Оказывается, что если в качестве μ выбрана левая мера Хаара в группе G, то из первого принципа следует, что мера ν является относительно инвариантной (см. (¹), стр. 257, задача 6). Если же еще выполнен и второй принцип, то ν становится правоинвариантной мерой, т. е. правой мерой Хаара, которая определяется однозначно с точностью до постоянного положительного множителя. Тем самым фид. р. P_x определено однозначно. При этом $p(\theta|x) = \Delta(x) \cdot q(\theta^{-1}x)$, где q — плотность распределения $P_{\theta_0} \in \mathscr{P}$, а $\Delta(x)$ — модулярная функция (см. (¹), стр. 256, задача 5).

Замечание. Для так определенного фид. р. справедливо равенство (фид. и доверительных вероятностей)

$$P_\theta\{\theta \in S(x)\} = P_x\{\theta \in S(x)\}$$

для всякого класса $\{S(x),\ x \in X\}$ измеримых (доверительных) множеств $S(x) \subseteq \Omega$, инвариантных относительно группы G, т. е. таких, что $g^*S(x) = S(gx)$ для любого $g \in G$. Это равенство может быть положено в основу определения фид. р. P_x: оно однозначно определяет P_x, если не требовать выполнения указанных двух приципов. Относительно определения фид. р. на основании так называемой центральной функции и частотной интерпретации фид. р. см. (²).

Теперь можно ввести фид. р. выборочной переменной как распределение выборочной переменной, когда ненаблюдаемый параметр θ имеет фид. р., соответствующее некоторому наблюдению x.

Пример. Пусть $x_1, \ldots, x_n$ — независимые наблюдения над сл. в. из r-мерной нормальной совокупности $N(\mu, A)$.

Случай 1. A известна, μ неизвестна. $\mathscr{P}$ — семейство распределений достаточных статистик $\bar{x} = \frac{1}{n}(x_1 + \ldots + x_n)$, соответствующих различным значениям μ; X — эвклидово пространство E_r; G — группа сдвигов в E_r.

Случай 2. A неизвестна, μ известна и $\mu = 0$. Обозначим через G^- группу нижних треугольных матриц размерности $r \times r$ с положительными элементами на главной диагонали. Пусть $\mathscr{P}$ — семейство распределений достаточных статистик t, соответствующих различным значениям $a \in G^-$; здесь t и a определяются однозначно (почти всюду) требованием: $aa' = A,\ tt' = T = \sum_1^n x_k x_k'$; $\quad a,\ t \in G^-$; $n \geqslant r$. $\quad X = \Omega = G = G^-$.

Случай 3. A неизвестна, μ неизвестна, $\mathscr{P}$ — семейство распределений достаточных статистик $(s, \bar{x})$, соответствующих различным $(a, \mu) \in G^- \times E_r$; здесь a и s определяются однозначно (почти всюду) требованием: $\quad aa' = A,\quad ss' = S = \frac{1}{n-1} \sum_1^n (x_k - \bar{x})(x_k - \bar{x})'$; $\qquad a,\quad s \in G^-$,

$X = \Omega = G^- \times E_r$. $G = \{[a, \mu] \mid a \in G^-, \mu \in E_r\}$ — группа с групповой операцией $[a, \mu][s, \bar{x}] = [as, a\bar{x} + \mu]$.

Пусть $W^-(r, n, B)$ — распределение, сосредоточенное на множестве положительно определенных матриц размерности $r \times r$, с плотностью

$$p(A) = \gamma_0(r,n) \frac{|B|^{(n-r-1)/2} d_-(B)}{|A|^{n/2} d_-(A)} \exp\left\{-\frac{n}{2} \operatorname{tr}(A^{-1}B)\right\},$$

где $d_-(A)$ — произведение главных миноров. Через $S^-(r, n)$ обозначим распределение на E_r с плотностью

$$p(t) = \gamma_1(r, n) \left[1 + \frac{(t, t)}{n}\right]^{-(n+1)/2} \frac{[1 + (t, t)/n]^{(r-1)/2}}{\prod\limits_{1}^{r-1} [1 + (t,t)_k/n]},$$

где $t = (t_1, \ldots, t_r)$; $(t, t)_k = t_1^2 + \ldots + t_k^2$. Если сл. в. t имеет распределение $S^-(r, n)$ и $a \in G^-$, то распределение сл. в. at обозначим через $K^-(r, n, A)$, где $A = aa'$. Фид. случайный ненаблюдаемый параметр и фид. выборочную переменную будем снабжать знаком звездочка сверху. Тогда в случае 1

$$\mu^* \in N\left(\bar{x}, \frac{1}{n} A\right); \quad x^* \in N\left(\bar{x}, \frac{n+1}{n} A\right)$$

(знак $\in$ означает, что, например, x^* имеет распределение $N\left(\bar{x}, \frac{n+1}{n} A\right)$).

В случае 2

$$A^* \in W^-(r, n, \hat{A};) \quad x^* - \mu \in K^-(r, n, \hat{A}); \hat{A} = \frac{1}{n} \sum_{1}^{n} (x_k - \mu)(x_k - \mu)'.$$

В случае 3

$$\sqrt{n}(\mu^* - \bar{x}) \in K^-(r, n-1, S); \quad A^* \in W^-(r, n-1, S);$$
$$x^* - \bar{x} \in K^-\left(r, n-1, \frac{n+1}{n} S\right).$$

Московский государственный университет
им. М. В. Ломоносова

Поступило
1 IX 1969

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] П. Халмош, Теория меры, М., 1953. [2] D. A. S. Fraser, Biometrika, 48, 261 (1961).