Доклады Академии наук СССР
1971. Том 198, № 2

УДК 519.25

МАТЕМАТИКА

В. Н. СУДАКОВ

ЗАМЕЧАНИЕ О МАРГИНАЛЬНОЙ ДОСТАТОЧНОСТИ

(Представлено академиком Ю. В. Линником 20 XI 1970)

Пусть в конечномерном пространстве $\mathbf{R}^n = \{(x_1, \ldots, x_n)\}$ задано семейство эквивалентных борелевских вероятностных мер $\{P_\theta, \theta \in \Theta\}$. Измеримая функция $y = f(x_1, \ldots, x_n)$ называется достаточной статистикой, а определяемое ею измеримое разбиение ξ_f называется достаточным разбиением, если условные вероятностные меры $P_{y,\theta}$ на элементах C_y разбиения ξ_f не зависят от θ для почти всех по распределению f значений y. Пусть ξ_k, $k = 1, \ldots, n$,— разбиения, порождаемые координатными функционалами $f_k(x_1, \ldots, x_n) = x_k$, и пусть статистика $y = f(x_1, \ldots$ $\ldots, x_n)$ такова, что условные распределения каждой координаты, т. е. меры на пространствах $(C_y, P_{y,\theta}) / \xi_k \equiv (\mathbf{R}^1, P_{y,\theta} / \xi_k)$, $k = 1, \ldots, n$, для почти всех значений y не зависят от θ (маргинальная достаточность). В. С. Хузурбазар высказал предположение, что в случае, когда распределения P_θ соответствуют повторной выборке, маргинальная достаточность статистики влечет ее достаточность. Появившийся в 1968 г. препринт И. К. Гхоша ([1]) анонсировал доказательство этой гипотезы, однако автор настоящей заметки и его коллеги встретились с затруднениями при восстановлении полного доказательства по работе И. К. Гхоша. Ниже излагается краткое доказательство несколько более общего (при $n = 2$) утверждения.

Как, в частности, отмечено в ([1]), достаточно рассмотреть случай, когда Θ содержит два элемента.

Теорема 1. *Пусть P и Q — две взаимно абсолютно непрерывные борелевские вероятностные меры в $\mathbf{R}^n$, отвечающие независимой выборке. Пусть статистика $y = f(x_1, \ldots, x_n)$ такова, что для почти всех по распределению f значений y выполнены условия $(C_y, P_y) / \xi_k = (C_y, Q_y) / \xi_k$, $k = 1, \ldots, n$, где ξ_k — координатные разбиения, а P_y и Q_y — условные вероятностные меры на элементе C_y разбиения ξ_f.*

Тогда f — достаточная статистика.

В случае $n = 2$ и только в этом случае вместо независимости выборки достаточно потребовать выполнения условия $dQ / dP = p_1(x_1) p_2(x_2)$.

Лемма 1. *Пусть P и Q — две взаимно абсолютно непрерывные вероятностные меры на пространстве $(\Omega, \mathfrak{A})$ и ξ — измеримое разбиение, для которого существуют системы $\{P_C\}$ и $\{Q_C\}$ условных вероятностных мер (C — элемент разбиения ξ). Если $dQ / dP = p(\omega)$, то на почти каждом элементе $C \in \xi$ меры P_C и Q_C взаимно абсолютно непрерывны и $dQ_C / dP_C = k(C) p(\omega)$ ($\omega \in C$, $k(C)$ — значение плотности $d(Q / \xi) / d(P / \xi)$ на элементе C).*

Лемма 2. *Пусть U и V — такие конечные положительные борелевские меры в $\mathbf{R}^n$, что для некоторого l*

$$U\{x = (x_1, \ldots, x_n) \in \mathbf{R}^n : \sum x_k \leqslant l\} = 0,$$
$$V\{x = (x_1, \ldots, x_n) \in \mathbf{R}^n : \sum x_k \geqslant l\} = 0,$$
$$(\mathbf{R}^n, U) / \xi_k = (\mathbf{R}^n, V) / \xi_k \quad k = 1, \ldots, n,$$

и в случае $n > 2$ $\displaystyle\int_{-\infty}^{\infty} |x_k| \, U(dx) < \infty.$

Тогда $U = V = 0$.

Доказательство теоремы 1. Пусть сперва статистика f такова, что для некоторого y_0 $P(C_{y_0}) > 0$. Введем на множестве $C = C_{y_0}$ условные вероятности P_C и Q_C; тогда $dQ_C / dP_C = (P(C) / Q(C)) p_1(x_1) \dots p_n(x_n)$, где $p_k(x_k)$ — соответствующие плотности маргинальных распределений для Q относительно P. Рассмотрим отображение φ: $\mathbf{R}^n \to \mathbf{R}^n$, $(x_1, \dots, x_n) \mapsto (\ln p_1(x_1), \dots, \ln p_n(x_n))$. Отображение φ переводит меры P и Q в меры $\widetilde{P} = P\varphi^{-1}$ и $\widetilde{Q} = Q\varphi^{-1}$, а меры P_C и Q_C перейдут в некоторые меры $\widetilde{P}_C = P_C\varphi^{-1}$ и $\widetilde{Q}_C = Q_C\varphi^{-1}$, причем $d\widetilde{P}_C / d\widetilde{P} \leqslant (P(C))^{-1}$, $d\widetilde{Q}_C / d\widetilde{P}_C = (P(C) / Q(C)) \exp \sum x_k$.

Пусть теперь $\widetilde{Q}_C - \widetilde{P}_C = U - V$, где U и V — дизъюнктные неотрицательные меры. Докажем, что меры U и V имеют (абсолютный) первый момент. Действительно, если $\{U_s\}$ — семейство условных вероятностных мер на гиперплоскостях $\{x = (x_1, \dots, x_n) : \sum x_k = s\}$, а $\alpha(ds)$ — соответствующая фактор-мера на множестве таких гиперплоскостей, то

$$\widetilde{P}_C = \int_{-\ln R}^{\infty} U_s \alpha(ds) (R \exp s - 1)^{-1},$$

где $R = P(C) / Q(C)$, откуда

$$\int_{-\infty}^{\infty} \exp x_k \int_{-\ln R}^{\infty} U_s(dx) \alpha(ds) (R \exp s - 1)^{-1} \leqslant (P(C))^{-1} \quad (k = 1, \dots, n),$$

а отсюда вытекает, что

$$\int_{-\infty}^{0} |x_k| U(dx) < \infty \quad (k = 1, \dots, n).$$

Аналогично

$$\int_{0}^{\infty} |x_k| V(dx) < \infty \quad (k = 1, \dots, n).$$

Если маргинальные распределения P_C и Q_C совпадают, то совпадают и маргинальные распределения мер U и V, поэтому существование левосторонних первых моментов меры U и правосторонних моментов меры V влечет применимость леммы 2, и в этом случае теорема доказана: $P_C = Q_C$. В том случае, когда $P(C_y) = 0$ для почти всех y, как и прежде, с учетом леммы 1, для почти каждого y получаем

$$\int_{-\ln R}^{\infty} \alpha_y(ds) U_{y,s} (R_y \exp s - 1)^{-1} = \widetilde{P}_y = P_{C_y}\varphi^{-1},$$

где R_y — значение плотности распределения μ статистики f по мере P относительно распределения ее же по мере Q в точке y. Поскольку

$$\int_{-\infty}^{\infty} \widetilde{P}_y \mu(dy) = \widetilde{P} \quad \text{и} \quad \int_{-\infty}^{\infty} \exp x_k \widetilde{P}(dx) = 1,$$ для μ-почти всех значений y

выполнено условие

$$\int_{-\infty}^{\infty} \exp x_k \widetilde{P}_y(dx) < \infty,$$

откуда, как и раньше, следует существование левосторонних моментов для U_y и, следовательно, применимость леммы 2.

Можно привести пример, показывающий, что при $n > 2$ условие $dQ / dP = kp_1(x_1) \dots p_n(x_n)$ не обеспечивает достаточности каждой маргинально достаточной статистики.

Ленинградское отделение
Математического института им. В. А. Стеклова
Академии наук СССР

Поступило
14 XI 1970

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] J. K. Ghosh, Indian Statistical Institute, Tech. Report Math. Stat., 11, 1968.