Доклады Академии Наук СССР
1953. Том XCI, № 5

МАТЕМАТИКА

И. И. ГИХМАН

НЕКОТОРЫЕ ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ДЛЯ УСЛОВНЫХ РАСПРЕДЕЛЕНИЙ

(Представлено академиком А. Н. Колмогоровым 19 VI 1953)

Ряд задач математической статистики, рассмотренных в последнее время ([6-9]), привел к вопросу об исследовании условных распределений максимального и минимального членов последовательности сумм независимых случайных слагаемых. В настоящей заметке указывается, что возникающие здесь вопросы можно сравнительно просто и в общем виде исследовать с помощью метода верхних и нижних функций, впервые примененным для аналогичных целей И. Г. Петровским ([1]), А. Н. Колмогоровым ([2]) и А. Я. Хинчиным ([3]).

Теорема 1. *Пусть $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ — последовательность независимых одинаково распределенных случайных величин, имеющих конечные моменты* 1-го *и* 2-го *порядков, причем $M\xi_k = 0$, $M\xi_k^2 = \sigma^2$, и удовлетворяющих одному из следующих двух условий*:

a) *величины ξ_k имеют плотность распределения $f(x)$, где $f(x)$ — функция с ограниченным изменением*;

b) *величины ξ_k являются решетчатыми.*

Положим

$$\eta_k = \frac{1}{\sigma\sqrt{n}} \sum_{r=1}^{k} \xi_r,$$

$$M(n) = \max\{\eta_0, \eta_1, \ldots, \eta_n\},\ m(n) = -\min\{\eta_0, \eta_1, \ldots, \eta_n\}$$

и пусть z_n — одно из возможных значений величины η_n.

Тогда при $n \to \infty$, $z_n \to z$ условная функция распределения пары $\{m(n), M(n)\}$ при гипотезе $\eta_n = z_n$ стремится к $S(x, y)$,

$$S(x, y) = \sum_{-\infty}^{\infty} \exp\{-2r(x+y)[r(x+y)-z]\} -$$

$$-\exp\{-2(ry+(r-1)x)(ry+(r-1)x-z)\}. \qquad (1)$$

Доказательство этой теоремы основано на следующем замечании: если фиксировать величину η_n, положив $\eta_n = z_n$, то последовательность $\eta_0, \eta_1, \ldots, \eta_n$ превращается в простую цепь Маркова. В силу этого следует ожидать, что предельное распределение пары $\{M(n), -m(n)\}$ совпадает с распределением максимума и минимума случайной функции соответствующего непрерывного во времени марковского процесса.

Для составления соответствующего дифференциального уравнения нужны значения главных частей условных моментов величины ξ_i при гипотезе $\eta_n = z_n$. Пусть $\alpha_k(n, \varepsilon, z)$ — условное математическое ожидание величины $\chi_\varepsilon(\xi_1)\xi_1^k$ $(k = 0, 1, 2)$ при гипотезе $\eta_n = z$, где $\chi_\varepsilon(\xi) = 1$, если $|\xi| < \varepsilon\sqrt{n}$, и $\chi_\varepsilon(\xi) = 0$ в противном случае.

Лемма 1. *Если выполнены предположения теоремы* 1, *то*

$$1 - \alpha_0(n, \varepsilon, z) = \frac{\eta(\varepsilon)}{n} + o\left(\frac{1}{n}\right),$$

$$\alpha_1(n, \varepsilon, z) = \frac{z}{\sqrt{n}} + \frac{\eta(\varepsilon)}{\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right),$$

$$\alpha_2(n, \varepsilon, z) = \sigma^2 + \eta(\varepsilon) + o(1),$$

где оценки $o(1)$, $o\left(\frac{1}{n}\right)$ *имеют место при фиксированном* $\varepsilon > 0$ *равномерно относительно* z *в любом конечном интервале* $|z| < A$, *а величина* $\eta(\varepsilon) \to 0$ *при* $\varepsilon \to 0$.

Доказательство этой леммы, родственной ряду результатов, полученных А. Я. Хинчиным (см., например, [4]), основано на применении формулы Байесса и предельных локальных теорем для сумм независимых одинаково распределенных слагаемых.

Используя эту лемму и уже упомянутый метод мажорантных функций, легко получить, что $P_n(x, y / \alpha, z_n)$ — условная вероятность при гипотезе $\eta_n = z_n$ совмещения событий $\max\limits_{k \leqslant \alpha n} \eta_k < y$, $-\min\limits_{k \leqslant \alpha n} \eta_k < x$ $(0 \leqslant x,\ 0 \leqslant y,\ \alpha < 1)$ при $n \to \infty$, $z_n \to z$ — стремится к $u(0, 0 / z, \alpha)$, где $u(t, s / z, \alpha)$ удовлетворяют уравнению

$$\frac{\partial u}{\partial t} + \frac{z - s}{1 - t}\frac{\partial u}{\partial s} + \frac{1}{2}\frac{\partial^2 u}{\partial s^2} = 0 \tag{2}$$

и граничным условиям $u(\alpha, s / z, \alpha) = 1$ при $-x < s < y$ и $u(t, -x/z, \alpha) = u(t, y / z, \alpha) = 0$ для $0 \leqslant t < \alpha$. Далее, очевидно, существует $\lim u(t, s / z, \alpha) = u_0(t, s / z)$.

Нетрудно показать, что $\lim\limits_{n \to \infty} P_n(x, y / z, 1) = u_0(0, 0 / z)$.

Решив уравнение (2) при соответствующих граничных условиях, легко получить для $u_0(0, 0 / z)$ выражение (1).

Замечание. Теорему 1 можно обобщить, рассматривая вероятность осуществления неравенств

$$\varphi\left(\frac{k}{n}\right) < \eta_k < \psi\left(\frac{k}{n}\right),$$

где $\varphi(t)$, $\psi(t)$ — непрерывно дифференцируемые функции, $\varphi(0) < 0 < \psi(0)$, $\varphi(t) < \psi(t)$ $(0 \leqslant t \leqslant 1)$. Предельное значение условной вероятности соответствующего события при гипотезе $\eta_n = z_n$ получается так же, как и в теореме 1, только в этом общем случае от решения уравнения (2) следует требовать обращения в 0 на кривых $s = \varphi(t)$, $s = \psi(t)$.

В качестве одного из применений теоремы 1 укажем на следующее. В некоторых вопросах прикладной статистики рассматривают величины

$$\zeta_k = \eta_k - \frac{k}{n}\eta_n, \quad M'(n) = \max_{0 \leqslant k \leqslant n} \zeta_k, \quad m'(n) = -\min_{0 \leqslant k \leqslant n} \zeta_k.$$

Пользуясь соответствием между случайными процессами и суммированием случайных величин, Феллер [6] указал предельное распределение

ление пары $\{m'(n),\ M'(n)\}$ при $n \to \infty$. Полное доказательство соответствующей предельной теоремы может быть непосредственно получено из замечания к теореме 1. Предположим, что условия теоремы 1 выполнены. Тогда событие $m'(n) < x,\ M'(n) < y$ при гипотезе $\eta_n = z_n$ равносильно совместному выполнению неравенств $-x + \frac{k}{n} z_n < \eta_k < y + \frac{k}{n} z$. Замечание к теореме 1 приводит к следующему.

Теорема 2. *Если $n \to \infty$, $z_n \to z$, то условная функция распределения случайных величин $\{m'(n),\ M'(n)\}$ при гипотезе $\eta_n = z_n$ стремится к пределу*

$$S'(x, y) = 1 + \sum_1^{\infty} 2 \exp\{-2r^2 (x + y)^2\} - \exp\{-2 (ry + (r-1) x)^2\} - \\ - \exp\{-2 ((r-1) y + rx)^2\}, \tag{3}$$

не зависящему от z. К тому же пределу сходится безусловное распределение пары $\{m'(n),\ M'(n)\}$.

Остановимся еще на так называемом принципе инвариантности, тесно связанном с предыдущими теоремами. Обозначим через $\zeta(t)$ случайную функцию марковского процесса, определяемого дифференциальным уравнением («Первое дифференциальное уравнение Колмогорова»)

$$\frac{\partial u}{\partial t} - \frac{s}{1-t} \frac{\partial u}{\partial s} + \frac{1}{2} \frac{\partial^2 u}{\partial s^2} = 0, \quad 0 \leqslant t < 1. \tag{4}$$

В явном виде этот процесс можно определить как процесс, для которого совместная n-мерная функция распределения последовательности $\zeta(t_1), \zeta(t_2), \ldots, \zeta(t_n)$ $(0 < t_1 < t_2 < \ldots < t_n < 1)$ является нормальной с дисперсионной матрицей $a_{ik} = t_i (1 - t_k)$ $(i \leqslant k)$.

Пусть $f(x)$ — произвольная ограниченная и измеримая по Лебегу функция. Нетрудно доказать, что интеграл

$$I(f) = \int_0^1 f[\zeta(t)]\, dt \tag{5}$$

существует в смысле сходимости в среднем соответствующих интегральных сумм.

Лемма 2. *Пусть $f(x)$ — произвольная ограниченная и измеримая по Лебегу функция и величины ξ_k удовлетворяют условиям теоремы 1. Тогда условное распределение суммы*

$$I_n(f) = \frac{1}{n} \sum_1^n f(\zeta_k) \tag{6}$$

при гипотезе $\eta_n = z_n$ (z_n — допустимое значение величины η_n) слабо сходится при $n \to \infty$ и $z_n \to z$ к распределению функционала (5), *где $\zeta(t)$ — случайная функция марковского процесса, управляемого уравнением* (4).

Доказательство этой леммы можно построить следующим образом. Сначала рассматривается случай функций $f(x)$, удовлетворяющих условию Липшица. Для таких функций лемма легко следует из замечания, что разность характеристических функций сумм $\frac{1}{n} \sum_{r=1}^{n} f(\zeta_{rm})$ и $\frac{1}{nm} \sum_{r=1}^{nm} f(\zeta_r)$ (соответствующих условным распределениям при гипотезе $\eta_{n\cdot m} = z_{n\cdot m}$), при $n \to \infty$ стремится к 0 равномерно относительно m. Последнее же легко получить, используя лемму 1.

Доказательство для общего случая может быть получено путем аппроксимации произвольной функции $f(x)$ двумя последовательностями $\{f_n^+(x)\}$, $\{f_n^-(x)\}$ функций, удовлетворяющих условию Липшица и почти всюду сходящихся к $f(x)$, из которых первая монотонно убывающая, а вторая монотонно возрастающая.

Легко доказать, что $I(f_n^+) - I(f_n^-)$ при $n \to \infty$ сходится в среднем к 0.

Возьмем в качестве примера функцию $f(x) = 1$ при $x > 0$ и $f(x) = 1$, если $x \leqslant 0$. Условимся откладывать суммы $\sigma_k = \xi_1 + \ldots + \xi_k$, рассматривая их как функции аргумента $t = k/n$, в некоторой прямоугольной системе координат. Соединим точки $(0, 0) = (\sigma_0, 0)$ и $(\sigma_n, 1)$ отрезком L. В нашем примере (6) дает отношение числа тех точек $(\sigma_k, k/n)$, которые лежат выше отрезка L, к общему числу точек n. В том случае, когда величины ξ_k принимают значения ± 1 с одинаковой вероятностью, распределение сумм (6) при гипотезе $\eta_n = 0$ известно ([6,7]).

Лемма 2 позволяет сформулировать следующую общую теорему.

Теорема 3. *Пусть ν_n обозначает число положительных членов последовательности*

$$\sum_1^k \xi_r - \frac{k}{n} \sum_1^n \xi_r, \quad k = 1, 2, \ldots, n-1,$$

и $\{\xi_k\}$ удовлетворяют условиям теоремы 1.

Тогда при $n \to \infty$ и $z_n \to z$ условное распределение отношения ν_n / n при гипотезе $\eta_n = z_n$ слабо сходится к равномерному распределению.

Частный случай теоремы 3, соответствующий случаю $z = 0$ и величинам ξ_k, принимающим значения, кратные 1, был недавно доказан при несколько более жестких предположениях в работе ([8]).

Поступило
29 IV 1953

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] И. Г. Петровский, Math. Ann., **109**, 425 (1934). [2] А. Н. Колмогоров, Изв. АН СССР, отд. матем. и ест. наук, 363 (1933). [3] А. Я. Хинчин, Асимптотические законы теории вероятностей, 1936. [4] А. Я. Хинчин, Матем. сборн., **12**:2 (1943). [5] Б. В. Гнеденко, А. Н. Колмогоров, Предельные распределения для сумм независимых случайных слагаемых, 1950. [6] K. L. Chung, W. Feller, Proc. Nat. Acad. Sci., 35 (1949). [7] Б. В. Гнеденко, В. С. Михалевич, ДАН, **85**, 1 (1952). [8] M. Lipchitz, Proc. Am. Math. Soc., 3, No. 4 (1952). [9] W. Feller, Ann. Math. Stat., **22**, No. 3 (1951).