Доклады Академии Наук СССР
1949. Том LXIX, № 4

МАТЕМАТИКА

Г. М. МАНИЯ

ОБОБЩЕНИЕ КРИТЕРИЯ А. Н. КОЛМОГОРОВА ДЛЯ ОЦЕНКИ ЗАКОНА РАСПРЕДЕЛЕНИЯ ПО ЭМПИРИЧЕСКИМ ДАННЫМ

(Представлено академиком А. Н. Колмогоровым 7 X 1949)

Пусть $x_1, x_2, \ldots, x_n$ — совокупность независимых величин с общим непрерывным законом распределения $F(x)$. Пусть, далее, $x_1^*, x_2^*, \ldots, x_n^*$ представляет ту же последовательность, но в порядке их величин. Будем называть эмпирической функцией распределения ступенчатую функцию $s_n(x)$:

$$S_n(x) = \begin{cases} 0 & \text{для} \quad x < x_1^*, \\ k/n & \text{для} \quad x_k^* \leqslant x < x_{k+1}^*, \\ 1 & \text{для} \quad x \geqslant x_n^*. \end{cases}$$

Обозначим

$$D_n = \sup_{-\infty < x < \infty} |S_n(x) - F(x)|,$$

$$D_n^+ = \sup_{-\infty < x < \infty} \{S_n(x) - F(x)\}.$$

Согласно известной теореме, доказанной А. Н. Колмогоровым (1), при каждой $\lambda \geqslant 0$ и произвольной непрерывной функции распределения $F(x)$

$$P\left\{D_n \leqslant \frac{\lambda}{\sqrt{n}}\right\} \xrightarrow[n\to\infty]{} \Phi(\lambda) = 1 - 2\sum_{\nu=1}^{\infty} (-1)^{\nu-1} e^{-2\nu^2\lambda^2}. \tag{1}$$

Один из результатов Н. В. Смирнова (2) заключается в установлении асимптотической формулы для распределения D_n^+:

$$P\left\{D_n^+ \leqslant \frac{\lambda}{\sqrt{n}}\right\} \xrightarrow[n\to\infty]{} 1 - e^{-2\lambda^2}. \tag{2}$$

Мы обобщаем критерии соответствия А. Н. Колмогорова и Н. В. Смирнова, рассматривая максимум уклонения для определенного участка $(0 < \theta_1 < \theta_2 < 1)$ роста функции $F(x)$.

Определим две случайные величины:

$$D_n^+(\theta_1, \theta_2) = \sup_{\theta_1 \leqslant F(x) \leqslant \theta_2} \{S_n(x) - F(x)\},$$

$$D_n(\theta_1, \theta_2) = \sup_{\theta_1 \leqslant F(x) \leqslant \theta_2} |S_n(x) - F(x)|.$$

Полученные нами результаты могут быть сформулированы в виде следующих двух теорем:

Теорема 1. *Пусть $F(x)$ — непрерывная функция распределения каждой из независимых величин x_i $(i=1, 2, \ldots, n)$, $\theta_1^{(n)}=\frac{m_1}{n}=\theta_1+o\left(\frac{1}{\sqrt{n}}\right)$, и $\theta_2^{(n)}=\frac{m_2}{n}=\theta_2+o\left(\frac{1}{\sqrt{n}}\right)$, $0<\theta_1<\theta_2<1$.*

Тогда

$$P\left\{D_n^+(\theta_1^{(n)}, \theta_2^{(n)}) \leqslant \frac{\lambda}{\sqrt{n}}\right\} \xrightarrow[n\to\infty]{} \Phi(\theta_1, \theta_2; \lambda),$$

где

$$\Phi^+(\theta_1, \theta_2; \lambda)=\frac{1}{2\pi\sqrt{1-R^2}}\int_{-\infty}^{a}\int_{-\infty}^{b} e^{-1/2\,\overline{\theta}(z_1, z_2)}\,dz_1\,dz_2-$$

$$-\frac{e^{-2\lambda^2}}{2\pi\sqrt{1-R^2}}\int_{-\infty}^{a'}\int_{-\infty}^{b'} e^{-1/2\,\theta(z_1, z_2)}\,dz_1\,dz_2,$$

$$a=\frac{\lambda}{\sqrt{\theta_1(1-\theta_1)}}, \quad b=\frac{\lambda}{\sqrt{\theta_2(1-\theta_2)}}, \quad a'=\frac{\lambda-2\lambda\theta_1}{\sqrt{\theta_1(1-\theta_1)}}, \quad b'=\frac{\lambda-2\lambda(1-\theta_2)}{\sqrt{\theta_2(1-\theta_2)}},$$

$$R=\sqrt{\frac{\theta_1(1-\theta_2)}{\theta_2(1-\theta_1)}},$$

$$\theta(z_1, z_2)=\frac{1}{1-R^2}[z_1^2+2Rz_1z_2+z_2^2], \quad \overline{\theta}(z_1, z_2)=\frac{1}{1-R^2}[z_1^2-2Rz_1z_2+z_2^2].$$

Функция $\Phi^+(\theta_1, \theta_2; \lambda)$ может быть представлена также в следующем виде:

$$\Phi^+(\theta_1, \theta_2; \lambda)=\sum_{n=0}^{\infty}\frac{(+R)^n}{n!}\Phi^{(n)}(a)\,\Phi^{(n)}(b)-e^{-2\lambda^2}\sum_{n=0}^{\infty}\frac{(-R)^n}{n!}\Phi^{(n)}(a')\,\Phi^{(n)}(b').$$

Здесь $\Phi^{(n)}(x)$ есть производная n-го порядка от нормального интеграла

$$\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-z^2/2}\,dz.$$

При наиболее интересном частном случае, когда $\theta_1=1-\theta_2=\theta$, будем иметь

$$\Phi^+(\theta; \lambda)=\frac{1}{2\pi\sqrt{1-R^2}}\int_{-\infty}^{c}\int_{-\infty}^{c} e^{-1/2\,\overline{\theta}(z_1, z_2)}\,dz_1\,dz_2-$$

$$-\frac{e^{-2\lambda^2}}{2\pi\sqrt{1-R^2}}\int_{-\infty}^{c'}\int_{-\infty}^{c'} e^{-1/2\,\theta(z_1, z_2)}\,dz_1\,dz_2,$$

где

$$c=\frac{\lambda}{\sqrt{\theta(1-\theta)}}, \quad c'=\frac{\lambda-2\lambda\theta}{\sqrt{\theta(1-\theta)}}.$$

Отсюда при $\theta=0$ мы получаем (2).

Теорема 2. *При условиях теоремы* 1

$$P\{D_n(\theta_1^{(n)}, \theta_2^{(n)}) \leqslant \lambda n^{-1/2}\} \xrightarrow[n\to\infty]{} \Phi(\theta_1, \theta_2; \lambda),$$

причем

$$\Phi(\theta_1, \theta_2; \lambda) = \frac{1}{2\pi\sqrt{1-R^2}} \int\limits_{-a}^{a}\int\limits_{-b}^{b} e^{-1/2\theta(z_1, z_2)}\, dz_1\, dz_2 -$$

$$-\frac{2\sum\limits_{k=1}^{\infty}(-1)^{k-1} e^{-2k^2\lambda^2}}{2\pi\sqrt{1-R^2}} \int\limits_{-\alpha_k}^{\alpha_k}\int\limits_{-\beta_k}^{\beta_k} e^{-1/2\theta(z_1, z_2)}\, dz_1\, dz_2,$$

где

$$\alpha_k = \frac{\lambda - 2k\lambda\theta_1}{\sqrt{\theta_1(1-\theta_1)}}, \quad \beta_k = \frac{\lambda - 2k\lambda(1-\theta_2)}{\sqrt{\theta_2(1-\theta_2)}}.$$

В другом виде функция $\Phi(\theta_1, \theta_2; \lambda)$ *может быть представлена следующим образом:*

$$\Phi(\theta_1, \theta_2; \lambda) = \sum_{n=0}^{\infty} \frac{(-R)^n}{n!} \Phi^{(n)}(a)\, \Phi^{(n)}(b) -$$

$$-2\sum_{k=1}^{\infty}(-1)^{k-1} e^{-2k^2\lambda^2} \sum_{n=0}^{\infty} \frac{(-R)^n}{n!} \Phi^{(n)}(\alpha_k)\, \Phi^{(n)}(\beta_k).$$

В частности, когда $\theta_1 = 0$, $\theta_2 = 1$, *получаем* (1), *т. е. случай, рассмотренный впервые А. Н. Колмогоровым.*

В случае $\theta_1 = 1 - \theta_2 = \theta$ *имеем более компактное и симметричное выражение для предельной функции.*

Предлагаемый метод даст возможность теоретически решить вопрос о пригодности теоретического закона в тех границах, где материал, которым мы располагаем, более надежен для проводимого сравнения.

Доказательства теорем 1 и 2 основываются на теоремах непрерывности для производящих функций и трансформаций Лапласа.

Московский городской
педагогический институт
им. В. П. Потемкина

Поступило
7 X 1949

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] А. Н. Колмогоров, Giorn. d. Att., 4, 83 (1933). [2] Н. В. Смирнов, Матем. сборн., 6 (48), 3 (1939). [3] W. Feller, Ann. of Math. Statistics, 19, 2, 177 (1948).