Е. Л. Е м е л ь я н о в а, Л. П. У с о л ь ц е в (Самара, СГАУ, СамГУ). **Использование эффекта «увеличения» числа независимых наблюдений при проверке нормальности распределения по выборке малого объема.**

В соответствии с идеями Р. Мизеса [1], природа математической статистики такова, что она позволяет при интерпретировании своих основополагающих понятий выходить за жесткие рамки аксиоматики классической теории вероятностей. В частности, не входя в противоречие со здравым смыслом, можно предположить, что последовательности независимых наблюдений любой реальной случайной величины обладают теми же фундаментальными свойствами, что и совершенно абстрактные последовательности случайных чисел. В этом случае становится вполне объяснимой феноменология описываемого далее эффекта «увеличения» числа независимых наблюдений случайной величины.

Пусть требуется проверить гипотезу нормальности распределения случайной величины X, если известна некоторая отвечающая ей случайная выборка $\overline{X} = (X_1, X_2, \ldots, X_n)$ малого объема n. Возьмем произвольно целое число $q \geqslant 2$ и какую-либо равномерно распределенную в промежутке $[0,1)$ последовательность случайных чисел

$$\alpha_m = 0, \delta_1^{(m)} \delta_2^{(m)} \delta_3^{(m)} \ldots \quad (m = 1, 2, \ldots), \tag{1}$$

записанных в q-ичной системе счисления, т. е. такую последовательность, что для любого промежутка $\Delta \subset [0,1)$ длины $|\Delta| < 1$ при $P \to \infty$ выполняется соотношение

$$N_P(\Delta) = |\Delta| P + o\,(P), \tag{2}$$

где $N_P(\Delta)$ — количество чисел α_m, $m = 1, 2, \ldots, P$, попавших в промежуток Δ. При построении чисел (1) с $q = 10$ можно воспользоваться, например, табл. 7 из книги [2, с. 240]; для решения стандартных практических задач этого вполне достаточно. Используя последовательность (1) в рамках методики, предложенной в работах [3] и [4], мы и привносим в процесс обработки случайной величины X эффект «увеличения» числа независимых наблюдений.

Возьмем произвольный промежуток $\Delta \subset [0,1)$ длины $|\Delta| = 1/2$ и определим функцию $f(t)$, $0 \leqslant t < 1$, равенствами

$$f(t) = \chi_\Delta(t) = \begin{cases} 1 & \text{при } \{t\} \in \Delta, \\ 0 & \text{при } \{t\} \notin \Delta, \end{cases} \quad 0 \leqslant t < \infty. \tag{3}$$

Сопоставим каждому элементу X_i выборки $\overline{X}$ последовательность дробных долей $\{\alpha_i q^x\}$, $x = 0, 1, 2, \ldots$, и, взяв произвольно целое число $k \geqslant 2$, образуем k подвыборок $\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_k$ выборки $\overline{X}$, относя к $\overline{X}_s$ $(s = 1, 2, \ldots, k)$ элементы X_i выборки $\overline{X}$ с номерами i, для которых $\{\alpha_i q^{s-1}\} \in \Delta$. Если v_s — объем подвыборки $\overline{X}_s$, $s = 1, 2, \ldots, k$, то в силу соотношений (3) и (2), при всех $s = 1, 2, \ldots, k$ будет

$$v_s = \sum_{i=1}^{n} f\Big(\alpha_i q^{s-1}\Big) = \frac{n}{2} + o\,(n), \tag{4}$$

когда $n \to \infty$, т. е. $v_s \simeq [n/2]$, причем последнее приближенное равенство тем точнее, чем меньше модуль остаточного члена в соотношении (2).

В силу соотношения (4), подвыборки $\overline{X}_1, \overline{X}_2, \ldots, \overline{X}_k$ можно рассматривать как почти независимые случайные выборки, отвечающие случайной величине X, причем «качество» независимости определяется скоростью стремления к нулю остаточного члена в соотношении (2). В связи с этим, важную роль здесь могут сыграть построенные в теории чисел удобные для практического применения последовательности с очень быстрым равномерным распределением в промежутке $[0,1)$, ведущие себя как почти независимые случайные величины в соответствующем вероятностном пространстве.

Возьмем теперь произвольно r промежутков $\Delta_j \subset [0,1)$ с длинами $|\Delta_j| = 1/2$, $j = 1, 2, \ldots, r$, и повторим наши рассуждения, беря вместо промежутка Δ промежутки $\Delta_1, \Delta_2, \ldots, \Delta_r$. В результате мы получим kr практически независимых случайных выборок $\overline{X}_{sj}$ $(s = 1, 2, \ldots, k,\ j = 1, 2, \ldots, r)$, т. е. возникает классическая задача о проверке гипотезы о нормальности распределения достаточно большой совокупности попарно независимых малых выборок. Решение же этой задачи хорошо известно (см., например, [5, с. 287]).

СПИСОК ЛИТЕРАТУРЫ

1. *Мизес Р.* Вероятность и статистика. М.: ГИЗ, 1930.
2. *Чистяков В. П.* Курс теории вероятностей. М.: Наука, 1982.
3. *Емельянова Е. Л.*, *Усольцев Л. П.* Теоретико-числовая модификация статистики Пирсона. — Вестник СамТУ, серия «Физ.-матем. науки», 2005, в. 38, с. 15–18.
4. *Емельянова Е. Л.* О группировке эмпирических данных с использованием равномерно распределенных последовательностей случайных чисел. — Обозрение прикл. и промышл. матем., 2008, т. 15, в. 1, с. 89–90.
5. *Смирнов Н. В.*, *Дунин-Барковский И. В.* Курс теории вероятностей и математической статистики. М.: Наука, 1965.