

В. П. Котельников (Ростов-на-Дону, РВИРВ). **О проверке согласования генеральных и эмпирических распределений.**

Для проверки согласования генеральных и эмпирических распределений используют критерии Пирсона (χ^2), Колмогорова и Смирнова (ω^2), которые имеют свои особенности. Критерии Пирсона и Колмогорова не применимы при малых выборках. Критерий ω^2 не требует группировки опытных данных по разрядам, как критерий Пирсона, и может применяться как при малых, так и больших выборках. В критериях Колмогорова и ω^2 кроме вида генерального распределения должны быть указаны априорные значения всех параметров.

1. Предлагается отказаться от группировки опытных данных по разрядам и традиционной проверки согласования распределений по критериям χ^2 , Колмогорова, ω^2 и их модификациям. Вместо этого при объемах выборок $n > 5$ будем проверять гипотезу об адекватности генерального распределения с числом параметров m к эмпирическому распределению $G_i = i/n$, $i = 1, \dots, n$, методом регрессионного анализа. Для этого вычислим оценки дисперсий эмпирической функции распределения $s^2 = (n-1)^{-1} \sum_{i=1}^n (G_i - \bar{G})^2$, $\bar{G} = n^{-1} \sum_{i=1}^n G_i$, и адекватности $s_a^2 = (n-m)^{-1} \sum_{i=1}^n (G_i - \tilde{G}_i)^2$, где $\tilde{G}_i$ — предсказанное i -е значение генерального распределения. При уровне значимости α проверяем гипотезу $H_0: \sigma^2 = \sigma_a^2$ против альтернативной гипотезы $H_1: \sigma^2 > \sigma_a^2$. В качестве статистики критерия для проверки гипотезы H_0 используем $F = s^2/s_a^2$.

При нормальных распределениях независимых случайных величин для расчета квантилей используют F -распределение Фишера. Когда эти допущения не выполняются, в [1] предложен компьютерный метод определения критической точки $F_{1-\alpha}$, соответствующей уровню значимости α . Гипотеза об адекватности уравнения регрессии (генерального распределения) не противоречит эмпирическому распределению, если $F > F_{1-\alpha}$.

Если априорные значения параметров генерального распределения неизвестны, то их можно определить методом наименьших квадратов (МНК) в пакете Mathcad. Задача нелинейной аппроксимации решается путем применения нового алгоритма поиска значений параметров уравнения регрессии, минимизирующих сумму квадратов отклонений опытных точек от кривой регрессии. При решении задачи не требуется, как раньше, вычислять частные производные сглаживающего уравнения, эмпирические моменты аргументов и функции, а также решать системы нелинейных алгебраических уравнений.

2. Если случайные величины имеют сложные, малоизученные и неясные законы распределения, то предлагается в качестве генерального распределения использовать SN -распределение с интегральной функцией $G(x) = \int_{-\infty}^{\Psi(x)} S_k(t) dt$, где $S_k(t)$ — плотность стандартного распределения Стьюдента, $-\infty < t < \infty$, $k \geq 1$; $\Psi(x)$ — функция, обратная к монотонной функции одного случайного аргумента t и четырех параметров (a, b, γ, η)

$$x(t) = a + \frac{b-a}{\sqrt{2\pi}} \int_{-\infty}^{(t-\gamma)/\eta} \exp \left\{ -\frac{\xi^2}{2} \right\} d\xi, \quad \eta > 0, \quad -\infty < \gamma < \infty.$$

Параметры (a, g, γ, η, k) SN -распределения легко определяются МНК в пакете Mathcad. Это распределение существует в интервале (a, b) и в основном отражает физический смысл признаков реальных объектов и процессов. Кроме того, оно перекрывает области существования известных семейств распределений: бета, Джонсона (типов S_B , S_L , S_U), двойного нормального и др. Гипотеза о генеральном SN -распределении практически всегда не будет противоречить эмпирическому распределению, поэтому его не следует проверять на адекватность.

СПИСОК ЛИТЕРАТУРЫ

1. *Котельников В. П., Денисов О. Е.* Проверка адекватности уравнения регрессии при нарушении допущений критерия Фишера. — Обозрение прикл. и промышл. матем., 2004, т. 11, в. 4, с. 843–844.