

Г. В. Мартынов (Москва, ИППИ РАН). **Статистика Андерсона–Дарлинга и другие.**

Более шестидесяти лет критерий Андерсона–Дарлинга остается одним из наиболее применяемым среди всех критериев Крамера–Мизеса (омега-квадрат). В этой статистике равномерный эмпирический процесс, заданный на интервале $[0, 1]$ умножается на весовую функцию $\psi(t) = (t(1-t))^{-1/2}$. Использование весовых функций приводит к перераспределению чувствительности критерия к отклонениям функции распределения наблюдаемой случайной величины от гипотетической на различных ее участках. Для практики могут представлять критерии омега-квадрат и с другими весовыми функциями

$$\omega_n^2 = n \int_0^1 \psi^2(t)(F_n(t) - t)^2 dt,$$

где $F_n(t)$ — эмпирическая функция распределения. Из следующего дифференциального уравнения, полученного Андерсоном и Дарлингом,

$$h''(t) + \lambda \psi^2(t)h(t) = 0, \quad h(0) = h(1) = 0, \quad (1)$$

находятся собственные значения λ_i и собственные функции $\varphi_i(t) = h_i(t)\psi(t)$ для ковариационной функции взвешенного эмпирического процесса.

Для статистики Андерсона–Дарлинга уравнение (1) принимает вид

$$t(1-t)h''(t) + \lambda h(t) = 0, \quad h(0) = h(1) = 0, \quad (2)$$

откуда Андерсон и Дарлинг в 1952 г. нашли, что $\lambda_k = k(k+1)$, а собственные функции выражаются через ассоциированные с полиномами Лежандра полиномы Феррера. Можно заметить также, что уравнение (2) представляет собой частный случай гипергеометрического уравнения, и его решение легко выразить через гипергеометрическую функцию Гаусса следующим образом

$$\lambda_k = k(k+1), \quad h_k(t) = t \cdot {}_2F_1(-k, k+1; 2; t), \quad k = 1, 2, \dots$$

при прежних собственных значениях. Аналогичные результаты могут быть получены и для других весовых функций. Например, для $\psi(t) = \sqrt{t(1-t)}$ уравнение для собственных значений есть

$${}_1F_1\left(\frac{1}{4} - \frac{\sqrt{a}}{16}, \frac{1}{2}, \frac{\sqrt{a}}{4}\right) \cdot {}_1F_1\left(\frac{3}{4} - \frac{\sqrt{a}}{16}, \frac{3}{2}, \frac{\sqrt{a}}{4}\right) = 0.$$

Здесь ${}_1F_1(a; b; t)$ — обобщенная гипергеометрическая функция Куммера.

В таблице содержатся и значения квантилей предельного распределения статистики с $\psi(t) = (t(1-t))^a$, $a > -1$, при нескольких значениях a . Статистика Крамера–Мизеса с весовой функцией $\psi(t) = t^\beta$ рассмотрена Деувэлсом и Мартыновым в 2003 г. В [1] вычислена таблица квантилей статистик Крамера–Мизеса с весовыми функциями вида $\psi(t) = t^\alpha(1-t)^\beta$, $\alpha > -1$, $\beta > -1$, для различных пар α и β . Все значения квантилей вычислены численными методами без использования моделирования.

Таблица. Квантили взвешенной статистики Крамера–Мизеса

$p \backslash a$	-0,70	-0,50	-0,25	0	0,5	1	2	3
0,90	4,3776	1,9330	0,7945	3,4731	0,7213	1,5834	0,8245	4,5257
0,95	5,4914	2,4924	1,0437	4,6136	0,9709	2,1479	1,1284	6,2259
0,99	8,2378	3,8781	1,6606	7,4346	1,5872	3,5397	1,8769	10,411
0,995	9,4680	4,4982	1,9362	8,6939	1,8620	4,1603	2,2105	12,277
S	1	1	1	10	10	100	1000	100000

Для получения квантиля статистики следует значения из таблицы делить на величину S внизу ячеек.

Работа выполнена в ИППИ РАН при содействии гранта Российского Научного Фонда (Проект РНФ 14-50-00150)

СПИСОК ЛИТЕРАТУРЫ

1. Мартынов Г. В. Статистика Андерсона–Дарлинга и «обратная» к ней. — Информационные процессы (Электронный журнал ИППИ РАН), 2015, т. 15, № 3, с. 377–388. (Публикуется в Journal of Communications Technology and Electronics, 2016, v. 61, № 6.)