

Алгоритмы точного вычисления распределений статистики Пирсона и результаты численных экспериментов

Филина Марина Викторовна

Математический институт им. В.А. Стеклова РАН (Москва), Россия

e-mail: mfilina@mi.ras.ru

В работе проводится исследование точных распределений статистики Пирсона для различных полиномиальных схем, в том числе для нецентрального случая.

Как известно, распределение статистики Пирсона с числом исходов N и числом испытаний T , стремящимся к бесконечности, обычно аппроксимируется распределением χ^2 с $N - 1$ степенями свободы. В работах [?] приведен способ вычисления точных распределений разделимых статистик, который был реализован для случая статистики Пирсона в виде программ на языке C++.

Алгоритм позволяет за приемлемое время вычислять распределение для равновероятного случая с числом исходов до нескольких сотен и числом испытаний до нескольких тысяч (ограничения связаны с размером памяти для хранения массивов данных). В неравновероятном случае с рациональными вероятностями исходов размеры допустимых параметров в несколько раз меньше. Неравновероятный случай с произвольными вероятностями исходов сложнее в реализации ввиду многократного увеличения объема необходимой оперативной памяти для счета. Также был рассмотрен случай нецентрального распределения статистики Пирсона, когда реальные вероятности исходов отличаются от теоретических. Некоторые результаты исследования точности аппроксимации распределения статистики Пирсона распределением хи-квадрат были опубликованы в [2], [3].

Вычислительные эксперименты выявили некоторые интересные свойства разности между точной функцией распределения статистики Пирсона и соответствующей функцией распределения хи-квадрат:

- в графиках разностей проявляется «затухающая волна» с двумя локальными максимумами и двумя локальными минимумами;
- максимальные значения разности приблизительно обратно пропорциональны T ;
- при значениях аргумента, соответствующих квантилям, близким к 1, величины хвостов распределения статистики Пирсона оказываются больше соответствующих хвостов распределений χ^2 (при некоторых сочетаниях параметров — в разы);
- при фиксированном параметре нецентральности семейство точных функций распределений статистики Пирсона оказывается довольно широким, и аппроксимация функцией нецентрального распределения хи-квадрат или функцией нормального распределения (см. [4]) может оказаться неудовлетворительной.

Теоретическое обоснование экспериментальных результатов пока не получено. Алгоритмы точного вычисления функции распределения статистики Пирсона можно использовать для уточнения выбора ее критических значений.

Список литературы

- [1] ЗУБКОВ А. М. Методы расчета распределений сумм случайных величин. // Труды по дискретной математике — т. 5 — М.: Физматлит, 2002. — с. 51–60.
- [2] FILINA M. V., ZUBKOV A. M. Exact computation of Pearson statistics distribution and some experimental results. // Austrian Journal of Statistics. — vol. 37 (2008). — num. 1. — pp. 129–135.
- [3] FILINA M. V., ZUBKOV A. M. Tail Properties of Pearson Statistics Distributions. // Austrian Journal of Statistics. — vol. 40 (2011). — num. 1&2. — pp. 47–54.
- [4] BROFFITT J. D., RANGLES R. H. A Power Approximation for the Chi-Square Goodness-of-Fit Test: Simple Hypothesis Case. // JASA. — 1977. — vol. 72, No. 359. — pp. 604–607.