ОБ ИССЛЕДОВАНИИ МОЩНОСТИ КРИТЕРИЕВ ТИПА χ^2

Бушакова А.Д., Лемешко Б.Ю.
НГТУ, Новосибирск
E-mail: kaniala@yandex.ru

В настоящее время в прикладных задачах статистического анализа в качестве моделей законов распределения вероятностей реальных случайных величин используется несколько десятков законов распределения. Естественно, что все многообразие случайных величин не может быть покрыто этими законами распределений. Правомерность применения того или иного закона распределения в качестве модели для описания наблюдаемых случайных величин определяется используя критерии согласия.

Существует ряд критериев типа χ^2 для проверки гипотез о согласии эмпирического распределения с теоретическим. Рассматриваются критерий χ^2 Пирсона и критерий Джапаридзе-Никулина [1].

Статистика критерия χ^2 Пирсона задается выражением: $\mathrm{X}^2(\theta) = n\sum_{i=1}^{k}\frac{(n_i/n - p_i(\theta))^2}{p_i(\theta)}$, где k - число интервалов группирования, n - количество наблюдений, n_i - количество наблюдений, попавших в i-й интервал, $p_i(\theta)$ - вероятность попадания в этот интервал.

Критерий Джапаридзе-Никулина отличается от критерия Пирсона только при проверке сложных гипотез. Его статистика имеет вид [1]:

$$U^2(\theta) = \mathrm{X}^2(\theta) - nL^T(\theta)J^{-1}(\theta)L(\theta),\ L(\theta) = (l_1(\theta),\ldots,l_s(\theta)),\ l_j(\theta) = \sum_{i=1}^{k}\frac{n_i}{np_i(\theta)}\frac{\partial p_i(\theta)}{\partial\theta_j},$$

где s - число параметров закона, оцениваемых методом максимального правдоподобия по негруппированным данным, $J(\theta)$ - информационная матрица Фишера.

Однако при любом группировании происходит потеря информации об исходном законе распределения, что влияет на мощность критериев. Эти потери связанны с потерями в информации Фишера. В случае скалярного параметра задача асимптотически оптимального группирования (АОГ), связанная с минимизацией такого вида потерь, решается однозначно [2].

В случае векторного параметра мы имеем дело с информационной матрицей, и критерии минимизации потерь могут быть различными. В [2] задача АОГ была решена для случая D-оптимального группирования, когда в качестве функционала рассматривался определитель информационной матрицы по группированным данным. Однако не очевидно, что это наилучший вариант.

Возникает необходимость, во-первых, решения задач АОГ, обеспечивающих оптимум другим функционалам от информационной матрицы Фишера. Во-вторых, необходимость исследования мощности критериев при полученных вариантах АОГ.

Поэтому были рассмотрены две постановки задач АОГ. В задачах *A*- и *E*-оптимального группирования максимизировался след и наименьшее собственное число матрицы Фишера по группированным данным для ряда законов распределения (нормального, логистического, Вейбулла, Коши, экстремальных значений). Граничные точки интервалов группирования получены в виде, инвариантном относительно параметров закона. Построенные таблицы АОГ могут использоваться в задачах оценивания параметров по группированным данным и в критериях согласия.

Мощность критерия χ^2 Пирсона исследовалась для различных пар конкурирующих гипотез и критериев оптимальности группирования.

Исследование распределений статистики критерия Джапаридзе-Никулина при справедливости проверяемой гипотезы H_0 показало, что оно очень быстро сходится к предельному χ^2-распределению с $k-s-1$ степенями свободы [1].

Критерий Джапаридзе-Никулина рассматривался при различных парах конкурирующих гипотез. Мощность критерия Джапаридзе-Никулина исследовалась при *D*-оптимальном и равновероятном группировании.

Для сравнения рассматривался еще один способ группирования, предложенный в [4], который может использоваться при проверке гипотез при заданной конкурирующей гипотезе H_1. В качестве границ интервалов выбираются точки пересечения соответствующих плотностей (разбиение на интервалы Неймана-Пирсона [4]). Такое разбиение рекомендуется к применению в случае отсутствия возможности решать оптимизационную задачу.

Результатами исследований показано, что при *A*-оптимальном группировании мощность критерия χ^2 Пирсона несколько превосходит мощность его же в случае *D*-оптимального группирования. Применение *E*-оптимального группирования в критериях согласия неперспективно. Применение интервалов Неймана-Пирсона на практике является оправданным, когда нет возможности найти оптимальное решение.

Работа выполнена при поддержке РФФИ (проект № 09-01-0056а), АВП "Развитие научного потенциала высшей школы" (проект № 2.1.2/11855) и ФЦП Минобрнауки РФ “Научные и научно-педагогические кадры инновационной России”.

Литература

1 Джапаридзе К.О., Никулин М.С. Об одном видоизменении стандартной статистики Пирсона // Теория вероятностей и ее применения, 1974. Т.19. № 4. – С.886-888.

2 Денисов В.И., Лемешко Б.Ю., Цой Е.Б. Оптимальное группирование, оценка параметров и планирование регрессионных экспериментов: Монография. В 2-х ч. – Новосибирск: Новосиб. гос. техн. ун-т, 1993. – 347 с.

3 Воинов В.Г. Об оптимальных свойствах критерия Рао-Робсон-Никулина // Заводская лаборатория. Диагностика материалов. 2006. – Т.72. – №3. – С.65-70.

4 Greenwood P.E., Nikulin M.S. A Guide to Chi-Squared Testing. New York: John Wiley & Sons, Inc. 1996. – 280 p.