

Исследование влияния законов распределения ошибок на распределения статистик в линейном регрессионном анализе¹

Лемешко Б.Ю., Трушина Е.А.

НГТУ, г. Новосибирск, e-mail: headrd@fpm.ami.nstu.ru

Существующий аппарат проверки статистических гипотез в линейном регрессионном анализе разработан, основываясь на предположении о нормальности отклонений регрессии. Вопросы же проверки адекватности регрессионных моделей для случаев, когда ошибки наблюдений подчиняются другим законам, практически не исследованы. В работе исследовались распределения статистики критерия отношения правдоподобия при ошибках отклика, подчиняющихся распределениям: логистическому, Коши и экспоненциальному семейству распределений при различных значениях параметра формы.

Рассматривается классическая модель линейной регрессии

$$Y = X\theta + e, \quad (1)$$

где Y – вектор наблюдений размерности $(n \times 1)$, X – матрица независимых переменных размерности $(n \times m)$, θ – вектор оцениваемых параметров размерности $(m \times 1)$, e – вектор случайных отклонений системы размерности $(n \times 1)$.

Для нахождения оценок неизвестных параметров системы (1) используется метод максимального правдоподобия. Проверяемая гипотеза имеет вид:

$$H: \hat{\theta} = r. \quad (2)$$

Исследуемая статистика критерия отношения правдоподобия

$$\max_{\theta \in \Theta_H} f(Y, \theta, \lambda) / \max_{\theta \in \Theta} f(Y, \theta, \lambda)$$

при проверке гипотезы (2) принимает вид:

$$Q = \frac{(r - \hat{\theta})^T (X^T X)(r - \hat{\theta})}{ms^2}, \quad (3)$$

где s^2 – оценка выборочной дисперсии отклонений регрессии.

Исследования показали, что распределения статистики отношения правдоподобия существенно зависят от закона распределения ошибок. Исследуемые распределения статистики отношения правдоподобия хорошо описываются Бета-распределением II рода. Для различных значений числа наблюдений n и количества неизвестных параметров линейной регрессии m найдены значения параметров Бета-распределения II рода, аппроксимирующие распределения статистики. Построенные модели распределений статистики (3) могут использоваться при проверке адекватности регрессионных моделей при законах ошибок наблюдений, подчиняющихся распределению Коши и экспоненциальному семейству распределений с различными значениями параметра формы.

¹ Работа выполнена при поддержке Российского фонда фундаментальных исследований (проект 00-01-00913)

1. Демиденко Е.З. Линейная и нелинейная регрессия. – М.: Финансы и статистика, 1981. – 302 с.