

## Расширение области применения критериев типа Граббса проверки наблюдений на выбросы. Исследование распределений статистик и построение процентных точек<sup>1</sup>

Лемешко Б.Ю., Лемешко С.Б.  
НГТУ, г. Новосибирск, E-mail: skyer@phenix.bulat.info

Наиболее часто для проверки на аномальность наблюдений, принадлежащих выборкам из нормальной генеральной совокупности, применяют простые критерии Граббса [1-3]. Использование именно этих критериев предусматривает стандарт [4], представляющий собой аутентичный текст соответствующего международного стандарта ИСО 5725. Статистики критерия Граббса [1-3] предусматривают возможность проверки на наличие в выборке либо одного аномального наблюдения (наименьшего или наибольшего), либо двух (двух наименьших в выборке или двух наибольших). В то же время вид статистик позволяет естественным образом расширить возможности критерия за счет проверки гипотез о возможной принадлежности к аномальным измерениям большего числа выборочных наблюдений. В этих целях необходимо лишь исследовать распределения статистик соответствующего вида.

В данном случае исследованы распределения статистик типа Граббса при проверке на аномальность: **а)** одновременно наименьшего и наибольшего значения; **б)** одновременно трех максимальных (трех минимальных) значений. Построены соответствующие таблицы нижних процентных точек.

Пусть  $X_1, X_2, \dots, X_n$  – наблюдаемая выборка,  $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$  – построенный по ней вариационный ряд. Гипотеза  $H_0$  предполагает, что все наблюдения принадлежат одному закону.

**а)** Статистика для проверки на аномальность одновременно минимального  $X_{(1)}$  и максимального  $X_{(n)}$  выборочных значений принимает вид:

$$G_{1,n} = \frac{S_{1,n}^2}{S_0^2}, \quad (1)$$

где  $S_0^2 = \frac{1}{n-1} \sum_{j=1}^n (X_j - \bar{X})^2$ ,  $\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$ ,  $S_{1,n}^2 = \sum_{j=2}^{n-1} (X_j - \bar{X}_{1,n})^2$ ,

$\bar{X}_{1,n} = \frac{1}{n-2} \sum_{j=2}^{n-1} X_j$ . Оба значения считаются выбросами при заданном уровне

значимости  $\alpha$ , если вычисленное по выборке значение статистики (1) окажется ниже критического:  $G_{1,n} < G_{1,n,\alpha}$ .

<sup>1</sup> Работа выполнена при финансовой поддержке Минобразования РФ (проект № Т02-3.3-3356)

б) В случае проверки на аномальность одновременно трех минимальных или трех максимальных выборочных значений конкурирующая гипотеза  $H_1$  предполагает, что некоторому другому закону принадлежат  $X_{(1)}$ ,  $X_{(2)}$  и  $X_{(3)}$  (либо  $X_{(n-2)}$ ,  $X_{(n-1)}$  и  $X_{(n)}$ ). Статистики для проверки на аномальность одновременно трех минимальных или трех максимальных выборочных значений формируются в соответствии с соотношениями:

$$G_{1,2,3} = \frac{S_{1,2,3}^2}{S_0^2}, \quad (2)$$

$$G_{n-2,n-1,n} = \frac{S_{n-2,n-1,n}^2}{S_0^2}, \quad (3)$$

где  $S_{1,2,3}^2 = \sum_{j=4}^n (X_j - \bar{X}_{1,2,3})^2$ ,  $\bar{X}_{1,2,3} = \frac{1}{n-3} \sum_{j=4}^n X_j$ ,  $\bar{X}_{n-2,n-1,n} = \frac{1}{n-3} \sum_{j=1}^{n-3} X_j$ ,

$$S_{n-2,n-1,n}^2 = \sum_{j=1}^{n-3} (X_j - \bar{X}_{n-2,n-1,n})^2.$$

Статистики (2) и (3) распределены одинаково. Все три измерения считаются выбросами, если значение соответствующей статистики окажется ниже критического:  $G_{1,2,3} < G_{3,\alpha}$  или  $G_{n-2,n-1,n} < G_{3,\alpha}$ .

Таблицы процентных точек для одного минимального (максимального) значения и для двух минимальных (максимальных) были получены Граббсом в работах [1-3]. В стандарте [4] неверно указаны значения уровней значимости в соответствующей таблице (табл. 5). Эта неточность в стандарте, которая может приводить к пропуску аномальных наблюдений, было ранее отмечена нашими коллегами (Лях К.Н., Аврунев О.Е.) при предварительном исследовании распределений статистик критериев Граббса в системе Matlab.

В данной работе область применения критериев расширена за счет построения таблиц процентных точек (для объемов выборок  $5 < n < 150$ ) статистик типа Граббса при проверке на выброс одновременно трех максимальных (трех минимальных) значений и одновременно минимального и максимального значений в выборке. Кроме этого исследованы распределения статистик критериев Граббса при отклонениях наблюдаемого закона от нормального.

Подробности результатов исследований статистик критериев типа Граббса и построенные в данной работе таблицы процентных точек доступны по адресу [http://www.ami.nstu.ru/~headrd/seminar/Kontrol\\_Q/grubbs\\_1.htm](http://www.ami.nstu.ru/~headrd/seminar/Kontrol_Q/grubbs_1.htm).

1. Frank E. Grubbs. Sample Criteria for Testing Outlying observations // Ann. Math. Statist, 1950. – Vol. 21. – №. 1. – P.27-58.
2. Frank E. Grubbs. Procedures for Detecting Outlying Observations in Samples // Technometrics, 1969. – Vol. 11. – №. 1. – P.1-21

3. Frank E. Grubbs, Glenn Beck. Extension of sample sizes and percentage points for significance tests of outlying observations // *Technometrics*, 1972. – Vol. 14. – №. 4. – P.847-854.
4. ГОСТ Р ИСО 5725-2–2002. Точность (правильность и прецизионность) методов и результатов измерений. Часть 2. М.: Изд-во стандартов. – 51 с.