S©MR

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ
Siberian Electronic Mathematical Reports
http://semr.math.nsc.ru
ISSN 1813-3304

Том 22, № 1, стр. 154–168 (2025) УДК 519.23
https://doi.org/10.33048/semi.2025.22.012 MSC 62F03, 62F05

АНАЛОГ КРИТЕРИЯ ХАРКЕ–БЕРА ДЛЯ СЛУЧАЯ, КОГДА ПРОВЕРЯЕТСЯ ГИПОТЕЗА О РАСПРЕДЕЛЕНИИ СТЬЮДЕНТА

А.В. Логачёв, С.Е. Хрущев

Представлено Н.С. Аркашовым

Abstract: The Jarque–Bera test is commonly used in statistics and econometrics to test the hypothesis that sample elements adhere to a normal distribution with an unknown mean and variance. This paper proposes a modification of this criterion, which allows for testing hypotheses concerning the case where the sample comes from a skewed Student's distribution with an unknown number of degrees of freedom. The paper also provides power estimates for the constructed criterion, obtained through simulation of samples of various sizes.

Keywords: Jarque–Bera test; Student's distribution; Monte Carlo simulation.

1 Введение

Будем считать, что все рассматриваемые случайные элементы заданы на некотором вероятностном пространстве $(\Omega, \mathfrak{F}, \mathbf{P})$. Везде далее $\mathbf{E}$ и $\mathbf{D}$ будут означать математическое ожидание и дисперсию по мере $\mathbf{P}$,

---

Logachov, A.V., Khrushchev, S.E., Jarque – Bera Criterion Analog for Testing the Hypothesis of Student's Distribution.

Работа выполнена при поддержке РНФ, проект 24-21-00087.

Поступила 9 декабря 2024 г., опубликована 19 марта 2025 г.

соответственно. Пусть случайные величины $X, X_1, \ldots, X_n$, $n \in \mathbb{N}$ независимы и одинаково распределены. Предполагается, что их распределение неизвестно.

Нас будет интересовать критерий, который позволит проверять гипотезу о том, что элементы выборки имеют, вообще говоря, смещенное распределение Стьюдента. Этот критерий будет являться аналогом следующего критерия, предложенного C. Jarque и A.K. Bera (см., например, [2]–[4]) для проверки гипотезы о том, что элементы выборки имеют нормальное распределение.

Проверяется гипотеза
$H_0:\ X_i \sim N(a, \sigma^2)$, $1 \le i \le n$, *(a, σ^2 неизвестны) против гипотезы*
$H_1:\ X_i \nsim N(a, \sigma^2)$, $1 \le i \le n$.
Используется статистика

$$JB = n\left(\frac{S^{*2}}{6} + \frac{(K^* - 3)^2}{24}\right),$$

где n — объем выборки; S^* и K^* это выборочные асимметрия и эксцесс, соответственно, т.е.

$$S^* = \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{X})^3}{\hat{\sigma}^3}, \quad K^* = \frac{\frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{X})^4}{\hat{\sigma}^4},$$

$$\hat{\sigma} = \left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2\right)^{\frac{1}{2}}, \quad \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

Если верна гипотеза H_0, то при $n \to \infty$ распределение случайной величины JB сходится к $\chi^2(2)$. Поэтому при достаточно большом объеме выборки можно использовать следующее решающее правило, *задаем уровень значимости p, если $JB < \chi^2_{1-p}(2)$* (здесь $\chi^2_{1-p}(2)$ квантиль уровня $1-p$ распределения $\chi^2(2)$), *то гипотеза H_0 принимается, иначе отвергается.*

Отметим, что если в качестве альтернативы рассмотреть "более простую"гипотезу

$H_1:\ X_i$ *имеет распределение из семейства распределений Пирсона отличное от нормального,*
то критерий Харке – Бера будет состоятельным (т.е., если для элементов выборки справедлива гипотеза H_1, то статистика JB стремится по вероятности к ∞ при $n \to \infty$).

Напомним, что распределение случайной величины X принадлежит семейству распределений Пирсона, если его плотность $f(x)$ удовлетворяет следующему дифференциальному уравнению с разделяющимися переменными

$$\frac{df(x)}{dx} = \frac{b_0 + (x - a)}{b_1 + b_2(x - a) + b_3(x - a)^2} f(x),$$

где

$$a = \mathbf{E}X,\ b_1 = \sigma\frac{4K-3S}{10K-12S-18},\ b_0 = b_2 = \sigma\sqrt{S}\frac{K+3}{10K-12S-18},$$
$$b_3 = \frac{2K-3S+6}{10K-12S-18},\ \sigma^2 = \mathbf{D}X,$$

S, K — асимметрия и эксцесс случайной величины X, соответственно.

Заметим, что семейство распределений Пирсона весьма богато, в него входят, например, такие распределения, как показательное, гамма, бета, Стьюдента, нормальное. Если известно, что случайная величина имеет распределение из семейства Пирсона, то конкретный его тип определяется значениями асимметрии и эксцесса [5], см. рисунок 1. Именно

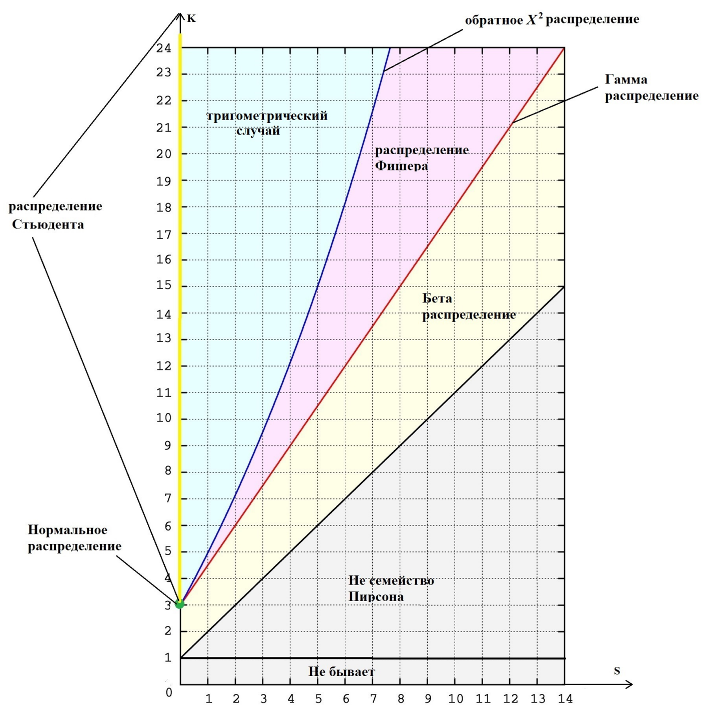

Рис. 1. Зависимость типа распределения от асимметрии S и эксцесса K

благодаря этому свойству, как отмечалось выше, критерий Харке – Бера является состоятельным на этом семействе. Отметим, что этот критерий на практике часто применяется в эконометрике для проверки гипотезы о нормальности распределения случайных возмущений в моделях линейной регрессии. Также отметим работу [6], в которой рассмотрены модификации критерия Харке – Бера для случаев, когда: известно математическое ожидание и неизвестна дисперсия; неизвестно математическое ожидание и известна дисперсия; известны и математическое ожидание и дисперсия.

Как упоминалось выше, нас будет интересовать аналог критерия Харке – Бера, который позволит проверять гипотезу о том, что элементы выборки имеют, вообще говоря, смещенное распределение Стьюдента. В разделе 2 настоящей работы нами построен следующий критерий (см. следствие 2).

Проверяется гипотеза

$H_0: \ X_i \sim T_a(k)$, $1 \leq i \leq n$, *(a и $9 \leq k \leq \infty$ неизвестны) против гипотезы*

$H_1: \ X_i$, $X_i \nsim T_a(k)$, $9 \leq k \leq \infty$, $1 \leq i \leq n$.

Используется статистика

$$\widehat{JB} := n\left(\frac{A_{1,n}^2}{|B_1|} + \frac{A_{2,n}^2}{|B_2|}\right),$$

где n — объем выборки, $A_{1,n} = A_{1,n}(S^*, \hat{\sigma})$, $A_{2,n} = A_{2,n}(K^*, \hat{\sigma})$, $B_1 = B_1(\hat{\sigma})$, $B_2 = B_2(\hat{\sigma})$ (точные формулы приведены в разделе 2), $T_a(k)$ — смещенное ($\mathbf{E}X = a$, где a, вообще говоря, необязательно равно нулю) распределение Стьюдента с k степенями свободы (полагается, что $T_a(\infty)$ это $N(a, 1)$).

Если гипотеза H_0 верна, то при $n \to \infty$ распределение случайной величины $\widehat{JB}$ сходится к $\chi^2(2)$ (см. раздел 2, следствие 1). Поэтому при достаточно большом объеме выборки можно использовать следующее решающее правило, *задаем уровень значимости p, если $\widehat{JB} < \chi^2_{1-p}(2)$, то гипотеза H_0 принимается, иначе отвергается.*

Таким образом, как и в классическом критерии Харке – Бера, распределение построенной статистики сходится к $\chi^2(2)$ при верной основной гипотезе, и для ее определения используются первые четыре выборочных момента. Предлагаемый критерий не является состоятельным при сложной конкурирующей гипотезе H_1 приведенной выше. При этом он будет состоятельным (см. раздел 2, замечание 1), если в качестве альтернативы рассмотреть "более простую"гипотезу

$H_1: \ X_i$ *имеет распределение такое, что или* $\mathbf{E}(X_i - a)^3 \neq 0$ *или* $(2 - \sigma^2)\mathbf{E}(X_i - a)^4 - 3\sigma^4 \neq 0$, $1 \leq i \leq n$.

Следует отметить, что предложенный критерий также может быть использован для проверки гипотезы о том, что в моделях линейной регрессии случайные возмущения имеют распределение Стьюдента. В частности, при проведении статистического анализа медицинских данных часто используется классическая линейная модель с нормальным распределением ошибок. Это объясняется простотой и удобством применения данной модели, а также наличием специализированных статистических пакетов, необходимых для выполнения расчетов. При этом проверка выполнения предпосылки о том, что случайные возмущения распределены нормально, зачастую игнорируется. В тех случаях, когда ошибки имеют распределение Стьюдента, диагностика модели будет значительно отличаться от диагностики классических линейных моделей с нормальными ошибками (см., например, [1]). Таким образом, обоснованный выбор модели при анализе реальных данных с учетом проверки распределения ошибок является важным этапом при проведении статистического анализа.

Оставшаяся часть работы состоит из четырех разделов. Во втором разделе приводятся основные результаты (предельная теорема и, как следствие из нее, критерий проверки соответствующих статистических гипотез). Третий раздел посвящен апробации построенного статистического критерия с помощью симуляций. Четвертый раздел посвящен доказательствам результатов раздела 2, пятый раздел содержит вспомогательные результаты.

Везде далее символ $\xrightarrow[n\to\infty]{d}$ будет означать сходимость по распределению. Также будем использовать обозначения $a := \mathbf{E}X$, $\sigma^2 := \mathbf{D}X$.

2 Основные результаты

Положим

$$A_{1,n} := \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^3, \quad A_{2,n} := \frac{1}{n}\sum_{i=1}^{n}\left((2-\hat\sigma^2)(X_i - \overline{X})^4 - 3\hat\sigma^4\right),$$

Теорема 1. *Пусть случайные величины $X, X_1, \ldots, X_n$, $n \in \mathbb{N}$ независимы, одинаково распределены и удовлетворяют условиям*

1) $\mathbf{E}X^8 < \infty$;

2) $\mathbf{E}(X-a)^{2k+1} = 0,\ 1 \le k \le 3$;

3) $(2-\sigma^2)\mathbf{E}(X-a)^4 - 3\sigma^4 = 0$.

Тогда последовательность случайных векторов $\sqrt{n}(A_{1,n}, A_{2,n})$ слабо сходится к случайному вектору (W_1, W_2), который имеет нормальное распределение с $\mathbf{E}W_1 = \mathbf{E}W_2 = 0$ и ковариационной матрицей

$$\Sigma = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix},$$

где

$$D_1 = \mathbf{E}\left((X-a)^3 - 3(X-a)\sigma^2\right)^2,$$

$$D_2 = \mathbf{E}((2-\sigma^2)(X-a)^4 -$$

$$-(6\sigma^2 + \mathbf{E}(X-a)^4)(X-a)^2 + \sigma^2\mathbf{E}(X-a)^4 + 3\sigma^4)^2.$$

Напрямую из теоремы 1 следует

Теорема 2. *Пусть выполнены условия теоремы 1, тогда*

$$n\left(\frac{A_{1,n}^2}{D_1} + \frac{A_{2,n}^2}{D_2}\right) \xrightarrow[n\to\infty]{d} \xi \sim \chi^2(2).$$

Заметим, что если случайная величина X имеет смещенное распределение Стьюдента с девятью и более степенями свободы (т.е. плотность распределения имеет вид $f(x) = \frac{\Gamma(\frac{k+1}{2})}{\sqrt{\pi k}\Gamma(\frac{k}{2})}(1+\frac{(x-a)^2}{k})^{-\frac{k+1}{2}}$, $k \geq 9$), то выполнены условия теоремы 1 и

$$\mathbf{E}(X-a)^4 = \frac{3\sigma^4}{2-\sigma^2}, \quad \mathbf{E}(X-a)^6 = \frac{15\sigma^6}{(2-\sigma^2)(3-2\sigma^2)},$$

$$\mathbf{E}(X-a)^8 = \frac{105\sigma^8}{(2-\sigma^2)(3-2\sigma^2)(4-3\sigma^2)}.$$

В этом случае

$$D_1 = \mathbf{E}\left((X-a)^3 - 3(X-a)\sigma^2\right)^2 = \frac{15\sigma^6}{(2-\sigma^2)(3-2\sigma^2)} - \frac{18\sigma^6}{2-\sigma^2} + 9\sigma^2,$$

$$\begin{aligned} D_2 &= \mathbf{E}\left((2-\sigma^2)(X-a)^4 - (6\sigma^2 + \mathbf{E}(X-a)^4)(X-a)^2 + \sigma^2\mathbf{E}(X-a)^4 + 3\sigma^4\right)^2 \\ &= \mathbf{E}\left((2-\sigma^2)(X-a)^4 - \frac{12\sigma^2 - 3\sigma^4}{2-\sigma^2}(X-a)^2 + \frac{6\sigma^4}{2-\sigma^2}\right)^2 \\ &= \sigma^8\left(\frac{105(2-\sigma^2)}{(3-2\sigma^2)(4-3\sigma^2)} + \frac{27\sigma^4 - 180\sigma^2 + 360}{(2-\sigma^2)^3} - \frac{90(4-\sigma^2)}{(2-\sigma^2)(3-2\sigma^2)}\right) \end{aligned}$$

и справедливо следующее следствие из теоремы 2

Следствие 1. *Пусть случайные величины $X, X_1, \ldots, X_n$, $n \in \mathbb{N}$ независимы и имеют смещенное распределение Стьюдента $T_a(k)$ с $k \geq 9$ (возможно $k = \infty$, т.е. нормальное распределение $N(a,1)$) степенями свободы. Тогда*

$$n\left(\frac{A_{1,n}^2}{B_1} + \frac{A_{2,n}^2}{B_2}\right) \xrightarrow[n\to\infty]{d} \xi \sim \chi^2(2),$$

где

$$B_1 := \hat\sigma^6\left(\frac{15}{(2-\hat\sigma^2)(3-2\hat\sigma^2)} + 9 - \frac{18}{2-\hat\sigma^2}\right),$$

$$B_2 := \hat\sigma^8\left(\frac{105(2-\hat\sigma^2)}{(3-2\hat\sigma^2)(4-3\hat\sigma^2)} + \frac{27\hat\sigma^4 - 180\hat\sigma^2 + 360}{(2-\hat\sigma^2)^3} - \frac{90(4-\hat\sigma^2)}{(2-\hat\sigma^2)(3-2\hat\sigma^2)}\right).$$

Положим

$$\widehat{JB} := n\left(\frac{A_{1,n}^2}{|B_1|} + \frac{A_{2,n}^2}{|B_2|}\right).$$

Из следствия 1 следует следующий статистический критерий.

Следствие 2. *Пусть задан уровень значимости p.*

Проверяется гипотеза

H_0 : $X_i \sim T_a(k)$, $1 \leq i \leq n$, математическое ожидание a и число степеней $9 \leq k \leq \infty$ неизвестны (здесь мы полагаем, что $T_a(\infty)$ это $N(a,1)$) против альтернативы

H_1 : X_i, $1 \leq i \leq n$ имеет другое распределение.

Из следствия 1 следует, что при достаточно большом объеме выборки можно использовать решающее правило: задаем уровень значимости p, если $\widehat{JB} < \chi^2_{1-p}(2)$, то гипотеза H_0 принимается, иначе отвергается.

Замечание 1. *Отметим, что приведенное выше решающее правило является состоятельным критерием, в случае "более простой" конкурирующей гипотезы*

H_1 : X_i *имеет распределение такое, что или* $\mathbf{E}(X_i - a)^3 \neq 0$ *или* $(2-\sigma^2)\mathbf{E}(X_i - a)^4 - 3\sigma^4 \neq 0$, $1 \leq i \leq n$.

3 Моделирование

В этом разделе мы исследуем эффективность работы построенного теста для проверки гипотезы о том, что выборка взята из несмещенного распределения Стьюдента $T(k)$. Эффективность проверялась с помощью оценки вероятности ошибки первого рода и оценки мощности при моделировании различных распределений методом Монте-Карло. Моделирование проводилось в программном обеспечении R версии 4.2.3. Мы использовали выборки следующих размеров (малые, средние и большие): n =25, 50, 75, 100, 150, 200, 250, 500 и 1000 для случая сложной основной гипотезы ($H_0 : X_i \sim T_a(k)$, где a и число степеней свободы неизвестны). Всюду при тестировании мы использовали универсальное значение уровня значимости $p = 0,05$. В таблицах ниже приведены результаты симуляций, где всюду в столбцах F указано моделируемое распределение.

В Таблице 1 моделируемое распределение соответствует основной гипотезе (напомним, что стандартное нормальное распределение, обозначаемое $N(0,1)$, является предельным распределением для распределения Стьюдента при стремлении числа степеней свободы к бесконечности). Для расчета эмпирической вероятности ошибки первого рода α мы генерировали 1000 выборок размера n, взятых из указанного в столбце F распределения. Эмпирическая вероятность ошибки первого рода — это доля отклоненных критерием основных гипотез. В Таблице 1 в некотором смысле продемонстрирована скорость сходимости к предельному распределению тестовой статистики из Теоремы 2.

n	F	α	F	α	F	α	F	α
25	$T(10)$	0,086	$T(15)$	0,085	$T(20)$	0,072	$N(0,1)$	0,088
50	$T(10)$	0,053	$T(15)$	0,081	$T(20)$	0,080	$N(0,1)$	0,074
75	$T(10)$	0,055	$T(15)$	0,070	$T(20)$	0,067	$N(0,1)$	0,091
100	$T(10)$	0,059	$T(15)$	0,061	$T(20)$	0,074	$N(0,1)$	0,072
150	$T(10)$	0,043	$T(15)$	0,074	$T(20)$	0,071	$N(0,1)$	0,071
200	$T(10)$	0,056	$T(15)$	0,067	$T(20)$	0,054	$N(0,1)$	0,064
250	$T(10)$	0,057	$T(15)$	0,059	$T(20)$	0,050	$N(0,1)$	0,058
500	$T(10)$	0,054	$T(15)$	0,061	$T(20)$	0,068	$N(0,1)$	0,051
1000	$T(10)$	0,049	$T(15)$	0,051	$T(20)$	0,052	$N(0,1)$	0,053

Таблица 1. Эмпирическое значение вероятности ошибки первого рода α, где $N(0,1)$ — стандартное нормальное распределение.

В Таблице 2 моделируемое распределение не соответствует основной гипотезе. Для расчета оценки мощности β (эмпирической мощности) мы генерировали 1000 выборок размера n, взятых из указанного в столбце F распределения, отличного от распределения Стьюдента. В этом случае, эмпирическая мощность — это доля отклоненных критерием основных гипотез. Мы рассмотрели случаи, в которых моделируемое распределение F принадлежит семейству распределений Пирсона, а также случаи, когда распределение F не относится к этому семейству.

n	F	β	F	β	F	β
25	$N(0,2)$	0,807	Mix	0,884	$U(-3;3)$	0,390
50	$N(0,2)$	0,939	Mix	0,980	$U(-3;3)$	0,636
75	$N(0,2)$	0,985	Mix	0,998	$U(-3;3)$	0,762
100	$N(0,2)$	0,999	Mix	1,000	$U(-3;3)$	0,867
150	$N(0,2)$	0,999	Mix	1,000	$U(-3;3)$	0,950
200	$N(0,2)$	1,000	Mix	1,000	$U(-3;3)$	0,986
250	$N(0,2)$	1,000	Mix	1,000	$U(-3;3)$	0,995
500	$N(0,2)$	1,000	Mix	1,000	$U(-3;3)$	1,000
1000	$N(0,2)$	1,000	Mix	1,000	$U(-3;3)$	1,000

Таблица 2. Оценка мощности β, где $N(a,\sigma)$ — нормальное распределение со средним значением a и стандартным отклонением σ, Mix — равновесная смесь распределений $N(0,1)$ и $N(0,3)$, $U(-3;3)$ — равномерное распределение на интервале $(-3;3)$.

Анализ Таблицы 2 говорит о том, что уже при малых объемах выборки построенный критерий обладает хорошей мощностью, особенно хорошо работает для альтернативных распределений $N(0,2)$ (из семейства распределения Пирсона) и равновесной смеси распределений $N(0,1)$ и $N(0,3)$ (не принадлежащей семейству распределений Пирсона). Равномерное распределение на интервале $(-3;3)$ в известном смысле также можно отнести к семейству распределений Пирсона (для него оно является предельным).

4 Доказательство теоремы 1

Д о к а з а т е л ь с т в о. Из леммы 2 следует, что

$$\lim_{n\to\infty}\left(\sqrt{n}A_{1,n}-\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}((X_i-a)^3-3(X_i-a)\sigma^2)\right)\right)=0 \text{ п.н.}, \quad (1)$$

применяя лемму 3, получаем

$$\lim_{n\to\infty}\sqrt{n}\left(\hat{\sigma}^4-\frac{2\sigma^2}{n}\sum_{i=1}^{n}(X_i-a)^2+\sigma^4\right)=0 \text{ п.н.}, \quad (2)$$

используя лемму 4, получаем

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big((2-\hat\sigma^2)(X_i-\overline{X})^4- \\ (2-\sigma^2)(X_i-a)^4+((X_i-a)^2-\sigma^2)\mathbf{E}\left((X-a)^4\right)\big)=0 \quad \text{(п.н.)} \tag{3}$$

Из формул (1)–(3) следует, что предельное распределение последовательности случайных векторов $\sqrt{n}(A_{1,n},A_{2,n})$ такое же как у последовательности

$$(A'_{1,n},A'_{2,n}):=\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}((X_i-a)^3-3(X_i-a)\sigma^2),\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big((2-\sigma^2)(X_i-a)^4\right.$$
$$\left.-(6\sigma^2+\mathbf{E}(X-a)^4)(X_i-a)^2+\sigma^2\mathbf{E}(X-a)^4+3\sigma^4\Big)\right).$$

Из условий 2), 3) следует, что

$$\mathbf{E}((X-a)^3-3(X-a)\sigma^2)=0,$$
$$\mathbf{E}((2-\sigma^2)(X_i-a)^4-(6\sigma^2+\mathbf{E}(X-a)^4)(X_i-a)^2+\sigma^2\mathbf{E}(X-a)^4+3\sigma^4)=0.$$

Поэтому, в силу центральной предельной теоремы последовательность $(A'_{1,n},A'_{2,n})$ сходится к случайному вектору (W_1,W_2), координаты которого имеют совместное нормальное распределение.

Найдем параметры этого распределения. Из условий 1), 2) следует, что

$$\mathbf{D}W_1=\mathbf{E}((X-a)^3-3(X-a)\sigma^2)^2,$$
$$\mathbf{D}W_2=\mathbf{E}((2-\sigma^2)(X_i-a)^4-(6\sigma^2+\mathbf{E}(X-a)^4)(X_i-a)^2+\sigma^2\mathbf{E}(X-a)^4+3\sigma^4)^2,$$
$$\mathbf{cov}(W_1,W_2)=\mathbf{E}\Big(((X-a)^3-3(X-a)\sigma^2)$$
$$\times((2-\sigma^2)(X_i-a)^4-(6\sigma^2+\mathbf{E}(X-a)^4)(X_i-a)^2+\sigma^2\mathbf{E}(X-a)^4+3\sigma^4)\Big)=0.$$

5 Вспомогательные результаты

Лемма 1. *Пусть выполнены условия*

1) $\mathbf{E}X^{4k-2}<\infty$ *для некоторого фиксированного* $k\in\mathbb{N}$*;*

2) $\mathbf{E}(X-a)^{2r-1}=0$ *для всех* $1\le r\le k$*.*

Тогда

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}((X_i-\overline{X})^{2k}-(X_i-a)^{2k})=0 \text{ п.н.}$$

Д о к а з а т е л ь с т в о. Легко видеть, что

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}((X_i-\overline{X})^{2k}-(X_i-a)^{2k})$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\left(X_i - a - \frac{1}{n} \sum_{j=1}^{n} (X_j - a) \right)^{2k} - (X_i - a)^{2k} \right)$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{r=0}^{2k-1} C_{2k}^{r} (X_i - a)^r \left(\frac{-1}{n} \sum_{j=1}^{n} (X_j - a) \right)^{2k-r}$$

$$= \left(\frac{1}{n^{1-\frac{1}{4k}}} \sum_{j=1}^{n} (X_j - a) \right)^{2k}$$

$$+ \sum_{r=1}^{2k-1} C_{2k}^{r} \left(\frac{1}{\sqrt{n} n^{\frac{2k-r}{3}}} \sum_{i=1}^{n} (X_i - a)^r \right) \left(\frac{-1}{n^{\frac{2}{3}}} \sum_{j=1}^{n} (X_j - a) \right)^{2k-r}.$$

Из условий 1), 2) и закона повторного логарифма следует, что случайная последовательность находящаяся в первой скобке, сходятся при $n \to \infty$ к нулю п.н. ействительно, из условий 1), 2) следует, что

$$\mathbf{E}(X_j - a)^2 < \infty, \ \mathbf{E}(X_j - a) = 0,$$

значит, в силу закона повторного логарифма справедливо равенство

$$\limsup_{n\to\infty} \frac{1}{\sqrt{2n \ln(\ln n) \cdot \mathbf{E}(X_j - a)^2}} \left| \sum_{j=1}^{n} (X_j - a) \right| = 1 \text{ п.н.} \tag{4}$$

Поэтому, при любом фиксированном $k \in \mathbb{N}$ для последовательности расположенной в первой скобке будем иметь

$$\lim_{n\to\infty} \frac{1}{n^{1-\frac{1}{4k}}} \left| \sum_{j=1}^{n} (X_j - a) \right| \le \limsup_{n\to\infty} \frac{1}{\sqrt{2n \ln(\ln n) \cdot \mathbf{E}(X_j - a)^2}} \left| \sum_{j=1}^{n} (X_j - a) \right|$$

$$\times \limsup_{n\to\infty} \frac{\sqrt{2 \ln(\ln n) \cdot \mathbf{E}(X_j - a)^2}}{n^{\frac{1}{2}-\frac{1}{4k}}} = 1 \cdot 0 = 0 \text{ п.н.} \tag{5}$$

Аналогичными рассуждениями можно показать, что последовательность находящаяся в третей скобке также сходится при $n \to \infty$ к нулю п.н.

Из условий 1), 2) следует, что

$$\mathbf{E}(X_i - a)^{4k-2} < \infty, \ \mathbf{E}(X_i - a)^{2k-1} = 0.$$

Значит, используя закон повторного логарифма и рассуждения, аналогичные приведенным в формулах (4), (5) можем заключить, что при $r = 2k - 1$ последовательность стоящая во второй скобке сходится к нулю п.н. при $n \to \infty$. Для $1 \le r \le 2k - 2$ выполнено неравенство $\sqrt{n} n^{\frac{2k-r}{3}} \ge n^{\frac{7}{6}}$. Поэтому, в силу усиленного закона больших чисел, случайная последовательность расположенная во второй скобке также сходится при $n \to \infty$ к нулю п.н. для таких r.

Лемма 2. *Пусть выполнены условия*

1) $\mathbf{E}X^{4k} < \infty$ *для некоторого фиксированного* $k \in \mathbb{N}$*;*

2) $\mathbf{E}(X-a)^{2r-1} = 0$ *для всех* $1 \le r \le k$*.*

Тогда

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} ((X_i - \overline{X})^{2k+1} - (X_i - a)^{2k+1} + (2k+1)(X_i - a)\mathbf{E}(X-a)^{2k}) = 0 \text{ п.н.}$$

Д о к а з а т е л ь с т в о. Легко видеть, что

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} ((X_i - \overline{X})^{2k+1} - (X_i - a)^{2k+1} + (2k+1)(X_i - a)\mathbf{E}(X-a)^{2k})$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\left(X_i - a - \frac{1}{n}\sum_{j=1}^{n}(X_j - a) \right)^{2k+1} \right.$$

$$\left. -(X_i - a)^{2k+1} + (2k+1)(X_i - a)\mathbf{E}(X-a)^{2k} \right)$$

$$= (2k+1)\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - a)^{2k} \left(\frac{-1}{n} \sum_{j=1}^{n}(X_j - a) \right)$$

$$+ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{r=0}^{2k-1} C_{2k+1}^{r} (X_i - a)^{r} \left(\frac{-1}{n} \sum_{j=1}^{n}(X_j - a) \right)^{2k-r+1}$$

$$+ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (2k+1)(X_i - a)\mathbf{E}(X-a)^{2k}$$

$$= \left(\frac{-1}{n^{1-\frac{1}{4k+2}}} \sum_{j=1}^{n}(X_j - a) \right)^{2k+1}$$

$$+ \sum_{r=1}^{2k-1} C_{2k+1}^{r} \left(\frac{1}{\sqrt{n} n^{\frac{2k-r+1}{3}}} \sum_{i=1}^{n} (X_i - a)^{r} \right) \left(\frac{-1}{n^{\frac{2}{3}}} \sum_{j=1}^{n}(X_j - a) \right)^{2k-r+1}$$

$$-(2k+1) \left(\frac{1}{\sqrt{n}} \sum_{j=1}^{n}(X_j - a) \right) \left(\frac{1}{n} \sum_{i=1}^{n}(X_i - a)^{2k} \right)$$

$$+ \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (2k+1)(X_j - a)\mathbf{E}(X-a)^{2k}.$$

Первых два слагаемых сходятся при $n \to \infty$ к нулю п.н. по причинам полностью аналогичным тем, которые приведены в конце доказательства леммы 1. Запишем третье и четвертое слагаемое в следующем виде

$$(2k+1)\left(\frac{1}{\sqrt{n}}\sum_{j=1}^{n}(X_j-a)\right)\left(\frac{1}{n}\sum_{i=1}^{n}(\mathbf{E}(X-a)^{2k}-(X_i-a)^{2k})\right)$$

$$=(2k+1)\left(\frac{1}{n^{\frac{3}{4}}}\sum_{j=1}^{n}(X_j-a)\right)\left(\frac{1}{n^{\frac{3}{4}}}\sum_{i=1}^{n}(\mathbf{E}(X-a)^{2k}-(X_i-a)^{2k})\right).$$

Из условий 1), 2) следует, что

$$\mathbf{E}(X_j-a)^2<\infty,\ \mathbf{E}(X_j-a)=0,\ \mathbf{E}(\mathbf{E}(X-a)^{2k}-(X_i-a)^{2k})^2<\infty.$$

Поэтому, используя закон повторного логарифма и рассуждения, аналогичные приведенным в формулах (4),(5), можем заключить, что все находящиеся в скобках последовательности сходятся при $n \to \infty$ к нулю п.н.

Лемма 3. *Пусть* $\mathbf{E}X^4<\infty$*, тогда для любого* $k\in\mathbb{N}$

$$\lim_{n\to\infty}\sqrt{n}\left(\hat{\sigma}^{2k}-\frac{\sigma^{2k-2}k}{n}\sum_{i=1}^{n}(X_i-a)^2+(k-1)\sigma^{2k}\right)=0 \ \textit{п.н.}$$

Д о к а з а т е л ь с т в о. Положим

$$Y_n:=\sqrt{n}\hat{\sigma}^{2k}-\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-a)^2\right)^k.$$

Будем иметь

$$Y_n=\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\right)^k-\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-a)^2\right)^k$$

$$=\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}((X_i-\overline{X})^2-(X_i-a)^2)\right)$$

$$\times\sum_{r=0}^{k-1}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\right)^{k-1-r}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-a)^2\right)^r=:g_{1,n}\sum_{r=0}^{k-1}f_{k,r,n}.$$

Из леммы 1 и усиленного закона больших чисел следует, что случайные последовательности $g_{1,n}$ и $f_{k,r,n}$, $0\le r\le k-1$ сходятся п.н. к 0 и $\sigma^{2(k-1)}$, соответственно. Поэтому

$$\lim_{n\to\infty}Y_n=0 \text{ п.н.}, \tag{6}$$

и нам достаточно показать, что

$$\lim_{n\to\infty}\sqrt{n}\left(\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-a)^2\right)^k-\frac{\sigma^{2k-2}k}{n}\sum_{i=1}^{n}(X_i-a)^2+(k-1)\sigma^{2k}\right)=0 \text{ п.н.} \tag{7}$$

Легко видеть, что

$$\sqrt{n}\left(\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-a)^2\right)^k-\frac{\sigma^{2k-2}k}{n}\sum_{i=1}^{n}(X_i-a)^2+(k-1)\sigma^{2k}\right)$$

$$=\sqrt{n}\left(\left(\frac{1}{n}\sum_{i=1}^{n}((X_i-a)^2-\sigma^2)+\sigma^2\right)^k-\frac{\sigma^{2k-2}k}{n}\sum_{i=1}^{n}(X_i-a)^2+(k-1)\sigma^{2k}\right)$$

$$=\sqrt{n}\sum_{r=0}^{k-2}C_k^r\left(\frac{1}{n}\sum_{i=1}^{n}((X_i-a)^2-\sigma^2)\right)^{k-r}\cdot\sigma^{2r}$$

$$+\sqrt{n}k\left(\frac{1}{n}\sum_{i=1}^{n}((X_i-a)^2-\sigma^2)\right)\sigma^{2(k-1)}$$

$$+\sqrt{n}\sigma^{2k}-\sqrt{n}\frac{\sigma^{2k-2}k}{n}\sum_{i=1}^{n}(X_i-a)^2+\sqrt{n}(k-1)\sigma^{2k}$$

$$=\sum_{r=0}^{k-2}C_k^r\left(\frac{1}{n^{1-\frac{1}{2(k-r)}}}\sum_{i=1}^{n}((X_i-a)^2-\sigma^2)\right)^{k-r}\cdot\sigma^{2r}.$$

Из условия $\mathbf{E}X^4<\infty$ следует, что

$$\mathbf{E}((X_i-a)^2-\sigma^2)^2<\infty.$$

Поэтому, используя закон повторного логарифма и рассуждения, аналогичные приведенным в формулах (4),(5), можем заключить, что случайные последовательности расположенные в скобках сходятся при $n\to\infty$ к нулю п.н. для всех $0\le r\le k-2$. Следовательно, справедлива формула (7).

Лемма 4. *Пусть* $\mathbf{E}X^8<\infty$, *тогда*

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\big((2-\hat{\sigma}^2)(X_i-\overline{X})^4$$

$$-(2-\sigma^2)(X_i-a)^4+((X_i-a)^2-\sigma^2)\mathbf{E}(X-a)^4\big)=0 \text{ п.н.}$$

Д о к а з а т е л ь с т в о. Легко видеть, что

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(2-\hat{\sigma}^2)(X_i-\overline{X})^4=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(2-\frac{1}{n}\sum_{j=1}^{n}((X_j-\overline{X})^2-\sigma^2)-\sigma^2\right)(X_i-\overline{X})^4$$

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(2-\hat{\sigma}^2)(X_i-\overline{X})^4$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(2 - \frac{1}{n} \sum_{j=1}^{n} ((X_j - \overline{X})^2 \pm (X_j - a)^2 \pm \sigma^2) \right) (X_i - \overline{X})^4$$

$$= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (2 - \sigma^2)(X_i - \overline{X})^4 - \frac{1}{\sqrt{n}} \sum_{j=1}^{n} ((X_j - \overline{X})^2 - (X_j - a)^2) \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^4$$

$$- \frac{1}{\sqrt{n}} \sum_{j=1}^{n} ((X_j - a)^2 - \sigma^2) \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^4 = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (2 - \sigma^2)(X_i - \overline{X})^4$$

$$- \frac{1}{\sqrt{n}} \sum_{j=1}^{n} ((X_j - \overline{X})^2 - (X_j - a)^2) \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^4$$

$$- \frac{1}{n} \sum_{j=1}^{n} ((X_j - a)^2 - \sigma^2) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} ((X_i - \overline{X})^4 - (X_i - a)^4)$$

$$- \frac{1}{n^{\frac{3}{4}}} \sum_{j=1}^{n} ((X_j - a)^2 - \sigma^2) \frac{1}{n^{\frac{3}{4}}} \sum_{i=1}^{n} ((X_i - a)^4 - \mathbf{E}(X - a)^4)$$

$$- \frac{1}{\sqrt{n}} \sum_{j=1}^{n} ((X_j - a)^2 - \sigma^2) \mathbf{E}(X - a)^4. \tag{8}$$

Из условия $\mathbf{E}X^8 < \infty$ и леммы 1 следует, что

$$\lim_{n \to \infty} \frac{2 - \sigma^2}{\sqrt{n}} \sum_{i=1}^{n} ((X_i - \overline{X})^4 - (X_i - a)^4) = 0 \text{ п.н.}, \tag{9}$$

используя лемму 1 и усиленный закон больших чисел, получаем

$$\lim_{n \to \infty} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} ((X_j - \overline{X})^2 - (X_j - a)^2) \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^4 = 0 \text{ п.н.}, \tag{10}$$

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=1}^{n} ((X_j - a)^2 - \sigma^2) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} ((X_i - \overline{X})^4 - (X_i - a)^4) = 0 \text{ п.н.}, \tag{11}$$

применяя условие $\mathbf{E}X^8 < \infty$ и закон повторного логарифма, получаем

$$\lim_{n \to \infty} \frac{1}{n^{\frac{3}{4}}} \sum_{j=1}^{n} ((X_j - a)^2 - \sigma^2) \frac{1}{n^{\frac{3}{4}}} \sum_{i=1}^{n} ((X_i - a)^4 - \mathbf{E}(X - a)^4) = 0 \text{ п.н.} \tag{12}$$

Утверждение леммы 4 следует из (8)–(12).

Авторы благодарят рецензента за полезные замечания и рекомендации, которые помогли улучшить настоящую работу.

References

[1] F. Cysneiros, G. Paula, *Restricted methods in symmetrical linear regression models*, Comput. Stat. Data Anal., **49**:3 (2005), 689–708. Zbl 1429.62285

[2] C.M. Jarque, A.K. Bera, *Efficient tests for normality, homoscedasticity and serial independence of regression residuals*, Econom. Lett., **6**:3 (1980), 255–259.

[3] C.M. Jarque, A.K. Bera, *Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence*, Econom. Lett., **7**:4 (1981), 313–318.

[4] C.M. Jarque, A.K. Bera, *A test for normality of observations and regression residuals*, Int. Stat. Rev., **55**:2 (1987), 163–172. Zbl 0616.62092

[5] K. Pearson, *Mathematical contributions to the theory of evolution. Second supp'lement to a memoir on skew variation*, Lond. Phil. Trans. (A), **216** (1916), 429–457. JFM 46.1495.08

[6] V. Glinskiy, Y. Ismayilova, S. Khrushchev, A. Logachov, O. Logachova, L. Serga, A. Yambartsev, K. Zaykov, *Modifications to the Jarque–Bera Test*, Mathematics, **12**:16 (2024), Paper No. 2523.

Artem Vasilhevich Logachov
Dep. of Computer Science in Economics, Novosibirsk State Technical University, pr. K. Marksa, 20,
630073, Novosibirsk, Russia
Email address: omboldovskaya@mail.ru

Sergey Evgenievich Khrushchev
Lab. of Applied Inverse Problems, Sobolev Institute of Mathematics, pr. Koptyuga, 4,
630090, Novosibirsk, Russia
Email address: hrushew@rambler.ru