Math-Net.Ru

Общероссийский математический портал

А. П. Ковалевский, Е. В. Шаталин, Асимптотика сумм остатков однопараметрической линейной регрессии, построенной по порядковым статистикам, *Теория вероятн. и ее примен.*, 2014, том 59, выпуск 3, 452–467

DOI: <https://doi.org/10.4213/tvp4579>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 52.6.47.48

26 апреля 2023 г., 15:27:17

© 2014 г. КОВАЛЕВСКИЙ А. П.^{*}, ШАТАЛИН Е. В.^{**}

**АСИМПТОТИКА СУММ ОСТАТКОВ ОДНОПАРАМЕТРИЧЕСКОЙ
ЛИНЕЙНОЙ РЕГРЕССИИ, ПОСТРОЕННОЙ ПО ПОРЯДКОВЫМ
СТАТИСТИКАМ¹⁾**

В однопараметрической регрессионной модели предполагается, что регрессор составлен из порядковых статистик, построенных по выборке, не зависящей от регрессионных ошибок. По остаткам регрессии строится эмпирический мост и доказывается его слабая сходимость к центрированному гауссовскому процессу. Изучаются статистические критерии, основанные на доказанной сходимости.

Ключевые слова и фразы: порядковые статистики, эмпирический мост, регрессионные остатки, линейная регрессия.

1. Введение и основной результат. Наиболее простая (однопараметрическая) модель линейной регрессии имеет вид

$$Y_i = \theta X_i + \varepsilon_i, \quad i = 1, \dots, n,$$

где X_i (значения регрессора) обычно предполагаются фиксированными (неслучайными) величинами, $\theta \in \mathbf{R}$ — неизвестный параметр регрессии, $\varepsilon_1, \dots, \varepsilon_n$ (регрессионные ошибки) — независимые одинаково распределенные случайные величины с нулевым математическим ожиданием и конечной ненулевой дисперсией σ^2 . Отличие модели, рассматриваемой в настоящей работе, состоит в том, что в качестве X_i выступают случайные величины, а именно порядковые статистики, построенные по набору $\xi_1, \dots, \xi_n$, т.е. $X_i = \xi_{i:n}$. Случайные величины $\xi_1, \dots, \xi_n$ предполагаются независимыми, одинаково распределенными с функцией распределения F и не зависящими от случайных величин $\varepsilon_1, \dots, \varepsilon_n$. Заметим, что независимость $\xi_1, \dots, \xi_n$ и $\varepsilon_1, \dots, \varepsilon_n$ не является само собой разумеющейся и

^{*}Новосибирский государственный технический университет, Новосибирск, Россия; e-mail: pandorra@ngs.ru

^{**}Институт математики им. С. Л. Соболева СО РАН, Новосибирск, Россия; e-mail: sh.e.v.89@list.ru

¹⁾ Работа первого автора выполнена при поддержке РФФИ (проект № 13-01-00661) и Программы стратегического развития НГТУ (проект № 2.1.1-2013). Работа второго автора выполнена при поддержке Правительства Новосибирской области (стипендия Правительства НСО).

на практике нуждается в проверке: в предположениях совместной нормальности достаточно проверять гипотезу о равенстве коэффициента корреляции нулю, а в общем случае может быть применен критерий независимости типа хи-квадрат.

Предложенная в работе модель возникает всякий раз, когда анализируется двумерная выборка и предполагается линейная зависимость одной ее компоненты (отклика) от другой (регрессора) с точностью до случайной регрессионной ошибки. Упорядочение значений регрессора по возрастанию естественно возникает при графическом изображении регрессионной зависимости. Например, для исследования зависимости массы тела человека W_i от роста H_i в [1] предложена модель пропорциональности массы тела квадрату роста. Отметим, что пропорциональность предполагается для лиц одной возрастной группы и одного пола. В [2] на основании масштабных исследований показано, что эта зависимость является наилучшей в классе степенных зависимостей. Эту пропорциональность можно проинтерпретировать в виде двух различных регрессионных моделей: $\ln W_i = \ln(\theta H_i^2) + \varepsilon_i$ и $W_i = \theta H_i^2 + \varepsilon_i$. Отметим, что рост и массу тела индивидуума можно считать случайными величинами. При этом, как было отмечено выше, необходимо проверить независимость роста H_i и корректирующего фактора ε_i .

Первая из этих моделей после замены переменных $Y_i = \ln(W_i/H_i^2)$, $\alpha = \ln \theta$ приводит к модели выборки $Y_i = \alpha + \varepsilon_i$, а вторая является моделью однопараметрической линейной регрессии. Для того чтобы проверить соответствие каждой модели реальным данным, предлагается упорядочить наблюдения по неубыванию величин H_i .

Другой пример — модель радиоактивного распада $C_i = C \exp(-\alpha T_i + \varepsilon_i)$ в радиоуглеродном анализе археологических памятников (см. [3]). Здесь C — начальная концентрация изотопа (известная и одинаковая для всех образцов), C_i — концентрация в образце, возраст которого предполагается равным T_i . Логарифмируя, приходим к рассматриваемой в работе модели. Здесь также образцы упорядочиваются по возрасту для проверки гипотезы о том, что возраст T_i в каждом случае определен верно.

Если модель неправильно описывает данные, то значения Y_i будут систематически уклоняться от регрессионной кривой, и это отклонение можно выявить суммированием регрессионных остатков (разностей между наблюдаемыми и прогнозируемыми значениями). Для изучения значимости этих отклонений необходимо знать предельное распределение процесса центрированных и нормированных частичных сумм регрессионных остатков. Этот процесс называется эмпирическим мостом и определен ниже. Если для описания данных предложено несколько моделей (как в примере с зависимостью массы тела от роста), то вычисление достигнутых уровней значимости позволяет выбрать модель, наилуч-

шим образом описывающую данные. Эмпирический мост для модели выборки слабо сходится к стандартному броуновскому мосту, а сходимость эмпирического моста в модели однопараметрической линейной регрессии требует доказательства. При этом, как будет доказано ниже, предельный гауссовский процесс отличается от стандартного броуновского моста.

Неизвестный параметр регрессионной модели θ обычно оценивают по методу наименьших квадратов, получая оценку $\hat{\theta} = \overline{XY}/\overline{X^2}$, где $\overline{V} = (1/n) \sum_{i=1}^n V_i$ обозначает выборочное среднее случайных величин $V_1, \dots, V_n$. Эта оценка совпадает с оценкой максимального правдоподобия в случае нормальной регрессии, т.е. в случае, когда ε_1 имеет нормальное распределение. На основании регрессионной модели строятся прогнозные значения $\hat{Y}_i = \hat{\theta}X_i$. Остатками линейной регрессии называют случайные величины $\hat{\varepsilon}_i = Y_i - \hat{Y}_i$.

Приведем определение эмпирического моста, а также некоторые связанные с ним факты. Эмпирический мост — это кусочно-линейная случайная ломаная $Z_n^0 = \{Z_n^0(t), 0 \leq t \leq 1\}$ с узлами в точках

$$\left(\frac{k}{n}, \frac{\hat{\Delta}_k - (k/n)\hat{\Delta}_n}{\sqrt{\hat{\sigma}^2 n}}\right),$$

где $\hat{\Delta}_k = \hat{\varepsilon}_1 + \dots + \hat{\varepsilon}_k$, $k = 1, \dots, n$, $\hat{\Delta}_0 = 0$, $\hat{\sigma}^2 = \overline{\hat{\varepsilon}^2} - (\overline{\hat{\varepsilon}})^2$.

При условии сходимости $\hat{\sigma}^2 \xrightarrow{P} \sigma^2$ при $n \rightarrow \infty$ слабые пределы в $C(0, 1)$ (пространстве непрерывных на $[0, 1]$ функций) эмпирического моста и случайной ломаной $\hat{Z}_n^0$, построенной по точкам

$$\left(\frac{k}{n}, \frac{\hat{\Delta}_k - (k/n)\hat{\Delta}_n}{\sigma\sqrt{n}}\right),$$

совпадают.

Отметим, что $\hat{Z}_n^0$ получается непрерывным в равномерной метрике на $[0, 1]$ преобразованием $x^0(t) = x(t) - tx(1)$, отображающим x в x^0 , случайной ломаной $\hat{Z}_n$, построенной по точкам

$$\left(\frac{k}{n}, \frac{\hat{\Delta}_k}{\sigma\sqrt{n}}\right).$$

Отыскание предельного процесса для $\hat{Z}_n$ в случае, когда в качестве регрессора используется набор $\{\xi_{1:n}, \dots, \xi_{n:n}\}$ порядковых статистик, построенных по выборке из некоторого распределения, является содержанием данной работы.

В случае, когда в качестве регрессора используется вектор значений неслучайной гладкой функции в равноотстоящие моменты времени, предельный процесс для $\hat{Z}_n$ изучен в [4]. Это центрированный гауссовский

процесс с ковариационной функцией

$$K_f(s, t) = \min \{s, t\} - \int_0^s \int_0^t g(x, y) dx dy, \quad s, t \in [0, 1],$$

где функция $g(x, y)$ определяется через регрессор.

В работе [5] рассматривается многопараметрическая регрессионная модель, в которой в качестве регрессора выступают векторы вида $(1, i/n, \dots, (i/n)^p)^T$, т.е. значения степенных функций в равноотстоящие моменты времени. Для этой модели проверяется нулевая гипотеза, состоящая в том, что вектор параметров регрессии не зависит от времени и равен β_0 . В качестве альтернативы рассматривается гипотеза, состоящая в том, что до некоторого неизвестного момента времени k^* параметр равен β_0 , а в момент времени k^* происходит изменение параметра, и все оставшееся время он равняется β_A . Проверка описанной гипотезы строится на основе критерия, использующего статистику, построенную на основе сумм квадратов регрессионных остатков.

Задача проверки нормальности регрессионных ошибок решена в [6]. Здесь анализ регрессионных остатков приводит к другому предельному гауссовскому процессу. В статье показано, что распределение статистики омега-квадрат от этого предельного процесса вычисляется в явном виде.

Перейдем теперь к рассмотрению основного результата данной статьи. Всюду далее предполагается, что $X_i = \xi_{i:n}$, $i = 1, \dots, n$, — порядковые статистики, где случайные величины $\xi_1, \dots, \xi_n$ независимы, одинаково распределены с функцией распределения F и не зависят от случайных величин $\varepsilon_1, \dots, \varepsilon_n$.

Пусть $GL_n(t)$ — случайная кусочно-линейная кривая, построенная по точкам $(k/n, S_k/n)$, где $S_k = \sum_{i=1}^k \xi_{i:n}$, $k = 1, \dots, n$, $S_0 = 0$. Эта случайная ломаная называется *эмпирической обобщенной кривой Лоренца* (см. [7]). Обозначим через $GL_n^0(t) = GL_n(t) - tGL_n(1)$ случайную ломаную, построенную по точкам

$$\left(\frac{k}{n}, \frac{nS_k - kS_n}{n^2} \right).$$

В работе [8] была доказана сходимость $\|GL_n - GL_F\|_\infty \rightarrow 0$ п.н., где $GL_F(t) = \int_0^t F^{-1}(s) ds$ — *теоретическая обобщенная кривая Лоренца* (см. [7]), $\|h\|_\infty = \sup_x |h(x)|$ — супремум-норма в $C(0, 1)$ заданной функции $h(x)$, $F^{-1}(s) = \sup\{x: F(x) < s\}$ — квантильное преобразование (обобщенная обратная функция) функции распределения $F(x)$.

Тогда в силу непрерывности в равномерной метрике в $C(0, 1)$ преобразования, отображающего x в x^0 , где $x^0(t) = x(t) - tx(1)$, получаем (см. [9, § 5]) $\|GL_n^0 - GL_F^0\|_\infty \rightarrow 0$ п.н., где $GL_F^0(t) = GL_F(t) - tGL_F(1)$.

Символом $\Rightarrow$ будем обозначать слабую сходимость (сходимость по распределению) в соответствующем пространстве. Так, в теореме ниже $\Rightarrow$ обозначает слабую сходимость в пространстве $C(0, 1)$, снабженном равномерной метрикой (см. [9, гл. 2, § 8, с. 82]).

Теорема. Если $0 < \mathbf{E} \xi_1^2 < \infty$, то справедливы следующие утверждения:

1) $\widehat{Z}_n \Rightarrow \widehat{Z}_F$, где $\widehat{Z}_F$ — центрированный гауссовский процесс с ковариационной функцией

$$K_F(t, s) = \min\{t, s\} - \frac{GL_F(s)GL_F(t)}{\mathbf{E} \xi_1^2}, \quad s, t \in [0, 1];$$

2) $Z_n^0 \Rightarrow \widehat{Z}_F^0$, где $\widehat{Z}_F^0$ — центрированный гауссовский процесс с ковариационной функцией

$$K_F^0(t, s) = \min\{t, s\} - ts - \frac{GL_F^0(s)GL_F^0(t)}{\mathbf{E} \xi_1^2}, \quad s, t \in [0, 1].$$

В п. 2 работы изучены статистические критерии типа хи-квадрат и омега-квадрат, основанные на доказанной в теореме сходимости. Эти критерии позволяют на практике проверять соответствие однопараметрической регрессионной модели реальным данным. Там же проводится сравнение предлагаемого статистического критерия с известным F -тестом, а также обсуждаются другие критерии, касающиеся проверки регрессионных моделей. В п. 3 проводится доказательство теоремы, которому предшествуют формулировки и доказательства вспомогательных утверждений. Доказательства используют известные утверждения об асимптотике порядковых статистик и равномерной интегрируемости. Эти утверждения приведены в приложении (п. 4).

2. Обсуждение основного результата. Для того чтобы применить основной результат работы к анализу соответствия модели исследуемым данным, нужны алгоритм оценивания ковариационной функции и построенная на его основе статистика, распределение которой при выполнении основной гипотезы сходится к известному распределению.

Мы построим статистику, слабо сходящуюся к распределению хи-квадрат с произвольным наперед заданным числом степеней свободы d . Однако критерий, построенный на ее основе, не является состоятельным при достаточно широком классе альтернатив. Для построения состоятельного критерия будем строить критерий типа омега-квадрат, предельное распределение для которого удастся вычислить в ряде частных случаев.

Формулы для ковариационной функции в теореме включают неизвестные функции $GL_F(t)$ и $GL_F^0(t)$. При практическом применении эти

функции необходимо заменить на их эмпирические аналоги $GL_n(t)$ и $GL_n^0(t)$.

В качестве оценки ковариационной функции выберем

$$\widetilde{K}_F^0(t, s) = \min\{t, s\} - ts - \frac{GL_n^0(s)GL_n^0(t)}{X^2}, \quad s, t \in [0, 1].$$

Тогда имеет место равномерная по $s, t \in [0, 1]$ сходимость

$$|\widetilde{K}_F^0(t, s) - K_F^0(t, s)| \rightarrow 0 \quad \text{п.н.}$$

(см. [7]).

Пусть $d > 0$ — целое число. Обозначим

$$\vec{q}_n = \left(Z_n^0\left(\frac{1}{d+1}\right), \dots, Z_n^0\left(\frac{d}{d+1}\right) \right),$$

тогда

$$\vec{q}_n \Rightarrow \vec{q}_F = \left(\widehat{Z}_F^0\left(\frac{1}{d+1}\right), \dots, \widehat{Z}_F^0\left(\frac{d}{d+1}\right) \right).$$

Пусть A обозначает ковариационную матрицу вектора $\vec{q}_F$. Если $\det A \neq 0$, то, как известно, $q_n A^{-1} q_n^T \Rightarrow \chi_d^2$ при $n \rightarrow \infty$, где χ_d^2 — распределение хи-квадрат с d степенями свободы.

Отметим, что сходимость будет иметь место и при замене матрицы A на матрицу $\tilde{A}_n = (\tilde{a}_{ij})_{i,j=1}^d$, элементы которой получены заменой элементов матрицы A на их оценки:

$$\tilde{a}_{ij} = \min \left\{ \frac{i}{d+1}, \frac{j}{d+1} \right\} - \frac{ij}{(d+1)^2} - \frac{GL_n^0(i/(d+1))GL_n^0(j/(d+1))}{X^2}.$$

В самом деле, так как $\vec{q}_n \Rightarrow \vec{q}_F$, а $\tilde{A}_n \rightarrow A$ п.н., то $q_n(\tilde{A}_n)^{-1} q_n^T \Rightarrow \chi_d^2$. Таким образом, мы получили следующее утверждение.

Следствие 1. Если $0 < \mathbf{E} \xi_1^2 < \infty$, $\det A \neq 0$, то $q_n(\tilde{A}_n)^{-1} q_n^T$ сходится слабо к случайной величине, имеющей распределение хи-квадрат с d степенями свободы.

Отметим отличие предлагаемого подхода от известного F -теста проверки гипотезы о регрессионной модели (см. [10, гл. 3, § 3.7, с. 109]). F -тест использует коэффициент детерминации R^2 — долю объясненной выборочной дисперсии: если эта доля достаточно велика, то модель принимается. Предлагаемый в настоящей работе подход позволяет отвергнуть модель на основании анализа последовательных сумм регрессионных остатков: модель может отвергаться при сколь угодно близком к единице значении R^2 . Это показывает следующий пример, в котором каждое значение отклика отличается от исходной модели на некоторую величину $c \neq 0$.

Пример 1. Пусть последовательности $\{\xi_i\}$ и $\{\varepsilon_i\}$ удовлетворяют стохастическим предположениям предыдущего пункта, причем ξ_1 имеет непрерывное распределение с медианой q , т.е. $F(q) = 1/2$. Пусть

$$h(t) = \begin{cases} \theta t - c, & t < q; \\ \theta t + c, & t \geq q, \end{cases}$$

и $Y_i = h(\xi_{i:n}) + \varepsilon_i$, $i = 1, \dots, n$. По-прежнему $\hat{Y}_i = \hat{\theta}\xi_{i:n}$. Имеет место сходимость $\bar{Y} \rightarrow \theta \mathbf{E}\xi_1$ п.н. Отметим, что $GL_F(1) = \mathbf{E}\xi_1$. Выборочная дисперсия

$$\frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2$$

в силу УЗБЧ сходится п.н. к величине

$$\begin{aligned} c^2 + \sigma^2 + \theta^2 \mathbf{D}\xi_1 - 2c\theta \left(GL_F\left(\frac{1}{2}\right) - \frac{\mathbf{E}\xi_1}{2} \right) \\ + 2c\theta \left(GL_F(1) - GL_F\left(\frac{1}{2}\right) - \frac{\mathbf{E}\xi_1}{2} \right) = c^2 + \sigma^2 + \theta^2 \mathbf{D}\xi_1 - 4c\theta GL_F^0\left(\frac{1}{2}\right). \end{aligned}$$

Необъясненная выборочная дисперсия

$$\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

сходится по вероятности к $c^2 + \sigma^2$ (доказательство следует из леммы 5, приведенной в п. 3).

В силу этого коэффициент детерминации R^2 при $n \rightarrow \infty$ сходится по вероятности к величине

$$1 - \frac{c^2 + \sigma^2}{c^2 + \sigma^2 + \theta^2 \mathbf{D}\xi_1 - 4c\theta GL_F^0(1/2)}$$

и может быть сделан сколь угодно близким к 1 выбором соответствующих параметров. В то же время последовательные суммы остатков в этой модели растут асимптотически линейно, и $q_n(\tilde{A}_n)^{-1} q_n^T \xrightarrow{P} +\infty$ для любого $d > 0$ при условии $\det A \neq 0$.

З а м е ч а н и е. Отметим, что сходимость к распределению хи-квадрат доказана без каких-либо дополнительных предположений о распределении F по сравнению с основной теоремой. Более того, для доказательства следствия используется только часть теоремы, касающаяся сходимости конечномерных распределений, установленной в лемме 3. Но критерий, основанный на предлагаемой статистике, использует лишь значения эмпирического моста в конечном числе точек и является несостоятельным. Построение состоятельных критериев наталкивается на

проблему нахождения предельного распределения при справедливости основной гипотезы. Наиболее разработан вопрос о предельном распределении для критериев типа омега-квадрат. Итак, будем использовать критерий, основанный на статистике

$$\omega_n^2 = \int_0^1 (Z_n^0(x))^2 dx.$$

Эта статистика слабо сходится к $\omega_F^2 = \int_0^1 (\widehat{Z}_F^0(x))^2 dx$, где гауссовский процесс $\widehat{Z}_F^0$ определен в формулировке теоремы.

Будем предполагать, что известна функция распределения F с точностью до параметра масштаба: $F(x) = G(x/\sigma)$, где $\sigma > 0$, G — известная функция распределения с нулевым математическим ожиданием и дисперсией, равной 1. Тогда $GL_G(1) = 0$, и ковариационная функция предельного гауссовского процесса полностью известна, так как

$$GL_F^0(x) = \int_0^x \sigma G^{-1}(t) dt - \sigma x GL_G(1) = \sigma GL_G(x),$$

$$K_F^0(s, t) = \min\{s, t\} - st - GL_G(s)GL_G(t).$$

В [11, теорема 3.2] доказано, что если функция $A(s, t) = 2(K_F^0(s, t))^2$ имеет ограниченные частные производные 4-го порядка на единичном квадрате за вычетом диагонали, то распределение ω_F^2 в этом случае можно вычислить приближенно, заменяя интеграл суммой

$$\omega_F^{2,m} = \frac{1}{m} \sum_{i=1}^m \left(\widehat{Z}_F^0 \left(\frac{i-1/2}{m} \right) \right)^2,$$

при этом дисперсия погрешности аппроксимации эквивалентна cm^{-2} при $m \rightarrow \infty$, где константа c отыскивается в явном виде. В [11, § 3.3] разработан эффективный алгоритм вычисления распределения случайной суммы $\omega_F^{2,m}$.

Другой, классический, подход к вычислению распределения ω_F^2 состоит в нахождении собственных чисел интегрального оператора с ядром $K_F^0(s, t)$ и использования формулы Смирнова. Вычисления осуществляются для конкретных распределений. Ряд результатов такого рода для предельного распределения статистики хи-квадрат при проверке сложных гипотез получен Г. В. Мартыновым в книге [12]. Как мы увидим, эти результаты применимы к отысканию распределения ω_F^2 . Рассмотрим два случая: центрированного нормального распределения и сдвинутого показательного распределения.

Пример 2. Пусть ξ_1 имеет нормальное распределение с параметрами $0, \sigma^2$. Тогда $F(x) = \Phi(x/\sigma)$, где $\Phi(x)$ — функция распределения стандартного нормального закона. Обозначим $\varphi(x) = \Phi'(x)$, тогда

$$GL_F(x) = -\varphi(\Phi^{-1}(x)), \quad K_F^0(s, t) = \min\{s, t\} - st - \varphi(\Phi^{-1}(s))\varphi(\Phi^{-1}(t)).$$

Таким образом, мы получили ковариационную функцию вида K_1 , где $\varkappa_1(t) = \varphi(\Phi^{-1}(t))$ (см. [12, гл. 2, § 3, с. 34]). В [12, гл. 2, § 3, с. 37] приведена формула для определителя Фредгольма $D_1(\lambda)$, а в [12, гл. 3, § 3, с. 56] приведена формула Смирнова, которая имеет место при некоторых дополнительных ограничениях. Однако теорема 2 (см. [12, гл. 3, § 3, с. 57]) дает достаточные условия для ее справедливости, которые выполняются, в частности, для рассматриваемой корреляционной функции. Наконец, распределение функционала ω_F^2 в этом случае приведено в таблице 3 (см. [12, приложение, с. 65]).

Пример 3. Пусть ξ_1 имеет сдвинутое показательное распределение с плотностью

$$f_{\xi_1}(x) = \begin{cases} \alpha e^{-\alpha x - 1}, & \text{если } x \geq -1/\alpha, \\ 0, & \text{если } x < -1/\alpha. \end{cases}$$

Тогда

$$GL_F^0(x) = \frac{1}{\alpha} (1 - x) \ln(1 - x),$$

$$K_F^0(s, t) = \min\{s, t\} - st - (1 - s) \ln(1 - s)(1 - t) \ln(1 - t).$$

Распределение функционала ω_F^2 в этом случае получено в [13] (см. также [12, приложение, табл. 12, с. 71]).

3. Сходимость к предельному процессу.

Лемма 1. Пусть $0 \leq s < t \leq 1$. Если $\mathbf{E} \xi_1^2 < \infty$, то

$$\frac{1}{\sqrt{n}} \left(\sum_{i=[ns]+1}^{[nt]} \xi_{i:n} \varepsilon_i - \sum_{i=[ns]+1}^{[nt]} \varepsilon_i \mathbf{E} \xi_{i:n} \right) \xrightarrow{P} 0.$$

Доказательство. Для начала заметим, что, поскольку существует $\mathbf{E} \xi_1$, существуют $\mathbf{E} \xi_{i:n}$ (см. [14, с. 40]). Далее, в силу неравенства Чебышёва достаточно доказать, что $(1/n) \mathbf{D}(\sum_{i=[ns]}^{[nt]} (\xi_{i:n} - \mathbf{E} \xi_{i:n}) \varepsilon_i) \rightarrow 0$ при $n \rightarrow \infty$.

В силу того, что случайные величины $\{\varepsilon_i\}$ независимы и не зависят от $\{\xi_{i:n}\}$, а $\mathbf{E} \varepsilon_1 = 0$, дисперсия суммы $\mathbf{D}(\sum_{i=[ns]}^{[nt]} (\xi_{i:n} - \mathbf{E} \xi_{i:n}) \varepsilon_i)$ равна сумме дисперсий, и

$$\begin{aligned} \frac{1}{n} \mathbf{D} \left(\sum_{i=[ns]}^{[nt]} (\xi_{i:n} - \mathbf{E} \xi_{i:n}) \varepsilon_i \right) &= \frac{1}{n} \sum_{i=[ns]}^{[nt]} \mathbf{D}(\xi_{i:n} - \mathbf{E} \xi_{i:n}) \mathbf{D} \varepsilon_i \\ &= \frac{\sigma^2}{n} \sum_{i=[ns]}^{[nt]} \mathbf{D} \xi_{i:n} \leq \frac{\sigma^2}{n} \sum_{i=1}^n \mathbf{D} \xi_{i:n}. \end{aligned}$$

Последнее выражение стремится к нулю с ростом n в силу предложения 8 (см. п. 4). Таким образом, лемма 1 доказана.

Сходимость конечномерных распределений процесса $\widehat{Z}_n$ будет доказана с помощью следующей леммы.

Лемма 2. Если $0 < \mathbf{E}\xi_1^2 < \infty$, то для любых $0 = t_0 < t_1 < \dots < t_k = 1$, $k \geq 1$, случайный вектор

$$\frac{1}{\sigma\sqrt{n}} \left(\sum_{i=[nt_{j-1}]+1}^{[nt_j]} \varepsilon_i, j = 1, \dots, k; \sum_{i=1}^n \varepsilon_i \xi_{i:n} \right) \quad (1)$$

слабо сходится к $(k+1)$ -мерному нормальному вектору $(\zeta_1, \dots, \zeta_{k+1})$ с нулевым математическим ожиданием и ковариациями

$$\mathbf{E}\zeta_i \zeta_j = \begin{cases} 0, & i \neq j, i < k+1, j < k+1; \\ t_i - t_{i-1}, & i = j < k+1; \\ GL_F(t_i) - GL_F(t_{i-1}), & i < k+1, j = k+1; \\ \mathbf{E}\xi_1^2, & i = j = k+1. \end{cases}$$

Доказательство. Заменяя в (1) случайные величины $\xi_{i:n}$ на их математические ожидания $a_{in} = \mathbf{E}\xi_{i:n}$, получаем случайный вектор

$$\frac{1}{\sigma\sqrt{n}} \left(\sum_{i=[nt_{j-1}]+1}^{[nt_j]} \varepsilon_i, j = 1, \dots, k; \sum_{i=1}^n \varepsilon_i a_{in} \right). \quad (2)$$

В силу леммы 1 разность между векторами (1) и (2) сходится по вероятности к нулю. Следовательно, достаточно показать, что вектор (2) сходится по распределению к вектору ζ .

Пусть $\mathbf{i}$ обозначает мнимую единицу. Характеристическая функция случайного вектора (2) имеет вид

$$\begin{aligned} \varphi(u_1, \dots, u_k, u_{k+1}) &= \prod_{j=1}^k \mathbf{E} \exp \left(\frac{\mathbf{i}}{\sigma\sqrt{n}} \left(u_j \sum_{i=[nt_{j-1}]+1}^{[nt_j]} \varepsilon_i + u_{k+1} \sum_{i=1}^n \varepsilon_i a_{in} \right) \right) \\ &= \prod_{j=1}^k \prod_{i=[nt_{j-1}]+1}^{[nt_j]} \mathbf{E} \exp \left(\frac{\mathbf{i}}{\sigma\sqrt{n}} ((u_j + u_{k+1} a_{in}) \varepsilon_i) \right). \end{aligned}$$

Справедливо следующее разложение характеристической функции:

$$\mathbf{E} \exp \left(\frac{\mathbf{i} t \varepsilon_i}{\sigma\sqrt{n}} \right) = 1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right).$$

Обозначим $\alpha_{in} = (u_j + u_{k+1} a_{in})^2 / (2n)$. Так как, в силу предложения 6 (см. п. 4), $a_{in} = o(\sqrt{n})$ равномерно по i , $1 \leq i \leq n$, то $\alpha_{in} = o(1)$

равномерно по i , $1 \leq i \leq n$, и

$$\begin{aligned} \varphi(u_1, \dots, u_{k+1}) &\rightarrow \prod_{j=1}^k \lim_{n \rightarrow \infty} \prod_{i=[nt_{j-1}]+1}^{[nt_j]} (1 - \alpha_{in}) \\ &= \prod_{j=1}^k \lim_{n \rightarrow \infty} \exp \left(\sum_{i=[nt_{j-1}]+1}^{[nt_j]} \ln(1 - \alpha_{in}) \right). \end{aligned}$$

Так как $-x(1 + o(1)) \leq \ln(1 - x) \leq -x$ при $x \rightarrow 0$, то для любого $\varepsilon > 0$ найдется n_0 такое, что для всех $n \geq n_0$ выполнены неравенства $-\alpha_{in}(1 + \varepsilon) \leq \ln(1 - \alpha_{in}) \leq -\alpha_{in}$. В силу последнего соотношения

$$\lim_{n \rightarrow \infty} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} \ln(1 - \alpha_{in}) = \lim_{n \rightarrow \infty} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} \alpha_{in}.$$

Покажем, что последний предел существует. Имеем:

$$\begin{aligned} \lim_{n \rightarrow \infty} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} \alpha_{in} &= \frac{1}{2} u_{k+1}^2 \lim_{n \rightarrow \infty} \sum_{i=1}^n \frac{a_{in}^2}{n} + \frac{1}{2} (t_j - t_{j-1}) u_j^2 \\ &\quad + u_j u_{k+1} \lim_{n \rightarrow \infty} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} \frac{a_{in}}{n}. \end{aligned}$$

Так как $(1/n) \sum_{i=1}^n \mathbf{E} \xi_{i:n}^2 = \mathbf{E} \xi_1^2$, то

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n a_{in}^2 = \mathbf{E} \xi_1^2 - \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{D} \xi_{i:n} = \mathbf{E} \xi_1^2$$

в силу предложения 8 (см. п. 4).

Так как (см. [8])

$$\eta_n = \frac{1}{n} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} \xi_{i:n} \rightarrow GL_F(t_j) - GL_F(t_{j-1}) \quad \text{п.н.,}$$

то для сходимости

$$\frac{1}{n} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} a_{in} \rightarrow GL_F(t_j) - GL_F(t_{j-1})$$

достаточно равномерной интегрируемости последовательности $\{\eta_n\}$, которая имеет место в силу предложения 3 (см. п. 4).

Таким образом, существование $\lim_{n \rightarrow \infty} \sum_{i=[nt_{j-1}]+1}^{[nt_j]} \alpha_{in}$ доказано.

Итак,

$$\varphi(u_1, \dots, u_k, u_{k+1}) \rightarrow \exp \left(-\frac{1}{2} u_{k+1}^2 \mathbf{E} \xi_1^2 \right) \\ \times \prod_{j=1}^k \exp \left(-\frac{1}{2} (t_j - t_{j-1}) u_j^2 - u_j u_{k+1} (GL_F(t_j) - GL_F(t_{j-1})) \right).$$

Следовательно, вектор (2) сходится по распределению к вектору ζ . Лемма 2 доказана.

Докажем сходимост конечномерных распределений процесса $\widehat{Z}_n$.

Лемма 3. Если $0 < \mathbf{E} \xi_1^2 < \infty$, то конечномерные распределения процесса $\widehat{Z}_n$ сходятся к конечномерным распределениям процесса $\widehat{Z}_F$.

Доказательство. Остатки регрессии $\widehat{\varepsilon}_i$ допускают следующее представление:

$$\widehat{\varepsilon}_i = Y_i - \widehat{\theta} \xi_{i:n} = \theta \xi_{i:n} + \varepsilon_i - \frac{\overline{X(\theta X + \varepsilon)}}{\overline{X^2}} X_i = \varepsilon_i - \frac{\xi_{i:n}}{\sqrt{n}} \frac{\sum_{i=1}^n \xi_{i:n} \varepsilon_i}{\sqrt{n}} \frac{1}{\xi^2}. \quad (3)$$

Вследствие предыдущей леммы $(1/\sqrt{n}) \sum_{i=1}^n \xi_{i:n} \varepsilon_i$ сходится по распределению к нормальной случайной величине с нулевым математическим ожиданием и дисперсией $\sigma^2 \mathbf{E} \xi_1^2$, а $\overline{\xi^2} \rightarrow \mathbf{E} \xi_1^2$ п.н. согласно УЗБЧ.

Согласно предложению 5 (см. п. 4) случайные величины $\widehat{\varepsilon}_i - \varepsilon_i$ сходятся по вероятности к нулю равномерно по $1 \leq i \leq n$.

Так как

$$\widehat{Z}_n(t) = \frac{\widehat{\Delta}_{[nt]}}{\sigma \sqrt{n}} + \left(t - \frac{[nt]}{n} \right) \frac{\widehat{\varepsilon}_{[nt]+1}}{\sigma \sqrt{n}},$$

то

$$\left| \widehat{Z}_n(t) - \frac{\widehat{\Delta}_{[nt]}}{\sigma \sqrt{n}} \right| \leq \left| \frac{\widehat{\varepsilon}_{[nt]+1}}{\sigma \sqrt{n}} \right| \leq \left| \frac{\varepsilon_{[nt]+1}}{\sigma \sqrt{n}} \right| + \left| \frac{\widehat{\varepsilon}_{[nt]+1} - \varepsilon_{[nt]+1}}{\sigma \sqrt{n}} \right| \xrightarrow{p} 0.$$

Следовательно, для любых $0 = t_0 < t_1 < \dots < t_{k-1} < t_k = 1$ слабая сходимость вектора $(\widehat{Z}_n(t_1), \dots, \widehat{Z}_n(t_k))$ эквивалентна слабой сходимости вектора

$$\frac{1}{\sigma \sqrt{n}} (\widehat{\Delta}_{[nt_1]}, \dots, \widehat{\Delta}_{[nt_k]}).$$

В силу (3) имеет место представление

$$\frac{\widehat{\Delta}_{[nt]}}{\sigma \sqrt{n}} = \frac{\Delta_{[nt]}}{\sigma \sqrt{n}} - \frac{\sum_{i=1}^n \varepsilon_i \xi_{i:n}}{\sigma \sqrt{n}} \frac{S_{[nt]}}{n \overline{\xi^2}}. \quad (4)$$

Так как $S_{[nt]}/n \rightarrow GL_F(t)$ п.н., $\overline{\xi^2} \rightarrow \mathbf{E} \xi_1^2$ п.н., то согласно лемме 2

$$\frac{1}{\sigma \sqrt{n}} (\widehat{\Delta}_{[nt_1]}, \dots, \widehat{\Delta}_{[nt_k]}) \Rightarrow \left(\sum_{i=1}^j \zeta_i - \zeta_{k+1} \frac{GL_F(t_j)}{\mathbf{E} \xi_1^2}, j = 1, \dots, k \right).$$

Итак, конечномерные распределения слабо сходятся к гауссовскому закону с нулевым математическим ожиданием и ковариациями ($t_1 < t_2$)

$$\begin{aligned} \text{cov}(\widehat{Z}_F(t_1), \widehat{Z}_F(t_2)) &= \mathbf{E} \left(\zeta_1 - \zeta_{k+1} \frac{GL_F(t_1)}{\mathbf{E} \xi_1^2} \right) \left(\zeta_1 + \zeta_2 - \zeta_{k+1} \frac{GL_F(t_2)}{\mathbf{E} \xi_1^2} \right) \\ &= \mathbf{E} \xi_1^2 - \frac{1}{\mathbf{E} \xi_1^2} (GL_F(t_1) \mathbf{E} (\zeta_1 + \zeta_2) \zeta_{k+1} + GL_F(t_2) \mathbf{E} \zeta_1 \zeta_{k+1}) \\ &\quad + \frac{GL_F(t_1) GL_F(t_2)}{(\mathbf{E} \xi_1^2)^2} \mathbf{E} \zeta_{k+1}^2 = t_1 - \frac{GL_F(t_1) GL_F(t_2)}{\mathbf{E} \xi_1^2}. \end{aligned}$$

Лемма 3 доказана.

Лемма 4. Пусть $0 < \mathbf{E} \xi_1^2 < \infty$, тогда семейство распределений $\{\widehat{Z}_n(t), 0 \leq t \leq 1\}$ относительно компактно.

Доказательство. Заметим, что, так как пространство $C(0, 1)$ сепарабельно, то, в силу теоремы Прохорова (см. [9, гл. 1, § 6]), относительная компактность и плотность семейства распределений в нашем случае эквивалентны. Таким образом, для доказательства утверждения леммы достаточно доказать плотность семейства распределений, порожденного последовательностью (4). Обозначим $k = [nt]$. Достаточно показать только плотность семейства распределений, порожденного вторым слагаемым, т.е.

$$\frac{\sqrt{n}(\widehat{\theta} - \theta)}{\sigma} \frac{S_k}{n},$$

так как плотность семейства распределений, порожденного последовательностью $\Delta_k/(n^{1/2}\sigma)$, была показана в [9] при доказательстве принципа инвариантности Донскера–Прохорова.

Для доказательства плотности семейства, порожденного (4), достаточно показать (см. [9, теорема 8.3]), что для произвольных $\varepsilon > 0$, $\eta > 0$ найдутся $0 < \delta < 1$, $n_0 \in \mathbf{N}$ такие, что

$$\frac{1}{\delta} \mathbf{P} \left\{ \sup_{t \leq s \leq t+\delta} \left| \frac{\sqrt{n}(\widehat{\theta} - \theta)}{\sigma} \frac{S_{[ns]} - S_{[nt]}}{n} \right| \geq \varepsilon \right\} \leq \eta \quad (5)$$

для всех $n > n_0$, $0 \leq t \leq 1$.

Докажем соотношение (5). Заметим, что в силу леммы 2

$$\frac{\sqrt{n}(\widehat{\theta} - \theta)}{\sigma} \Rightarrow \frac{\zeta}{\sqrt{\mathbf{E} \xi_1^2}},$$

и (см. [8])

$$\sup_{t \leq s \leq t+\delta} \left| \frac{S_{[ns]} - S_{[nt]}}{n} \right| \rightarrow \sup_{t \leq s \leq t+\delta} |GL_F(s) - GL_F(t)| \quad \text{п.н.,}$$

где ζ — случайная величина, имеющая стандартное нормальное распределение, а $GL_F(x)$, как и выше, — обобщенная кривая Лоренца.

Обозначим $A(t, \delta) = \sup_{t \leq s \leq t+\delta} |GL_F(s) - GL_F(t)|$. Тогда достаточно показать, что для произвольных $\varepsilon > 0$, $\eta > 0$ существует $0 < \delta < 1$ такое, что

$$\mathbf{P} \left\{ \frac{|\zeta|}{\sqrt{\mathbf{E} \xi_1^2}} A(t, \delta) \geq \varepsilon \right\} \leq \frac{\eta \delta}{2}, \quad 0 \leq t \leq 1. \quad (6)$$

В силу [15, с. 192, лемма 2] левая часть в (6) не превосходит $2e^{-x^2/2}/(x\sqrt{2\pi})$, где $x = \varepsilon\sqrt{\mathbf{E} \xi_1^2}/A(t, \delta)$. Согласно неравенству Коши–Буняковского,

$$A(t, \delta) = \sup_{t \leq s \leq t+\delta} \left| \int_t^s F^{-1}(x) dx \right| \leq \sup_{t \leq s \leq t+\delta} \int_t^s |F^{-1}(x)| dx \leq \sqrt{\delta \mathbf{E} \xi_1^2}.$$

Ясно, что требуемое δ всегда найдется. Лемма 4 доказана.

Лемма 5. Пусть $\widehat{\sigma}^2 = \widehat{\varepsilon}^2 - (\widehat{\varepsilon})^2$. Если $\mathbf{E} \xi_1^2 < \infty$, то $\widehat{\sigma}^2 \xrightarrow{P} \sigma^2$.

Доказательство. Так как $(\widehat{\theta} - \theta) \xrightarrow{P} 0$ вследствие леммы 2, то $\widehat{\varepsilon} = (\theta - \widehat{\theta})\overline{X} + \overline{\varepsilon} \xrightarrow{P} 0$, $\widehat{\varepsilon}^2 = \overline{\varepsilon}^2 - (\overline{X}\varepsilon)^2/\overline{X}^2 \xrightarrow{P} \sigma^2$. Лемма 5 доказана.

Доказательство теоремы. Первое утверждение теоремы является следствием теоремы 8.1 в [9, гл. 2, § 8, с. 82], лемм 3 (сходимость конечномерных распределений) и 4 (относительная компактность).

В силу непрерывности в равномерной метрике в $C(0, 1)$ преобразования, отображающего x в x^0 , где $x^0(t) = x(t) - tx(1)$, и леммы 5, из утверждения 1) теоремы вытекает утверждение 2). Теорема доказана.

4. Приложение. В данном приложении представлен ряд известных утверждений, используемых в настоящей работе.

Введем обозначения $\xi^+ = \max\{0, \xi\}$ и $\xi^- = \min\{0, \xi\}$. Тогда $\xi = \xi^+ + \xi^-$.

Предложение 1. Справедливы следующие утверждения:

1) (см. [16, задача 21.16]). Пусть $\xi_n \Rightarrow \xi$ при $n \rightarrow \infty$. Тогда следующие утверждения эквивалентны:

- а) последовательность $\{\xi_n\}$ равномерно интегрируема;
- б) $\mathbf{E} |\xi_n| < \infty$, $\mathbf{E} |\xi| < \infty$ и $\mathbf{E} \xi_n^+ \rightarrow \mathbf{E} \xi^+$, $\mathbf{E} \xi_n^- \rightarrow \mathbf{E} \xi^-$ при $n \rightarrow \infty$;

2) (см. [17, гл. II, § 6, с. 235, теорема 5]). Пусть $0 \leq \xi_n \rightarrow \xi$ п.н. и $\mathbf{E} \xi_n < \infty$. Тогда $\mathbf{E} \xi_n \rightarrow \mathbf{E} \xi < \infty$ в том и только том случае, когда семейство случайных величин $\{\xi_n\}_{n \geq 1}$ равномерно интегрируемо.

Предложение 2. Пусть $\xi_1, \xi_2, \dots$ — независимые одинаково распределенные случайные величины, $S_n = \xi_1 + \dots + \xi_n$. Тогда из сходимости $S_n/n \rightarrow a$ п.н. следует, что $\mathbf{E} |\xi_1| < \infty$ и $\{S_n/n\}$ равномерно интегрируема.

Предложение 2 следует из усиленного закона больших чисел Колмогорова, предложения 1 и соотношений $S_n = \xi_1^+ + \dots + \xi_n^+ - (-\xi_1^- - \dots - \xi_n^-)$, $(\xi_1^+ + \dots + \xi_n^+)/n \rightarrow c = \mathbf{E} \xi^+ < \infty$, $\mathbf{E}((\xi_1^+ + \dots + \xi_n^+)/n) = \mathbf{E} \xi^+$.

Предложение 3. Пусть $\{\xi_{i:n}\}$ — порядковые статистики, построенные по выборке $\{\xi_1, \dots, \xi_n\}$ с конечным математическим ожиданием, и $0 \leq c_1 < c_2 \leq 1$. Тогда $\{\xi_{n:n}/n\}$ и $\{(1/n) \sum_{i=[nc_1]+1}^{[nc_2]} \xi_{i:n}\}$ равномерно интегрируемы.

Предложение 3 следует из неравенств

$$|\xi_{n:n}| \leq \sum_{i=1}^n |\xi_i|, \quad \sum_{i=[nc_1]+1}^{[nc_2]} |\xi_{i:n}| \leq \sum_{i=1}^n |\xi_i|$$

и предложения 2.

Предложение 4. Если $\mathbf{E} |\xi_1| < \infty$, то $\xi_{n:n}/n \rightarrow 0$ п.н., $\mathbf{E} \xi_{n:n}/n \rightarrow 0$.

Предложение 4 является широко известным фактом, который легко доказывается с помощью леммы Бореля–Кантелли (см. [17, гл. II, § 10, с. 327, лемма 1]) и предложения 3.

Подставляя $\xi_i = |\eta_i|^\alpha$ в предложение 4, получаем следующий результат.

Предложение 5. Если $\mathbf{E} |\eta_1|^\alpha < \infty$, $\alpha > 0$, то $|\eta_{n:n}|^\alpha/n \rightarrow 0$ п.н., $\mathbf{E} |\eta_{n:n}|^\alpha/n \rightarrow 0$, $\eta_{n:n}/n^{1/\alpha} \rightarrow 0$ п.н., $\eta_{i:n}/n^{1/\alpha} \rightarrow 0$ п.н. равномерно по $1 \leq i \leq n$.

Применяя неравенство Йенсена к последнему утверждению, получаем следующее предложение.

Предложение 6. Если $\mathbf{E} |\eta_1|^\alpha < \infty$, $\alpha \geq 1$, то $\mathbf{E} \eta_{i:n} = o(n^{1/\alpha})$ равномерно по $1 \leq i \leq n$.

Предложение 7 (см. [18, теорема 1]). Пусть $\xi_1, \xi_2, \dots$ — независимые одинаково распределенные случайные величины, $\mathbf{E} |\xi_1| < \infty$ и $g(x)$ — вещественнозначная непрерывная функция такая, что $|g(x)| \leq h(x)$, где функция $h(x)$ выпуклая и $\mathbf{E} h(\xi_1) < \infty$. Тогда

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n g(\mathbf{E} \xi_{i:n}) = \mathbf{E} g(\xi_1).$$

Применяя предложение 7 с $g(x) = x^2$ и вспоминая, что $(1/n) \sum_{i=1}^n \mathbf{E} \xi_{i:n}^2 = \mathbf{E} \xi_1^2$, получаем следующий результат.

Предложение 8. Пусть $\xi_1, \xi_2, \dots$ — независимые одинаково распределенные случайные величины, $\mathbf{E} \xi_1^2 < \infty$, тогда

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \mathbf{D} \xi_{i:n} = 0.$$

Авторы благодарят С.Г.Фосса за многочисленные обсуждения, позволившие существенно улучшить изложение, Е.А.Бакланова и А.И.Саханенко за полезные ссылки.

Авторы искренне благодарят анонимного рецензента за ряд замечаний, касающихся практических применений основного результата, а также за ссылки на неизвестные ранее авторам работы.

СПИСОК ЛИТЕРАТУРЫ

1. *Quetelet A.* Recherches sur le poids de l'homme aux differens âges. — Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles, 1832, v. 7, p. 1–46.
2. *Keys A., Fidanza F., Karvonen M. J., Kimura N., Taylor H. L.* Indices of relative weight and obesity. — J. Chronic Diseases, 1972, v. 25, № 6-7, p. 329–343.
3. *Arnold J. R., Libby W. F.* Age determinations by radiocarbon content: checks with samples of known age. — Science, 1949, v. 110, № 2869, p. 678–680.
4. *MacNeill I. B.* Limit processes for sequences of partial sums of regression residuals. — Ann. Probab., 1978, v. 6, № 4, p. 695–698.
5. *Aue A., Horvath L., Huskova M., Kokoszka P.* Testing for change in polynomial regression. — Bernoulli, 2008, v. 14, № 3, p. 637–660.
6. *Муганцева Л. А.* Проверка нормальности в схемах одномерной и многомерной линейной регрессии. — Теория вероятн. и ее примен., 1977, т. 22, в. 3, с. 603–614.
7. *Gastwirth J. L.* A general definition of the Lorenz curve. — Econometrica, 1971, v. 39, p. 1037–1039.
8. *Goldie C. M.* Convergence theorems for empirical Lorenz curves and their inverses. — Adv. Appl. Probab., 1977, v. 9, p. 765–791.
9. *Биллингсли П.* Сходимость вероятностных мер. М.: Наука, 1977, 351 с.
10. *Дугерти К.* Введение в эконометрику. М.: ИНФРА-М, 1999, 416 с.
11. *Deheuvels P., Martynov G. V.* Cramér-von Mises-type tests with applications to tests of independence for multivariate extreme-value distributions. — Comm. Statist. Theory Methods, 1996, v. 25, № 4, p. 871–908.
12. *Мартынов Г. В.* Критерии омега-квадрат. М.: Наука, 1978, 79 с.
13. *Durbin J., Knott M., Taylor C. C.* Components of Cramer-von Mises Statistics. II. — J. Roy. Statist. Soc. Ser. B, 1975, v. 37, p. 216–237.
14. *Дэйвид Г.* Порядковые статистики. М.: Наука, 1979, 335 с.
15. *Феллер В.* Введение в теорию вероятностей и ее приложения, т. 1. М.: Мир, 1984, 527 с.
16. *Коршунов Д. А., Фосс С. Г., Эйсымонт И. М.* Сборник задач и упражнений по теории вероятностей. СПб.: Лань, 2004, 191 с.
17. *Ширяев А. Н.* Вероятность, т. 1. М.: МЦНМО, 2004, 519 с.
18. *Hoeffding W.* On the distribution of the expected values of the order statistics. — Ann. Math. Statist., 1953, v. 24, № 1, p. 93–100.

Поступила в редакцию
5.V.2012

Исправленный вариант
6.V.2014