

УДК 519.25, 004.622, 004.67, 311.313

АНАЛИТИЧЕСКИЕ ИССЛЕДОВАНИЯ РАСПРЕДЕЛЕНИЯ РЕЗУЛЬТАТОВ ЕГЭ НА ПРИМЕРЕ РЯЗАНСКОЙ ОБЛАСТИ

Р. В. Хруничев, к.т.н., доцент кафедры ЭВМ РГРТУ, Рязань, Россия;
orcid.org/0009-0006-1442-0649, e-mail: hrunichev_robert@mail.ru

*Приводится статистический и аналитический анализ выборок больших объёмов с результатами единого государственного экзамена (ЕГЭ) выпускников школ Рязанской области в период с 2019 по 2024 годы по ключевым для инженерного образования общеобразовательным предметам – математика, физика, информатика, а также русский язык как обязательный предмет для всех выпускников 11 классов. **Целью работы** ставится задача подтвердить или опровергнуть гипотезу о нормальном распределении результатов ЕГЭ по отдельным предметам за указанный период, установить наличие систематических тенденций в распределении результатов ЕГЭ. Предварительно выборочным методом определить основные характеристики – математическое ожидание и среднее квадратическое отклонение случайной величины, необходимые для проверки гипотезы о характере распределения результатов ЕГЭ.*

Ключевые слова: анализ результатов ЕГЭ, статистический анализ, выборочный анализ, характеристики выборки, нормальное распределение, проверка статистической гипотезы, выборки ЕГЭ большого объёма, данные результатов ЕГЭ.

DOI: 10.21667/1995-4565-2025-91-233-240

Введение

Рассматривается решение статистической задачи подтверждения или опровержения гипотезы о нормальном распределении результатов ЕГЭ выпускников школ Рязанского региона в период с 2019 по 2024 год по основным для инженерного образования общеобразовательным предметам. Для решения данной задачи проводится статистический анализ выборок большого объёма с результатами ЕГЭ, определяются основные характеристики выборок – математическое ожидание и среднее квадратическое отклонение (ско). На основе анализа выборок по каждому предмету за ряд лет приводится вывод о наличии систематических тенденций в распределении результатов ЕГЭ.

Теоретическая часть

В соответствии с условиями центральной предельной теоремы [1] и графическим методом анализа на основе метода гистограмм [2] выдвинем предположение о том, что результаты ЕГЭ при достаточно больших объёмах выборочных совокупностей должны подчиняться нормальному закону распределения с параметрами: m – математическое ожидание (среднее значение) результата ЕГЭ по предмету и σ – стандартное отклонение (ско) результатов ЕГЭ от среднего.

Сформулируем гипотезу **H0**: результаты ЕГЭ распределены нормально, альтернативная к **H0** гипотеза **H1**: результаты ЕГЭ не подчиняются нормальному закону распределения.

Зададимся стандартным значением уровня значимости $\alpha = 0,05$ и на этом уровне проверим справедливость выдвинутой гипотезы **H0** с помощью критерия согласия χ^2 (Пирсона):

$$\chi^2 = \sum_{i=1}^k \frac{(n_i - n'_i)^2}{n'_i}, \quad (1)$$

где n_i – эмпирические частоты, полученные в ходе интервального анализа выборок, а n_i' – теоретические частоты, соответствующие нормально распределенной величине для аналогичных интервалов [3].

Теоретические частоты n_i' определяются с помощью значений функции Лапласа по выражению:

$$n_i' = n \cdot p_i = n \cdot \left[\Phi\left(\frac{x_{i+1} - \bar{x}}{S}\right) - \Phi\left(\frac{x_i - \bar{x}}{S}\right) \right], \quad (2)$$

где n – объём выборочной совокупности; x_i и x_{i+1} – границы соответствующего интервала группировки; $\bar{x}$ – среднее по выборке значение параметра (результата ЕГЭ), S – выборочное значение среднего квадратического отклонения [3].

Аналитические исследования

Исследование основано на статистических данных, полученных в ходе проведения приёмных кампаний в период с 2019 по 2024 год. С целью чистоты исследования анализу подвергались результаты ЕГЭ абитуриентов, поступивших по общему конкурсу по результатам ЕГЭ. При этом минимальные пороги по каждому из предметов установлены в соответствии с минимальными баллами, определяемыми Минобрнауки на уровне 2024 года.

Таблица 1 – Объёмы выборочных совокупностей для анализа ЕГЭ по годам и предметам
Table 1 – Sample populations volume to analyze the results of Unified State Exam in subjects for 2019-2024

Год\Предмет	Русский язык	Математика (профиль)	Физика	Информатика
2019	520	513	515	–
2020	548	542	546	–
2021	757	747	620	225
2022	762	757	512	333
2023	671	666	406	341
2024	696	688	367	381

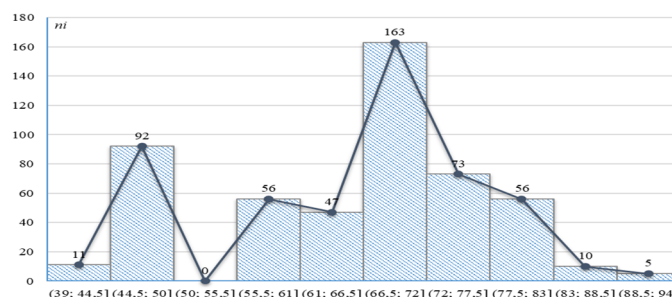
Сформулируем нулевую гипотезу: выборочные данные результатов ЕГЭ X подчиняются нормальному закону распределения – $H_0: X \sim N(m; \sigma^2)$, где m – математическое ожидание результатов ЕГЭ, а σ – среднее квадратическое отклонение, альтернативная гипотеза H_1 : закон распределения отличается от нормального.

Для каждой выборочной совокупности составлены интервальные статистические ряды на основе следующих параметров:

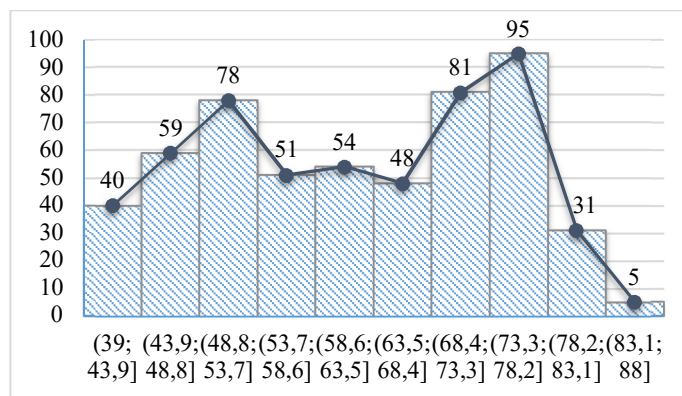
– число интервалов, определяемое формулой Стерджесса, $m = 1 + \log_2 n = 1 + 3,322 \cdot \lg n$;

– шаг выборки $h = \frac{x_{\max} - x_{\min}}{m}$.

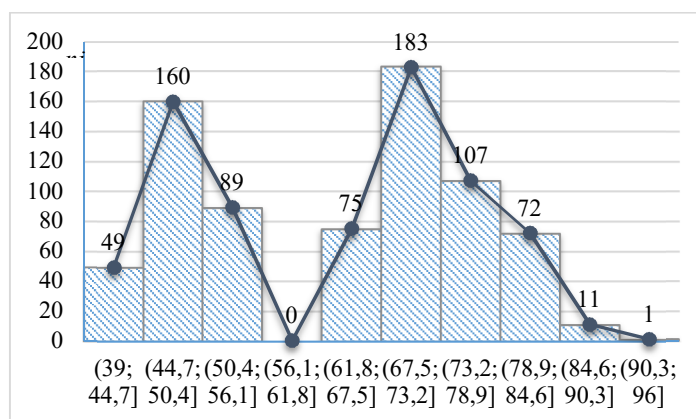
Построим гистограммы распределения выборочных совокупностей по основным для инженерного образования предметам – математике и физике (рисунки 1-2).



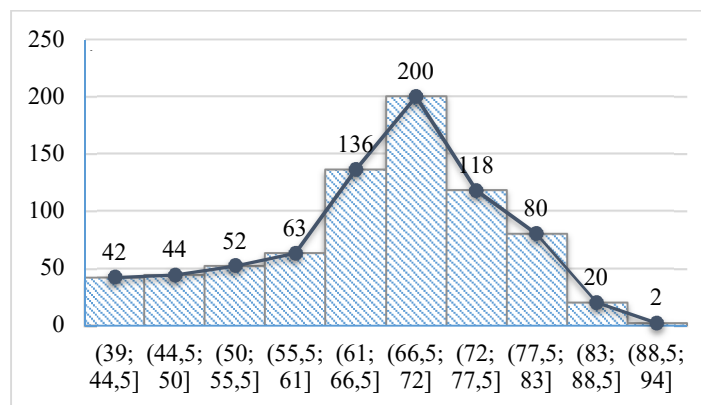
а (а) – математика 2019 г.



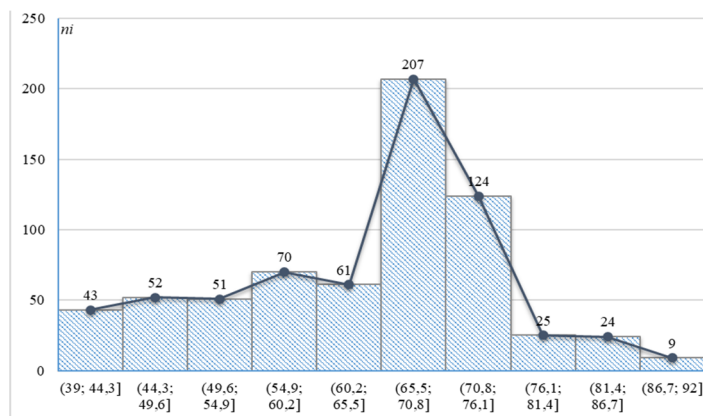
б (b) – математика 2020 г.



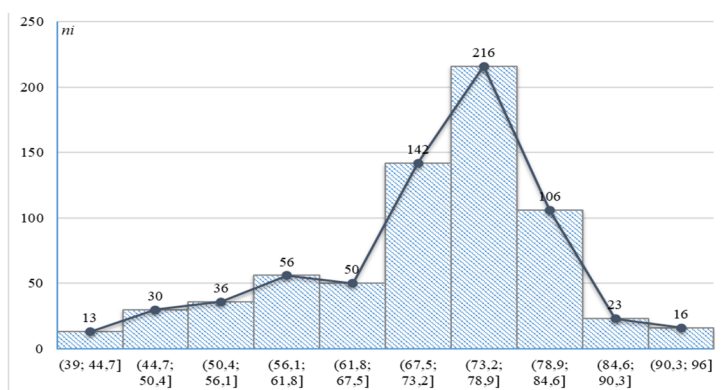
в (c) – математика 2021 г.



г (d) – математика 2022 г.

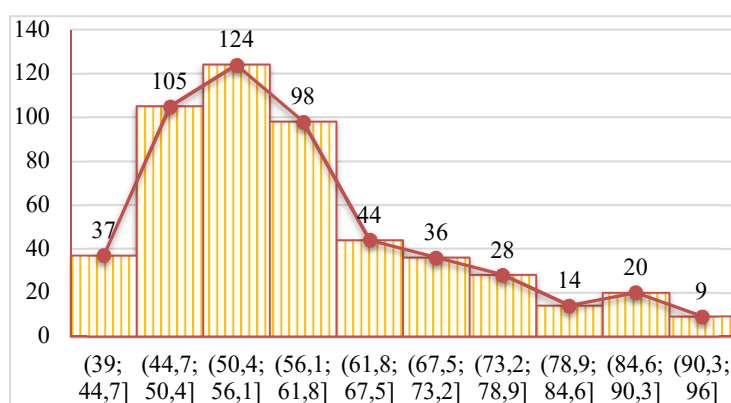


д (e) – математика 2023 г.

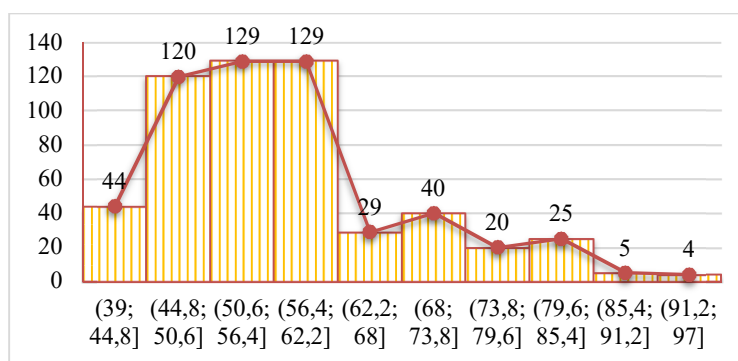


е (f) – математика 2024 г.

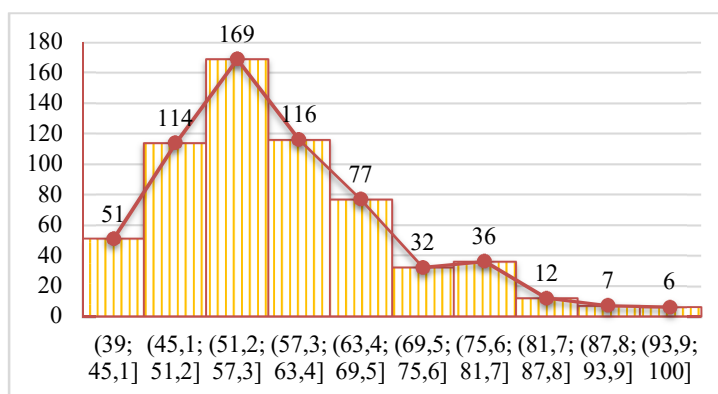
Рисунок 1 – Гистограмма и полигон частот: математика 2019 (а) – 2024 (е) год
 Figure 1 – Histogram and frequency polygon: mathematics of 2019 (a) – 2024 (f) year



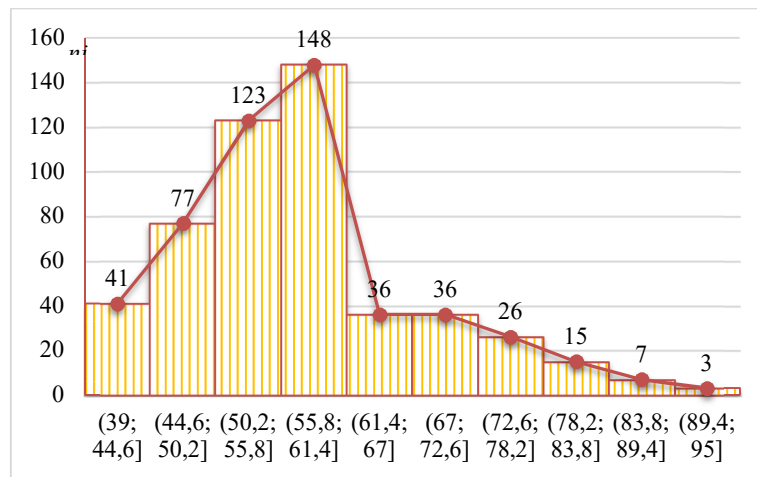
а (a) – физика 2019 г.



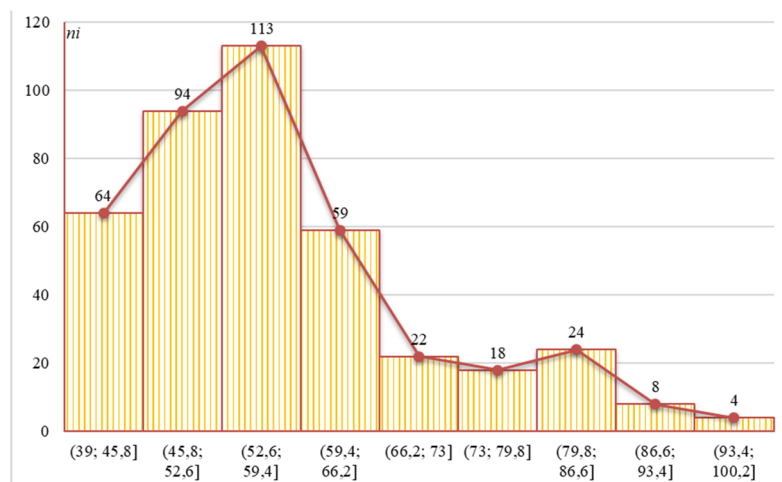
б (b) – физика 2020 г.



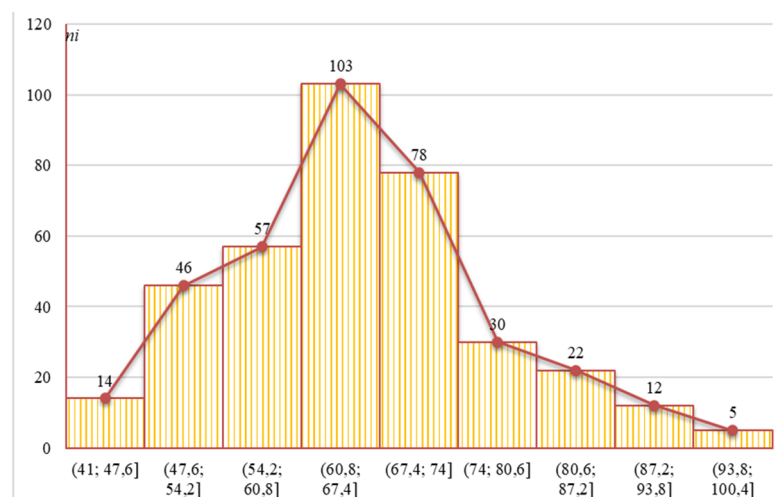
в (c) – физика 2021 г.



г (d) – физика 2022 г.



д (e) – физика 2023 г.



е (f) – физика 2024 г.

Рисунок 2 – Гистограмма и полигон частот: физика 2019 (а) – 2024 (е) год
Figure 2 – Histogram and frequency polygon: Physics 2019 (a) – 2024 (e) year

На основе проведенного анализа выборок получены их числовые характеристики, необходимые для проверки гипотезы о нормальном распределении результатов ЕГЭ (таблица 2).

Таблица 2 – Наблюдаемые значения математического ожидания (m) и среднего квадратического отклонения (σ) по выборочным данным результатов ЕГЭ

Table 2 – Observed values of mathematical expectation (m) and mean square deviation (σ) according to sample data of USE results

Год\Предмет	Русский язык		Математика (профиль)		Физика		Информатика	
	параметр							
	m	σ	m	σ	m	σ	m	σ
2019	76,21	10,22	65,53	11,52	58,56	12,21	–	–
2020	74,51	10,94	61,95	12,21	57,25	11,05	–	–
2021	75,03	11,79	63,18	13,17	58,56	11,34	65,71	12,35
2022	74,5	10,56	65,92	10,92	57,5	10,26	67,81	12,98
2023	74,08	10,9	64,21	11,1	57,59	12,47	66,07	12,73
2024	67,71	11,16	71,23	10,84	65,65	11,22	63,47	12

В соответствии с общепринятыми рекомендациями [2, 3] для вычисления наблюдаемого значения критерия согласия χ^2 интервалы, число наблюдений в которых менее 5, объединены с соседними.

С помощью выражений (1) и (2) найдём наблюдаемые значения критерия Пирсона (таблица 3). Для уровня значимости $\alpha = 0,05$ и числа степеней свободы $r = k-3$, где k – количество интервалов после объединения, найдем критические значения критерия χ^2 (таблица 3).

Таблица 3 – Наблюдаемые и критические значения критерия χ^2 проверки гипотезы о нормальном распределении результатов ЕГЭ

Table 3 – Observed and critical values of criterion χ^2 to test the hypothesis of USE results normal distribution

Год\Предмет	Русский язык		Математика (профиль)		Физика		Информатика	
	$\chi^2_{\text{набл}}$	$\chi^2_{\text{крит}}$	$\chi^2_{\text{набл}}$	$\chi^2_{\text{крит}}$	$\chi^2_{\text{набл}}$	$\chi^2_{\text{крит}}$	$\chi^2_{\text{набл}}$	$\chi^2_{\text{крит}}$
2019	3,8	14,1	276,7	12,6	136	14,1	–	
2020	42,2	12,6	121,8	14,1	127,1	12,1	–	
2021	28,3	14,1	352,3	11,1	118,5	14,1	11,05	11,1
2022	11,5	12,6	99,3	12,6	85,1	12,6	25,6	12,6
2023	23,86	14,1	167,39	14,1	75,5	11,1	43,6	11,1
2024	22,6	12,6	155,29	14,1	25,3	12,6	20,5	11,1

Из приведенных статистических данных и произведенных расчётов видно, что гипотеза **H0** о нормальном распределении (выполнение условия принятия нулевой гипотезы $\chi^2_{\text{набл}} < \chi^2_{\text{крит}}$) **не** отвергается для результатов ЕГЭ по русскому языку за 2019 и 2022 годы, а также по информатике за 2021 год. Во всех остальных случаях гипотезу **H0** следует отвергнуть.

При более детальном анализе полученных результатов становится очевидным, что наименее сильно отклонены от нормального распределения результаты ЕГЭ по русскому языку и информатике, в то время как результаты по физике и профильной математике критически далеки от него.

Анализируя гистограммы результатов ЕГЭ по физике, отметим, что наблюдается ярко выраженная правосторонняя асимметрия, свидетельствующая о том, что доля абитуриентов способных получить на ЕГЭ балл выше среднего, значительно меньше тех, кто получает балл меньше среднего. Справедливыми в данной ситуации кажутся вопросы о качестве подготовки в школах и о соответствии контрольно-измерительных материалов ЕГЭ по физике существующей подготовке.

В то же время в результатах ЕГЭ по профильной математике начиная с 2022 года наблюдается значительная левосторонняя асимметрия, что указывает на несбалансированность контрольно-измерительных материалов ЕГЭ.

Заключение

Приведён статистический анализ данных результатов ЕГЭ по необходимым для получения инженерного образования общеобразовательным предметам – математике, физике, информатике и русскому языку, полученным в ходе проведения приёмных кампаний в период с 2019 по 2024 годы.

На основе условий центральной предельной теоремы и графического метода анализа гистограмм была выдвинута гипотеза о нормальном распределении результатов ЕГЭ, которая на основе критерия согласия Пирсона нашла подтверждение за 2019 и 2022 годы по русскому языку и за 2021 год по информатике.

Наиболее критически оцениваются полученные данные о результатах ЕГЭ по физике и профильной математике, поскольку здесь наблюдается значительное отклонение от нормального распределения. При этом возникает вопрос о сбалансированности заданий ЕГЭ по профильной математике, где начиная с 2022 года наблюдается значительная левосторонняя асимметрия и об уровне подготовки в школах по физике, где в случае соответствия контрольно-измерительных материалов ФГОС наблюдается значительная правосторонняя асимметрия. Очевидно, что большинство выпускников школ не могут справиться с ЕГЭ по физике на уровне выше среднего.

Понимание ситуации сложилось и в высших эшелонах власти, важность повышения качества естественно-научного образования [4] отражена в комплексном плане мероприятий по повышению качества математического и естественно-научного образования на период до 2030 года.

Библиографический список

1. Центральная предельная теорема // Национальный научно-образовательный портал «Большая российская энциклопедия»: [сайт]. 2025. URL: <https://bigenc.ru/c/tsentral-naia-predel-naia-teorema-56c807> (дата обращения 15.01.2025 г.).
2. Кулаичев А.П. Методы и средства комплексного статистического анализа данных: учебное пособие / А.П. Кулаичев. 5-е изд., перераб. и доп. Москва : ИНФРА-М, 2025. 484 с. (Высшее образование). DOI 10.12737/25093.
3. Хруничев Р.В. Прикладные статистические методы анализа : учеб. пособие / Р.В. Хруничев. Рязан. гос. радиотехн. ун-т. Рязань, 2023. 80 с.
4. Правительство России [Электронный ресурс] / Распоряжение Правительства РФ от 19 ноября 2024 г. № 3333-р «Комплексный план мероприятий по повышению качества математического и естественно-научного образования на период до 2030 года». Режим доступа: <http://government.ru/docs/all/156334/> (Дата обращения: 12.01.2025 г.).

UDC 519.25, 004.622, 004.67, 311.313

ANALYTICAL STUDIES OF USE RESULTS DISTRIBUTION USING THE EXAMPLE OF RYAZAN REGION

R. V. Khrunichev, Ph.D. (Tech.), associate professor, department of electronic computers, RSREU, Ryazan, Russia;
orcid.org/0009-0006-1442-0649, e-mail: hkrunichev_robert@mail.ru

The statistical and analytical analysis of large-volume samples with the results of unified state exam (USE) of school graduates of Ryazan region in the period from 2019 to 2024 in key general education subjects for engineering education - mathematics, physics, computer science, as well as Russian as a compulso-

ry subject for all graduates of grades 11 is carried out. **The aim of the work** is to confirm or refute the hypothesis of normal distribution of USE results for individual subjects over the specified period, to establish the presence of systematic trends in USE results distribution. Using pre-selected method, the authors are aimed to determine the main characteristics –mathematical expectation and average square deviation of a random variable, necessary to test the hypothesis about the nature of USE results distribution.

Keywords: USE results analysis, statistical analysis, sample analysis, selection characteristics, normal distribution, statistical hypothesis verification, large-volume USE samples, USE results data.

DOI: 10.21667/1995-4565-2025-91-233-240

References

1. The Central limit theorem. *The National scientific and educational portal «The Great Russian Encyclopedia»*: [website]. 2025. URL: <https://bigenc.ru/c/tsentral-naia-predel-naia-teorema-56c807> (accessed 15.01.2025).
2. **Kulaichev A.P.** *Methods and means of complex statistical data analysis : textbook*. A.P. Kulaichev. 5th ed., revised and add. Moscow: INFRA-M, 2025. 484 p. (Higher education). DOI 10.12737/25093.
3. **Khrunichev R.V.** *Applied statistical methods of analysis : textbook. the manual*. R.V. Khrunichev. Ryazan State Radio Engineering un-T. Ryazan, 2023. 80 p.
4. The Government of Russia [Electronic resource]. Decree of the Government of the Russian Federation dated November 19, 2024, no. 3333-р «*Comprehensive action Plan for improving the quality of mathematical and natural science education for the period up to 2030*». Access mode: <http://government.ru/docs/all/156334> / (Date of request: 12/01/2025).