

УДК 65-656.222.6:519.872.8
DOI 10.51955/2312-1327_2025_3_162

АНАЛИЗ РАСПРЕДЕЛЕНИЯ СТАТИСТИЧЕСКИХ ДАННЫХ О ВАГОНОПОТОКАХ. ЧАСТЬ I

Анна Константиновна Мозалевская,
orcid.org/0000-0002-0590-176X,
соискатель

Иркутский государственный университет путей сообщения,
ул. Чернышевского, 15
Иркутск, 664074, Россия
mozalevskay@mail.ru

Екатерина Викторовна Маловецкая,
orcid.org/0000-0003-1549-5336,
кандидат технических наук, доцент
Иркутский государственный университет путей сообщения,
ул. Чернышевского, 15
Иркутск, 664074, Россия
Katerina8119@mail.ru

Роман Сергеевич Большаков,
orcid.org/0000-0002-1187-5932,
кандидат технических наук, доцент
Иркутский государственный университет путей сообщения,
ул. Чернышевского, 15
Иркутск, 664074, Россия
bolshakov_rs@mail.ru

Аннотация. Рассматриваются особенности анализа колебаний вагонопотоков в направлении портов Дальнего Востока с учётом прохождения через стыковые пункты смежных железных дорог, входящих в границы Восточного полигона. При этом поступающий вагонопоток как массив данных представлен в виде временного ряда. Оценивается распределение вагонопотока при помощи статистических гипотез, поскольку подбор вида распределения при работе с временными рядами является важным этапом анализа. Произведена проверка соответствия гипотез о виде распределения с фактическими значениями статистических данных вагонопотоков. Анализируется массив данных эксплуатационных показателей «погрузка» и «выгрузка». Исследованы статистические данные по объему погруженных и выгруженных вагонов, проходящих по Восточному полигону через стыковые пункты по единой методике с оценкой ежемесячных данных за пять лет. Для обработки статистических данных об изменении вагонопотоков и анализе характера их распределения в качестве основного инструмента используется совокупность статистических параметров.

Ключевые слова: погрузка, выгрузка, анализ вагонопотоков, закон распределения, параметрические критерии, непараметрические критерии, критерии нормальности.

ANALYSIS OF DISTRIBUTION OF STATISTICAL DATA ABOUT WAGON FLOWS. PART I

*Anna K. Mozalevskaya,
orcid.org/0000-0002-0590-176X,
applicant*
*Irkutsk State Transport University,
15, Chernyshevskogo
Irkutsk, 664074, Russia
mozalevskay@mail.ru*

*Ekaterina V. Malovetskaya,
orcid.org/0000-0003-1549-5336,
Candidate of technical sciences, docent
Irkutsk State Transport University,
15, Chernyshevskogo
Irkutsk, 664074, Russia
Katerina8119@mail.ru*

*Roman S. Bolshakov,
orcid.org/0000-0002-1187-5932,
Candidate of technical sciences, docent
Irkutsk State Transport University,
15, Chernyshevskogo
Irkutsk, 664074, Russia
bolshakov_rs@mail.ru*

Abstract. The features of the analysis of fluctuations in wagon flows in the direction of the ports of the Far East are considered, taking into account the passage through the junction points of adjacent railways entering the borders of the Eastern polygon. At the same time, the incoming wagon traffic is presented as a time series as an array of data. The distribution of wagon flow is estimated using statistical hypotheses. Selection the type of distribution when working with a volume of data is an important stage of the analysis. The data set of operational indicators "loading" and "unloading" is analyzed. Statistical data on the volume of wagon traffic passing through the Eastern polygon through junction points was studied using a unified methodology with a sample of five years. To process static data on changes in wagon flows and form predictive models, the complex of statistical parameters is used as the main tool.

Key words: loading, unloading, analysis of wagon flows, distribution law, parametric criteria, non-parametric criteria, normality criteria.

Введение

Разработка эффективных транспортных моделей и методик исследований позволяет осуществить научную поддержку транспортной отрасли, особенно в период геополитической и макроэкономической трансформации. Анализ объемов вагонопотоков является важной темой в области транспортной логистики и управления эксплуатационной работой железнодорожного транспорта.

Основные принципы неравномерности движения поездов были заложены в конце 1960 г. А.К. Угрюмовым. Поэтапный переход исследований от вопросов неравномерности вагонопотоков к их прогнозированию произошел в результате эволюции научных исследований и практических разработок в области организации транспортных потоков [Угрюмов, 1968]. Этот процесс

был обусловлен необходимостью более точного и эффективного управления как грузовыми, так и пассажирскими потоками на железнодорожном транспорте.

Основы прогнозирования железнодорожных транспортных потоков были заложены в работах Правдина Н.В. [Правдин и др., 1980] в середине 1980-х годов.

Аналитические исследования вагонопотоков – многогранная область, охватывающая как традиционные методы, так и современные технологии. Научные исследования в период с 1980-х годов показывают эволюцию подходов к прогнозированию, включая учет экономических, социальных и технологических факторов. Научные работы таких авторов, как М.А. Горелик, В.Г. Галабурда, В.Я. Негрей, В.А. Персианов и другие, сформировали основу для дальнейших исследований и практического применения в области железнодорожного транспорта [Гольц, 1982; Горелик, 1982; Правдин и др., 1980].

На сегодняшний день методы прогнозирования основаны на совокупности различных подходов. Это сочетание маркетинговых, балансовых, логистических исследований с использованием информационных технологий.

Проводимые исследования подчеркивают важность интеграции различных подходов и технологий для достижения более точных и надежных прогнозов вагонопотоков.

Напряженная работа холдинга ОАО «РЖД» по складированию, обработке и транспортировке различных видов грузов напрямую связана с возможностями имеющейся инфраструктуры железнодорожного транспорта, поэтому одними из основных стратегических целей компании являются уменьшение размеров издержек инфраструктуры, повышение пропускной способности железнодорожных линий за счёт модернизации её объектов, а также другие мероприятия, направленные на улучшение клиентаориентированности и формирования эффективной логистической составляющей [Бычков и др., 2023; Задачи..., 2023; Кабанов и др., 2022, Лапидус, 2023, Стратегия развития..., 2008; Modeling..., 2021].

Принятие управленческих решений о проведении масштабных реконструктивных мероприятий существующих объектов инфраструктуры сопряжено с проведением трудоемких технологических и экономических расчётов, направленных на оценку эффекта от их внедрения [Бышляго и др., 2018; Технология построения..., 2019; Domojrova et al., 2021; Eremenko et al., 2023]. Однако в большинстве случаев расчёты производятся на основании текущих данных, либо информации за небольшой промежуток времени, что, в свою очередь, не всегда позволяет учесть перспективу изменения вагонопотоков. Для решения такой задачи необходимо использование существующих методов прогнозирования [Маловецкая и др., 2023; Маловецкая и др., 2022; Malovetskaya et al., 2023].

Подходящей методикой для формирования устойчивых прогнозов для железнодорожного транспорта является применение временных рядов [Мамаев и др., 2016]. В некоторых случаях поиск закона распределения исходных

статистических данных является достаточно сложной задачей, требующей использования дополнительных математических средств.

В предлагаемой статье рассматривается анализ статистических данных железнодорожных вагонопотоков для оценки закона распределения.

Материалы и методы

Как известно, математические методы в транспортных перевозках, в том числе железнодорожных, используются для построения моделей перевозочных процессов, в частности, для проверки соответствия аналитических исследований реальным объемам перевозок. В транспортной прогнозной статистике в последние годы получены положительные результаты главным образом благодаря тому, что при проведении исследований имеется представление о специфике транспортных задач, а выбор математического аппарата отвечал специфике рассматриваемого аналитического материала.

Поскольку специфике железнодорожных перевозок отвечает теория вероятностей, авторы статьи для своих исследований широко использовали математическую статистику, являющуюся связующим звеном между статистическими данными (историческими данными изменения основных параметров перевозочного процесса) и аппаратом теории вероятности.

В связи с этим авторами был введен специальный аппарат исследования, позволяющий эффективно изучать специфические черты распределения вагонопотоков при погрузке и выгрузке, извлекать из них необходимую информацию и давать им соответствующее объяснение, удовлетворяющее современным представлениям о транспортной статистике. Указанный аппарат включал в себя обработку наблюдений методами математической статистики с последующей интерпретацией полученных данных на основе синтеза статистических и транспортных положений. При этом теоретические положения транспортного распределения были переведены на язык статистических гипотез, и произведена проверка соответствия этих гипотез фактическим данным. Материалом для исследования послужили статистические данные, накопленные за период с 2018 по 2022 гг. по объему вагонопотока, проходящего по Восточному полигону ОАО «РЖД» по стыковым пунктам. Исследования по каждому стыковому пункту проводились по единой методике с исторической выборкой за пять лет. Трудоёмкие вычислительные операции, связанные с обработкой материала, выполнены в программном пакете Statistica. Для проведения исследований были выбраны железнодорожные стыковые пункты, расположенные в границах Восточного полигона [Маловецкая и др., 2023].

Для оценки ряда распределения были определены основные статистические параметры [Кремер, 2010].

Средняя взвешенная (выборочная средняя) определена выражением:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}. \quad (1)$$

Далее медиана по следующей формуле

$$Me = \frac{x_{\frac{n}{2}} + x_{\frac{n+1}{2}}}{2}. \quad (2)$$

На следующем этапе необходимо определить абсолютные показатели вариации. В частности, размах вариации представляет собой выражение

$$R = x_{\max} - x_{\min}. \quad (3)$$

Далее найдём среднее линейное отклонение:

$$d = \frac{\sum |x_i - \bar{x}| f_i}{\sum f_i}. \quad (4)$$

Дисперсия определяется выражением

$$D = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i}. \quad (5)$$

Формула несмещённой (исправленной) оценки дисперсии имеет вид

$$S^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i - 1}. \quad (6)$$

Найдём среднее квадратическое отклонение.

$$\sigma = \sqrt{D}. \quad (7)$$

Оценка среднеквадратического отклонения примет следующий вид

$$s = \sqrt{S}. \quad (8)$$

Следующим шагом является нахождение относительных показателей вариации, первым из которых является коэффициент вариации

$$v = \frac{\sigma}{\bar{x}}. \quad (9)$$

Далее определим линейный коэффициент вариации (относительное линейное отклонение)

$$Kd = \frac{d}{\bar{x}}. \quad (10)$$

Выражение для коэффициента осцилляции имеет вид

$$Kr = \frac{R}{\bar{x}}. \quad (11)$$

На основании полученных параметров можно определить форму распределения рассматриваемых статистических данных, то есть показатели формы распределения, к которым можно отнести степень асимметрии:

$$A_s = \frac{M_3}{s^3}, \quad (12)$$

где M_3 – центральный момент третьего порядка, s – среднеквадратическое отклонение.

Оценка существенности показателя асимметрии определяется следующим образом:

$$s_{As} = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}}. \quad (13)$$

Выражение для определения структурного коэффициента асимметрии Пирсона имеет вид:

$$A_{SP} = x - M_{o\sigma}. \quad (14)$$

Эксцесс определяем следующим образом:

$$E_x = \frac{M_4}{S^4} - 3. \quad (15)$$

Оценка существенности эксцесса при учёте статистики проводится при помощи выражения

$$\frac{E_x}{S_{Ex}}, \quad (16)$$

где S_{Ex} – средняя квадратическая ошибка коэффициента эксцесса:

$$S_{Ex} = 24n(n-2)(n-3)(n+1)(n+3)(n+5). \quad (17)$$

Выражения для определения вероятности выхода за нижнюю и верхнюю границы распределения случайной величины запишем в виде

$$P(\chi_{n-1}^2 < h_H) = \frac{\gamma}{2}, \quad (18)$$

$$P(\chi_{n-1}^2 \geq h_B) = 1 - P(\chi_{n-1}^2 < h_H). \quad (19)$$

Число степеней свободы массива данных определим по следующей формуле

$$k = n - 1. \quad (20)$$

Далее запишем выражения для определения случайных ошибок дисперсии нижней и верхней границы распределения соответственно:

$$t_H = \frac{(n-1)S^2}{h_H}, \quad (21)$$

$$t_B = (n-1) \frac{S^2}{h_H}. \quad (22)$$

После чего необходимо определить доверительный интервал для среднеквадратического отклонения случайной величины

$$S^*(1-q) < \sigma < S^*(1+q). \quad (23)$$

Выражение для идентификации доверительного интервала для генеральной доли распределения имеет вид

$$2\Phi(t_{kp}) = \gamma. \quad (24)$$

После преобразования (24) найдём

$$\Phi(t_{kp}) = \frac{\gamma}{2}. \quad (25)$$

Представленный математический аппарат позволяет проанализировать исследуемые данные колебаний вагонопотоков с установлением законов распределения для дальнейшего построения адекватных прогнозных моделей.

Дискуссия

В своём исследовании авторы исходили из того положения, что решение рассматриваемого вопроса зависит не только от получения объективных сведений о закономерностях изменения объемов проходящего вагонопотока во времени, но и от успешного извлечения из него необходимой информации о распределении транспортных потоков.

На примере показателя «Выгрузка» одной из рассматриваемых железных дорог за исторический период с 2018 по 2022 год оценивается статистическая обработка результатов измерений с использованием вариационного ряда величин значений с учётом критериев согласия и надежности оценки, основанной на вероятностных характеристиках [Мамаев и др., 2016].

Краеугольным камнем транспортной статистики по праву считается оценка нормальности распределения [Ghosh, 2020]. В статистике тесты на нормальность используются для того, чтобы определить, насколько хорошо моделируется набор данных в исследовании нормальным распределением, и рассчитать вероятность того, что случайная переменная, лежащая в основе набора данных, будет нормально распределена [Razali et al., 2011].

Погрузка и выгрузка всех видов грузов суммарно выражается в ежемесячном количестве вагонов X , прошедших через стыковой пункт на рассматриваемой железной дороге. Объем выборки составил $n = 60$ (12 календарных месяцев за период 5 лет). В работе было исследовано 4 стыковых пункта железных дорог Восточного полигона, анализ массивов данных проводился отдельно по эксплуатационным показателям – «Погрузка» и «Выгрузка».

Выбор метода статистического анализа возможен только после оценки характеристик распределения данных и сравнения вариабельности в выборках. Авторами были выбраны параметрические методы статистического анализа, что соответствует первоначальному предположению, что массив данных соответствует нормальному закону распределения. В анализе было выбрано использование параметрических критериев (t -критерий Стьюдента, критерий Пирсона χ^2 и другие), так как они обладают большей значимостью по сравнению с непараметрическими методами и позволяют уменьшить вероятность ошибки второго рода.

Результаты

На основании известных положений по определению статистических показателей найдём их значения [Кремер, 2010]. Используются следующие статистические параметры: x_i – количество выгруженных вагонов за каждый месяц рассматриваемого периода, f_i – частота появления события, S – накопленная частота, x_{cp} – средневзвешенное значение количества выгруженных вагонов, f – относительная частота.

$$\text{Средняя взвешенная (выборочная средняя): } \bar{x} = \frac{1855665}{60} = 30927.8.$$

$$\text{Медиана: } x_i = 31636.$$

Найдём численное значение размаха вариации: $R = 35935 - 24531 = 11404$.

По формуле (4) определим среднее линейное отклонение:

$$d = \frac{143638}{60} = 2394.$$

Таким образом, отличие случайных величин друг от друга варьируется в среднем на 2394.

При помощи выражения (5) найдём значение дисперсии:

$$D = \frac{485594753.25}{60} = 8093245.888.$$

Её несмещенную оценку определим с учётом выражения (6):

$$S = \frac{485594753.25}{59} = 8230419.547$$

Далее определим среднее квадратическое отклонение.

$$\sigma = \sqrt{8093245.888} = 2844.863,$$

которое показывает отличие случайных величин ряда от средних показателей 30927.8 в среднем на 2844.863.

Далее проведём оценку среднеквадратического отклонения по формуле (8):

$$s = \sqrt{8230419.547} = 2868.871.$$

$$\text{Коэффициент вариации: } v = \frac{2844.863}{30927.8} 100\% = 9.2\%.$$

Таким образом, достоверность результата подтверждается значением v меньше 30% при однородности свойств структуры данных со слабой вариацией.

Определим относительное линейное отклонение:

$$Kd = \frac{2394}{30927.8} 100\% = 7.74\%.$$

$$\text{Коэффициент осцилляции: } Kr = \frac{11404}{30927.8} 100\% = 36.87\%.$$

$$M_3 = -\frac{477353514521.13}{60} = -7955725242.02,$$

$$A_s = \frac{-7955725242.02}{(2844.863)^3} = -0,346.$$

Имеет место наличие левосторонней асимметрии.

$$s_{As} = \sqrt{\frac{6(60-2)}{(60+1)(60+3)}} = 0.301.$$

Распределение обладает несущественной асимметрией ($\frac{-0.346}{0.301} = 1.15 < 3$)

Структурный коэффициент асимметрии Пирсона:

$$A_{sp} = 30927.8 - 316362844.863 = -0.25.$$

Эксцесс в этом случае равен: $E_x = \frac{M_4}{s^4} - 3$.

При $E_x < 0$ распределение является плосковершинным относительно нормального:

$$M_4 = 9.0201748664926E+15/60 = 1.5033624777488E+14 ,$$

$$E_x = 1.5033624777488E+142844.8634 - 3 = 2.2952 - 3 = -0.7 .$$

Существенность эксцесса оценивается при помощи статистики $\frac{E_x}{s_{Ex}}$.

Если полученное значение > 3 , то это указывает на существенное отклонение от нормального закона распределения.

$$s_{Ex} = 24 \cdot 60(60-2)(60-3)(60+1)2(60+3)(60+5) = 0.559$$

$$\frac{E_x}{s_{Ex}} = \frac{-0.7}{0.559} = -1.252$$

В данном случае отклонение не существенно.

Возможность выхода за нижнюю границу распределения определим следующим образом

$$P(\chi^2_{n-1} < h_H) = 0.023 .$$

Далее в соответствии с выражениями (21) – (24) найдём:

$$t_H = 59 \cdot 2868.871283.29768 = 5829631.19 ,$$

$$t_B = 59 \cdot 2868.871237.48485 = 12954427.01 ,$$

$$\chi^2(59; 0.023) = 83.29768 ,$$

$$\chi^2(59; 0.977) = 37.48485 .$$

Таким образом надежность покрытия S_2 составляет $\alpha = 0.046$ ($\gamma = 95.4\%$).

Доверительный интервал $q = q(0.954; 60)$ определится следующим образом

$$2868.871(1 - 0.188) < \sigma < 2868.871(1 + 0.188)$$

$$2329.523 < \sigma < 3408.219 ,$$

что показывает то, что надежность покрытия параметра σ $\gamma = 0.954$.

Определим доверительный интервал для генеральной доли распределения случайной величины

$$\Phi(t_{kp}) = \frac{0.954}{2} = 0.477 .$$

После чего по таблице функций Лапласа определим значение $\Phi(t_{kp}) = 0.477$ с учётом t_{kp}

$$t_{kp}(\gamma) = (0.477) = 2 .$$

Таким образом, вероятность равенства долей будет находиться в обозначенных интервалах с вероятностью 0.954.

Выполнены тесты нормальности, доступные в статистическом программном пакете «Statistica». По результатам анализа определены доступные тесты нормальности в данном программном обеспечении:

графические методы (построение гистограммы и Q-Q графика), критерий Колмогорова-Смирнова и критерий χ^2 Пирсона (рис. 1).

Среднее	Медиана	Минимум	Максим.	Дисперсия	Ст.откл.	Коэф.Вар.	Асимметрия	Стд.ош. Асимметрия	Эксцесс	Стд.ош. Эксцесс
30927,75	31364,50	24531,00	35935,00	8230420	2868,871	9,276041	-0,354462	0,308694	-0,660199	0,608492

Рисунок 1 – Общие описательные статистические показатели массива данных «Выгрузка» на одной из рассматриваемых железных дорог, полученные с использованием программы «Statistica»

Анализ критериев описательной статистики показывает, что значения средней взвешенной и медианы примерно равны, что говорит о симметричности ряда. Коэффициент вариации меньше 30%, что тоже свидетельствует о нормальном распределении. Асимметрия отрицательная, следовательно, имеется сдвиг гистограммы в сторону больших по величине значений, но величина сдвига незначительна. В этом случае асимметрия незначительная, что указывает на нормальное распределение. Эксцесс также имеет малое отрицательное значение, что свидетельствует о плосковершинности распределения. Также график «Огива-Гальтона» (рис. 2) имеет плавное очертание, что дает возможность сделать предположения о нормальности распределения.

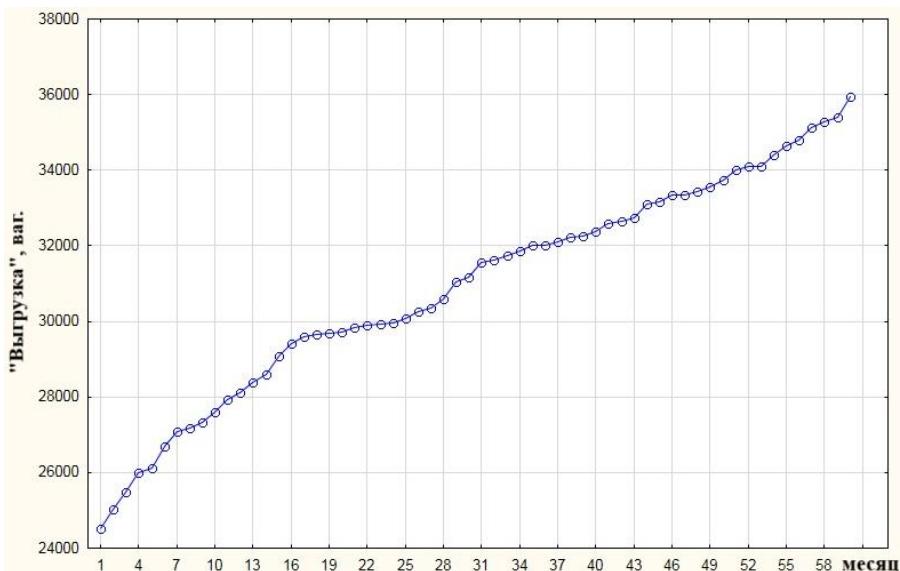


Рисунок 2 – График «Огива-Гальтона» «Выгрузка» на рассматриваемой железной дороге

Из графика, представленного на рисунке 3, видно, что распределение похоже на кривую нормального распределения со смещением влево.

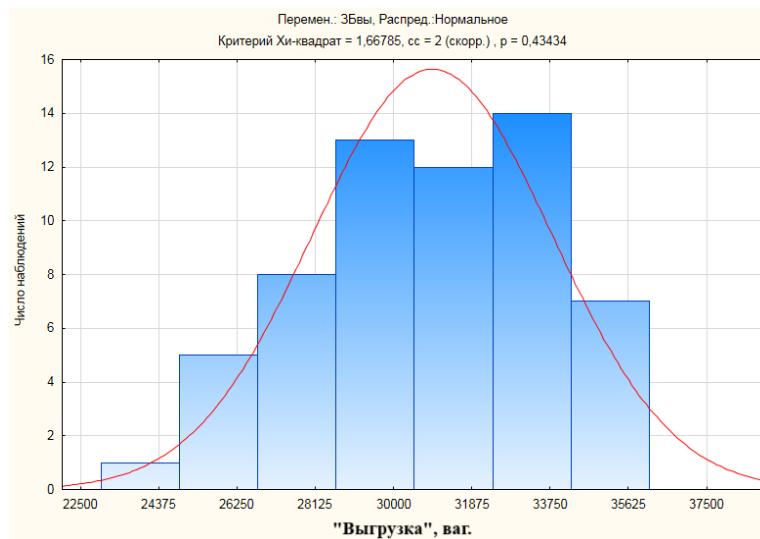


Рисунок 3 – Гистограмма расчета критерия χ^2 -квадрат «Выгрузка» на железной дороге

Графические методы субъективны и не могут предоставить убедительных доказательств того, что предположение о нормальности закона распределения выполняется. Поэтому необходимы численные статистические тесты, чтобы подтвердить вывод, сделанный с помощью графических методов [Yap et al., 2011; Yazici et al., 2007].

В исследовании размер выборки $n=60$ достаточен для применения теста Колмогорова-Смирнова (без поправки Лиллиефорса) и критерия χ^2 . В связи с этим необходима оценка нормальности распределения по критерию согласованности статистического и теоретического распределений χ^2 Пирсона:

$$\chi^2 = n \sum_{i=1}^l \frac{(w_i - p_i)^2}{p_i}, \quad (26)$$

где w_i – относительные частоты; p_i – вероятности, полученные по некоторому теоретическому закону распределения.

В данном случае $\chi^2 = 1,67$, вероятность нулевой гипотезы $p = 0,4343$. Минимальный стандартный уровень значимости обычно принимается на уровне 5%. Результаты статистического анализа показывают, что уровень значимости равен 43%, что свидетельствует об истинности гипотезы, а по критерию Пирсона распределение соответствует нормальному.

Верхняя Граница	Перемен.: ЗБвы, Распред.:Нормальное (Книга1 ПОГРУЗКА-ВЫГРУЗКА) Хи-квадрат = 1,66785, cc = 2 (скорр.) , p = 0,43434								
	Наблюд. Частота	Кумул. Наблюд.	Процент Наблюд.	Кумул. % Наблюд.	Ожидаем. Частота	Кумул. Ожидаем.	Процент Ожидаем.	Кумул. % Ожидаем.	Наблюд.- Ожидаем.
<= 24875,00000	1	1	1,66667	1,6667	1,04626	1,04626	1,74377	1,7438	-0,04626
26750,00000	5	6	8,33333	10,0000	3,31357	4,35983	5,52261	7,2664	1,68643
28625,00000	8	14	13,33333	23,3333	8,30517	12,66500	13,84196	21,1083	-0,30517
30500,00000	13	27	21,66667	45,0000	13,77923	26,44423	22,96538	44,0737	-0,77923
32375,00000	12	39	20,00000	65,0000	15,13777	41,58200	25,22961	69,3033	-3,13777
34250,00000	14	53	23,33333	88,3333	11,01250	52,59450	18,35416	87,6575	2,98750
36125,00000	7	60	11,66667	100,0000	5,30407	57,89857	8,84012	96,4976	1,69593
< бесконеч.	0	60	0,00000	100,0000	2,10143	60,00000	3,50239	100,0000	-2,10143

Рисунок 4 – Таблица расчета критерия χ^2 «Выгрузка» на рассматриваемой железной дороге

Далее проведём анализ нормальности распределения по критерию Колмогорова. Количество интервалов в гистограмме принимаем $\sqrt{60} = 7,75 \approx 8$.

Критерий Колмогорова равен:

$$D_{\text{факт}} = 0,0879 \leq D_{0,05} = \frac{1,36}{\sqrt{60}} = 0,1755.$$

Так как наблюдаемое значение меньше значения критического, принимается нулевая гипотеза. На рисунке 5 приведена гистограмма с учётом проверки по критерию Колмогорова.

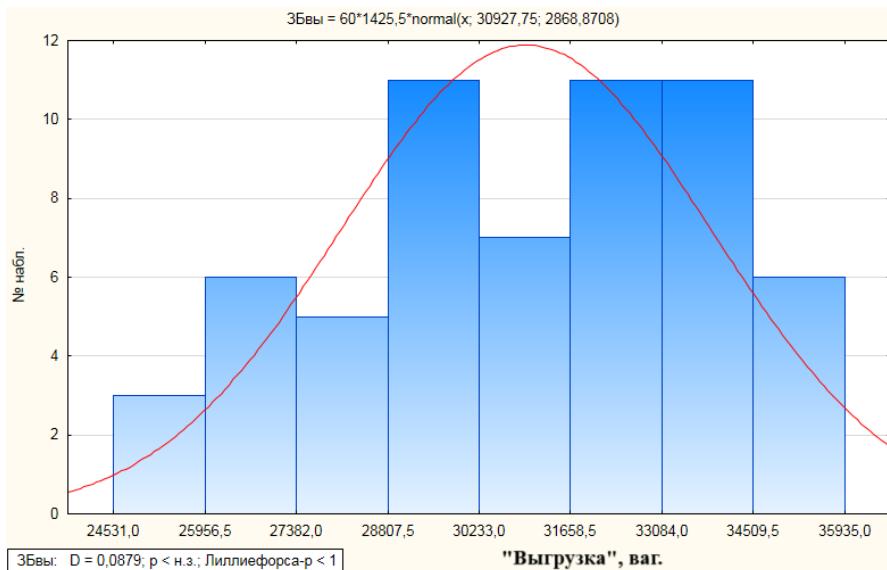


Рисунок 5 – Гистограмма параметра «Выгрузка» вагонов на стыковом пункте на рассматриваемой железной дороге

Установлено, что по критерию Колмогорова и по критерию χ^2 Пирсона, рассматриваемое распределение является нормальным, что подтверждает предположение, выдвинутое по косвенным наблюдениям (с учётом описательной статистики).

Показатели «выгрузки» вагонов на стыковом пункте рассматриваемой железной дороги подчиняются нормальному закону распределения с плотностью вероятности

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (27)$$

где Mx и σ^2 – параметры распределения (математическое ожидание и дисперсия).

Таким образом, анализ значений полученных статистических параметров удовлетворяет критериям, необходимым для признания рассмотренного распределения в нормальных пределах.

Заключение

В исследовании проведен анализ данных поступления вагонопотоков на стыковые пункты Восточного полигона за период пять лет.

При этом поступающий вагонопоток рассматривается в качестве временного ряда. Рассматриваемый объём данных является достаточно специфичной совокупностью параметров, характерных для железнодорожного транспорта. В ходе проведения исследований определены статистические параметры массива данных погрузки железнодорожных вагонопотоков, представлен пошаговый алгоритм определения принадлежности распределения опытных данных соответствующему закону распределения. Оценивается распределение грузопотока при помощи статистических гипотез, поскольку подбор вида распределения при работе с временными рядами является важным этапом исследования.

Был осуществлен анализ статистических данных по всем стыковым пунктам Восточного полигона. В представленной статье отражен ход исследований на примере массива данных по одному стыковому пункту. По итогам анализа было установлено, что не по всем выборкам подтверждается гипотеза о нормальном законе распределения. Нулевая гипотеза не подтвердилась по параметру «погрузка», о чем свидетельствует ряд характерных особенностей. Поскольку параметры рассматриваемых стыковых пунктов достаточно различны в связи с исторически сложившейся спецификой работы (география расположения, количество примыканий, структура и объемы входящего и исходящего вагонопотока) и специфичны, это предполагает дальнейшее изучение и анализ влияния различных условий на характер процессов, управляющих распределением объемов пропуска (обмена) грузовых поездов между железными дорогами.

Правильный выбор распределения позволяет более точно в дальнейшем моделировать поведение временного ряда и выявлять его закономерности. Разные модели и методы прогнозирования требуют различных предположений о распределении данных. На последующих этапах исследования, при использовании ARIMA или GARCH моделей важно знать, какое распределение лучше всего описывает ошибки модели. С этой целью во второй части научной статьи будет проведен поиск функции распределения для стыковых пунктов, не прошедших проверки на нулевую гипотезу. Также будут рассмотрены практические применения данных заключений о законах распределения при построении адекватных моделей прогноза вагонопотоков.

Библиографический список

- Бычков И. В.* Применение теории массового обслуживания для моделирования Улан-Баторской железной дороги / И. В. Бычков, М. Л. Жарков, А. Л. Казаков // Вычислительные технологии. 2023. Т. 28, № 6. С. 17-36. DOI 10.25743/ICT.2023.28.6.003. EDN JKTSVJ.
- Бышляго А. А.* Об организации пропуска соединенных поездов в целях повышения провозной и пропускной способностей Байкало-Амурской магистрали / А. А. Бышляго, А. В. Дудакова // Молодая наука Сибири. 2018. № 1(1). С. 1-7. EDN YCOKST.
- Гольц Г. А.* Исследование закономерностей развития перевозок в СССР. Проблемы прогнозирования и оптимизации работы транспорта. М.: Наука, 1982. С. 23-63.
- Горелик М. А.* Прогнозирование перспективных грузопотоков на морском транспорте // Проблемы прогнозирования и оптимизации работы транспорта. М.: Наука, 1982. С. 63-78.
- Задачи и перспективы развития научных исследований в рамках сотрудничества между ОАО

«РЖД» и Российской академией наук / Н. А. Махутов, Б. М. Лапидус, М. М. Гаденин, Е. Ю. Титов // Железнодорожный транспорт. 2023. № 7. С. 6-11. EDN ZAFGPF.

Кабанов А. Б. Научный подход к планированию железнодорожных грузовых перевозок / А. Б. Кабанов, А. Т. Осьминин // Железнодорожный транспорт. 2022. № 8. С. 12-16. EDN OZHUWT.

Кремер Н. Ш. Теория вероятностей и математическая статистика. 3-е изд., перераб. и доп. М.: ЮНИТИ-ДАНА, 2010. 551 с.

Кроль Н. В. Технология построения оптимального маршрута при организации мультимодальных пассажирских перевозок с учетом выбора места пребывания / Н. В. Кроль, А. С. Полетаев, Р. Ю. Упры, А. В. Дудакова // Современные технологии. Системный анализ. Моделирование. 2019. № 2(62). С. 109-118. DOI 10.26731/1813-9108.2019.2(62).109-118. EDN HZCFRZ.

Лапидус Б. М. Задачи опережающего развития российских железных дорог // Железнодорожный транспорт. 2023. № 2. С. 4-14. EDN HYHML.

Маловецкая Е. В. Возможности повышения эффективности перевозочного процесса на основе построения комплексных прогнозных моделей загрузки инфраструктуры / Е. В. Маловецкая, А. К. Мозалевская // T-Comm: Телекоммуникации и транспорт. 2023. Т. 17, № 7. С. 38-46. DOI 10.36724/2072-8735-2023-17-7-38-46. EDN IAEWSG.

Маловецкая Е. В. Оценка влияния неравномерностей на перевозочный процесс / Е. В. Маловецкая, А. К. Мозалевская // Проблемы безопасности на транспорте: Материалы XII Международной научно-практической конференции, посвященной 160-летию Белорусской железной дороги. В 2-х частях, Гомель, 24–25 ноября 2022 года. Гомель: Белорусский государственный университет транспорта, 2022. С. 207-209. EDN BRTMRC.

Мамаев Э. А. Прогнозирование контейнерооборота морских портов мира / Э. А. Мамаев, Е. С. Вальчук // Вестник Ростовского государственного университета путей сообщения. 2016. № 2(62). С. 94-101. EDN WFEHQZ.

Правдин Н. В. К вопросу построения теоретических моделей колебаний транспортных потоков / Н. В. Правдин, В. Я. Негрей, А. И. Шелеляев // Проблемы перспективного развития железнодорожных станций и узлов: Межвузовский сборник научных статей / Белорусский институт инженеров железнодорожного транспорта. Гомель: Белорусский институт инженеров железнодорожного транспорта, 1980. С. 3-20. EDN YNHDQN.

Правдин Н. В. Прогнозирование пассажирских потоков / Н. В. Правдин, В. Я. Негрей. М.: Транспорт, 1980. 224 с.

Стратегия развития железнодорожного транспорта в Российской Федерации до 2030 года: от 17 июня 2008 г. № 877-р: утверждена распоряжением Правительства Российской Федерации. – 2008. – 17 июня. Доступно по адресу: <https://mintrans.gov.ru/documents/1/1010> (доступно 30 апреля 2025).

Угрюмов А. К. Неравномерность движения поездов. М.: Транспорт, 1968. 112 с.

Domojrova A. Predicting for train movement based on historical statistical data / A. Domojrova, R. Upur, M. Eremenko // Journal of Physics: Conference Series: International Conference on IT in Business and Industry, ITBI 2021, Novosibirsk, 12–14 мая 2021 года. Vol. 2032. IOP Publishing Ltd: IOP Publishing Ltd, 2021. P. 012036. DOI 10.1088/1742-6596/2032/1/012036. EDN QBLNGM.

Eremenko M. N. Stability of train traffic in the event of failures and the mechanism of phantom traffic jams on the railway / M. N. Eremenko, R. Yu. Upur, A. D. Domojrova // International Scientific and Practical Conference "Railway Transport and Technologies" (RTT-2021) : Collection of conference materials. Volume 2624, Ekaterinburg, 24–25 ноября 2021 года. Vol. 2624, Issue 1. USA: AIP PUBLISHING, 2023. P. 040018. DOI 10.1063/5.0133388. EDN PTKFMQ.

Ghosh S. Importance of Normality Testing, Parametric and Non-Parametric Approach, Association, Correlation and Linear Regression (Multiple & Multivariate) of Data in Food & Bio-Process Engineering // In: Mathematical and Statistical Applications in Food Engineering. CRC Press; 2020. P. 112-126. DOI 10.1201/9780429436963-9.

Malovetskaya E. Prospects for Comprehensive Forecasts When Assessing the Load of Railway Transport Infrastructure / E. Malovetskaya, E. Voskresenskaya, A. Mozalevskaya // Finance, Economics, and Industry for Sustainable Development (ESG 2023) : Proceedings of the 4th International Scientific Conference on Sustainable Development, St. Petersburg, 19–20 октября 2023 года. Cham: Springer Nature Switzerland AG, 2024. P. 217-225. EDN GLELMN.

Modeling of railway stations based on queuing networks / I. Bychkov, A. Kazakov, A. Lempert, M. Zharkov // Applied Sciences (Switzerland). 2021. Vol. 11, No. 5. DOI 10.3390/app11052425. EDN UUVMVI.

Razali N. M. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests / N. M. Razali, Y. B. Wah // Journal of Statistical Modeling and Analytics. 2011. № 2(1). Pp. 21-33.

Yap B. W. Comparisons of various types of normality tests / B. W. Yap, C. H. Sim // Journal of Statistical Computation and Simulation. 2011. № 81(12). Pp. 2141-2155. DOI 10.1080/00949655.2010.520163.

Yazici B. A comparison of various tests of normality / B. Yazici, S. Yolacan // Journal of Statistical Computation and Simulation. 2007. № 77(2). Pp. 175-183. DOI 10.1080/10629360600678310.

References

- Bychkov I. V., Kazakov A. L., Lempert A. A., Zharkov M. L.* (2021). Modeling of railway stations based on queuing networks. *Applied Sciences (Switzerland)*. 5(11). DOI 10.3390/app11052425.
- Bychkov I. V., Zharkov M. L., Kazakov A. L.* (2023). Application of queuing theory for modeling the Ulaanbaatar railway. *Computational technologies*. 6(28): 17-36. (In Russian)
- Byshlyago A. A., Dudakova A. V.* (2018). On the organization of the passage of connected trains in order to increase the carrying and throughput capacities of the Baikal-Amur Mainline. *Young Science of Siberia*. 1(1): 1-7.
- Domojirova A. D., Upyr R. Yu., Eremenko M. N.* (2021). Predicting for train movement based on historical statistical data. *Journal of Physics: Conference Series*, Novosibirsk. 012036. DOI 10.1088/1742-6596/2032/1/012036.
- Eremenko M. N., Upyr R. Yu., Domojirova A. D.* (2021). Stability of train traffic in the event of failures and the mechanism of phantom traffic jams on the railway. *International Scientific and Practical Conference "Railway Transport and Technologies" (RTT-2021)*. (1)2624: 040018. DOI 10.1063/5.0133388.
- Ghosh S, Mitra J.* (2020). Importance of Normality Testing, Parametric and Non-Parametric Approach, Association, Correlation and Linear Regression (Multiple & Multivariate) of Data in Food & Bio-Process Engineering. In: *Mathematical and Statistical Applications in Food Engineering*. CRC Press; 2020; 112-126.
- Golts G. A.* (1982). Studying the patterns of transport development in the USSR. *Problems of forecasting and optimization of transport operations*. Moscow: Nauka, 1982. pp. 23-63. (In Russian)
- Gorelik M. A.* (1982). Forecasting promising cargo flows in maritime transport. *Problems of forecasting and optimization of transport operations*. Moscow: Nauka, 1982. pp. 63-78.
- Kabanov A. B., Osminin A. T.* (2022). Scientific approach to planning rail freight transportation. *Railway transport*. 18: 12-16. (In Russian)
- Kremer N. Sh.* (2010). Probability Theory and Mathematical Statistics. 3rd ed., revised and enlarged. Moscow: UNITY-DANA, 2010. 551 p. (In Russian)
- Krol N. V., Poletaev A. S., Upyr R. Yu., Dudakova A. V.* (2019). Technology for constructing an optimal route when organizing multimodal passenger transportation, taking into account the choice of place of stay. *Modern technologies. Systems analysis. Modeling*. 2 (62): 109-118. DOI 10.26731/1813-9108.2019.2(62).109-118.
- Lapidus B. M.* (2023). Tasks of advanced development of Russian railways. *Railway transport*. 2: 4-14.

- Makhutov N. A., Lapidus B. M., Gadenin M. M., Titov E. Yu.* (2023). Tasks and Prospects for the Development of Scientific Research within the Framework of Cooperation between Russian Railways and the Russian Academy of Sciences. *Railway Transport*. 7: 6-11.
- Malovetskaya E. V., Mozalevskaya A. K.* (2022). Assessing the Impact of Irregularities on the Transportation Process. *Problems of Transport Safety: Proceedings of the XII International Scientific and Practical Conference Dedicated to the 160th Anniversary of the Belarusian Railway*. In 2 parts, Gomel, November 24–25, 2022. Gomel: Belarusian State University of Transport. 207-209.
- Malovetskaya E. V., Mozalevskaya A. K.* (2023). Possibilities of Improving the Efficiency of the Transportation Process Based on the Construction of Comprehensive Forecast Models of Infrastructure Loading. *T-Comm: Telecommunications and Transport*. 7(17): 38–46. DOI 10.36724/2072-8735-2023-17-7-38-46.
- Malovetskaya E. V., Voskresenskaya E. V., Mozalevskaya A. K.* (2023). Prospects for comprehensive forecasts when assessing the load of railway transport infrastructure. *Finance, Economics, and Industry for Sustainable Development (ESG 2023)*. 217-225.
- Mamaev E. A., Valchuk E. S.* (2016). Forecasting container turnover of world seaports. *Bulletin of the Rostov State University of Railway Engineering*. 2 (62): 94-101.
- Pravdin I. V., Negrey V. Ya.* (1980). Forecasting passenger flows. Moscow: *Transport*, 1980. 224 p.
- Pravdin N. V., Negrey V. Ya., Shelelyaev A. I.* (1980). On the issue of constructing theoretical models of fluctuations in traffic flows. *Problems of perspective development of railway stations and junctions*. 3-20.
- Razali N. M., Wah Y. B.* (2011). Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics*. 2(1): 21-33.
- Strategy for the Development of Railway Transport in the Russian Federation until 2030: of June 17, 2008 No. 877-r: approved by the order of the Government of the Russian Federation (2008). Available at: <https://mintrans.gov.ru/documents/1/1010> (accessed 30 April 2025).
- Ugryumov A. K.* (1968). The unevenness of train movement. Moscow: *Transport*, 1968. 112 p.
- Yap B. W., Sim C. H.* (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*. 81(12): 2141-2155.
- Yazici B., Yolacan S.* (2007). A comparison of various tests of normality. *Journal of Statistical Computation and Simulation*. 77(2): 175-183.