УДК 338.27

***С.Н. Ладыженко, А.А. Жарич, Н.Н. Терещенко, А.В. Волошин***
Институт торговли и сферы услуг, Сибирский федеральный университет, Красноярск,
email: avvoloshin@sfu-kras.ru

# ПРОГНОЗИРОВАНИЕ ДОХОДОВ ПРЕДПРИЯТИЙ ГОСТЕПРИИМСТВАКРАСНОЯРСКОГОКРАЯСИСПОЛЬЗОВАНИЕМ МНОГОФАКТОРНОГО РЕГРЕССИОННОГО АНАЛИЗА

**Ключевые слова:** прогнозирование, многофакторный регрессионный анализ, сфера гостеприимства, доходы предприятий гостеприимства, экономическое моделирование, регрессионная модель, экономика региона, Красноярский край.

В работе исследуются ключевые факторы, влияющие на доходы предприятий гостеприимства Красноярского края. Для достижения поставленной цели был использован многофакторный регрессионный анализ данных с применением программного обеспечения Minitab. В статье рассмотренная получившаяся регрессионная модель, способная с достаточной точностью прогнозировать выбранный фактор.

***S.N. Ladyzhenko, A.A. Zharich, N.N. Tereshchenko, A.V. Voloshin***
Institute of Trade and Services, Siberian Federal University, Krasnoyarsk,
email: avvoloshin@sfu-kras.ru

# FORECASTING THE INCOME OF HOSPITALITY ENTERPRISES IN THE KRASNOYARSK REGION USING MULTIFACTOR REGRESSION ANALYSIS

**Keywords:** forecasting, multivariate regression analysis, hospitality industry, income of hospitality enterprises, economic modeling, regression model, regional economy, Krasnoyarsk region.

The paper examines the key factors influencing the income of hospitality enterprises in the Krasnoyarsk Territory. To achieve this goal, multivariate regression data analysis was used using Minitab software. The article examines the resulting regression model, which is capable of predicting the selected factor with sufficient accuracy.

Сфера гостеприимства представляет собой одну из ключевых отраслей экономики, оказывающую значительное влияние на социально-экономическое развитие стран и регионов. Она охватывает широкий спектр услуг, включая гостиничное и ресторанное обслуживание, туризм, организацию мероприятий и другие виды деятельности, связанные с приёмом и обслуживанием гостей. Важность этой отрасли заключается не только в создании рабочих мест и улучшении качества жизни, но и в её вкладе в развитие инфраструктуры, привлечение инвестиций и стимулирование роста смежных отраслей, таких как транспорт и розничная торговля.

Доходы предприятий гостеприимства составляют значительную долю в структуре национальных и региональных экономик, что свидетельствует о важности этой отрасли для экономической стабильности. Эти доходы оказывают влияние на уровень занятости, налоговые поступления и благосостояние местных сообществ. Особенно это проявляется в странах с развитой туристической отраслью, где её вклад в валовый внутренний продукт может достигать значительных величин.

Прогнозирование доходов предприятий гостеприимства имеет важное значение не только для внутренних нужд самих организаций, но и для экономики в целом. Точные прогнозы позволяют государственным органам и аналитическим центрам эффективно планировать бюджетные поступления, разрабатывать стратегии развития туристической отрасли и поддерживать развитие необходимой инфраструктуры. Более того, прогнозирование доходов позволяет оценить возможное воздействие внешних факторов, таких как экономические кризисы, политическая нестабильность или природные катастрофы, на данный

сектор, что способствует выработке мер по минимизации рисков и обеспечению устойчивости отрасли в условиях неопределённости.

## Цель исследования

Целью данного исследования является анализ факторов, влияющих на доходы предприятий гостеприимства, а также разработка модели прогнозирования этих доходов с учётом экономических, социальных и внешних факторов. Это позволит улучшить стратегическое планирование и повысить устойчивость сектора гостеприимства к изменениям внешней среды.

## Материалы и методы исследования

Материалами исследования стали данные официальной статистики, предоставленные Федеральной службой государственной статистики России (Росстат), включающие показатели доходов предприятий гостеприимства, а также информацию о влияющих на них экономических и социальных факторах. Для анализа использовался регрессионный многофакторный анализ, проведённый с использованием программного обеспечения Minitab. Этот метод позволяет оценить влияние нескольких независимых переменных на зависимую переменную, что способствует точному прогнозированию доходов предприятий гостеприимства с учётом различных факторов.

## Результаты исследования и их обсуждение

Исследование было организовано в несколько последовательных этапов, каждый из которых был направлен на достижение поставленных целей, обеспечивая системный подход к анализу и прогнозированию доходов предприятий гостеприимства.

На первом этапе было осуществлено комплексное сбор данных, охватывающих период с 2013 по 2023 годы. В качестве материалов исследования использовалась официальная статистика, предоставленная Федеральной службой государственной статистики России (Росстат). В процессе сбора данных были учтены не только непосредственно экономические показатели, такие как доходы предприятий гостеприимства, но и различные социально-экономические факторы, которые могли оказывать влияние на данный сектор. Принцип сбора данных заключался в учёте даже тех факторов, которые на первый взгляд не имели бы прямого отношения к прогнозируемому показателю. Этот подход был выбран с учетом того, что зависимость между факторами может быть сложной и не всегда очевидной, а скрытые взаимосвязи между переменными могут существенно повлиять на точность прогноза [2,3].

На втором этапе исследования был проведён детализированный анализ собранных данных с использованием программного обеспечения Minitab, которое позволило провести ряд статистических процедур для выявления ключевых факторов, влияющих на доходы предприятий гостеприимства. В первую очередь был использован анализ временных рядов, который позволил оценить изменения показателей во времени и выявить долгосрочные тренды. Также был проведён анализ автокорреляции, который позволил оценить зависимость между значениями показателей на различных временных интервалах. Для более глубокого понимания структуры данных был применён метод автокорреляции остатков с использованием метода разностей, который позволил выявить возможные скрытые зависимости и исключить мультиколлинеарность между факторами. Важным этапом второго шага было построение корреляционной матрицы, которая служила для определения силы и направленности связи между различными факторами и целевым показателем (доходами предприятий гостеприимства). Это позволило на основе статистических данных выделить те переменные, которые оказывают наибольшее влияние на прогнозируемый результат [1,5].

На третьем этапе анализа была проведена фильтрация факторов, которые не оказывали значительного влияния на модель. Из анализа были исключены те переменные, чья корреляция с доходами предприятий гостеприимства оказалась статистически незначимой, а также те, чье влияние было негативным или противоречило логике модели. Данный этап был необходим для уточнения структуры модели и повышения

её адекватности реальным условиям. Исключение нерелевантных факторов позволило сузить круг переменных, которые в дальнейшем использовались для построения итоговой модели, что также способствовало улучшению точности прогнозирования.

На заключительном, четвёртом этапе было построено математическое моделирование на основе оставшихся значимых факторов. Для этого была использована методология многомерного регрессионного анализа, что позволило учесть взаимодействие различных переменных и выявить их влияние на доходы предприятий гостеприимства в условиях многозадачности. Построенная модель была проверена на её адекватность с помощью различных статистических критериев, таких как тесты на значимость коэффициентов, анализ остатков и проверка на мультиколлинеарность. Результаты этого этапа позволили получить стабильную и достоверную модель, которая способна эффективно прогнозировать доходы предприятий гостеприимства на основе актуальных факторов, что имеет важное значение для дальнейших исследований и практического применения в области экономического планирования и прогнозирования [9].

В рамках настоящего исследования была построена и оценена регрессионная модель, описывающая зависимость целевой переменной ICAS (доходы коллективных средств размещения от предоставляемых слуг) от пяти независимых факторов: NOSCA (число ночёвок в коллективных средствах размещения), GRP (ВРП Красноярского края), GRPPC (ВРП Красноярского края на душу), AS (средняя заработная плата) и DER (курс доллара). Приведенные результаты позволяют всесторонне оценить качество модели, определить степень влияния каждого фактора, а также выявить проблемы, которые могут ограничивать применимость полученных выводов.

Основным показателем для оценки значимости регрессионной модели является F-критерий, полученный в результате анализа дисперсии (ANOVA). В данной модели значение F-критерия составило 113,80, что является весьма высоким значением, а соответствующее p-значение равно 0,000. Это указывает на то, что модель статистически значима на любом традиционном уровне значимости (например, 1%, 5% или 10%). Данный результат позволяет утверждать, что совокупность факторов NOSCA, GRP, GRPPC, AS и DER вносит значимый вклад в объяснение вариации зависимой переменной ICAS. Это означает, что выбранные независимые переменные в совокупности эффективно описывают динамику ICAS, что подтверждает корректность подхода к построению модели.

Коэффициент детерминации ($R^2$) является одним из ключевых показателей качества регрессионной модели. В рассматриваемой модели $R^2$ составил 99,13%. Это указывает на то, что модель объясняет 99,13% вариации зависимой переменной ICAS. Такое высокое значение коэффициента детерминации свидетельствует о том, что зависимость между ICAS и факторами NOSCA, GRP, GRPPC, AS и DER практически полностью учитывается данной моделью. Однако стоит учитывать, что значение $R^2$ может искусственно увеличиваться с ростом числа факторов, включенных в модель.

Для более точной оценки объясняющей способности модели используется скорректированный коэффициент детерминации ($R^2_{adj}$), который учитывает число независимых переменных в модели. В данной модели $R^2_{adj}$ составляет 98,26%. Это значение немного ниже, чем R2, что объясняется корректировкой на число факторов. Тем не менее, оно также свидетельствует о высокой объясняющей способности модели. Снижение значения $R^2_{adj}$ по сравнению с $R^2$ невелико, что подтверждает целесообразность включения всех пяти факторов в модель. Это также указывает на то, что объем выборки достаточно велик для обеспечения надежности оценок коэффициентов.

Предсказательная способность модели оценивается через коэффициент детерминации $R^2_{pred}$, который показывает, насколько хорошо модель объясняет вариацию зависимой переменной на новых данных. В данной модели $R^2_{pred}$, составляет 71,76%. Это значение указывает на то, что модель способна достаточно эффективно прогнозировать значения ICAS. Однако оно существенно ниже,

чем $R^2$ и $R^2_{adj}$, что свидетельствует о необходимости осторожного применения модели для прогнозирования на новых данных. Разница между $R^2_{adj}$ и $R^2_{pred}$ может быть связана с проблемой мультиколлинеарности или наличием факторов, которые влияют на ICAS, но не учтены в модели.

Среднеквадратичная ошибка модели (*S*) составила 226727. Этот показатель отражает среднее отклонение фактических значений ICAS от предсказанных моделью. Чем меньше значение S, тем точнее модель. Хотя данное значение не позволяет непосредственно сравнивать модели с разными зависимыми переменными, оно демонстрирует приемлемую точность для интерпретации и прогнозирования ICAS.

Для проверки автокорреляции остатков использовалась статистика Дарбина-Уотсона, значение которой в данной модели составило 2,40365. Это значение находится в пределах, близких к 2, что указывает на отсутствие значимой автокорреляции остатков. Таким образом, ошибки модели являются независимыми, что подтверждает корректность выбранного подхода и отсутствие нарушений основного предположения регрессионного анализа.

Включение факторов NOSCA, GRP, GRPPC, AS и DER в модель позволило объяснить значительную часть вариации ICAS. Однако не все переменные оказались статистически значимыми на традиционном уровне значимости 5%. Среди факторов только NOSCA показала статистически значимое влияние на ICAS с p-значением 0,027. Другие факторы, такие как GRP, GRPPC и AS, не достигли уровня значимости 5%, но их p-значения (0,094, 0,089 и 0,085 соответственно) близки к значимости. Это может свидетельствовать о потенциальном влиянии данных факторов, что требует дальнейшего анализа. DER оказался статистически незначимым с p-значением 0,405.

На основе данных коэффициентов регрессии можно оценить влияние каждого фактора на зависимую переменную ICAS. Уравнение регрессии выглядит следующим образом:

$$ICAS = -1136010 + 0{,}668 \times NOSCA + 4{,}55 \times GRP - 10{,}87 \times GRPPC + 50{,}4 \times AS - 10540 \times DER. \quad (1)$$

Каждый из коэффициентов указывает на ожидаемое изменение ICAS при изменении соответствующего фактора на единицу при прочих равных условиях. Положительные коэффициенты NOSCA, GRP и AS указывают на положительное влияние данных факторов, тогда как отрицательные коэффициенты GRPPC и DER отражают обратную связь.

Таким образом, общая характеристика модели позволяет утверждать, что она обладает высокой объясняющей способностью и статистической значимостью. Однако обнаруженные проблемы, такие как мультиколлинеарность и частичная незначимость некоторых факторов, требуют дальнейшего анализа и возможной доработки модели.

Коэффициент при переменной NOSCA равен 0,668, а его p-значение составляет 0,027. Это позволяет утверждать, что данный фактор оказывает статистически значимое положительное влияние на ICAS на уровне значимости 5%. Таким образом, увеличение NOSCA на единицу приводит к увеличению ICAS в среднем на 0,668 единицы.

Коэффициент при переменной GRP равен 4,55, что указывает на положительное влияние данного фактора на ICAS. Однако p-значение составляет 0,094, что превышает уровень значимости 5%, хотя и близко к нему. Это позволяет предположить, что GRP может иметь влияние на ICAS, но для подтверждения требуется либо увеличение объема выборки, либо более детальное исследование.

Коэффициент при переменной GRPPC равен -10,87, что говорит о наличии отрицательной связи между данным фактором и ICAS. Однако его p-значение равно 0,089, что также превышает уровень значимости 5%, но близко к значимости. Это может указывать на потенциальное влияние GRPPC, которое требует дополнительного анализа.

Коэффициент при переменной AS равен 50,4, что предполагает положительное влияние данного фактора на ICAS. Однако p-значение составляет 0,085, что превышает уровень значимости 5%. Как и в случае с GRP и GRPPC, фактор AS близок к значимости, что делает необходимым дальнейшее исследование.

Коэффициент при переменной DER равен -10540, указывая на отрицатель-

ное влияние данного фактора. Однако p-значение равно 0,405, что значительно превышает уровень значимости 5%. Это позволяет сделать вывод о статистической незначимости влияния DER на ICAS.

Одна из ключевых проблем, обнаруженных в данной регрессионной модели, связана с мультиколлинеарностью независимых переменных. Мультиколлинеарность возникает, когда между объясняющими переменными существует сильная корреляция, что приводит к искажению оценки регрессионных коэффициентов. Это, в свою очередь, снижает интерпретируемость модели и затрудняет статистическую значимость отдельных факторов. В рамках представленной модели оценка мультиколлинеарности производилась с использованием коэффициента инфляции дисперсии (VIF), который позволяет определить уровень корреляции между объясняющими переменными.

NOSCA: VIF = 2,76 – значение находится в пределах нормы, что говорит об отсутствии значимой корреляции с другими факторами.

GRP: VIF = 572,03 – экстремально высокая мультиколлинеарность, что указывает на сильную корреляцию с другими переменными.

GRPPC: VIF = 350,08 – также экстремально высокая мультиколлинеарность.

AS: VIF = 27,82 – высокая мультиколлинеарность, которая может искажать оценки коэффициентов.

DER: VIF = 5,78 – умеренная мультиколлинеарность.

Высокие значения VIF для факторов GRP, GRPPC и AS являются серьезной проблемой, поскольку они могут приводить к нестабильности коэффициентов и снижать точность интерпретации результатов. Это свидетельствует о необходимости применения методов устранения мультиколлинеарности, таких как исключение избыточных факторов, использование метода главных компонент или регуляризационных методов.

Для проверки модели на наличие выбросов и необычных наблюдений был выполнен анализ остатков. Выявлено одно наблюдение, отмеченное как «Unusual X». Остаток для данного наблюдения составил -3525, а стандартный остаток равен -0,42. Несмотря на то, что данное наблюдение классифицировано как необычное, его влияние на модель незначительно, что подтверждается низким значением стандартного остатка.

Для расчёта доходов предприятий гостеприимства Красноярского края была использована регрессионная модель, формула которой описана выше (1). Подставив данные значения переменных в уравнение, можно получить прогнозируемые доходы для рассматриваемого региона.

$$-1136010 + 0{,}668 \times 3790753 + 4{,}55 \times 3490000 - 10{,}87 \times 1128042 + 50{,}4 \times 81056 - 10540 \times 85{,}81 = 8194681{,}46 \text{ (руб.)}$$

Анализ прогноза доходов предприятий гостеприимства показывает, что фактическое значение за исследуемый период незначительно превышает прогнозируемое значение на следующий период, что свидетельствует о стабильности отрасли, но также указывает на возможные риски, связанные с колебаниями в спросе и экономической ситуации. Это подчеркивает необходимость тщательного мониторинга и адаптации стратегий для поддержания конкурентоспособности и роста доходов.

Результаты анализа показывают, что разработанная модель регрессии имеет высокую объясняющую способность ($R^2$= 99,13%) и приемлемую предсказательную способность ($R^2_{pred}$ = 71,76%). Единственным статистически значимым фактором в модели является NOSCA, который положительно влияет на ICAS. Остальные факторы (GRP, GRPPC, AS, DER) не достигли уровня значимости 5%, хотя некоторые из них (GRP, GRPPC, AS) близки к значимости.

Основной проблемой модели является высокая мультиколлинеарность, особенно между факторами GRP, GRPPC и AS. Это требует дальнейшего пересмотра модели, включая сокращение числа переменных или применение методов регуляризации. Кроме того, диагностические тесты показали отсутствие автокорреляции остатков, что подтверждает надежность модели.

Таким образом, модель может быть использована для анализа влияния факторов на ICAS, но требует доработки для повышения интерпретируемости и точности прогнозов.

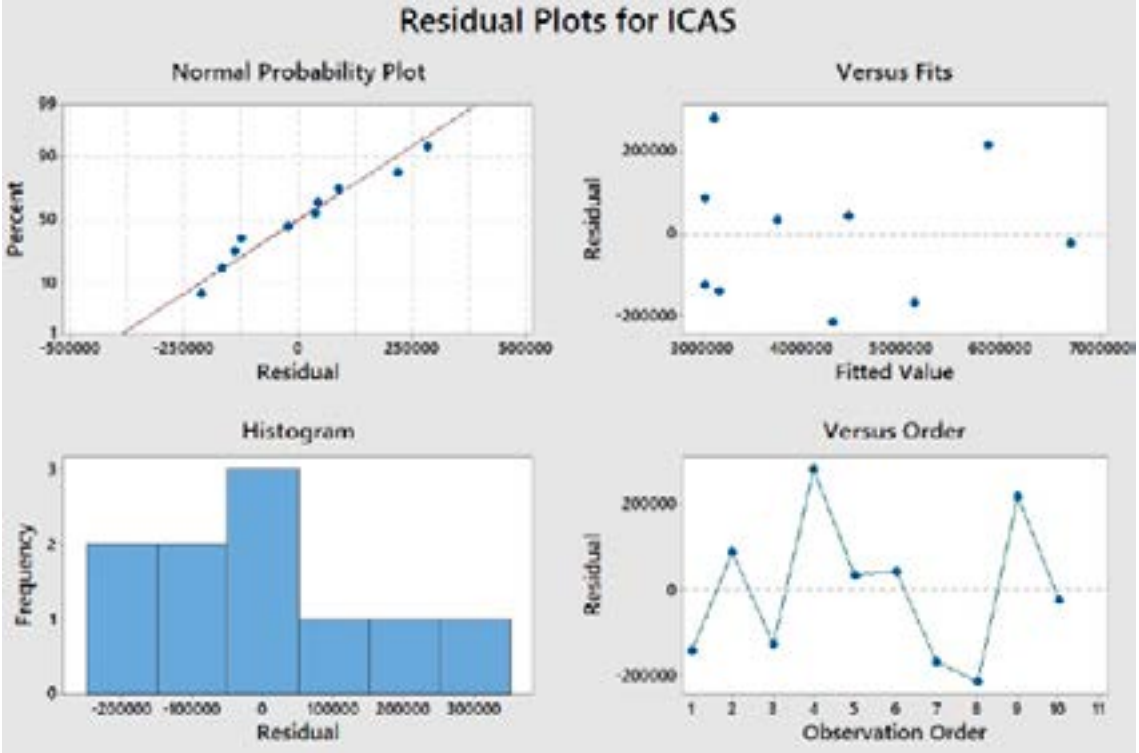


*Рис. 1. Графики остатков регрессионной модели*

Анализ графиков остатков является ключевым этапом проверки качества и корректности построенной регрессионной модели. Он позволяет выявить возможные систематические отклонения, оценить выполнение основных предпосылок регрессионного анализа, таких как линейность, равномерность дисперсии (гомоскедастичность) и отсутствие автокорреляции. В рамках исследования был проведён анализ визуализаций остатков для определения потенциальных недостатков модели, что позволит уточнить её структуру и повысить точность прогнозирования. График остатков представлен на рисунке 1.

На графике нормальности остатков (normal probability plot) точки в целом лежат близко к линии, что свидетельствует о допустимости предположения о нормальности распределения остатков. Однако можно заметить небольшие отклонения в крайних точках, что может указывать на присутствие небольших аномалий или отклонений от нормальности в хвостах распределения. Это требует дополнительного анализа.

На графике остатков против предложенных значений (Residuals versus Fits) остатки распределены относительно нулевой линии и не формируют четкой структуры, что свидетельствует о том, что проблема гетероскедастичности отсутствует. Однако на графике можно заметить небольшие скопления точек вблизи отдельных предсказанных значений, что может указывать на существование скрытых закономерностей или возможного влияния некоторых переменных, не учтенных в модели.

На гистограмме видно, что распределение остатков приближено к нормальному: оно имеет симметричный вид и центральный пик. Однако можно заметить небольшие отклонения от идеальной симметрии и присутствие длинных хвостов.

На графике остатков по порядку наблюдений (Residuals versus Order) видно, что остатки чередуются относительно нулевой линии, но отсутствует четкий тренд или периодичность. Это свидетельствует об отсутствии ярко выраженной автокорреляции.

Таким образом, представленные графики подтверждают, что модель удовлетворяет основным предположениям регрессионного анализа, и её результаты могут считаться надежными. Однако небольшие отклонения от нормальности и слабые намеки на скрытые закономерности в остатках.

## Выводы

В результате проделанной работы была разработана многофакторная регрессионная модель для прогнозирования доходов предприятий гостеприимства Красноярского края, основанная на официальной статистике и анализе ключевых экономических факторов. Модель показала свою способность достаточно точно описывать динамику доходов, выявляя влияние таких переменных, как уровень туристической активности, экономические показатели региона и социальные факторы. Полученная модель имеет значимость для дальнейшего анализа и прогнозирования в сфере гостеприимства, поскольку позволяет оценить влияние различных факторов на доходы и выявить потенциальные риски и возможности для развития сектора. Однако, несмотря на высокую достоверность прогноза, модель столкнулась с проблемой мультиколлинеарности между некоторыми факторами, что ограничивает её точность. Это указывает на необходимость доработки модели с целью устранения мультиколлинеарности и улучшения её прогностической способности в будущем.

*Библиографический список*

1. Багузова Л.В. Совершенствование механизма повышения эффективности ресурсного потенциала организаций высшего образования: дис. … канд. экон. наук. Красноярск, 2023. 248 с.

2. Барило Л.В. Факторы состояния и развития индустрии гостеприимства на примере Волгоградской области // Экономика и социум. 2022. № 12-1. С. 432-435.

3. Батырова Н.Т., Асан Д.С., Сарсенбаев А.А. Основные экономические факторы развития туризма и гостеприимства // Вестник науки. 2024. № 4 (73). С. 773-781.

4. Белоногова Е.В., Суслова Ю.Ю. Показатели оценки материального стимулирования труда работников с позиций заинтересованных пользователей // Российское предпринимательство. 2011. № 8-1. С. 119-124.

5. Волошин А.В. Формирование механизма повышения конкурентоспособности организаций высшего образования на рынке образовательных услуг: дис. … канд. экон. наук. Красноярск, 2017. 308 с.

6. Пранов Б.М. Модели динамики и прогнозирования в сфере гостеприимства // Вестник РМАТ. 2017. № 1. С. 24-27.

7. Суслова Ю.Ю. Качество жизни населения, как комплексный показатель оценки рыночной инфраструктуры города // Проблемы современной экономики. 2008. № 4 (28). С. 415-419.

8. Терещенко Н.Н., Есина О.Н., Ильина А.С. Эффективность деятельности предприятия торговли: монография / рец.: Ю.Ю. Суслова, Н.А. Симченко, С.М. Горлов, 2022. 216 с.

9. Ханк Д.Э., Уичерн Д.У., Райтс А.Д. Бизнес-прогнозирование / пер. с англ. М.: Вильямс, 2003. 665 с.

10. Шестов С.Н., Волошин А.В. Конкурентоспособность региона как предмет исследования экономической науки // Труд и социальные отношения. 2016. Т. 27. № 5. С. 27-36.

11. Voloshin A.V., Suslova J.J., Batrayeva E.A., Baguzova L.V., Ananeva N.V. Current state, tendencies and peculiarities of development in higher education: the case of Russia // European Research Studies Journal. 2018. V. 21. № 4. P. 502-511.

12. Suslova Yu.Yu., Voloshin A.V. The production potential of the agroindustrial complex of the siberian federal district – the basis for the development of the food market of the region // Journal of Siberian Federal University. Humanities and Social Sciences. 2017. V. 10. № 11. P. 1789-1797.

13. Zhironkin S., Voloshin A., Zhironkina O., Vöth S., Kayachev G. The principles of sustainable development as a form of structural transformation // E3S Web of Conferences. “5th International Innovative Mining Symposium, IIMS 2020”, 2020. P. 04011.