




Применение методов математической статистики для оценки инвестиционного потенциала региона




Ирина П. Болодурина ¹	ipbolodurina@yandex.ru	 0000-0003-0096-2587
Марина П. Болодурина ¹	bolodurina@inbox.ru	 0000-0002-6414-245X
Карина М. Абелгазина ¹	abelgazina99@mail.ru	 0000-0003-2596-6402

¹ Оренбургский государственный университет, пр-т Победы 13, г. Оренбург, 460018, Россия

Аннотация. Поскольку каждый субъект Российской Федерации обладает полными правами на ведение как межрегиональных, так и международных экономических отношений, то показатели регионального развития во многом зависят от реализуемой инвестиционной стратегии, в том числе от уровня инвестиционного потенциала. В этой связи в статье рассматривается актуальная задача оценки инвестиционного потенциала региона на основе социально-экономических показателей методами математической статистики, что позволяет получить объективные и достоверные результаты, а также обеспечивает масштабируемость используемой методики и возможность ее применения для анализа инвестиционного потенциала других регионов Российской Федерации. Использованы корреляционно-регрессионный анализ, проведено исследование линейной регрессионной модели на мультиколлинеарность. Применении многомерного статистического анализа показателей инвестиционной привлекательности Оренбургской области позволило выявить наиболее значимые факторы, оказывающие влияние на объем инвестиций в основной капитал региона, к которым относятся: объем работ, выполненных по виду экономической деятельности «Строительство»; среднедушевые денежные доходы населения; объем продукции обрабатывающего производства; удельный вес численность населения в трудоспособном возрасте в общей численности. На основании результатов корреляционно-регрессионного анализа сделан вывод о том, что на уровень объема инвестиций в основной капитал Оренбургской области наиболее значимое влияние оказывают такие сферы экономической деятельности, как сельское хозяйство, добыча полезных ископаемых и обрабатывающие производства. Привлечение инвестиций в выявленные отраслевые приоритеты инвестиционной привлекательности Оренбургской области обеспечит создание высокопроизводительных мест на территории региона, увеличение валового регионального продукта, а также даст мультипликативный эффект для развития других видов деятельности.

Ключевые слова: инвестиционный потенциал, корреляционно-регрессионный анализ, инвестиционная привлекательность, математическая статистика

Application of mathematical statistics methods to assess the investment potential of the region

Irina P. Bolodurina ¹	ipbolodurina@yandex.ru	 0000-0003-0096-2587
Marina P. Bolodurina ¹	bolodurina@inbox.ru	 0000-0002-6414-245X
Karina M. Abalgazina ¹	abelgazina99@mail.ru	 0000-0003-2596-6402

¹ Orenburg State University, 13 Pobedy Ave., Orenburg, 460018, Russia

Abstract. Since each subject of the Russian Federation has full rights to conduct both interregional and international economic relations, the indicators of regional development largely depend on the investment strategy being implemented, including the level of investment potential. In this regard, the article deals with the urgent task of assessing the investment potential of the region based on socio-economic indicators using mathematical statistics, which allows us to obtain objective and reliable results, and also ensures the scalability of the methodology used and the possibility of its application to analyze the investment potential of other regions of the Russian Federation. Correlation and regression analysis was used, a linear regression model was studied for multicollinearity. The use of multidimensional statistical analysis of the investment attractiveness indicators of the Orenburg region allowed us to identify the most significant factors influencing the volume of investments in the fixed capital of the region, which include: the amount of work performed by the type of economic activity "Construction"; average per capita monetary income of the population; the volume of manufacturing production; the proportion of the working-age population in the total number. Based on the results of correlation and regression analysis, it is concluded that the level of investment in fixed assets of the Orenburg region is most significantly influenced by such areas of economic activity as agriculture, mining and manufacturing. Attracting investments to the identified sectoral priorities of the investment attractiveness of the Orenburg region will ensure the creation of high-performance places in the region, increase the gross regional product, and also give a multiplier effect for the development of other activities.

Keywords: investment potential, correlation and regression analysis, investment attractiveness, mathematical statistics

Введение

Инвестиционная активность экономического субъекта во многом определяет вектор его экономического развития. Для грамотного построения приоритетов формирования основных направлений социально-экономического развития субъекта Российской Федерации необходимо вести учет особенностей его географического расположения, развития отраслей экономики,

ресурсов труда, наличия месторождений природных ресурсов и других компонентов инвестиционного потенциала. Современное состояние информационной среды позволяет использовать большое количество данных и показателей для подготовки и обоснования управленческих решений инвестиционного характера. Применение методов математической статистики позволяет повысить объективность, как используемых данных, так и полученных результатов.

Для цитирования

Болодурина И.П., Болодурина М.П., Абелгазина К.М. Применение методов математической статистики для оценки инвестиционного потенциала региона // Вестник ВГУИТ. 2022. Т. 84. № 1. С. 295–304. doi:10.20914/2310-1202-2022-1-295-304

For citation

Bolodurina I.P., Bolodurina M.P., Abalgazina K.M. Application of mathematical statistics methods to assess the investment potential of the region. *Vestnik VGUIT* [Proceedings of VSUET]. 2022. vol. 84. no. 1. pp. 295–304. (in Russian). doi:10.20914/2310-1202-2022-1-295-304

This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License

Материалы и методы

Подробный анализ существующих зарубежных и отечественных методик оценки инвестиционной привлекательности регионов проведен в статье Кособуцкой А.Ю. и Равунжинириной А.В., авторами отмечается большое разнообразие используемых показателей оценки и варьированности параметров в зависимости от целей оценки, а также в методах обработки информации и оценке результатов [1]. Множественность используемых параметров оценки инвестиционного потенциала территории отмечается и в зарубежных исследованиях [13, 14, 16].

Интеграционный подход и возможности его применения для оценки частных и общего инвестиционного потенциала региона представлены в трудах Болодуриной М.П. и Мишуровой А.И. [3, 15]. Авторы перечисленных выше исследований наряду с субъективными критериями оценки инвестиционного потенциала региона используют показатели национального рейтинга состояния инвестиционного климата в субъектах РФ [2, 8], аналитические данные рейтинговых агентств [4, 9], а также данные официальной статистики.

Статистические методы анализа в оценке инвестиционного потенциала региона применялись в работах Валерианова А.А. и его соавторов [5], Дмитриева Н.Д. [6] и других отечественных авторов. Экономико-математическое моделирование лежит в основе и многих зарубежных исследований, Куммел Р., Хенн Дж. И Линденбергер Д проводят моделирование распространения инвестиций и инноваций в зависимости от таких параметров как капитал, труд, энергия и креативность [16]. Результаты кластерного анализа при разработке инвестиционных рекомендаций представлены в трудах Лео В., Донгхун С. И других [20, 21].

Таким образом, затруднение вызывает не недостаток, а наоборот многочисленность,

используемых данных и методик их оценки. Оценка инвестиционного потенциала региона на основе социально-экономических показателей методами математической статистики позволит повысить достоверность объективность получаемых оценок и объективность сделанных на их основе выводов.

В настоящем исследовании для проведения многомерного статистического анализа показателей инвестиционной привлекательности Оренбургской области на основании статистических данных за одиннадцатилетний период [7], представленных в таблице 1 проведен корреляционный анализ социально-экономических показателей, характеризующих инвестиционную привлекательность региона.

В качестве исходных показателей использованы следующие:

ξ_1 – инвестиции в основной капитал (млн рублей);

ξ_2 – объем работ, выполненных по виду экономической деятельности «Строительство» (млн рублей);

ξ_3 – среднедушевые денежные доходы населения (в месяц рублей);

ξ_4 – оборот розничной торговли (млн рублей);

ξ_5 – уровень добычи полезных ископаемых (в фактически действовавших ценах; млн рублей);

ξ_6 – объем продукции обрабатывающего производства (в фактически действовавших ценах; млн рублей);

ξ_7 – продукция сельского хозяйства (млн рублей);

ξ_8 – удельный вес численности населения в трудоспособном возрасте в общей численности (проценты);

ξ_9 – объем инновационных товаров, работ, услуг (млн рублей);

ξ_{10} – валовой региональный продукт (в основных ценах, млн рублей).

Таблица 1.

Исходные выборочные данные

Table 1.

Initial sample data

Год Year	Показатель Indicator									
	ξ_1	ξ_2	ξ_3	ξ_4	ξ_5	ξ_6	ξ_7	ξ_8	ξ_9	ξ_{10}
2009	91268,0	25826,3	12008,1	136473,4	185875,0	117413,0	55288,3	63,4	6914,0	413395,5
2010	103648,0	31409,1	13557,1	157681,8	215311,0	151634,0	47707,8	63,6	11651,4	458145,4
2011	116658,0	35025,4	14892,0	187098,8	282235,0	174977,0	69144,2	60,6	14230,1	553320,9
2012	151250,0	41873,5	16541,9	214089,8	313118,0	211629,0	68019,0	63,7	9244,6	628563,6
2013	152877,0	41456,1	18389,9	240773,4	420880,0	192004,0	77641,4	63,4	8870,3	717014,8
2014	153979,0	47597,4	20699,9	269615,0	348081,0	230679,0	83840,4	60,4	6888,9	731277,7
2015	169243,0	60566,7	22943,1	278368,2	383955,0	274498,0	99108,4	62,5	14671,4	774962,1
2016	167278,5	49750,8	22144,8	274944,3	338587,0	274312,0	110753,6	61,9	25264,1	765333,3
2017	184877,2	62949,9	22909,8	293997,0	389692,4	304238,0	113038,3	60,0	23250,7	823856,4
2018	208105,1	78643,5	23384,8	307946,0	527943,1	348347,7	107971,1	61,0	28408,6	1058504,8
2019	211726,5	100988,0	24482,7	327692,6	561650,8	315272,8	113015,9	60,2	41663,7	1107155,3

Данный выбор обусловлен проведенным анализом работ по оценке инвестиционной привлекательности регионов РФ, в результате которого выявлены наиболее часто встречающиеся социально-экономические показатели, оказывающие влияние на уровень инвестиционной привлекательности. Кроме того, для исследования выбраны показатели, в наибольшей мере характеризующие экономическую деятельность Оренбургской области.

Точечное оценивание параметров многомерного вектора признаков сводится к оценке вектора математических ожиданий и ковариационной матрицы в виде вектора средних значений и выборочной ковариационной матрицы. С помощью пакета Statistica рассчитаны значения выборочных средних, дисперсий и средних квадратических отклонений для каждого рассматриваемого признака. Полученные результаты представлены в таблице 2. Для получения выборочной ковариационной матрицы использован табличный редактор MS Excel (настройка «Пакет анализа») [10].

$$R_k = \begin{pmatrix} 1 & 0,92 & 0,95 & 0,97 & 0,94 & 0,97 & 0,91 & -0,57 & 0,77 & 0,98 \\ 0,92 & 1 & 0,86 & 0,88 & 0,92 & 0,89 & 0,82 & -0,61 & 0,89 & 0,96 \\ 0,95 & 0,86 & 1 & 0,99 & 0,86 & 0,95 & 0,96 & -0,62 & 0,71 & 0,92 \\ 0,97 & 0,88 & 0,99 & 1 & 0,90 & 0,95 & 0,94 & -0,65 & 0,72 & 0,95 \\ 0,94 & 0,92 & 0,86 & 0,90 & 1 & 0,86 & 0,80 & -0,54 & 0,75 & 0,97 \\ 0,97 & 0,89 & 0,95 & 0,95 & 0,86 & 1 & 0,94 & -0,63 & 0,78 & 0,94 \\ 0,91 & 0,82 & 0,96 & 0,94 & 0,80 & 0,94 & 1 & -0,67 & 0,77 & 0,88 \\ -0,57 & -0,61 & -0,62 & -0,65 & -0,54 & -0,63 & -0,67 & 1 & -0,59 & -0,62 \\ 0,77 & 0,89 & 0,71 & 0,72 & 0,75 & 0,78 & 0,77 & -0,59 & 1 & 0,82 \\ 0,98 & 0,96 & 0,92 & 0,95 & 0,97 & 0,94 & 0,88 & -0,62 & 0,82 & 1 \end{pmatrix}.$$

Поскольку коэффициенты корреляции отражают взаимосвязь между двумя признаками под воздействием всех остальных, то следующим этапом исследования является оценка связи между парой признаков, очищенной от влияния всех остальных признаков.

С помощью пакета Statistica получена информация о значении множественного коэффициента корреляции $\hat{R}_{1/2-10} = 0,999993$, коэффициента детерминации $\hat{R}_{1/2-10} = 0,999986$, а также значение F -статистики, предназначенной для проверки гипотезы о незначимости коэффициента детерминации ($F = 7704,826$) и вероятности принятия данной гипотезы ($p = 0,008841$). Далее рассчитаны оценки частных коэффициентов корреляции.

Таблица 2.
Описательная статистика исследуемых признаков

Table 2.
Descriptive statistics of the studied features

Показатель Indicator	Mean	Variance	Std. Dev.
ξ_1	155537,3	1546966030	39331,5
ξ_2	52371,5	493681701	22218,9
ξ_3	19268,6	19390910,2	4403,5
ξ_4	244425,5	3954168087	62882,2
ξ_5	360666,2	13412522147	115812,4
ξ_6	235909,5	5397463417	73467,4
ξ_7	85957,1	581967392	24124,0
ξ_8	61,9	2,18274182	1,5
ξ_9	17368,9	121622620	11028,3
ξ_{10}	730139,1	47805879605	218645,6

Всего рассматривается 45 связей, из них 36 характеризуются прямой связью, а 9 – обратной. В пакете Statistica рассчитаны выборочные значения коэффициентов корреляции для каждой пары признаков. Таким образом, в результате расчетов получена оценка корреляционной матрицы вида:

Сравнивая по абсолютной величине оценки коэффициентов корреляции и соответствующих частных коэффициентов корреляции, можно сделать следующие выводы.

Поскольку $r_{1,2} = 0,92 < r_{1,2/3-10} = 0,997$, то такие показатели, как среднедушевые денежные доходы населения региона (ξ_3), оборот розничной торговли (ξ_4), уровень добычи полезных ископаемых (ξ_5), объем продукции обрабатывающего производства региона (ξ_6), продукция сельского хозяйства (ξ_7), удельный вес численности населения в трудоспособном возрасте в общей численности населения региона (ξ_8), объем инновационных товаров, работ, услуг (ξ_9), а также валовой региональный продукт (ξ_{10}) ослабляют (незначительно) взаимосвязь между парой показателей: инвестиции в основной капитал региона (ξ_1) и объем работ, выполненных

по виду экономической деятельности «Строительство» (ξ_2). Аналогичный результат можно наблюдать для таких пар показателей, как: инвестиции в основной капитал региона (ξ_1) и среднедушевые денежные доходы населения региона (ξ_3) (т. к. $r_{1,3} = 0,95 < r_{1,3/2,4-10} = 0,999$), инвестиции в основной капитал региона (ξ_1) и оборот розничной торговли (ξ_4) (т. к. $r_{1,4} = 0,97 < r_{1,4/2,3,5-10} = 0,999$), инвестиции в основной капитал региона (ξ_1) и уровень добычи полезных ископаемых (ξ_5) (т. к. $r_{1,5} = 0,94 < r_{1,5/2,4,6-10} = 0,980$), инвестиции в основной капитал региона (ξ_1) и объем продукции обрабатывающего производства региона (ξ_6) (т. к. $r_{1,6} = 0,97 < r_{1,6/2,5,7-10} = 0,999$), инвестиции в основной капитал региона (ξ_1) и продукция сельского хозяйства (ξ_7) (т. к. $r_{1,7} = 0,91 < r_{1,7/2,6,8-10} = 0,996$).

Так как $r_{1,8} = 0,57 < r_{1,8/2,7,9-10} = 0,999$, то такие показатели, как объем работ, выполненных по виду экономической деятельности «Строительство» (ξ_2), среднедушевые денежные доходы населения региона (ξ_3), оборот розничной торговли (ξ_4), уровень добычи полезных ископаемых (ξ_5), объем продукции обрабатывающего производства региона (ξ_6), продукция сельского хозяйства (ξ_7), а также объем инновационных товаров, работ, услуг (ξ_9) и валовой региональный продукт (ξ_{10}) ослабляют (значительно) взаимосвязь между парой показателей: инвестиции в основной капитал региона (ξ_1) и удельный вес численности населения в трудоспособном возрасте в общей численности населения региона (ξ_8). Подобный результат наблюдается и для пары показателей инвестиции в основной капитал региона (ξ_1) и объем инновационных товаров, работ, услуг (ξ_9) (здесь $r_{1,9} = 0,77 < r_{1,9/2,8,10} = 0,965$), так остальные показатели ($\xi_2, \xi_3, \xi_4, \xi_5, \xi_6, \xi_7, \xi_8, \xi_{10}$) значительно ослабляют связь между рассматриваемой парой показателей.

Для пары показателей инвестиции в основной капитал региона (ξ_1) и валовой региональный продукт (ξ_{10}) остальные показатели, такие, как объем работ, выполненных по виду экономической деятельности «Строительство» (ξ_2), среднедушевые денежные доходы населения региона (ξ_3), оборот розничной торговли (ξ_4), уровень добычи полезных ископаемых (ξ_5), объем продукции обрабатывающего производства региона (ξ_6), продукция сельского хозяйства (ξ_7), а также удельный вес численности населения в трудоспособном возрасте в общей численности населения региона (ξ_8) и объем инновационных товаров, работ, услуг (ξ_9) не оказывают никакого влияния, поскольку $r_{1,10} = 0,98 \approx r_{1,10/2-9} = 0,989$. Аналогично и для показателей объем инновационных товаров, работ, услуг (ξ_9) и валовой

региональный продукт (ξ_{10}), так как $r_{9,10} = 0,82 \approx r_{9,10/1-8} = 0,823$.

Для всех остальных пар признаков следует отметить, что оценки коэффициентов корреляции (по абсолютной величине) больше, чем оценки соответствующих частных коэффициентов корреляции, т. е. наличие других показателей усиливает взаимосвязь между исследуемой парой признаков.

С помощью пакета Statistica получена информация об остальных значениях множественного коэффициента корреляции, коэффициента детерминации, а также значениях F -статистики, предназначенной для проверки гипотезы о незначимости коэффициента детерминации и вероятности принятия данной гипотезы.

Дальнейшим этапом корреляционного анализа – установление значимости характеристик связи (коэффициентов корреляции, частных коэффициентов корреляции и коэффициентов детерминации) на заданном уровне значимости $\alpha = 0,05$. Это можно реализовать двумя способами.

В основе первого способа лежит сопоставление наблюдаемого значения статистики с критическим значением, в случае если $|t_{набл.}| > t_{кр.}$ (для коэффициентов корреляции и частных коэффициентов корреляции), $F_{набл.} > F_{кр.}$ (для коэффициентов детерминации), то с вероятностью ошибки 0,05 нулевая гипотеза о незначимости соответствующих характеристик связи отклоняется.

В основе второго способа лежит сопоставление значения вероятности принятия нулевой гипотезы о незначимости характеристики связи (наблюдаемого уровня значимости p) с заранее установленным уровнем значимости $\alpha = 0,05$, в случае, если $p < 0,05$, то на 5%-ом уровне значимости нулевая гипотеза отклоняется и делается заключение о статистической значимости соответствующих характеристик связи.

Для проверки значимости коэффициентов корреляции использованы результаты, полученные ранее в пакете Statistica.

Критическое значение определим с помощью функции СТЬЮДЕНТ.ОБР.2Х (Вероятность; Степени_свободы) пакета MS Excel [10].

Учитывая заданный уровень значимости $\alpha = 0,05$ и число степеней свободы $\nu = n - 2 = 11 - 2 = 9$, критическая точка принимает следующее значение $t_{кр} = 2,262157$.

Аналогичным образом на основе информации об оценках частных коэффициентов корреляции определена значимость частных коэффициентов корреляции. При $\alpha = 0,05$ и числе степеней свободы $\nu_1 = n - k = 11 - 10 = 1$, критическая точка принимает следующее значение $t_{кр} = 12,7062$.

Используя результаты оценивания множественных коэффициентов корреляции и коэффициентов детерминации, определена значимость коэффициентов детерминации

в таблице 3. Для этого рассчитано критическое значение с помощью функции F.ОБР.ПХ (Вероятность; Степени_свободы 1; Степени_свободы 2) табличного редактора MS Excel.

Таблица 3.

Результаты проверки гипотез о незначимости коэффициентов детерминации

Table 3.

Results of testing hypotheses about the insignificance of the coefficients of determination

Коэффициент детерминации Coefficient of determination	Оценка коэффициента детерминации Estimation of the coefficient of determination	$F_{набл}$	$F_{кр}$	Вероятность принятия нулевой гипотезы p Probability of acceptance of the null hypothesis p	Уровень значимости α Significance level α	Значимость Significant
$\hat{R}_{1/2-10}^2$	0,999986	7704,83	240,54	0,008841	0,05	+
$\hat{R}_{2/1,3-10}^2$	0,99983	653,81		0,030343		
$\hat{R}_{3/1-2,4-10}^2$	0,999992	13409,67		0,006702		
$\hat{R}_{4/1-3,5-10}^2$	0,999995	21037,03		0,005351		
$\hat{R}_{5/1-4,6-10}^2$	0,999252	148,43		0,063621		–
$\hat{R}_{6/1-5,7-10}^2$	0,999952	2300,46		0,016179		+
$\hat{R}_{7/1-6,8-10}^2$	0,999705	376,91		0,039955		
$\hat{R}_{8/1-7,9-10}^2$	0,999558	251,07		0,048943		
$\hat{R}_{9/1-8,10}^2$	0,994242	19,18		0,175493		–
$\hat{R}_{10/1-9}^2$	0,999882	944,96		0,025241		+

Учитывая заданный уровень значимости $\alpha = 0,05$ и число степеней свободы $v_1 = k - 1 = 10 - 1 = 9$, $v_2 = n - k = 11 - 10 = 1$, критическая точка принимает значение $F_{кр} = 240,5433$.

Дальнейший шаг корреляционного анализа – построение с вероятностью $\gamma = 0,95$ доверительных интервалов для значимых коэффициентов корреляции и частных коэффициентов корреляции с помощью z -преобразования Фишера.

Заключительным этапом корреляционного анализа является оценка уравнения регрессии. В качестве результативного признака выбран показатель ξ_1 – инвестиции в основной капитал региона. Признаки $\xi_2 - \xi_{10}$ выступают в качестве факторов, оказывающих влияние на объем инвестиций в основной капитал региона.

Таким образом, оценка уравнения регрессии имеет вид:

$$\hat{x}_1 = -262265,89 + 0,51x_2 - 11,77x_3 + 1,03x_4 + 0,05x_5 + 0,29x_6 + 0,36x_7 + 4803,58x_8 - 0,17x_9 - 0,06x_{10}$$

Полученное уравнение регрессии относительно признака ξ_1 адекватно выборочным данным (т. е. значимо), так как соответствующий коэффициент детерминации $\hat{R}_{1/2-10}^2$ является значимым.

Значимость коэффициентов уравнения регрессии определяют по итогам исследования значимости частных коэффициентов корреляции. Таким образом, среди коэффициентов уравнения регрессии значимыми оказались $\beta_{1,2}, \beta_{1,3}, \beta_{1,4}, \beta_{1,6}, \beta_{1,8}$ поскольку значимы соответствующие частные коэффициенты корреляции

$$\rho_{1,2/3-10}, \rho_{1,3/2,4-10}, \rho_{1,4/2-3,5-10},$$

$$\rho_{1,6/2-5,7-10}, \rho_{1,8/2-7,9-10}$$

Следовательно, корреляционно-регрессионный анализ показателей инвестиционной привлекательности Оренбургской области позволяет сделать вывод, что на объем инвестиций в основной капитал региона значимое влияние оказывают следующие факторы:

– объем работ, выполненных по виду экономической деятельности «Строительство» (млн рублей);

– среднедушевые денежные доходы населения области (в месяц рублей), оборот розничной торговли региона (млн рублей);

– объем продукции обрабатывающего производства (в фактически действовавших ценах; млн рублей);

– удельный вес численности населения в трудоспособном возрасте в общей численности Оренбургской области (проценты).

На основании полученного уравнения регрессии следует, что при прочих равных условиях при увеличении объема работ в регионе, выполненных по виду экономической деятельности «Строительство» на 100 млн рублей инвестиции в основной капитал области увеличатся на 51 млн рублей; с увеличением оборота розничной торговли на 100 млн рублей объем инвестиций в основной капитал региона увеличатся на 103 млн рублей; увеличение на 100 млн рублей объема продукции обрабатывающего производства приведет к росту инвестиций в основной капитал области на 29 млн рублей; с увеличением на 1% удельного веса численности населения в трудоспособном возрасте в общей численности населения региона инвестиции в основной капитал возрастут в среднем на 4803,58 млн рублей.

Поскольку в процессе проведения корреляционно-регрессионного анализа выявлено подозрение на наличие мультиколлинеарности, то далее проведем более подробное исследование полученной регрессионной модели.

Для исследования линейной регрессионной модели на мультиколлинеарность по исходным данным необходимо:

1. построить оценки коэффициентов линейной модели множественной регрессии методом наименьших квадратов (МНК) и провести анализ;
2. провести анализ построенной модели на мультиколлинеарность;
3. устранить мультиколлинеарность при ее наличии.

На основе информации об объеме инвестиций в основной капитал Оренбургской области, млн рублей (y), объеме работ, выполненных в регионе по виду экономической деятельности «Строительство», млн рублей (x_2), денежных доходов, приходящихся на душу населения области в месяц, руб. (x_3), обороте розничной торговли региона, млн рублей (x_5), уровне добычи полезных ископаемых, млн рублей (x_5), объеме продукции обрабатывающего производства области, млн рублей (x_6), объеме продукции сельского хозяйства области, млн рублей (x_7), численности трудоспособного населения, в% от общей численности населения субъекта (x_8), объеме инновационных товаров, работ и услуг, млн рублей (x_9), а также ВРП Оренбургской области, млн рублей (x_{10}) построена линейная регрессионная модель.

С помощью пакета Statistica рассчитана оценка множественного коэффициента корреляции $\hat{R}_{1/2-10} = 0,999993$. Далее проведено исследование на нормальный характер распределения регрессионных остатков. На рисунке 1 представлена гистограмма распределения регрессионных остатков.

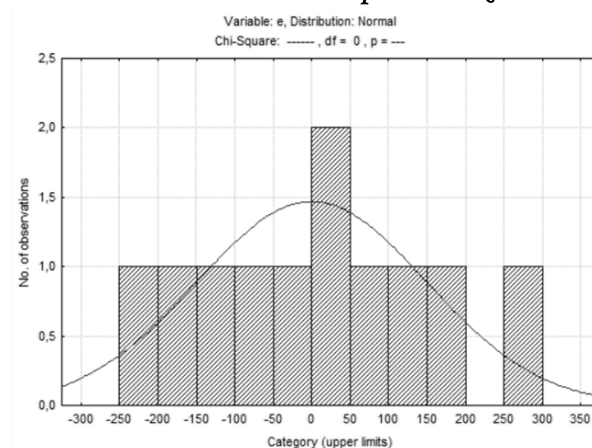


Рисунок 1. График распределения регрессионных остатков

Figure 1. Graph of the distribution of regression residuals

Согласно полученному графику, представленному на рисунке 1, можно предположить, что в разрезе большего временного интервала нулевая гипотеза о том, что распределение регрессионных остатков не отличается от нормального, может быть принята. Так как регрессионные остатки имеют нормальное распределение, то есть смысл проводить дальнейший анализ построенного уравнения множественной регрессии.

Отметим, что уравнение регрессии для выбранного признака адекватно выборочным данным – значимо, т. е. модель адекватна экспериментальным данным. Среди коэффициентов уравнения регрессии значимыми оказались коэффициенты при переменных x_2, x_3, x_4, x_6, x_8 . Среднеквадратические ошибки S_{b_j} при переменных x_5, x_{10} оказались того же порядка, что и коэффициенты регрессии при этих же переменных. Следовательно, коэффициенты при данных переменных могут иметь доверительный интервал, который содержит точку нуля.

Согласно полученной модели, можно сделать следующие выводы:

- при увеличении на 100 рублей в месяц среднедушевых доходов населения инвестиции в основной капитал региона уменьшатся на 1177 млн рублей (коэффициент при переменной x_2 имеет отрицательный знак);
- при увеличении объема инновационных товаров, работ, услуг на 100 млн рублей инвестиции в основной капитал региона уменьшатся на 17 млн рублей (коэффициент при переменной x_9 имеет отрицательный знак);
- при увеличении на 100 млн рублей валового регионального продукта объем инвестиций в основной капитал региона уменьшатся на 6 млн рублей (коэффициент при переменной x_{10} имеет отрицательный знак).

Полученные выводы противоречат экономическому смыслу.

Таким образом, все перечисленные признаки позволяют заподозрить наличие мультиколлинеарности между объясняющими переменными. Рассмотрим критерии по выявлению мультиколлинеарности.

Для начала проанализируем оценку матрицы парных коэффициентов корреляции между объясняющими переменными. Считается, что наличие значимых коэффициентов корреляции, по абсолютной величине превосходящих 0,7–0,8, свидетельствует о присутствии мультиколлинеарности.

На рисунке 2 представлена оценка матрицы парных коэффициентов корреляции.

На основе вычисленной матрицы есть основания подозревать тесную связь между всеми парами переменных, кроме x_1 и x_8 , x_2 и x_8 , x_3 и x_8 , x_4 и x_8 , x_5 и x_8 , x_6 и x_8 , x_7 и x_8 , x_9 и x_8 , x_{10} и x_8 .

Correlations										
Variable	X2	X3	X4	X5	X6	X7	X8	X9	X10	X1
X2	1.000000	0.861071	0.884353	0.919893	0.886989	0.824585	-0.610348	0.885786	0.959299	0.921652
X3	0.861071	1.000000	0.989780	0.856015	0.963073	0.958161	-0.621181	0.706815	0.918333	0.949943
X4	0.884353	0.989780	1.000000	0.903474	0.951990	0.943377	-0.650282	0.718577	0.949701	0.972980
X5	0.919893	0.856015	0.903474	1.000000	0.859274	0.797591	-0.536176	0.749624	0.973397	0.941518
X6	0.886989	0.953073	0.951990	0.859274	1.000000	0.938124	-0.629595	0.778504	0.937495	0.966965
X7	0.824585	0.958161	0.943377	0.797591	0.938124	1.000000	-0.669558	0.770804	0.880630	0.912213
X8	-0.610348	-0.621181	-0.650282	-0.536176	-0.629595	-0.669558	1.000000	-0.588219	-0.616945	-0.573921
X9	0.885786	0.706815	0.718577	0.749624	0.778504	0.770804	-0.588219	1.000000	0.822768	0.769588
X10	0.959299	0.918333	0.949701	0.973397	0.937495	0.880630	-0.616945	0.822768	1.000000	0.977208
X1	0.921652	0.949943	0.972980	0.941518	0.966965	0.912213	-0.573921	0.769588	0.977208	1.000000

Рисунок 2. Оценка матрицы парных коэффициентов корреляции

Figure 2. Estimation of the matrix of paired correlation coefficients

Более внимательное изучение этого вопроса достигается с помощью расчета значений коэффициентов детерминации каждой из объясняющих переменных $x^{(j)}$ по всем остальным переменным $x^{(j)} = (x^{(1)}, \dots, x^{(j-1)}, x^{(j+1)}, \dots, x^{(p)})$.

Получены следующие результаты оценивания коэффициентов детерминации:

$$R^2_{x_2/x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}} = 0,968525,$$

$$R^2_{x_3/x_2, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}} = 0,994205,$$

$$R^2_{x_4/x_2, x_3, x_5, x_6, x_7, x_8, x_9, x_{10}} = 0,995287,$$

$$R^2_{x_5/x_2, x_3, x_4, x_6, x_7, x_8, x_9, x_{10}} = 0,980859,$$

$$R^2_{x_6/x_2, x_3, x_4, x_5, x_7, x_8, x_9, x_{10}} = 0,960731,$$

$$R^2_{x_7/x_2, x_3, x_4, x_5, x_6, x_8, x_9, x_{10}} = 0,966417,$$

$$R^2_{x_8/x_2, x_3, x_4, x_5, x_6, x_7, x_9, x_{10}} = 0,655590,$$

$$R^2_{x_9/x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_{10}} = 0,916071,$$

$$R^2_{x_{10}/x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9} = 0,994329.$$

Анализ оценок коэффициентов детерминации показал наличие тесной линейной связи между всеми объясняющими переменными и остальными признаками.

Достаточным условием плохой обусловленности матрицы (наличия мультиколлинеарности) является большое значение числа обусловленности:

$$M = \frac{|\lambda_{\max}|}{|\lambda_{\min}|}, \quad (1)$$

где λ_{\max} – максимальное собственное число матрицы $X^T X$. Для вычисления собственных чисел матрицы $X^T X$ использована функция `eigenvals` из математического пакета Mathcad [12].

Таким образом, получено, что значение числа обусловленности составляет $M = 1,915 \cdot 10^8$. На основании полученного результата можно сделать вывод, что между всеми объясняющими переменными существует мультиколлинеарность.

В случае, если между объясняющими переменными существует частичная мультиколлинеарность, то оценки коэффициентов линейной модели, полученные по МНК, становятся неустойчивыми, незначительное изменение состава выборки или состава объясняющих переменных может вызвать кардинальное изменение модели, что делает модель непригодной для практических целей. Наиболее распространенные в таких случаях приемы оценивания параметров регрессионной модели: методы пошаговой регрессии, использование гребневой регрессии (ридж-регрессии), переход от первоначальных переменных к их главным компонентам. Все вышеприведенные методы реализуются в пакете Statistica.

На рисунке 3 представлено исследование регрессионных остатков после оценивания параметров линейной модели множественной регрессии методом пошагового исключения (Backward stepwise).

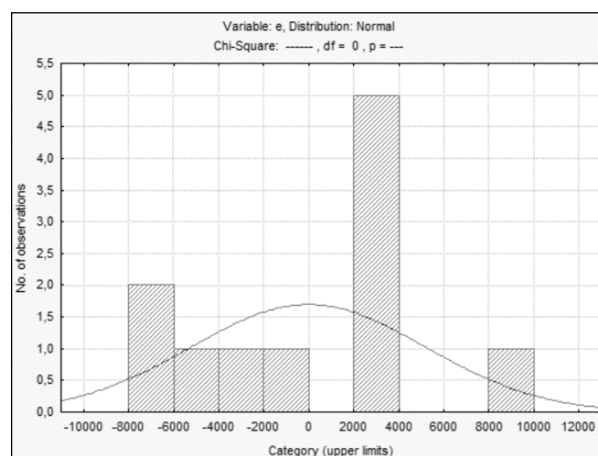


Рисунок 3. Гистограмма распределения регрессионных остатков

Figure 3. Histogram of the distribution of regression residuals

На основании полученного графика, можно предположить, что в разрезе большего временного интервала нулевая гипотеза о том, что распределение регрессионных остатков не отличается от нормального, может быть принята.

Обсуждение

В результате проведения пошаговой регрессии получена следующая оценка уравнения регрессии: $\hat{y} = 0,1433x_5 + 0,3281x_6 + 427,8377x_8$.

(0,0309) (0,0483) (96,2992)

Оценка уравнения регрессии значима, т. к. нулевая гипотеза отклонена; коэффициенты при переменных также значимы. Коэффициент детерминации составил 0,9991, т. е. 99,91% доли вариации результирующей переменной объясняется переменными x_5 , x_6 и x_8 , а 0,09% доли вариации, вероятно, объясняется неучтенными в модели факторами. Другими словами, на изменение объема инвестиций в основной капитал региона большее влияние оказывают такие показатели, как

- уровень добычи полезных ископаемых;
- объем продукции обрабатывающего производства;
- удельный вес численности населения в трудоспособном возрасте.

Согласно полученной регрессионной модели, можно сделать выводы о том, что увеличение уровня добычи полезных ископаемых на 100 млн рублей приводит к увеличению объема инвестиций в основной капитал региона на

14,33 млн рублей, а при увеличении на 100 млн рублей объема продукции обрабатывающего производства инвестиции в основной капитал региона увеличиваются на 32,81 млн рублей. Кроме того, объем инвестиций в основной капитал региона увеличиваются на 427,8377 млн рублей при увеличении на 1% удельного веса численности населения в трудоспособном возрасте в общей численности. С экономической точки зрения полученные выводы можно считать адекватными.

Заключение

На основании результатов корреляционно-регрессионного анализа можно сделать вывод о том, что на уровень объема инвестиций в основной капитал Оренбургской области наиболее значимое влияние оказывают такие сферы экономической деятельности, как сельское хозяйство, добыча полезных ископаемых и обрабатывающие производства. Следовательно, региональным органам управления инвестиционной деятельностью и экономического развития необходимо обратить внимание на реализацию инвестиционного потенциала именно этих отраслей экономики.

Привлечение инвестиций в выявленные отраслевые приоритеты инвестиционной привлекательности Оренбургской области обеспечит создание высокопроизводительных мест на территории региона, увеличение валового регионального продукта, а также даст мультипликативный эффект для развития других видов деятельности.

Литература

- 1 Кособуцкая А.Ю., Равуанжинирова А.В. Инвестиционная привлекательность региона: методики оценки // Вестник ВГУ. Серия: Экономика и управление. 2019. № 1. С. 32–37.
- 2 Литвинцева Г.П., Голдобина А.А. Факторы и пути повышения инвестиционной привлекательности региона // Идеи и идеалы. 2019. № 4–2. С. 243–266. doi: 10.17212/2075–0862–2019–11.4.2–243–266
- 3 Болодурин М.П., Мишурова А.И. Стандартизация оценки инвестиционной привлекательности туристско-рекреационного комплекса региона // Национальные интересы: приоритеты и безопасность. 2019. Т. 15. № 11 (380). С. 2108–2127. doi: 10.24891/ni.15.11.2108
- 4 Чимириш А.С. Анализ инвестиционной привлекательности регионов России // Международный журнал гуманитарных и естественных наук. 2019. № 5–4. С. 143–148. doi: 10.24411/2500–1000–2019–11035
- 5 Валерианов А.А., Дмитриева А.Г., Леванова Т.А. Статистические методы анализа в оценке инвестиционного потенциала региона // Вестник НГИЭИ. 2018. № 6 (85). С. 138–150.
- 6 Дмитриев Н.Д., Родионов Д.Г., Кубарский А.В. Формирование эконометрического инструментария для оценки инвестиционной привлекательности региона // Kant. 2020. № 4(37). С. 70–77. doi: 10.24923/2222–243X.2020–37.15
- 7 Официальная статистика: Территориальный орган Федеральной службы государственной статистики по Оренбургской области. URL: <https://orenstat.gks.ru/ofstatistics>
- 8 Национальный рейтинг состояния инвестиционного климата в субъектах РФ. Агентство «АСИ»: агентство стратегических инициатив по продвижению новых проектов. Москва: Агентство стратегических инициатив, 2021. URL: https://asi.ru/government_officials/rating/
- 9 Рейтинг инвестиционной привлекательности регионов RAEX за 2020 год: RAEX (РАЭК-Аналитика): рейтинговое агентство. Москва: РАЭК-Аналитика. 2020. URL: <https://raex-a.ru/ratings/regions/2020#graph>
- 10 Соловьёв И.В. Геодезия и прикладная информатика // Вестник МГТУ МИРЭА. 2014. №. 2. С. 3.
- 11 Моисеев Н.А. Сравнительный анализ эффективности методов устранения мультиколлинеарности // Учет и статистика. 2017. № 2 (46). С. 62–73.
- 12 Ерёмкина Д.В. Прикладная информатика для специалистов аграрного сектора // Агропродовольственная политика России. 2017. №. 9. С. 98–103.
- 13 Ali E.B., Amfo B. Comparing the values of economic, ecological and population indicators in High – and Low-Income Economies // Ekonomika regiona. 2021. №. 17(1). P. 72–85. doi: 10.17059/ekon. reg.2021–1–6


- 14 Ayres R., Voudouris V. The economic growth enigma: Capital, labour and useful energy? // Energy Policy. 2014. №. 64. P. 16–28.
- 15 Baltina A., Bolodurina M., Gorbatenko E. The application of information technologies in the development of corporate growth strategy // EMIT 2018 Internationalization of Education in Applied Mathematics and Informatics for HighTech Applications. Electronic data. 2018. №. 2093. P. 58–66.
- 16 Ramachandran K.M., Tsokos C.P. Mathematical statistics with applications in R. Academic Press, 2020.
- 17 Gabdrahmanova K.F., Izmailova G.R., Samigullina L.Z. Methods of mathematical statistics application in assessing the density of actual and forecasting distribution density of residual oil reserves // IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2020. V. 860. №. 1. P. 012008.
- 18 Zhang Z., Wu Y., Zhang R., Jiang P. et al. Novel transformer fault identification optimization method based on mathematical statistics // Mathematics. 2019. V. 7. №. 3. P. 288. doi: 10.3390/math7030288
- 19 Nodira S., Gulchekhra K., Nodira A., Maprat R. et al. Teaching Currently Using Interactive Methods in Problem" Probability Theory and Mathematical Statistics" // Central Asian journal of mathematical theory and computer sciences. 2021. V. 2. №. 4. P. 26-28.
- 20 Luo W. Application of improved clustering algorithm in investment recommendation in embedded system // Microprocessors and Microsystems. 2020. №. 75. P. 1–8. doi: 10.1016/j.micpro.2020.103066
- 21 Donghyun C. et al. Using genetic algorithm to support clustering-based portfolio optimization by investor information // Appl. Soft Comput. 2017. №. 61. P. 593–602.

References


- 1 Kosobutskaya A. Yu., Ravuaginirina A.V. Investment attractiveness of the region: assessment methods. Bulletin of the VSU. Series: Economics and Management. 2019. no. 1. pp. 32–37. (in Russian).
- 2 Litvintseva, G.P., Goldobina, A.A. Factors and ways to increase the investment attractiveness of the region. Ideas and Ideals. 2019. no. 4–2. pp. 243–266. doi: 10.17212/2075–0862–2019–11.4.2–243–266 (in Russian).
- 3 Bolodurina M.P., Mishurova A.I. Standardization of the assessment of the investment attractiveness of the tourist and recreational complex of the region. National interests: priorities and security. 2019. vol. 15. no. 11 (380). pp. 2108–2127. doi: 10.24891/ni.15.11.2108 (in Russian).
- 4 Chimiris A.S. Analysis of investment attractiveness of Russian regions. International Journal of Humanities and Natural Sciences. 2019. no. 5–4. pp. 143–148. doi: 10.24411/2500–1000–2019–11035 (in Russian).
- 5 Valerianov A.A., Dmitrieva A.G., Levanova T.A. Statistical methods of analysis in assessing the investment potential of the region. Bulletin of the NGIEI. 2018. no. 6 (85). pp. 138–150. (in Russian).
- 6 Dmitriev N.D., Rodionov, D.G., Kubarski A.V. Formation of econometric tools for evaluating the investment attractiveness of the region. Kant. 2020. no. 4 (37). pp. 70–77. doi: 10.24923/2222–243X.2020–37.15 (in Russian).
- 7 Official statistics: the Territorial body of Federal state statistics service of the Orenburg region. Available at: <https://orenstat.gks.ru/ofstatistics> (in Russian).
- 8 National rating of the investment climate in the subjects of the Russian Federation. Agency "ASI": agency for strategic initiatives to promote new projects. Moscow: Agency for Strategic Initiatives, 2021. Available at: https://asi.ru/government_officials/rating/ (in Russian).
- 9 Rating of investment attractiveness of RAEX regions for 2020: RAEX (RAEX-Analytics): rating agency. Moscow: RAEX-Analytics. 2020. Available at: <https://raex-a.ru/ratings/regions/2020#graph> (in Russian).
- 10 Solovyov I.V. Geodesy and applied informatics. Bulletin of MSTU MIREA. 2014. no. 2. pp. 3. (in Russian).
- 11 Moiseev N.A. Comparative analysis of the effectiveness of multicollinearity elimination methods. Accounting and statistics. 2017. no. 2 (46). pp. 62–73. (in Russian).
- 12 Eremina D.V. Applied informatics for specialists in the agricultural sector. Agro-food policy of Russia. 2017. no. 9. pp. 98-103. (in Russian).
- 13 Ali E.B., Amfo B. Comparing the values of economic, ecological and population indicators in High – and Low-Income Economies. Ekonomika regiona. 2021. no. 17(1). pp. 72–85. doi: 10.17059/ekon. reg.2021–1–6
- 14 Ayres R., Voudouris V. The economic growth enigma: Capital, labour and useful energy? Energy Policy. 2014. no. 64. pp. 16–28.
- 15 Baltina A., Bolodurina M., Gorbatenko E. The application of information technologies in the development of corporate growth strategy. EMIT 2018 Internationalization of Education in Applied Mathematics and Informatics for HighTech Applications. Electronic data. 2018. no. 2093. pp. 58–66.
- 16 Ramachandran K.M., Tsokos C.P. Mathematical statistics with applications in R. Academic Press, 2020.
- 17 Gabdrahmanova K.F., Izmailova G.R., Samigullina L.Z. Methods of mathematical statistics application in assessing the density of actual and forecasting distribution density of residual oil reserves. IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2020. vol. 860. no. 1. pp. 012008.
- 18 Zhang Z., Wu Y., Zhang R., Jiang P. et al. Novel transformer fault identification optimization method based on mathematical statistics. Mathematics. 2019. vol. 7. no. 3. pp. 288. doi: 10.3390/math7030288
- 19 Nodira S., Gulchekhra K., Nodira A., Maprat R. et al. Teaching Currently Using Interactive Methods in Problem" Probability Theory and Mathematical Statistics". Central Asian journal of mathematical theory and computer sciences. 2021. vol. 2. no. 4. pp. 26-28.
- 20 Luo W. Application of improved clustering algorithm in investment recommendation in embedded system. Microprocessors and Microsystems. 2020. no. 75. pp. 1–8. doi: 10.1016/j.micpro.2020.103066
- 21 Donghyun C. et al. Using genetic algorithm to support clustering-based portfolio optimization by investor information. Appl. Soft Comput. 2017. no. 61. pp. 593–602.

Сведения об авторах


Ирина П. Болодурина д.т.н., профессор, кафедра прикладной математики, Оренбургский государственный университет, пр-т Победы, 13, г. Оренбург, 460018, Россия, ipbolodurina@yandex.ru

 <https://orcid.org/0000-0003-0096-2587>

Марина П. Болодурина к.э.н., доцент, кафедра управления персоналом, сервиса и туризма, Оренбургский государственный университет, пр-т Победы 13, г. Оренбург, 460018, Россия, bolodurina@inbox.ru


 <https://orcid.org/0000-0002-6414-245X>

Карина М. Абелгазина студент, факультет математики и информационных технологий, Оренбургский государственный университет, пр-т Победы, 13, г. Оренбург, 460018, Россия, abelgazina99@mail.ru


 <https://orcid.org/0000-0003-2596-6402>

Information about authors


Irina P. Bolodurina Dr. Sci. (Engin.), professor, applied mathematics department, Orenburg State University, 13 Pobedy Av., Orenburg, 460018, Russia, ipbolodurina@yandex.ru

 <https://orcid.org/0000-0003-0096-2587>

Marina P. Bolodurina Cand. Sci. (Econ.), associate professor, personnel management, service and tourism department, Orenburg State University, 13 Pobedy Av., Orenburg, 460018, Russia, bolodurina@inbox.ru

 <https://orcid.org/0000-0002-6414-245X>

Karina M. Abelgazina student, applied mathematics department, Orenburg State University, 13 Pobedy Av., Orenburg, 460018, Russia, abelgazina99@mail.ru

 <https://orcid.org/0000-0003-2596-6402>

Вклад авторов

Все авторы в равной степени принимали участие в написании рукописи и несут ответственность за плагиат

Contribution

All authors are equally involved in the writing of the manuscript and are responsible for plagiarism

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Conflict of interest

The authors declare no conflict of interest.

Поступила 23/12/2021	После редакции 24/01/2022	Принята в печать 18/02/2022
Received 23/12/2021	Accepted in revised 24/01/2022	Accepted 18/02/2022