

К методологии проверки интегральных оценок социально-экономических объектов*

**Дмитрий Александрович АЛФЕРЬЕВ**

Вологодский научный центр Российской академии наук
Вологда, Российская Федерация
Санкт-Петербургский политехнический университета Петра Великого
Санкт-Петербург, Российская Федерация
e-mail: alferiev_1991@mail.ru
ORCID: 0000-0003-3511-7228; ResearcherID: I-8333-2016

**Александр Евгеньевич КРЕМИН**

Вологодский научный центр Российской академии наук
Вологодский институт права и экономики ФСИН России
Вологда, Российская Федерация
e-mail: akremin@inbox.ru
ORCID: 0000-0001-9405-2255; ResearcherID: I-8188-2016

**Дмитрий Григорьевич РОДИОНОВ**

Санкт-Петербургский политехнический университета Петра Великого
Санкт-Петербург, Российская Федерация
e-mail: drodionov@spbstu.ru
ORCID: 0000-0002-1254-0464; ResearcherID: A-9693-2017

**Дарья Сергеевна ВЕЛИЧЕНКОВА**

Санкт-Петербургский политехнический университета Петра Великого
Санкт-Петербург, Российская Федерация
e-mail: velichenkova_ds@spbstu.ru
ResearcherID: ABG-8363-2020

* Статья подготовлена при поддержке гранта Президента РФ (МК-2249.2020.6).

Для цитирования: К методологии проверки интегральных оценок социально-экономических объектов / Д.А. Алферьев, А.Е. Кремин, Д.Г. Родионов, Д.С. Величенкова // Экономические и социальные перемены: факты, тенденции, прогноз. 2021. Т. 14. № 6. С. 86–106. DOI: 10.15838/esc.2021.6.78.5

For citation: Alfer'ev D.A., Kremin A.E., Rodionov D.G., Velichenkova D.S. On the methodology of checking integral estimates of socio-economic objects. *Economic and Social Changes: Facts, Trends, Forecast*, 2021, vol. 14, no. 6, pp. 86–106. DOI: 10.15838/esc.2021.6.78.5

Аннотация. Достоверная и качественная оценка научно-технологического и инновационного развития территорий позволяет определить социально-экономические условия и спрогнозировать динамику экономического роста данного субъекта. Одним из наиболее распространенных подходов к оценке науки и инновационной деятельности, а также других социально-экономических объектов является применение интегральных показателей. Но из-за того, что оценивается объект нематериальный, а некая собирательная синтетическая категория, возникает вопрос, насколько точно подобная метрика характеризует то, что призвана описывать. В связи с этим актуальны вопросы, касающиеся разработки методик для проверки обозначенных выше положений. Цель исследования может быть обозначена как определение достоверности искусственно выведенных интегральных комплексных оценок, которые в свою очередь описывают различные процессы и явления социально-экономического характера. Научная новизна исследования заключается в разработке подхода к определению достоверности интегральных метрик на основе инструментария математической статистики. Посредством соответствующих методов (дисперсионный, корреляционный и регрессионный анализ) на примере индекса инновационного развития субъектов РФ и оценки научно-технологического потенциала территорий была предпринята попытка определить качество искусственно выведенных интегральных комплексных оценок, которые, по мнению их создателей, характеризуют различные проявления науки, инноваций и технологий. Полученные результаты представляют практическую значимость в отношении проводимых в российской экономике оценочных практик научной и инновационной сферы. Теоретическая значимость исследования может быть охарактеризована разработкой подхода, который может применяться к другим объектам социально-экономического характера. Сделан вывод о том, что интегральные оценки в аспекте их применения к гуманитарным направлениям являются крайне субъективным инструментом и могут быть корректно использованы при условии подтверждения связи с показателями, на которые они должны оказывать влияние или иметь от них сильный отклик.

Ключевые слова: научно-технологическое развитие, интегральная оценка, дисперсионный анализ, корреляционный анализ, регрессионный анализ, логит-функция, экономический рост, инновационная деятельность.

Введение

Прогнозирование экономического роста, развития и социальных последствий является важной задачей для государственных структур. С помощью данного инструмента они могут предсказать последствия планируемых действий и в соответствии с этим в правильном направлении скорректировать свою деятельность и политику, чтобы избежать серьезных социально-экономических потрясений или ускорить момент наступления каких-либо положительных событий.

Для оценки экономического и социального благосостояния какой-либо территории на протяжении многих лет выделяют такой показатель, как ВВП (аналогом для субъектов хозяйствования более низкого уровня является величина ВРП) [1; 2]. Он предназначен характеризовать динамику экономического роста. При оценке субъектов хозяйствования по данной метрике определяются наиболее развитые территории.

В соответствии с этим важной научной проблемой является поиск «рычагов» регулирования, то есть перед научным сообществом стоит задача выделить факторы, изменяя которые можно увеличить или снизить величину ВВП (ВРП), а следовательно, повлиять на экономический рост или развитие.

К ключевым факторам подобного рода относят научно-технологический (в более старых источниках «научно-технический») прогресс и сопряженную с ним инновационную деятельность предприятий и организаций [3; 4; 5]. По мнению ряда авторов [6–11], наука и технологии являются двигателем социально-экономического развития, следовательно, государству, для того чтобы повысить свою конкурентоспособность на международной арене, а также улучшить условия проживания граждан на конкретно взятой территории, стоит реализовывать политику, направленную на поддержку науки и исследовательского сектора [12; 13; 14].

В соответствии с этим разрабатываются различные экономико-математические модели [15; 16] (Р.М. Солоу [8], С. Ребело [9], К. Эрроу [17], П.М. Ромер [18], Р. Лукас [19], Д. Гроссман и Э. Хелпман [20], К. Фримен и Б.-А. Лундвалл [21], Ц. Грилихес [22]), которые в зависимости от выбранных и обоснованных учеными и исследователями показателей описывают функциональные связи между экономическим ростом и каким-либо экономическим объектом (затраты на науку; количество проводимых научно-исследовательских и опытно-конструкторских работ; человеческий капитал; предприятия, осуществляющие и реализующие инновационную деятельность; образовательные учреждения; динамика инновационных идей). Данные модели имеют как теоретическое обоснование на уровне гипотез, так и некоторую эмпирическую реализацию, что в определенной мере является их апробацией и подтверждением сформулированных идей.

Учеными в этом направлении также предпринимаются попытки комплексной оценки научно-технологических изменений в экономической среде. Для этого разрабатываются и рассчитываются различного рода «интегральные» оценки. Примером может служить индекс научно-технического потенциала региона, методика расчета которого опубликована в работах К.А. Задумкина и И.А. Кондакова [23], или оценка научно-технологического потенциала территорий, представленная в статье коллектива авторов под руководством К.А. Гулина [24]. Довольно серьезным и фундаментальным трудом является проект Рейтинга инновационного развития субъектов Российской Федерации, разработанный и публикуемый с 2012 по 2019 год коллективом авторов Института статистических исследований и экономики знаний, входящего в состав ВШЭ¹. Данная методика включает оценку 37 показателей, а также имеет значительный временной период расчета (7 лет).

В целом стоит отметить, что в вышеперечисленных работах в различной мере предпринимались попытки оценить совокупное влияние факторов научно-технологического и инновационного развития на социально-эко-

номический уровень территорий (России как отдельно взятого государства и округов или регионов, входящих в ее состав, представляющих аналогичные единицы меньшего размера). В результате комплексности полученных оценок их интерпретация представляет собой некую условную единицу, характеризующую некий обобщенный процесс или явление (в нашем случае научно-технологический потенциал, инновационное развитие и др.), но не имеющую явной количественной интерпретации, которая может быть получена в ходе проведения измерительных процедур. В результате возникает вопрос об адекватности и достоверности описываемых процессов такими интегральными оценками. Их достоверность может быть подтверждена только по истечении какого-либо времени, что существенно снижает значимость и прикладную применимость подобного рода моделей.

С потребностью оценивания объектов, которые характеризуются неоднородными параметрами, человек сталкивается достаточно часто. Чаще всего подобная оценка проводится интуитивно, как следствие, имеет место отрицательный итог. Использование самих интегральных оценок также сопряжено с рядом проблем, которые могут быть охарактеризованы следующим² [25; 26; 27]:

- необходимость учета веса и значимости каждого из параметров, включенного в состав общей оценки;
- необходимость задать способ перевода качественных оценок в количественный вид;
- распределение оцениваемых объектов в соответствующие группы, характеризующие величину выделенных в исследовании уровней;
- возможность сравнения оценок с теми, которые будут получены в будущем (социально-экономические показатели зачастую имеют непостоянную природу и с течением времени могут быть не актуальны в отличие от физических величин, которыми измеряются объекты естественных наук).

Описанные выше проблемные аспекты интегральных оценок также обозначены в ряде

¹ Рейтинг инновационного развития регионов. URL: <https://issek.hse.ru/rir>

² Минакер В.Е., Быховский М.В. Проблемы интегральных оценок технических систем // ТРИЗ-Саммит – 2006. Санкт-Петербург, 13–14 октября 2006 г. URL: <https://www.metodolog.ru/00821/00821.html>

отечественных научных работ, где предпринята попытка систематизации достоинств и недостатков подобного подхода относительно интерпретации объектов социально-экономического характера. Так, в статье Е.Н. Волковой на основе интегральной оценки сформирована и описана методология характеристики социально-экономического развития региона [28]. В работе Е.В. Ключниковой и Е.М. Шитова обозначены особенности построения интегральных оценок в соответствии с основными этапами моделирования: нормирование, агрегирование, взвешивание [29]. За рубежом проводятся аналогичные исследования. Одной из подобных работ является публикация коллектива ученых под руководством М.-С. Саиба [30], в которой авторами посредством интегрального показателя оценивают неравенство населения территорий относительно условий и факторов воздействия на здоровье.

Из современных работ особое внимание следует обратить на публикацию А.А. Сидорова [31]. Он скрупулезно систематизирует и описывает математическую природу интегрального подхода, в некоторой мере продолжая работы известного российского эконометриста С.А. Айвазяна [32; 33; 34]. В зарубежной среде также присутствуют исследования, связанные с математическим построением интегрального показателя. Так, в статье коллектива авторов под руководством П. Чжоу [35] предложен вариант агрегирования интегральной оценки на основе произведения скорректированных частных значений показателей, входящих в нее.

Несмотря на большое многообразие работ по использованию интегрального подхода применительно к социально-экономическим объектам, в них опускается один момент – правильно ли интегральные показатели характеризуют то, что предназначены описывать? Хотя практически во всех рассмотренных исследованиях оговаривается, что подобного рода оценки несут в себе изрядную долю неопределенного субъективизма.

В связи с этим в статье поставлена цель определить достоверность искусственно выведенных интегральных комплексных оценок посредством математико-статистических методов, которые, в свою очередь, описывают различные процессы и явления социально-экономического характера.

Под «достоверностью» обозначенных в работе оценок мы будем понимать их способность в явном виде описывать те процессы и явления, которые они должны характеризовать.

Для достижения поставленной цели в работе решен ряд задач:

- определены и описаны методы математической статистики, с помощью которых осуществлен поиск взаимосвязи между экономическим ростом и оценкой научно-технического развития;
- сформирована выборка данных для обозначенных в работе расчетов;
- проведены и описаны результаты расчетов, на основании которых были сформулированы соответствующие выводы и рекомендации.

Методы исследования

Алгоритм однофакторного дисперсионного анализа

В математической статистике дисперсионный анализ используется для исследования наличия влияния качественных факторов на значения количественного показателя. В нашем случае за результирующий показатель выступит y : ВРП, а за фактор x : $x \in (I_1 || I_2), (I_1 || I_2) \supset S_1 \& S_2 \& S_3 \& S_4$, где I_1 – российский региональный инновационный индекс, I_2 – комплексная оценка научно-технологического потенциала, а $\{S_n\}_{n=1}^4$ – входящие в их состав субиндексы.

При помощи однофакторного дисперсионного анализа попробуем определить, является ли статистически значимым различие y в субъектах Российской Федерации, отличающихся по признаку I , наблюдаемом на $k = 4$ уровнях (социально-экономические условия, научно-технический потенциал, инновационная деятельность, качество инновационной политики в одном случае и исследования и разработки, кадры, технологии и инновации в другом).

Алгоритм формирования выборок для проведения однофакторного дисперсионного анализа

Алгоритм отбора элементов четырех различных выборок в соответствии с уровнями I_1 сформируем на основе рейтингов субъектов Российской Федерации по субиндексам российского регионального инновационного индекса. Полученные учеными и исследователями ИСИЭЗ распределения включают в себя четыре

группы, которые можно охарактеризовать следующим образом (группы смоделированы авторами рассматриваемого исследования, но в них отсутствует явное обоснование пограничных значений; вариант возможной интерпретации представлен далее):

I группа – города федерального значения (Москва, Санкт-Петербург), имеющие наилучшие показатели по большинству статистических метрик, характеризующих социально-экономическое развитие. По своей природе их оценки разнятся с другими аналогичными объектами (региональные территориальные единицы) в несколько раз, из-за чего они явно выбиваются из общей картины распределения и выглядят как выбросы. В связи с этим при проведении однофакторного дисперсионного анализа они будут исключены из нашего исследования, что позволит оценить действительно равнозначные объекты между собой;

II группа – региональные территориальные субъекты, обладающие наилучшими значениями по оцениваемым признакам (за исключением объектов, входящих в I группу), зачастую превышающими средние общероссийские оценки. Подобные субъекты могут рассцениваться как те, которые в полной мере обладают наличием рассматриваемого признака;

III группа – регионы РФ, зачастую имеющие значения по S_n оценкам, меньшие, чем среднероссийские характеристики; данные территории только подбираются к качественному уровню исследуемых признаков I и поэтому не могут считаться в полной мере обладающими исследуемыми характеристиками, а значит, их следует исключить из дальнейших расчетов;

IV группа – территориальные единицы с наименьшими значениями по рассматриваемым признакам; их можно охарактеризовать сильным разбросом оценок в разрезе динамики, нестабильностью темпов роста и отсутствием явно выраженных трендов; в связи с этим в исследуемую выборку подобные объекты включены не будут.

Алгоритм отбора элементов четырех различных выборок в соответствии с уровнями I_2 сформируем на основе рейтингов субъектов Российской Федерации по субиндексам оценки

научно-технологического потенциала территорий. Полученные учеными и исследователями ВолНИЦ РАН распределения включают в себя пять групп (уровни: высокий, выше среднего, средний, ниже среднего, низкий), которые можно охарактеризовать как равномерно распределенные отмасштабированные до целого десятка ($I_2 \in [0; 10]$).

В соответствии с названным выше правилом включения наблюдений в дисперсионный анализ (по аналогии с рейтингом субъектов Российской Федерации) нас интересуют группы «среднего» и «выше среднего» уровня. В них, как и в предыдущем варианте, мы подразумеваем присутствие рассматриваемой интегральной комплексной оценки, следовательно, данные наблюдения должны найти отражение в изменении динамики ВРП. Территории, вошедшие в выборку с высоким уровнем, являются уникальными единичными объектами и своего рода представляют выбросы. Группы «ниже среднего» и «низкого» уровня включают в себя основную часть наблюдений и, по сути, представляют собой идентичные объекты без каких-либо характерных особенностей (на 2015 год их доли в выборке в соответствии с выделенными группами субиндексов составили 93,75; 95; 95 и 98,75% соответственно).

На этом алгоритм формирования выборок заканчивается, и далее продолжается алгоритм однофакторного дисперсионного анализа.

Осуществив выборку регионов РФ, определим количество объектов, включенных в каждый из k уровней ($k = 4$) как сумму элементов, входящих в рассматриваемую совокупность, и обозначим ее через m_n , где n – порядковый номер уровня.

Затем определим общее количество (m) вошедших во все выборки объектов (формула 1):

$$m = \sum_{n=1}^k m_n . \quad (1)$$

Далее рассчитаем среднее значение ВРП по каждой из сформированных групп k (формула 2):

$$\bar{y}_n = \frac{1}{m_n} \sum_{i=1}^{m_n} y_{ni} , \quad (2)$$

где y_{ni} – значение ВРП, соответствующее i региону в выборке n .

На следующем шаге найдем среднее значение результирующей переменной y по всем имеющимся значениям, входящим во все выборки n (формула 3):

$$\bar{y} = \frac{1}{m} \sum_{n=1}^k \sum_{i=1}^{m_n} y_{ni} \leftrightarrow \frac{1}{m} \sum_{n=1}^k \bar{y}_n m_n. \quad (3)$$

Дальнейший этап – поиск суммы квадратов отклонений результирующих оценок (y_{ni}) по выборкам от общего среднего (\bar{y}) (формула 4):

$$S^2 = \sum_{n=1}^k \sum_{i=1}^{m_n} (y_{ni} - \bar{y})^2. \quad (4)$$

Затем – сумма квадратов отклонений средних по группам (\bar{y}_n) от общего среднего значения (\bar{y}) (формула 5):

$$S_F^2 = \sum_{n=1}^k (\bar{y}_n - \bar{y})^2 m_n. \quad (5)$$

Далее рассчитаем остаточную сумму квадратов отклонений как сумму квадратов разности результирующих значений (y_{ni}) от средних значений (\bar{y}), входящих в одну и ту же выборку n (формула 6):

$$S_{\text{ост}}^2 = \sum_{n=1}^k \sum_{i=1}^{m_n} (y_{ni} - \bar{y}_n)^2. \quad (6)$$

Для проверки S^2 (формула 4) может быть использовано следующее равенство (формула 7):

$$S_F^2 + S_{\text{ост}}^2 = S^2. \quad (7)$$

Вычислим факторную дисперсию (формула 8):

$$\sigma_F^2 = \frac{S_F^2}{k-1}. \quad (8)$$

Вычислим остаточную дисперсию (формула 9):

$$\sigma_{\text{ост}}^2 = \frac{S_{\text{ост}}^2}{m-k}. \quad (9)$$

Далее найдем значение F_ϕ по формуле частного факторной дисперсии (σ_F^2) к остаточной ($\sigma_{\text{ост}}^2$) (формула 10):

$$F_\phi = \frac{\sigma_F^2}{\sigma_{\text{ост}}^2}. \quad (10)$$

Используя распределение Фишера – Снедекора, при заданном уровне значимости (α) и

двух степеней свободы (формулы 11 и 12) можно определить значение метрики F_k :

$$df1 = k - 1; \quad (11)$$

$$df2 = m - k. \quad (12)$$

Для поиска F_k могут применяться классические таблицы распределения Фишера – Снедекора, фигурирующие в справочниках по математической статистике, но при их использовании ученый может столкнуться с рядом проблем. Это отсутствие необходимых числовых значений, на которых строится исследование, и схожая ситуация, связанная с выбором уровня надежности (зачастую в справочной литературе присутствуют лишь $\alpha: \alpha = 0.1 || 0.05 || 0.01$). Подобную проблему в настоящий момент несложно решить, используя возможности вычислительных компьютерных программ (например, функция « F_k » библиотеки Python «Scipy»).

По итогу необходимо сравнить полученные значения F_ϕ & F_k (формула 13):

$$F_\phi > F_k \rightarrow A, \quad (13)$$

где A – утверждение о том, что исследуемые качественные признаки действительно оказывают влияние на величину результирующего показателя.

Алгоритм корреляционного анализа

При помощи корреляционного анализа определяют наличие статистической взаимосвязи между несколькими случайными величинами. Одним из таких инструментов является коэффициент корреляции Пирсона (r), который указывает на наличие линейной взаимосвязи. Данная величина принимает значения от -1 до 1 включительно и в случае отрицательного значения указывает на уменьшение результирующей оценки по отношению к увеличению факторной. В случае положительного значения метрики r ситуация диаметрально противоположная. Чем ближе значение r по абсолютной величине к единице, тем больше имеющаяся между показателями линейная связь. Стоит отметить, что при $r = 0$ линейная связь отсутствует, но не исключается возможность ее описания через нелинейную функцию или функционал. Коэффициент корреляции

между показателями, факторами или результирующим признаком (r_{xy}) может быть найден как отношение ковариации между ними (cov_{xy}) и произведением соответствующих средних квадратичных отклонений ($\sigma_x \& \sigma_y$) (формула 14):

$$r_{xy} = \frac{\text{cov}_{xy}}{\sigma_x \sigma_y},$$

где $\text{cov}_{xy} = M((x_n - \mu(x_n))(y_n - \mu(y_n))) \leftrightarrow$ (14)
 $\leftrightarrow M(x_n y_n) - \mu(y_n)\mu(x_n)$;

$M()$ – несмещенная оценка математического ожидания по выборке;

$\mu()$ – среднее значение исследуемых наблюдений.

Алгоритм построения модели множественной линейной регрессии

Модель множественной линейной регрессии является инструментом, используемым в многомерном статистическом анализе для описания связи признаков (причин) с каким-либо результатом или следствием. Ее общий вид представлен аналитической формулой (формула 15):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \dots, \quad (15)$$

где $\{\beta_n\}_{n=1}^k$ – коэффициенты регрессии, показывающие степень воздействия фактора на результирующий признак;

β_0 – свободный параметр модели, позволяющий оптимальным образом расположить кривую в пространстве таким образом, чтобы сумма квадратов отклонений (МНК) была наименьшей.

Для расчета коэффициентов β удобно использовать матричный метод поиска (формула 16):

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_z \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1n} \\ 1 & x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{z1} & x_{z2} & \dots & x_{zn} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \rightarrow (16)$$

$$\rightarrow \hat{B} = (X^T X)^{-1} X^T Y.$$

После построения самой модели необходимо оценить ее точность и значимость. Точность модели множественной линейной регрессии определяется при помощи коэффициента

детерминации (R^2), который может быть найден по следующей формуле (формула 17):

$$R^2 = \frac{SS_\phi}{SS}, \quad (17)$$

где $SS_\phi = \sum_{j=1}^z (\hat{y}_i - \bar{y})^2$ – объясненная сумма квадратов регрессии;

$SS = \sum_{j=1}^z (y_i - \bar{y})^2$ – общая сумма квадратов регрессии;

j – порядковый номер наблюдения, входящий в формируемую модель.

Чем ближе данная метрика к единице, тем лучше построенная модель аппроксимирует имеющиеся эмпирические данные.

Для сравнения точности идентичных моделей, различающихся по количеству регрессоров, используют другую метрику, так как коэффициент детерминации, рассчитанный способом по формуле 17, будет всегда лучше при большем количестве параметров. Для сравнения качества подобных регрессий следует использовать скорректированный R^2 (формула 18):

$$R_{adj}^2 = 1 - \frac{(1 - R^2) \cdot (z - 1)}{z - n - 1}. \quad (18)$$

На финальном этапе оценивается значимость построенной модели. Она определяется при помощи F -критерия Фишера. Для этого необходим требуемый уровень значимости (α) и рассчитываем следующие характеристики (формула 19):

$$SS_o = \sum_{j=1}^z (y_i - \hat{y}_i)^2 \leftrightarrow SS - SS_\phi$$

– остаточная сумма квадратов регрессии;

$$(df1_p = n) \& (df2_p = z - (n + 1))$$
 – число степеней свободы;

$$\left(MS_\phi = \frac{SS_\phi}{df1_p} \& MS_o = \frac{SS_o}{df2_p} \right) \rightarrow (F_k)_p = \frac{MS_\phi}{MS_o}.$$

Определение $(F_k)_p$ происходит аналогично процедуре по поиску (F_k) , только в этом случае вместо $df1$ (формула 11) и $df2$ (формула 12) берутся соответственно $df1_p \& df2_p$ (формула 19). Затем аналогично формуле 13 сравниваются новые полученные соответствующие значения.

Алгоритм определения значимости коэффициентов множественной линейной регрессии

Для того чтобы определить значимость коэффициентов множественной линейной регрессии, используют t -критерий Стьюдента. Он позволяет дать ответ на вопрос: «Могут ли коэффициенты, полученные в модели, быть интерпретированы?»

Наблюдаемые значения, полученные из построенной статистической модели, рассчитываются при помощи стандартной ошибки параметров (l_{β_0} & l_{β_n}) β_0 & β_n соответственно (формула 20):

$$t_{\beta_0} = \frac{\beta_0}{l_{\beta_0}} \text{ \& } t_{\beta_n} = \frac{\beta_n}{l_{\beta_n}}; t = t_{\beta_0} || t_{\beta_n}. \quad (20)$$

Стандартные ошибки могут быть найдены из формулы 16 через матрицу A (формула 21):

$$A = (X^T X)^{-1}. \quad (21)$$

Данная матрица является квадратной и определяется размером $(k + 1) \times (k + 1)$. Следовательно, ее диагональный элемент может быть обозначен как a_{nn} . Тогда (формула 22):

$$l_{\beta_0}^2 = MS_0 a_{00}, \quad (22)$$

$$l_{\beta_n}^2 = MS_0 a_{nn}, n = \overline{1, k}, k = 4.$$

Полученные в формуле 21 значения сравниваются с двусторонней критической точкой распределения Стьюдента – $t_k(\alpha; z - n - 1)$. При $|t| > t_k$ соответствующий параметр множественной линейной регрессии считается статистически значимым, нулевая гипотеза вида $H_0: \beta_0 || \beta_n = 0$ отвергается.

Алгоритм построения логистической регрессии

При помощи данного метода может быть определена потенциальная вероятность наступления какого-либо события, в нашем случае – вероятность роста динамики ВРП в зависимости от совокупного годового изменения k -факторов.

В общем виде модель выглядит следующим образом (формула 23):

$$\hat{P} = p_0 + p_1 \cdot \Delta x_1 + p_2 \cdot \Delta x_2 + \dots + p_n \cdot \Delta x_n + \dots, \quad (23)$$

где \hat{P} – предварительная оценка вероятности наступления определенного события;

$\{p_n\}_{n=0}^k$ – коэффициенты регрессии, аналогичные тем, которые характеризуют модель (формула 15);

$\{\Delta x_n\}_{n=1}^k$ – значения годового изменения соответствующего фактора.

Итоговое преобразование (\hat{P}') может быть осуществлено через функцию *sigmoid* и выглядит так (формула 24):

$$\hat{P}' = \text{sigmoid}(\hat{P}) = \frac{1}{(1 + e^{-\hat{P}})}. \quad (24)$$

Алгоритм преобразования годовой динамики ВРП в значения булевой функции

Если y – значение ВРП в настоящий момент, а y^{-1} – в предыдущий, то Δy – благоприятная динамика y по абсолютному отклонению от y^{-1} . Тогда может быть получено P (формула 25):

$$\Delta y \rightarrow P = \begin{cases} 1, & \text{если } \Delta y > 0 \\ 0, & \text{если } \Delta y \leq 0 \end{cases}. \quad (25)$$

В соответствии с обозначенными выше методами далее представлены необходимые статистические формы, которые были разработаны авторами для проведения соответствующих статистических оценок, а также расчетные результаты и вытекающие из этого комментарии, выводы и заключения.

Результаты

Дисперсионный анализ

Для проведения дисперсионного анализа по алгоритму, описанному в первом пункте «Методов исследования», был сформирован массив данных (табл. 1), где представлены значения ключевых субиндексов (социально-экономические условия, научно-технический потенциал, инновационная деятельность, качество инновационной политики), входящих в общую оценку «Рейтинга инновационного развития регионов», разработанного ИСИЭЗ, и аналогичных показателей (исследования и разработки, кадры, технологии, инновации) по оценке научно-технологического потенциала ВолНЦ РАН. Данные взяты в соответствии с последними актуальными расчётами, проводимыми на момент описанного в работе исследования за 2015 год.

Таблица 1. Часть апостериорного набора данных для проведения однофакторного дисперсионного анализа

Региональная территориальная единица	1. Российский региональный инновационный индекс ВШЭ				
	Валовой региональный продукт, млн руб. (y)	Социально-экономические условия инновационной деятельности (n = 1)	Научно-технический потенциал (n = 2)	Инновационная деятельность (n = 3)	Качество инновационной политики (n = 4)
Алтайский край	487903.3			0.4003	0.4948
Астраханская область	322303.0	0.4431			
Владимирская область	368489.2			0.3688	
Волгоградская область	740458.0				0.4828
Вологодская область	478893.0				0.4986
Ивановская область	180517.5		0.3968		
Иркутская область	1001717.6		0.3821		
Калининградская область	349818.6	0.4675			
Калужская область	339760.8	0.4829	0.4385		
Камчатский край	175404.8				0.4444
.....
Количество элементов выборки	m = 82*	m₁ = 28	m₂ = 16	m₃ = 15	m₄ = 23
2. Оценка научно-технологического потенциала ВолНЦ РАН					
		Исследования и разработки (n = 1)	Кадры (n = 2)	Технологии (n = 3)	Инновации (n = 4)
Количество элементов выборки	m = 11*	m₁ = 4	m₂ = 3	m₃ = 3	m₄ = 1

* Общее количество объектов, вошедших во все k выборки.

Составлено по: Рейтинг инновационного развития субъектов Российской Федерации. Выпуск 5. М.: НИУ ВШЭ, 2017. 260 с. URL: <https://issek.hse.ru/news/206633461.html>; Научно-технологический потенциал территорий и его сравнительная оценка / К.А. Гулин [и др.] // Проблемы развития территории. 2017. № 1 (87). С. 7–26. URL: <http://pdt.vssc.ac.ru/article/2102>; Федеральная служба государственной статистики. Валовый региональный продукт (в текущих основных ценах – всего (1998–2017 гг.). Обновлено 01.03.2019. URL: http://www.gks.ru/free_doc/new_site/vvpr/vpr98-17.xlsx (дата обращения 20.04.2021).

Таблица 2. Результаты, полученные в ходе проведения дисперсионного анализа

Наименование показателя	Единица измерения	Символьное обозначение	Расчетное значение	
			I_1	I_2
Среднее значение ВРП по группе S_1	млн руб.	\bar{Y}_1	1 183 444	1 204 667
Среднее значение ВРП по группе S_2	млн руб.	\bar{Y}_2	1 338 762	1 577 121
Среднее значение ВРП по группе S_3	млн руб.	\bar{Y}_3	1 810 813	4 743 038
Среднее значение ВРП по группе S_4	млн руб.	\bar{Y}_4	1 181 169	837 495
Среднее значение ВРП по всем объектам, отобраным в выборки	млн руб.	\bar{Y}	1 327 875	2 237 876
Значение расчетной статистики Фишера	усл. ед.	F_{ϕ}	0.4422	0.4837
Значение критической статистики Фишера	усл. ед.	F_k	2.7218*	4.3468*
* При $\alpha = 0.05$. Составлено по таблице 1.				

В каждую из четырех групп вошли те значения ВРП регионов, для которых соблюдалось условие попадания соответствующего индекса по группе значимости – II группа для методики ВШЭ и группы уровня «средний» и «выше среднего» для методики ВолНЦ РАН.

В соответствии с формулами 1–13 и апостериорным набором данных из таблицы 1 получены статистические метрики (табл. 2), которые могут быть использованы для интерпретации проведенного в работе дисперсионного анализа на предмет определения значимого влияния субиндексов рассматриваемых методик на уровень экономического роста территорий, выраженного в значениях показателя ВРП.

В ходе проведения дисперсионного анализа проверялась гипотеза о наличии воздействия и влияния искусственно выведенных интегральных оценок рейтинга российского регионального инновационного индекса и оценки научно-технологического потенциала на величину валового регионального продукта территорий (динамику экономического роста). При отборе территорий, где были зафиксированы соответствующие оценки по каждому из четырех субиндексов для каждой из методик, были соответственно сформированы восемь выборок, характеризующихся высоким проявлением процессов, присущих одной из четырех ключевых групп исследуемых рейтингов.

Расчетное значение F -статистики как в одном, так и в другом случае ($F_{\phi 1} = 0.4422$, $F_{\phi 2} = 0.4837$) оказалось в значительной мере меньше критического значения ($F_{k1} = 2.7118$, $F_{k2} = 4.3468$) при заданном уровне надежности в 5% и полученных соответствующих степенях свободы ($df_{1||2} = 3$ & $(df_{2_1} = 78 \parallel df_{2_2} = 7)$)

(см. табл. 2). Из этого может следовать вывод, что различия между группами регионов, вошедших в выборки по большим значениям (но не максимально возможным) ключевых индикаторов рассматриваемых методик являются статистически незначимыми в отношении различий по величине ВРП. В первую очередь это может указывать на два момента: либо данные индикаторы не оказывают существенного влияния на изменение экономического роста регионов, либо сравниваемые комплексные оценки являются взаимообусловленными.

Панель данных для проведения корреляционно-го анализа, построения множественной линейной регрессии и логит-модели

Для проведения выделенных анализов и вариантов моделирования была сформирована статистическая панель, сопоставляющая значения изменений динамики субиндексов Рейтинга инновационного развития субъектов Российской Федерации и оценки научно-технологического потенциала территорий РФ (Δn – столбец 8 и др. табл. 3) с фактом положительного роста динамики ВРП (столбец 5 табл. 3). Сформированные данные использованы при проведении корреляционного анализа, где была задействована информация по субиндексам в 2015 году (столбец 7 и др. табл. 3) и соответствующим данным, характеризующим значения ВРП в этот год (столбец 4 табл. 3). Этот же набор данных использован и при построении модели множественной линейной регрессии. Для создания логит-модели необходимо было установить факт роста ВРП (столбец 5 табл. 3), как положительного отклонения к предыдущему году при абсолютном отклонении факторов (столбец 8 и др. табл. 3), потенциально оказывающих влияние на этот процесс.

Таблица 3. Часть апостериорного набора данных для проведения корреляционного анализа, построения модели множественной линейной и логистической регрессии

Территория	ВРП, 2014 г. (млн руб.)		ВРП, 2015 г. (млн руб.)	Динамика роста	Соц.-эконом. усл. иннов. деят., 2014 (n = 1)	Соц.-эконом. усл. иннов. деят., 2015 (n = 1)	ΔЛ ₁ по абсолютной величине
	в текущих ценах	в ценах 2015 г.*					
1	2	3	4	5	6	7	8
Алтайский край	446023.8	479475.6	487903.3	↗ 1	0.3042	0.3038	-0.0207
Амурская область	232053.0	249457.0	277380.4	↗ 1	0.2893	0.3151	-0.0667
Архангельская область	542695.3	583397.4	627698.1	↗ 1	0.3317	0.2977	-0.0243
Астраханская область	296319.3	318543.2	322303.0	↗ 1	0.3997	0.4431	-0.0509
Белгородская область	619677.7	666153.5	693379.4	↗ 1	0.4049	0.3946	-0.0191
Брянская область	242722.4	260926.6	271782.5	↗ 1	0.3245	0.3446	0.0456
Владимирская область	328064.2	352669.0	368489.2	↗ 1	0.3755	0.3651	-0.0101
Волгоградская область	715409.6	769065.3	740458.0	0 ↘	0.3255	0.3435	-0.0947
Вологодская область	387211.7	416252.6	478893.0	↗ 1	0.3110	0.3191	-0.0215
Воронежская область	717667.2	771492.2	805969.6	↗ 1	0.4253	0.3993	-0.0391
.....

* Индекс дефлятор – 1,075 усл. ед.
Составлено по: Рейтинг инновационного развития субъектов Российской Федерации. Выпуск 4-5. URL: <https://issek.hse.ru/rir/>
Научно-технологический потенциал территорий и его сравнительная оценка / К.А. Гулин [и др.] // Проблемы развития территории. 2017. № 1 (87). С. 7–26. URL: <http://pdt.vssc.ac.ru/article/2102>
Федеральная служба государственной статистики. Индексы-дефляторы, в % к предыдущему году. Обновлено 01.04.2021. URL: http://www.gks.ru/free_doc/new_site/vvp/vvp-god/tab4a.xls (дата обращения 20.04.2021).

Корреляционный анализ

Матрица из *таблицы 4* характеризует линейную взаимосвязь между результирующим показателем экономического роста региона (ВРП) и индикаторами, являющимися ключевыми субиндексами Рейтинга инновационного развития субъектов Российской Федерации. Анализ данных проведен на основании статистики за 2015 год.

В соответствии с каждым из ключевых показателей Рейтинга инновационного развития субъектов РФ и величины ВРП, присущей рассматриваемым территориям, сделан ряд соответствующих выводов.

1) Валовой региональный продукт

Прослеживается линейная связь между социально-экономическими условиями деятельности. Обнаружен отклик от значений показателя, характеризующего научно-технический потенциал. Показатели, описывающие уровень инновационной деятельности и качество инновационной политики, линейно практически никак не отражаются на росте или снижении экономического роста регионов РФ.

2) Социально-экономические условия инновационной деятельности

Данный показатель сильно коррелирует со значениями показателя, характеризующего научно-технический потенциал регионов РФ. Этим может быть обусловлена высокая корреляционная

связь между ВРП и научно-техническим потенциалом. Прослеживается связь с показателями инновационной деятельности, с качеством инновационной политики – несущественная.

3) Научно-технический потенциал

Значения данного показателя коррелируют со значениями показателей, характеризующих инновационную деятельность регионов России. Прослеживается более сильная связь с качеством инновационной политики, чем у предыдущих показателей.

4) Инновационная деятельность и качество инновационной политики

Данные показатели в большей мере коррелируют между собой, нежели остальные.

Общий вывод

Из всего этого следует, что есть явно выраженная взаимообусловленность между рассматриваемыми метриками. Это подтверждает выводы, полученные в ходе проведения дисперсионного анализа, который показал отсутствие статистически значимой разницы между величиной ВРП объектов, входящих в разные категориальные группы по уровням инновационного развития регионов.

Матрица из *таблицы 5* характеризует линейную взаимосвязь между ВРП и ключевыми субиндексами рейтинга, оценивающего научно-технологический потенциал территорий (2015 г.).

Таблица 4. Результаты проведенного корреляционного анализа по методике ИСИЭЗ

	ВРП	n_1	n_2	n_3	n_4
ВРП	1	0.6669	0.4540	0.2652	0.1571
n_1	0.6669	1	0.5905	0.3935	0.2289
n_2	0.4540	0.5905	1	0.4007	0.2897
n_3	0.2652	0.3935	0.4007	1	0.4892
n_4	0.1571	0.2289	0.2897	0.4892	1

Составлено по таблице 3.

Таблица 5. Результаты корреляционного анализа, проведенного по методике ВолНЦ РАН

	ВРП	n_1	n_2	n_3	n_4
ВРП	1	0.6674	0.6862	0.4558	0.3051
n_1	0.6674	1	0.8034	0.8205	0.3481
n_2	0.6862	0.8034	1	0.6430	0.2631
n_3	0.4558	0.8205	0.6430	1	0.3783
n_4	0.3051	0.3481	0.2631	0.3783	1

Составлено по таблице 3.

1) Валовой региональный продукт

В соответствии с практикой эконометрического моделирования связи ВРП с показателями, которые выбраны в качестве предикторов, слабые и не могут быть в должной мере использованы для прогнозирования или интерпретации движения динамики результирующей оценки как линейного отклика от них.

2) Исследования и разработки

Группа показателей, имеющая наибольшие коррелирующие значения по всем исследуемым наборам субиндексов методики ВолНЦ РАН. Значительный ее уровень наблюдается для показателей групп «Технологии» (0.8) и «Инновации» (0.82). При этом сами эти группы между собой коррелируют значительно меньше (0.64), что, возможно, может быть успешно использовано для построения модели множественной линейной регрессии, где в качестве результирующей оценки выступают «Исследования и разработки». Однако это также указывает, что данные третьей группы в значительной мере дублируют друг друга.

3) Технологии

Слабая линейная связь с группой показателей «Инновации» и «Кадры».

4) Инновации

Практически отсутствующая линейная связь с последней группой показателей, характеризующих абстрактную систематизацию научных кадров.

Общий вывод

Наблюдается ситуация, аналогичная случаю с корреляционным анализом, проведенным по рейтингу оценки инновационного развития субъектов РФ. Как особый момент следует выделить связь показателей групп «Технологии» и «Инновации» с «Исследованиями и разработками». Тем не менее все так же присутствует сильная взаимообусловленность одних метрик с другими, используемыми для расчета конечной результирующей оценки, что указывает на необходимость применения эффективной методики поиска весов при формировании итогового расчета, либо снижения размерности количества предикторов за счет исключения несущественных факторов.

Множественная линейная регрессия

Таблица 6 отражает основные статистические метрики полученных результатов по по-

строенным в работе статистическим моделям множественной линейной регрессии, а именно параметры регрессионной модели, как базовые, так и модифицированные; коэффициенты детерминации и их скорректированную оценку; расчетную и критическую F -статистику Фишера при заданном уровне надежности 5% и оценку значимости полученных параметров моделей с использованием t -статистики.

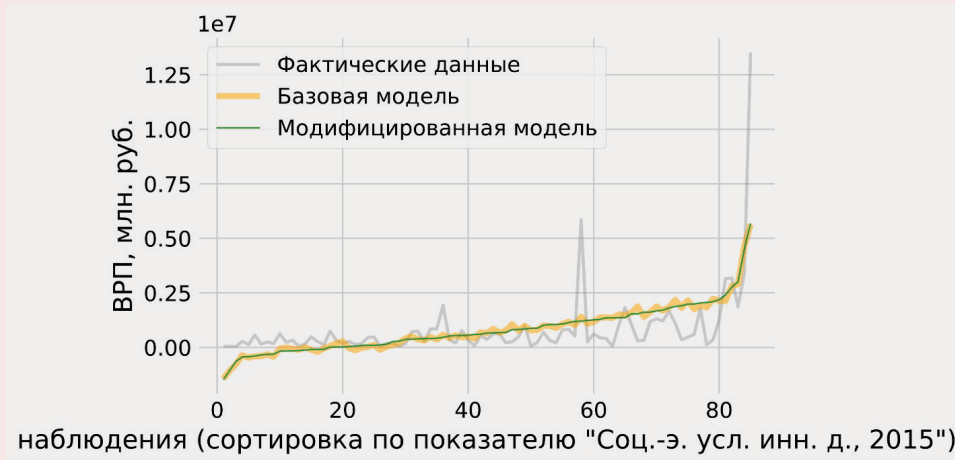
График фактических значений величины ВРП регионов в 2015 году, отсортированных по возрастанию значений субиндекса Рейтинга инновационного развития субъектов России (рис. 1), характеризующего социально-экономические условия для реализации и ведения инновационной деятельности в территориальных субъектах страны, представлен далее. Данное ранжирование обусловлено наличием в модели аппроксимирующей значения ВРП по единственному параметру, указанному выше. В связи с этим можно попробовать графически отобразить зависимость исследуемой величины (валовой региональный продукт) от конкретно установленного предиктора (интегральная оценка социально-экономических условий инновационной деятельности).

Так как значение F расчетной статистики ($(F_{\phi})_p = 16.3967$) больше критического значения $((F_{к})_p = 2.4859)$ при заданном уровне значимости в 5%, то гипотеза об одновременном равенстве нулю всех предикторов в регрессионной модели отвергается, т. е. построенная базовая статистическая модель является статистически значимой (см. табл. 6). Но при более детальном ее рассмотрении и проведении тестов для полученных параметров на t -статистику были идентифицированы смещенные оценки. Из пяти коэффициентов, включая «intercept», значимыми оказались лишь два: «intercept» и параметр β_1 , характеризующий социально-экономические условия инновационной деятельности. Из модифицированной модели были исключены незначимые параметры (β_2 – научно-технический потенциал; β_3 – инновационная деятельность; β_4 – качество инновационной политики), и получена новая статистическая модель, соответствующая условиям F и t статистики. Также удалось увеличить значение скорректированного R^2 с 0.42 до 0.44.

Таблица 6. Результаты регрессионного моделирования

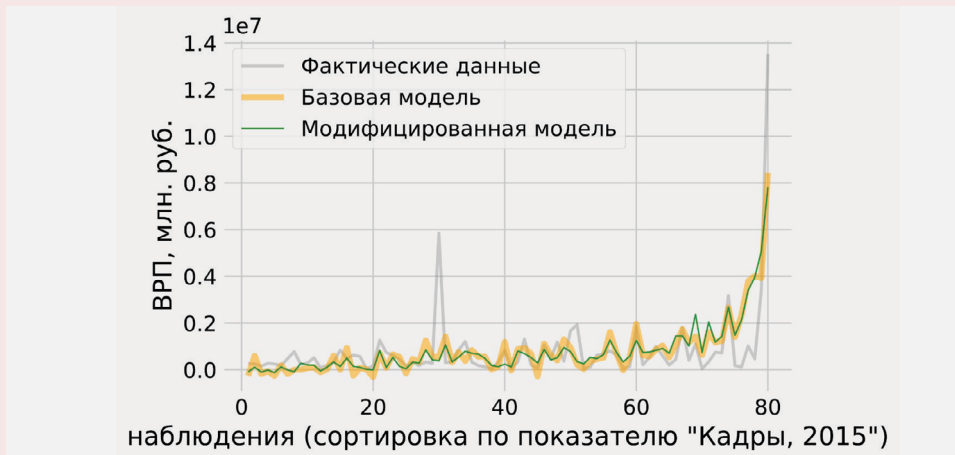
Наименование показателя	Символьное обозначение	Расчетное значение	
		I_1	I_2
1. Базовая регрессионная модель: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4$			
$y_{I_1} = -3964780.87 + 11310999.28x_1 + 1854358.59x_2 - 173667.55x_3 - 50769.69x_4$			
$y_{I_2} = -663095.39 + 564364.88x_1 + 533221.86x_2 - 335361.73x_3 + 201959.81x_4$			
intercept	β_0	-3 964 780.9	-663 095.39
Расчетная t -статистика для β_0	t_{β_0}	6.2250	2.7310
Социально-экономические условия инновационной деятельности / Исследования и разработки	β_1	11 310 999	564 364.88
Расчетная t -статистика для β_1	t_{β_1}	5.8615	3.0278
Научно-технический потенциал / Кадры	β_2	1 854 358.6	533 221.86
Расчетная t -статистика для β_2	t_{β_2}	0.9157	3.1671
Инновационная деятельность / Технологии	β_3	-173 667.55	-335 361.73
Расчетная t -статистика для β_3	t_{β_3}	0.1326	2.1266
Качество инновационной политики / Инновации	β_4	-50 769.69	201 959.81
Расчетная t -статистика для β_4	t_{β_4}	0.0564	1.4406
Критическая точка распределения Стьюдента	t_{κ}	1.9901	1.9921
Коэффициент детерминации	R^2	0.4505	0.5437
Скорректированный коэффициент детерминации	R_{adj}^2	0.4230	0.5194
F -критерий Фишера расчетный	$(F_{\Phi})_p$	16.3967	22.3439
Критическое значение распределения Фишера	$(F_{\kappa})_p$	2.4859	2.4936
2. Модифицированная регрессионная модель: $y'_{I_1} = \beta_0' + \beta_1' x_1$			
$y'_{I_1} = -3813320.76 + 12243723.31x_1$			
Новый intercept	β_0'	-3 813 320.76	
Расчетная t -статистика для β_0'	t_{β_0}'	6.5117	
Социально-экономические условия инновационной деятельности (несмещенная оценка)	β_1'	12 243 723.31	
Расчетная t -статистика для β_1'	t_{β_1}'	8.1531	
Новая критическая точка распределения Стьюдента	t_{κ}'	1.9890	
Новый коэффициент детерминации	$R^{2'}$	0.4447	
Новый скорректированный коэффициент детерминации	$R_{adj}^{2'}$	0.4380	
Новый F -критерий Фишера расчетный	$(F_{\Phi})_p'$	66.4736	
Новое критическое значение распределения Фишера	$(F_{\kappa})_p'$	3.9560	
3. Модифицированная регрессионная модель: $y'_{I_2} = \beta_0' + \beta_1' x_1 + \beta_2' x_2$			
$y'_{I_2} = -393469.1 + 346573.61x_1 + 543161.88x_2$			
Новый intercept	β_0'	-393 469.10	
Расчетная t -статистика для β_0'	t_{β_0}'	2.0592	
Исследования и разработки (несмещенная оценка)	β_1'	346 573.61	
Расчетная t -статистика для β_1'	t_{β_1}'	2.4416	
Кадры (несмещенная оценка)	β_2'	543 161.88	
Расчетная t -статистика для β_2'	t_{β_2}'	3.1554	
Новая критическая точка распределения Стьюдента	t_{κ}'	1.9912	
Новый коэффициент детерминации	$R^{2'}$	0.5090	
Новый скорректированный коэффициент детерминации	$R_{adj}^{2'}$	0.4962	
Новый F -критерий Фишера расчетный	$(F_{\Phi})_p'$	39.9059	
Новое критическое значение распределения Фишера	$(F_{\kappa})_p'$	3.1154	
Составлено по таблице 3.			

Рис. 1. Графическая визуализация регрессионного моделирования по методике ВШЭ



Составлено по таблице 3.

Рис. 2. Графическая визуализация регрессионного моделирования по методике ВолНЦ РАН



Составлено по таблице 3.

Аналогично предыдущему представлены результаты моделирования по методике ВолНЦ РАН (рис. 2). Сортировка результирующих и фактических значений ВРП проведена по показателю «Кадры» в соответствии с наибольшим значением, полученным при расчете его t -статистики (3.1554).

При первичном осмотре можно отметить, что регрессии, построенные по второму исследуемому показателю (I_2), более точно повторяют динамику фактических значений ВРП. Об этом свидетельствует и скорректированный коэффициент детерминации $R_{adj}^2(I_2) = 0.4962$ больший $R_{adj}^2(I_1) = 0.4380$.

Данный момент, возможно, обусловлен снижением размерности включенных во вторую методику показателей, из-за чего не так сильно выражено усреднение исследуемых оценок. В какой-то мере это является преимуществом методики ВолНЦ РАН по отношению к той, которую разработали исследователи ВШЭ.

По модифицированной регрессии y'_{I_2} в сравнении с y'_{I_1} можно также отметить, что она имеет большее количество степеней свободы (на один предиктор). Это дает ей соответствующую гибкость и независимость при прогнозировании результирующей оценки.

Проведенное исследование еще раз показало отсутствие значимого влияния факторов, фигурирующих в российском региональном инновационном индексе и индексе научно-технологического потенциала, которые, в свою очередь, обуславливают научно-технологическое развитие и инновационную деятельность в России, хотя в соответствии с положениями экономической теории они должны оказывать непосредственное воздействие на динамику экономического роста и развития субъекта хозяйствования. Из всего этого может следовать вывод о том, что искусственно выведенные показатели не характеризуют в должной мере научно-технологический потенциал территорий, инновационную деятельность и инновационную политику, проводимую на территории конкретно взятого региона.

Логистическая регрессия

В таблице 7 даны основные оценки по работанной и апробируемой в работе логит-модели, предсказывающей вероятность роста ВРП в соответствии с абсолютным изменением субиндексов рейтинга российского регионального инновационного индекса и оценки научно-технологического потенциала. В базовую логистическую регрессию вошли все четыре имеющихся в рейтинге ключевых показателя,

но из-за несостоятельности модели (незначимость по статистике Фишера и коэффициентов по включенным показателям в модель по t -статистике) ее модификация отсутствует.

В соответствии с полученными значениями F -теста (см. табл. 7) составленные модели являются статистически незначимыми, т. е. совокупное изменение параметров модели, характеризующих изменение социально-экономических условий инновационной деятельности, научно-технический потенциал, инновационную деятельность, качество инновационной политики или инноваций, технологий, кадров и исследований и разработок в регионах, не проявляется в изменении динамики ВРП.

Таким образом, можно сделать следующий вывод: рассмотренные искусственно выведенные оценки недостаточно эффективно отражают процессы и явления научно-технологического развития, которые, в свою очередь, должны прямо воздействовать на значения показателей экономического роста.

Относительно частных коэффициентов полученной модели можно отметить, что их значения в соответствии с t -тестом также в большинстве случаев получились несостоятельными. Единственным исключением является элемент «intercept», который не несет в себе

Таблица 7. Результаты регрессионного моделирования логит-модели

Наименование показателя	Символьное обозначение	Расчетное значение	
		I_1	I_2
Логистическая регрессионная модель: $\hat{P}' = \frac{1}{1 + e^{-(p_0 + p_1 \cdot \Delta x_1 + p_2 \cdot \Delta x_2 + p_3 \cdot \Delta x_3 + p_4 \cdot \Delta x_4)}}$			
$\hat{P}'_{I_1} = \frac{1}{1 + e^{-(0.7466 - 1.4252 \cdot \Delta x_1 + 1.8070 \cdot \Delta x_2 - 1.7907 \cdot \Delta x_3 + 0.3162 \cdot \Delta x_4)}}$			
$\hat{P}'_{I_2} = \frac{1}{1 + e^{-(0.7111 + 0.0775 \cdot \Delta x_1 + 0.1452 \cdot \Delta x_2 - 0.1321 \cdot \Delta x_3 + 0.1373 \cdot \Delta x_4)}}$			
intercept	p_0	0.7466	-0.7111
Расчетная t -статистика для p_0	t_{p_0}	13.4386	11.3365
Изменение социально-экономических условий инновационной деятельности	p_1	-1.4252	0.0775
Расчетная t -статистика для p_1	t_{p_1}	1.0734	0.3926
Изменение научно-технического потенциала	p_2	1.8070	0.1452
Расчетная t -статистика для p_2	t_{p_2}	1.2301	0.4324
Изменение инновационной деятельности	p_3	-1.7907	-0.1321
Расчетная t -статистика для p_3	t_{p_3}	1.8590	2.0185
Изменение качества инновационной политики	p_4	0.3162	0.1373
Расчетная t -статистика для p_4	t_{p_4}	0.4394	1.9119
Критическая точка распределения Стьюдента	t_{κ}	1.9908	1.9921
F -критерий Фишера расчетный	$(F_{\Phi})_p$	1.9378	1.4182
Критическое значение распределения Фишера	$(F_{\kappa})_p$	2.4889	2.4937
Составлено по таблице 3.			

какой-либо интерпретации, а отвечает за оптимизацию расположения графика полученной модели в пространстве. Это в очередной раз свидетельствует об отсутствии взаимосвязи между изменениями ВРП и показателями научно-технологического развития. Следует искать другие параметры для определения вероятности положительного изменения валового регионального продукта.

Также стоит отметить, что один из включенных в базовую модель \hat{P}'_{12} параметров, характеризующий «Технологии» (p_3), имеет t -статистику, приемлемую (2.0185) в сравнении с посчитанным критическим уровнем в размере 1.9921. Но при дальнейшем исключении из модели несущественных предикторов гипотеза о его отличии от нуля не подтвердилась. Значимым в этом случае остался лишь параметр «intercept», который характеризуется наличием сторонних условий, не включенных в модель, следовательно, не может быть адекватно интерпретирован.

Выводы

В заключение, относительно интегральных оценок, характеризующих совокупный уровень научно-технологического развития территории, можно сделать следующий вывод: подобные метрики в должной мере не характеризуют те процессы и явления, для оценки которых были предназначены. Их достоверность сомнительна и обусловлена тем, что зачастую исследуемый объект представляет собой искусственно выведенную категорию, которая учеными и исследователями трактуется по-разному.

В какой-то мере это сравнимо с диагностикой заболевания, на основании которой мы можем сказать, что пациент болен. Но чем

именно? Подобный подход ответа не дает. В результате не могут быть приняты соответствующие меры по устранению негативных эффектов.

Также стоит отметить специфику социально-экономических показателей. В отличие от естественных наук моделирование на подобных оценках не может быть воспроизведено для проверки в лабораторных условиях. В связи с этим достоверность и адекватность подобного рода моделей опирается лишь на теоретические положения, которые могут оказаться как правильными, так и опровергнутыми с течением времени.

В дополнение к этому необходимо отметить, что природа социально-экономических метрик непостоянна и изменчива. То есть один и тот же показатель в разное время может обозначать абсолютно различные вещи. Эти показатели узкопрофильны и характеризуют различные уникальные объекты, чего нельзя сказать о массе материального объекта, скорости движения, наличии определенного набора генов и многом другом, присущем объектам исследования естественных наук.

В значительной мере смещенность получаемых оценок обусловлена тем, что при их составлении используются дублирующие (сильно коррелирующие) показатели. При их значительной доле в итоговом расчете значение интегрального показателя получается усредненным и не отражает дифференциацию в совокупности исследуемых объектов.

Для того чтобы избежать подобного рода проблем, могут применяться рассмотренные в статье статистические методы. При этом они могут быть использованы как отдельно друг от

Рис. 3. Оценка достоверности интегральных оценок социально-экономических объектов

1. Привязка комплексной оценки к результирующему показателю на который будем оказывать воздействие посредством изменения интегральной оценки				
2.1. Однофакторный дисперсионный анализ	2.2. Корреляционный анализ	2.3. Множественная линейная регрессия	2.4. Логистическая регрессия изменения динамики	2.5. Другие методы статистического анализа и обработки данных

Источник: составлено авторами.

друга, так и в совокупности, дополняя и подтверждая соответствующие статистические гипотезы, либо же указывая на необходимость проведения дополнительных исследований (рис. 3).

В соответствии с рисунком 3 вначале должен быть определен устойчивый стабильный показатель, который фиксируется в статистических отчетностях без каких-либо изменений и потенциально будет использоваться в дальнейшем. Затем необходимо установить между ним и интегральной оценкой связь (как показано в работе, это может быть сделано посредством дисперсионного, корреляционного и регрессионного анализа), которая позволит дать оценку достоверности разработанной интегральной методики. Чем больше статистических тестов будет применено к оценке взаимосвязи между результирующей метрикой и интегральной и чем больше совпадений по итогу окажется между ними, тем более устойчивыми и достоверными будут выводы, которые могут быть получены при сопоставлении социально-экономических объектов, оцениваемых по интегральным методикам.

Немаловажными в решении данной проблемы также выступают методы, предназначенные для снижения размерности, количества иссле-

дуемых процессов и явлений. В итоговую модель следует включать существенные факторы, что показано на примере рассмотренных в статье методик.

Также стоит отметить подход, связанный с использованием узкопрофильных эконометрических моделей, учитывающих пространственное расположение исследуемых объектов; временной лаг проявления какого-либо события; нелинейный характер протекания процесса и др. Данные проблемы и вопросы не являются тривиальными и в большей мере применимы к метрикам и показателям, которые могут быть однозначно сняты с интересующего нас исследуемого объекта. В случае оценок на манер интегральных или комплексных подобный подход вряд ли будет иметь успех, так как включает вещи, которые в различном проявлении протекают независимо друг от друга и будут только путать исследователя при попытках математического моделирования.

Материалы публикации могут быть использованы специалистами в области научно-технологического развития, инновационной деятельности и продвижения технологий, а также учеными и исследователями, занимающимися статистической обработкой и анализом данных.

Литература

1. Абалкин Л.И. Логика экономического роста. М.: Ин-т экономики РАН, 2002. 228 с.
2. Проблемы экономического роста территории: монография. Т.В. Ускова [и др.]. Вологда: ИСЭРТ РАН, 2013. 170 с.
3. Shumpeter J.A. *Capitalism, Socialism and Democracy*. London & New York: Routledge, 2003. 460 p.
4. Foster R., Kaplan S. *Creative Destruction: Why Companies That Are Built to Last Underperform the Market – and How to Successfully Transform Them*. Crown Business, 2001. 384 p.
5. Шумпетер Й.А. Теория экономического развития. М.: Прогресс, 1982. 401 с.
6. Грасмик К. Инновационная активность фирм в условиях экономического кризиса // Проблемы теории и практики управления. 2017. № 2. С. 58–64. URL: <https://elibrary.ru/item.asp?id=28289539>
7. Корнаи Я. Инновации и динамизм: взаимосвязь систем и технического прогресса // Вопросы экономики. 2012. № 4. С. 4–31. DOI: 10.32609/0042-8736-2012-4-4-31
8. Solow R. Contribution to the theory of economic growth. *Quarterly Journal of Economics*, 1957, no. 70 (1), pp. 65–94.
9. Rebelo S. Long-run policy analysis and long-run growth. *Journal of Political Economy*, 1991, no. 3, pp. 500–521.
10. Комков Н.И. Научно-технологическое развитие: ограничения и возможности // Проблемы прогнозирования. 2017. № 5. С. 11–21. URL: <https://ecfor.ru/publication/nauchno-tehnologicheskoe-razvitie-ogranicheniya-i-vozmozhnosti/>
11. Черных С., Фролова Н. Об участии российского бизнеса в финансировании научно-технологической сферы (экономические и идеологические аспекты) // Общество и экономика. 2018. № 11. С. 86–97.

12. Росс А. Индустрии будущего. М.: АСТ, 2017. 351 с.
13. Шваб К. Четвертая промышленная революция. М.: Эксмо, 2018. 288 с.
14. Шваб К. Технологии четвертой промышленной революции. М.: Эксмо, 2018. 320 с.
15. Канева М.А., Унтура Г.А. Эволюция теорий и эмпирических моделей взаимосвязи экономического роста, науки и инноваций (часть 1) // Мир экономики и управления. 2017. Т. 17. № 4. С. 5–21. DOI: 10.25205/2542-0429-2017-17-4-5-21
16. Канева М.А., Унтура Г.А. Эволюция теорий и эмпирических моделей взаимосвязи экономического роста, науки и инноваций (часть 2) // Мир экономики и управления. 2018. Т. 18. № 1. С. 5–17. DOI: 10.25205/2542-0429-2018-18-1-5-17
17. Arrow K. Economic welfare and allocation of resources for invention. In: *The Rate and Direction of Inventive Activity*. Princeton Uni. Press, 1962. Pp. 609–625.
18. Romer P.M. Increasing returns and long-run growth. *Journal of Political Economy*, 1986, no. 94 (5), pp. 1002–1037.
19. Lucas R. On the mechanics of economic development. *Journal of Monetary Economics*, 1988, no. 22, pp. 3–42.
20. Grossman G.M., Helpman E. *Innovation and Growth in the Global Economy*. Cambridge: MIT Press, 1991. 384 p.
21. Freeman C. *Technology Policy and Economic Performance: Lessons from Japan*. London: Pinter, 1987. 155 p.
22. Griliches Z. Issues in assessing the contribution of research and development to productivity growth. *The Bell Journal of Economics*, 1979, no. 10, pp. 92–116.
23. Задумкин К.А., Кондаков И.А. Научно-технический потенциал региона: оценка состояния и перспективы развития: монография. Вологда: ИСЭРТ РАН, 2010. 205 с.
24. Научно-технологический потенциал территорий и его сравнительная оценка // К.А. Гулин [и др.]. Проблемы развития территории. 2017. № 1 (87). С. 7–26. URL: <http://pdt.vscs.ac.ru/article/2102>
25. Эпштейн Н.Д., Карманов М.В., Васильева А.В. Проблемы интегральной оценки и обеспечения демографической безопасности // Статистика и экономика. 2015. № 3. С. 233–237. DOI: 10.21686/2500-3925-2015-3-233-237
26. Тейлор Дж. Введение в теорию ошибок: пер. с англ. М.: Мир, 1985. 272 с.
27. Свешников А.А. Основы теории ошибок. Л.: Изд-во ЛГУ, 1972. 122 с.
28. Волкова Е.Н. Проблемы интегральной статистической оценки социально-экономического развития региона // Экономика, статистика и информатика. 2015. № 3. С. 170–175. URL: <https://core.ac.uk/download/pdf/234096888.pdf>
29. Ключникова Е.В., Шитова Е.М. Методические подходы к расчету интегрального показателя, методы ранжирования // ИнноЦентр. 2016. № 1 (10). URL: <http://innoj.tversu.ru/number10.html>
30. Saib M.-S., Caudeville J., Beauchamp M. et al. Building spatial composite indicators to analyze environmental health inequalities on a regional scale. *Environ Health*, 2015. Nj. 68. DOI: 10.1186/s12940-015-0054-3
31. Сидоров А.А. Методологический подход к интегральной оценке состояния и динамики многомерных объектов социально-экономической природы // Проблемы управления. 2016. № 3. С. 32–40. URL: http://pu.mtas.ru/archive/pu_316.php
32. Айвазян С.А. К методологии измерения синтетических категорий качества жизни населения // Экономика и математические методы. 2003. Т. 39. № 2. С. 33–53.
33. Айвазян С.А. Эмпирический анализ синтетических категорий качества жизни населения // Экономика и математические методы. 2003. Т. 39. № 3. С. 19–53.
34. Айвазян С.А., Степанов В.С., Козлова М.И. Измерение синтетических категорий качества жизни населения региона и выявление ключевых направлений совершенствования социально-экономической политики (на примере Самарской области и ее муниципальных образований) // Прикладная эконометрика. 2006. № 2. С. 18–84. URL: <https://ideas.repec.org/a/ris/apltrx/0086.html>
35. Zhou P., Ang B.W., Zhou D.Q. Weighting and aggregation in composite indicator construction: A multiplicative optimization approach. *Social Indicators Research*, 2010, no. 96 (1), pp. 169–181. DOI: 10.1007/s11205-009-9472-3

Сведения об авторах

Дмитрий Александрович Алферьев – кандидат экономических наук, научный сотрудник, Вологодский научный центр Российской академии наук (160014, Российская Федерация, г. Вологда, ул. Горького, д. 56а; e-mail: alferiev_1991@mail.ru); ассистент, Высшая инженерно-экономическая школа, Институт промышленного менеджмента, экономики и торговли, Санкт-Петербургский политехнический университет Петра Великого (194064, Российская Федерация, г. Санкт-Петербург, ул. Политехническая, д. 29)

Александр Евгеньевич Кремин – кандидат экономических наук, Вологодский научный центр Российской академии наук (160014, Российская Федерация, г. Вологда, ул. Горького, д. 56а; e-mail: akremin@inbox.ru); начальник отделения, Вологодский институт права и экономики ФСИН России (160026, Российская Федерация, г. Вологда, ул. Щетинина, д. 2)

Дмитрий Григорьевич Родионов – доктор экономических наук, профессор, директор, Высшая инженерно-экономическая школа, Институт промышленного менеджмента, экономики и торговли, Санкт-Петербургский политехнический университет Петра Великого (194064, Российская Федерация, г. Санкт-Петербург, ул. Политехническая, д. 29; e-mail: drodionov@spbstu.ru)

Дарья Сергеевна Величенкова – кандидат экономических наук, инженер 2 категории, Высшая инженерно-экономическая школа, Институт промышленного менеджмента, экономики и торговли, Санкт-Петербургский политехнический университет Петра Великого (194064, Российская Федерация, г. Санкт-Петербург, ул. Политехническая, д. 29; e-mail: velichenkova_ds@spbstu.ru)

Alfer'ev D.A., Kremin A.E., Rodionov D.G., Velichenkova D.S.

On the Methodology of Checking Integral Estimates of Socio-Economic Objects

Abstract. A reliable and high-quality assessment of scientific, technological, and innovative development of territories helps to define socio-economic conditions and forecast economic growth dynamics of a given subject. The usage of integral indicators is among the most popular approaches toward assessing science, innovative activity, and other socio-economic objects. However, since a collective synthetic category is estimated, accuracy of this metric's characterization of an intangible subject is uncertain. In this regard, issues related to the development of methodology for checking aforementioned provisions are relevant. The purpose of the study is to define the reliability of artificially derived integral complex estimates that in turn describe various socio-economic processes and phenomena. Scientific novelty of the research is to develop an approach to determining the reliability of integral metrics based on mathematical statistical tools. We attempted to determine the quality of artificially derived integrated estimates that, according to their creators, characterize various manifestations of science, innovation, and technology. We applied corresponding methods (variance, correlation, and regression analysis) using the Innovation Development Index of RF constituent entities and assessment of territories' scientific and technological potential. The results obtained are of practical importance in relation to the evaluation practices of the scientific and innovative sphere carried out in the Russian economy. The theoretical significance of the study is characterized by the development of an approach that can be applied to other socio-economic objects. We conclude that integral assessments become an extremely subjective tool when applied to humanitarian areas. They can be used correctly if there is a confirmed connection with the indicators: integral assessments should influence them or have a strong response from them.

Key words: scientific and technological development, integral estimate, analysis of variance, correlation analysis, regression analysis, logit function, economic growth, innovation activity.

Information about the Authors

Dmitrii A. Alfer'ev – Candidate of Sciences (Economics), Researcher, Vologda Research Center, Russian Academy of Sciences (56A, Gorky Street, Vologda, 160014, Russian Federation; e-mail: alferrev_1991@mail.ru); Assistant, Higher School of Engineering and Economics, Institute of Industrial Management, Economics, and Trade, Peter the Great St. Petersburg Polytechnic University (29, Polytechnicheskaya Street, Saint Petersburg, 194064, Russian Federation)

Aleksandr E. Kremin – Candidate of Sciences (Economics), Vologda Research Center of the Russian Academy of Sciences (56A, Gorky Street, Vologda, 160014, Russian Federation; e-mail: akremin@inbox.ru); head of section, Vologda Institute of Law and Economics, Federal Penitentiary Service of Russia (2, Schetinin Street, Vologda, 160026, Russian Federation)

Dmitrii G. Rodionov – Doctor of Sciences (Economics), Professor, director, Higher School of Engineering and Economics, Institute of Industrial Management, Economics, and Trade, Peter the Great St. Petersburg Polytechnic University (29, Polytechnicheskaya Street, Saint Petersburg, 194064, Russian Federation; e-mail: drodionov@spbstu.ru)

Darya S. Velichenkova – Candidate of Sciences (Economics), 2nd Category Engineer, Higher School of Engineering and Economics, Institute of Industrial Management, Economics, and Trade, Peter the Great St. Petersburg Polytechnic University (29, Polytechnicheskaya Street, Saint Petersburg, 194064, Russian Federation; e-mail: velichenkova_ds@spbstu.ru)

Статья поступила 23.04.2021.