

## **В. П. Сиротин**

канд. техн. наук, профессор кафедры статистических методов Государственного университета — Высшей школы экономики (Москва)

## **О. М. Кузьмин**

аспирант кафедры математической статистики и эконометрики Московского государственного университета экономики, статистики и информатики (МЭСИ)

# **МОДЕЛИРОВАНИЕ ИННОВАЦИОННОГО ПОТЕНЦИАЛА РЕГИОНА**

## **Введение**

Для характеристики инновационного потенциала могут быть использованы различные показатели: число организаций, выполняющих исследования и разработки; численность персонала, занятого исследованиями и разработками; объем отгруженных инновационных товаров. Однако наличие определенного числа организаций или квалифицированного персонала не означает, что их действия направлены на развитие инновационного потенциала. Показатель отгруженной инновационной продукции отражает, скорее, не потенциал, а уровень развития сферы исследований и разработок.

Наилучшим образом инновационный потенциал региональной экономики отражает *показатель внутренних затрат на исследования и разработки на душу населения (руб./чел.)*. Моделирование данного показателя способствует выявлению закономерностей развития инновационного потенциала. Действие этих закономерностей может быть различным в зависимости от уровня инновационного развития. Поэтому целесообразно произвести структурное моделирование и построение регрессионных моделей для каждой из однородных групп.

Выделение однородных групп объектов может производиться традиционными методами однозначного определения принадлежности каждого региона к определенной группе. Альтернативой этому служит нечеткое решающее правило, предполагающее выявление меры принадлежности объекта к каждой из выделенных групп.

При расчетах использованы данные Росстата по 83 регионам России за 2003, 2004 и 2005 гг. (Регионы России, 2007).

## **Параметрическое моделирование нечеткой структуры регионов по инновационному потенциалу**

Задача классификации экономических объектов, направленная на выявление структуры исследуемой генеральной совокупности, занимает одно из центральных мест среди задач статистического анализа (Айвазян, Мхитарян, 2001; Дубров, Мхитарян, Трошин, 2003; Мхитарян, Архипова, Дуброва, Сиро-

тин, 2008). От качества решения данной задачи во многом зависит возможность реализации дальнейших этапов статистического исследования. Проведение структурного анализа по одному показателю обуславливает необходимость наиболее полного использования содержащейся в нем информации и возможность применения аппарата расщепления смеси вероятностных распределений, обеспечивающего возможность выделения однородных групп объектов (Сиротин, Архипова, 2007). Данный метод не только позволяет использовать традиционные жесткие решающие правила для интерпретации результатов классификации, но и предоставляет возможность реализовать гибкие подходы, основанные на применении нечетких множеств (Сиротин, 2007).

При расщеплении смеси вероятностных распределений принципиальной задачей является выбор вида закона распределения. Поскольку исследуемый признак формируется под действием совокупности факторов, среди которых ввиду сложности изучаемого процесса отсутствуют явно доминирующие, а характер действия каждого из них на итоговый показатель можно считать мультипликативным, следует сделать вывод, что для однородной с рассматриваемой точки зрения группы регионов закон распределения признака будет логарифмически нормальным:

$$f(\ln y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}},$$

где  $\mu$  и  $\sigma$  — соответственно математическое ожидание и среднее квадратическое отклонение величины логарифма изучаемого признака.

При наличии  $k$  однородных групп в исследуемой совокупности закон распределения будет представлять собой смесь  $k$  логарифмически нормальных распределений:

$$f(\ln y) = \sum_{i=1}^k q_i f(\ln y, \mu_i, \sigma_i),$$

где  $q_i$  — доля объектов  $i$ -й группы в генеральной совокупности,  $\sum_{i=1}^k q_i = 1$ ,  $f(\ln y, \mu_i, \sigma_i)$  — плотность вероятности распределения  $i$ -й группы.

Вид гистограммы логарифма изучаемого признака «Внутренние затраты на исследования и разработки на душу населения» ( $\ln y$ ) в 2003 г. согласуется с предположениями о логарифмически нормальном законе распределения для отдельной группы объектов и о наличии в выборке различных однородных совокупностей (рис. 1). Для 2004 и 2005 гг. были выполнены аналогичные расчеты.

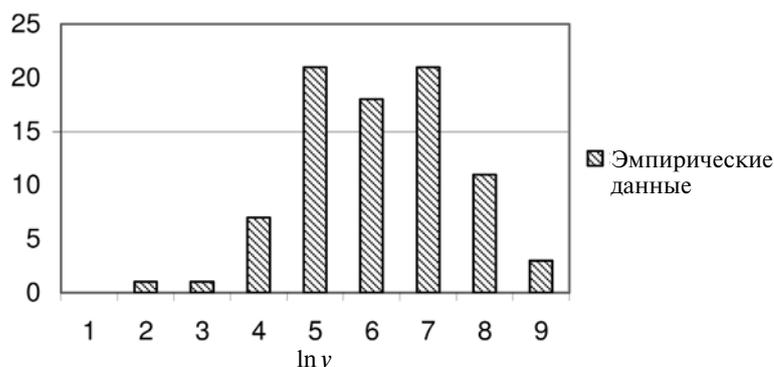


Рис. 1. Гистограмма распределения наблюдаемых значений признака (2003 г.)

Поскольку структура исследуемой совокупности достаточно стабильна, в качестве примера приводятся результаты структурного моделирования за 2003 г.

Наилучшим для аппроксимации и интерпретации результатов оказалось представление совокупности регионов в виде трех страт. В результате проведения численной процедуры оптимизации были получены максимально правдоподобные оценки параметров, приведенные в табл. 1.

Таблица 1

Максимально правдоподобные оценки параметров (2003 и 2005 гг.)

2003 г.			2005 г.				
$i$	$\hat{\mu}_i$	$\hat{\sigma}_i$	$\hat{q}_i$	$i$	$\hat{\mu}_i$	$\hat{\sigma}_i$	$\hat{q}_i$
1	5,087	1,234	0,73	1	5,299	1,160	0,71
2	6,787	0,480	0,24	2	7,142	0,481	0,26
3	8,273	0,261	0,03	3	8,706	0,267	0,03

Графические изображения теоретической и практической гистограмм приведены на рис. 2.

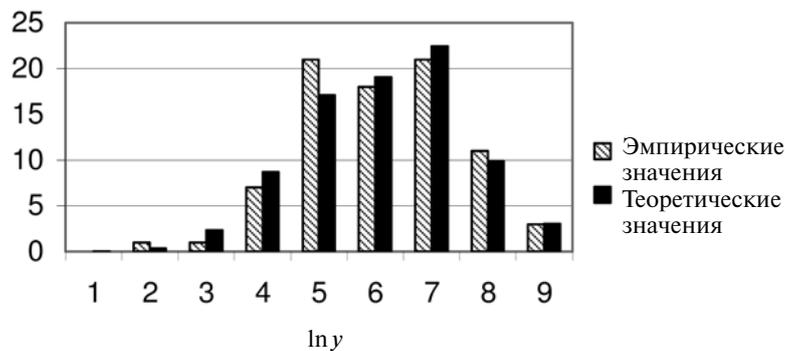


Рис. 2. Гистограммы эмпирического и теоретического распределений (2003 г.)

О достаточно хорошем качестве аппроксимации результатов свидетельствует согласие теоретического и эмпирического распределений. Коэффициент подобия распределений  $K_{\text{под}} = \sum_{\min} (P_i^T, P_i^Э)$  составил 93,2%.

Результаты моделирования плотности вероятности смеси распределений представлены в виде графиков на рис. 3.

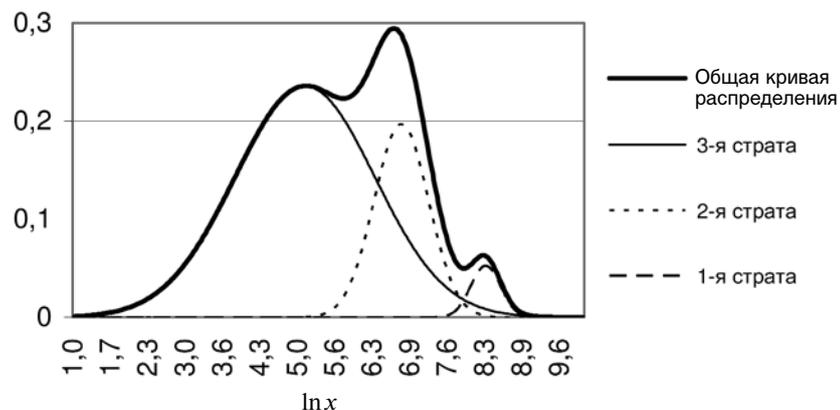


Рис. 3. Распределение регионов по инновационному потенциалу в 2003 г. и его декомпозиция

Для определения границ страт был использован *байесовский критерий минимума среднего риска ошибок классификации*, определяемый при отсутствии дополнительной априорной информации как абсциссы точек пересечения взвешенных кривых распределения соседних страт (Сиротин, Архипова, 2007). В 2003 г. регионы, у которых значение  $\ln u$  меньше 6,38, следует отнести к группе с низким инновационным потенциалом (3-я страта); регионы со значением  $\ln u$  в интервале от 6,38 до 7,86 — к группе с относительно высоким инновационным потенциалом (2-я страта), регионы со значением  $\ln u$  выше 7,86 — к группе передовых регионов (1-я страта).

Для анализа динамики структурных изменений аналогичные расчеты были проведены по данным за 2004 и 2005 гг. В результате был сделан вывод, что структура регионов стабильна и характеризуется высоким уровнем неоднородности.

В *первую группу* постоянно входили 4 *передовых региона* (Москва, Санкт-Петербург, Московская и Нижегородская области). За период с 2003 г. по 2005 г. средний по группе рассматриваемый показатель вырос с 3894 руб. на чел. до 5403 руб. на чел., т. е. на 38,7%, что свидетельствует о значительном росте инновационной активности в этих субъектах Российской Федерации.

Во *вторую группу* входили от 25 до 28 регионов с относительно высоким инновационным потенциалом. Главная причина выделения данных регионов во вторую группу заключается в том, что составляющие ее основу регионы обладают значительным экономическим потенциалом и возможностью осуществлять финансирование сферы исследований и разработок на высоком уровне. В рассматриваемый период в группе постоянными членами были 23 региона, поэтому можно утверждать, что состав группы стабилен. Средний темп роста показателя по группе за период с 2003 г. по 2005 г. составил 32% (он возрос с 1031 руб. на чел. до 1369 руб. на чел.), что несколько ниже результатов лидирующей группы с точки зрения темпов роста и существенно ниже в абсолютном выражении.

В *третью группу* входили от 51 до 54 регионов с низкими значениями рассматриваемого показателя: от 183 руб. на чел. до 226 руб. на чел. в 2003 и 2005 гг., что соответственно в 3 и 3,6 раза соответственно ниже среднероссийского уровня. На данном этапе экономического развития регионы, входящие в данную группу, не обладают существенным инновационным потенциалом и направляют имеющиеся ресурсы на решение других задач региональной экономики.

Результаты классификации за 2003—2005 гг. представлены на рис. 4, при этом как регионы с относительно высоким инновационным потенциалом на схеме отображены 23 субъекта, постоянно присутствовавшие в этой группе на протяжении всего рассматриваемого периода.

Таким образом, достаточно высокий уровень инновационного потенциала имеет около трети регионов России (рис. 5).

Необходимо отметить, что темпы увеличения инновационного потенциала соответствуют его уровню — наиболее высокий темп роста наблюдается в первых двух стратах (38,7 и 32,7% соответственно), что выше среднего значения по стране (29,9%). Темп роста в третьей страте (23,4%), в свою очередь, ниже данного показателя. Данный факт указывает на *наличие дивергенции* в изучаемой совокупности.

Для выявления причин, обуславливающих данное разбиение регионов России на группы по инновационному потенциалу, был проведен анализ его основных факторов:

- наличие квалифицированного персонала;



Рис. 4. Результаты классификации регионов России по инновационному потенциалу в 2003—2005 гг.

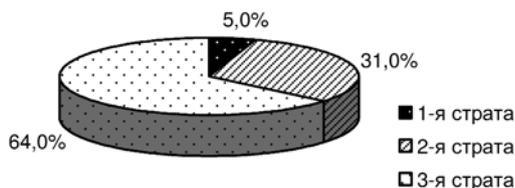


Рис. 5. Структура регионов России по инновационному потенциалу в 2005 г.

- уровень образования в регионе;
- развитие институциональной среды.

Соответственно для их характеристики были выбраны следующие статистические показатели, рассчитанные по данным Росстата (Регионы России, 2007):

- численность персонала, занятого исследованиями и разработками (на 1000 занятых);
- численность студентов высших учебных заведений (на 10 тыс. человек населения);
- число малых предприятий (на 10 тыс. человек населения).

Полученные результаты представлены на рис. 6.

Результаты анализа средних значений факторов по выделенным группам подтверждают, что во всех трех случаях наблюдается доминирование передовых по инвестиционному потенциалу регионов. Несмотря на то что показатели регионов второй группы более скромные, их значения существенно превосходят общероссийский уровень. Регионы со слабым инновационным потенциалом, в свою очередь, имеют наименьшие величины всех трех рассматриваемых факторов.

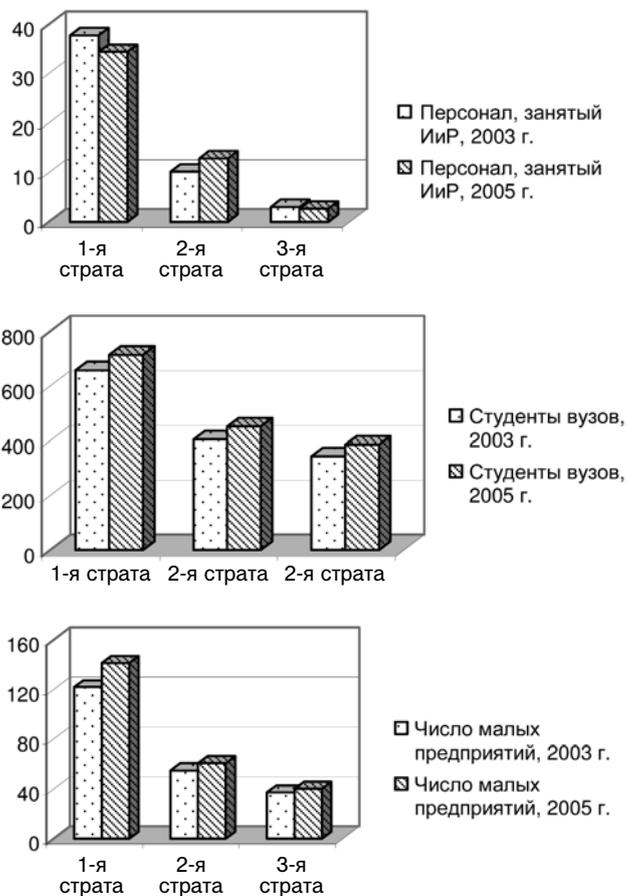


Рис. 6. Средние значения факторов инвестиционного потенциала в 2003 и 2005 гг. в выделенных группах регионов

### Моделирование инновационного потенциала регионов методом типологической регрессии на основе нечеткой классификации

В ходе анализа структуры различных экономических систем традиционные алгоритмы, используемые в стандартных пакетах прикладных статистических программ, как правило, предусматривают однозначное отнесение каждого объекта к определенному классу. Подобное разделение исследуемой совокупности соответствует структуре с жесткими границами. Однако в случаях, когда исследуемые объекты незначительно отличаются друг от друга по величинам, выбранным для структурного моделирования признаков, построение регрессионных моделей в однозначно определенных стратах может не отвечать требованию адекватности отражения реальных экономических явлений. Подобная ситуация наблюдается и при изучении регионального инновационного потенциала: значительное число регионов находится в «пограничном» состоянии (рис. 3), каждый из них может обладать характерными чертами различных страт, поэтому представляется нецелесообразным однозначно относить данные объекты к одной из них. Следовательно, в рассматриваемом случае необходимо реализовать *нечеткий подход к классификации регионов*, который в большей степени соответствует реальной региональной структуре.

Выделение классов с нечетко очерченными границами предполагает определение функции принадлежности каждого объекта к каждой из выделенных групп. *Функцию принадлежности* (membership function)  $i$ -го объекта к  $j$ -й группе можно определить как отношение значения взвешенной плотности вероятности данной страты к общей плотности вероятности для данного объекта:

$$mf_{ij} = \frac{q_j f_j(y_i; \theta_j)}{\sum_{j=1}^k q_j f_j(y_i; \theta_j)},$$

где  $\theta_j = (\mu_j; \sigma_j)$  — вектор параметров плотности вероятности  $f$  для  $j$ -й страты;  $q_j$  — удельный вес  $j$ -й страты в общем законе распределения.

Функция принадлежности показывает, что  $i$ -й объект на  $mf_{ij} \times 100\%$  относится к  $j$ -й группе. Изменение значений функции принадлежности обуславливает наглядное представление о динамике перехода объекта из одной страты в другую. Объекты, принадлежащие к данной страте в максимальной степени, образуют ядро страты. «Пограничные» объекты включаются одновременно в две страты с весами, соответствующими значениям функции принадлежности к ним этих объектов, что обеспечивает адекватность модели описываемому явлению.

Данный подход был использован для моделирования инновационного потенциала регионов *методом типологической регрессии*. На его основе были построены модели множественной регрессии в группах регионов с низким и относительно высоким инновационным потенциалом (3-я и 2-я страты соответственно) по данным 2003, 2004 и 2005 гг. Проведение моделирования в группе лидирующих регионов (1-я страта) является нецелесообразным по причине малого количества наблюдений. Для сравнения результатов моделирования, полученных с использованием нечеткого решающего правила, с результатами традиционной типологической регрессии в этих стратах также были построены регрессионные модели на основе строгой классификации.

В качестве зависимой переменной во всех случаях выступал логарифм внутренних затрат на исследования и разработки на душу населения ( $y^* = \ln y$ ), используемый для характеристики регионального инновационного потенциала, а в качестве регрессоров — рассмотренные выше показатели, определяющие инновационный потенциал региона:

- $x_1$  — численность персонала, занятого исследованиями и разработками (на 1000 занятых);
- $x_2$  — численность студентов высших учебных заведений (на 10 тыс. человек населения);
- $x_3$  — число малых предприятий (на 10 тыс. человек населения).

В *группе регионов с низким инновационным потенциалом* по данным 2003 г. на основе строгой классификации было построено уравнение регрессии:

$$\hat{y}^* = \underset{(8,94)}{3,0231} + \underset{(7,64)}{0,0304}x_1 + \underset{(2,59)}{0,0025}x_2, \quad R^2 = 0,60, \quad F_{\text{расч}} = 37,8.$$

В уравнении под коэффициентами приведены расчетные значения *t*-статистик.

Модели для 2004 и 2005 гг., полученные при однозначном отнесении объекта к группе с низким инновационным потенциалом, имеют такой же состав регрессоров и идентичные качественные характеристики.

Построенное для рассматриваемой группы регионов по данным 2003 г. регрессионное уравнение, в которое все наблюдения включались с *весами, соот-*

ветствующими значениям их функции принадлежности к выделенным группам, имеет следующий вид:

$$\hat{y}^* = 3,3293 + 0,0195x_1 + 0,0024x_2, \quad R^2 = 0,53, \quad F_{\text{расч}} = 44,1.$$

(11,35)                      (7,99)                      (2,90)

Коэффициент детерминации этого уравнения несколько ниже по сравнению с предыдущим. Это объясняется большей «размытостью» исходных данных в модели, полученной на основе нечеткой классификации.

Регрессионная модель инновационного потенциала регионов рассматриваемой группы, полученная при использовании нечеткой классификации, в 2004 г. аналогична приведенной выше, а в 2005 г. в нее дополнительно включается показатель, характеризующий развитие институциональной среды. Коэффициент при переменной становится статистически значимым на уровне  $\alpha = 0,087$ :

$$\hat{y}^* = 3,0219 + 0,0222x_1 + 0,0027x_2 + 0,0089x_3, \quad R^2 = 0,60, \quad F_{\text{расч}} = 39,4.$$

(9,99)                      (7,92)                      (3,72)                      (1,73)

Нечеткий подход позволяет идентифицировать данное изменение в структуре модели, которое свидетельствует о возрастании роли малых предприятий в формировании инновационного потенциала региона. При этом в последней модели несколько увеличивается коэффициент детерминации, что указывает на возрастание ее прогностической силы.

В группе регионов с относительно высоким инновационным потенциалом по результатам классификации, использующей строгое решающее правило, были построены следующие модели для 2003 и 2005 гг.:

$$\hat{y}^* = 6,4661 + 0,0039x_1, \quad R^2 = 0,39, \quad F_{\text{расч}} = 15,1.$$

(53,96)                      (3,89)

$$\hat{y}^* = 6,7078 + 0,0047x_1, \quad R^2 = 0,42, \quad F_{\text{расч}} = 15,7.$$

(50,62)                      (3,96)

В состав статистически значимых регрессоров в данных моделях включается лишь один показатель, что явно не обеспечивает адекватного отражения многоплановых экономических процессов, участвующих в формировании регионального инновационного потенциала. Достаточно слабая объясняющая сила этих моделей также отразилась в невысоких значениях коэффициентов детерминации.

При включении в состав рассматриваемой группы наблюдений в соответствии со значениями функции принадлежности модель для 2003 г. дополняется показателем  $x_3$ :

$$\hat{y}^* = 6,2535 + 0,0037x_1 + 0,0041x_3, \quad R^2 = 0,42, \quad F_{\text{расч}} = 17,2.$$

(47,73)                      (4,71)                      (1,87)

В приведенном уравнении регрессии коэффициент при переменной  $x_1$  значим на 1%-ном уровне, в то время как коэффициент при регрессоре  $x_3$  является статистически значимым на уровне  $\alpha = 0,067$ .

Модель для группы рассматриваемых регионов, получаемая при использовании нечеткого алгоритма классификации, характеризуется стабильной структурой и в 2005 г. имеет следующий вид:

$$\hat{y}^* = 6,4189 + 0,0042x_1 + 0,0056x_3, \quad R^2 = 0,55, \quad F_{\text{расч}} = 30,5.$$

(45,2)                      (6,21)                      (2,42)

В последней модели значительно увеличивается расчетная величина *t*-статистики коэффициента при переменной  $x_3$ , в результате чего он становится статистически значимым на уровне  $\alpha = 0,019$ . В модели также возрастает коэффициент детерминации, что указывает на увеличение ее прогностической силы.

### Заключение

Построение модели, адекватной экономическому явлению, объективно способствует выявлению основных тенденций его развития. Неявно выраженному разделению регионов на группы по инновационному потенциалу соответствует их нечеткая классификация, обеспечивающая возможность реализации нового подхода к типологической регрессии.

Проведенный анализ показывает, что важнейшей задачей в настоящее время является повышение инновационного потенциала в регионах с его низким уровнем, поскольку инновационный потенциал является важнейшим фактором обеспечения конкурентоспособности и развития региональной экономики. Ее решение возможно, в первую очередь, за счет повышения привлекательности инновационной деятельности в данных регионах, а также внедрения новых перспективных технологий производства, разработки новых продуктов и проведения экономической политики, направленной на активацию эндогенных факторов развития.

Необходимо способствовать дальнейшему развитию инновационного потенциала лидирующих наиболее конкурентоспособных регионов, поскольку именно в них возможно осуществлять внедрение передовых технологических достижений. Регионы этой группы в перспективе имеют возможность осуществить выход на глобальные рынки инновационных товаров и послужить опорой для развития инновационной сферы в других регионах, способствуя тем самым экономическому развитию страны в целом.

### Источники

- Айвазян С. А., Мхитарян В. С.* Прикладная статистика. Основы эконометрики. В 2 т. М., 2001.
- Дубров А. М., Мхитарян В. С., Трошин Л. И.* Многомерные статистические методы. М., 2003.
- Мхитарян В. С., Архипова М. Ю., Дуброва Т. А., Сиротин В. П.* Эконометрика: учебник для вузов. М., 2008.
- Регионы России: Социально-экономические показатели. 2005 / Статистический сборник. М., 2007.
- Сиротин В. П.* Стратификация профессорско-преподавательского состава университета по заработной плате // Вестник Самарского государственного университета. 2007. № 1 (27).
- Сиротин В. П., Архипова М. Ю.* Расщепление смеси вероятностных распределений в задачах моделирования социально-экономических процессов. М., 2007.