

О ПОДХОДАХ К РАБОТЕ СО СТАТИСТИКОЙ ПРИ МЕЖРЕГИОНАЛЬНЫХ СРАВНЕНИЯХ И ОЦЕНКЕ НЕОДНОРОДНОСТИ

Экономическое пространство России в значительной степени неоднородно. Сейчас становится очевидным, что различия между регионами не переходят в здоровую специализацию, а разрыв в уровнях развития экономики и жизни населения между лидирующими и отстающими субъектами продолжает нарастать – налицо системный характер проблемы. Для нахождения правильного ее решения, в первую очередь, должна быть проанализирована современная ситуация. То есть, проведена полная оценка неравномерности развития территории страны.

Актуальность этой проблемы увеличила интерес научного сообщества к публикациям на тему оценки неоднородности территорий и межрегиональных сравнений. Исследователями в них обычно анализируется либо единственный показатель в его вариации по субъектам, либо (что дает более адекватные результаты) множество показателей, агрегируемых в один синтетический.

В итоге мы имеем набор индексов, сопоставленных территориям так, что каждой территории соответствует один индекс и все они соизмеримы (как величины, полученные из соизмеримых данных посредством единого алгоритма преобразования). Наиболее интуитивным будет рассмотрение этих индексов как совокупности чисел, что можно представить, предположив, что все они являются значениями некоторой случайной величины. Тогда мы имеем ее случайную выборку и можем пользоваться методами классической статистики.

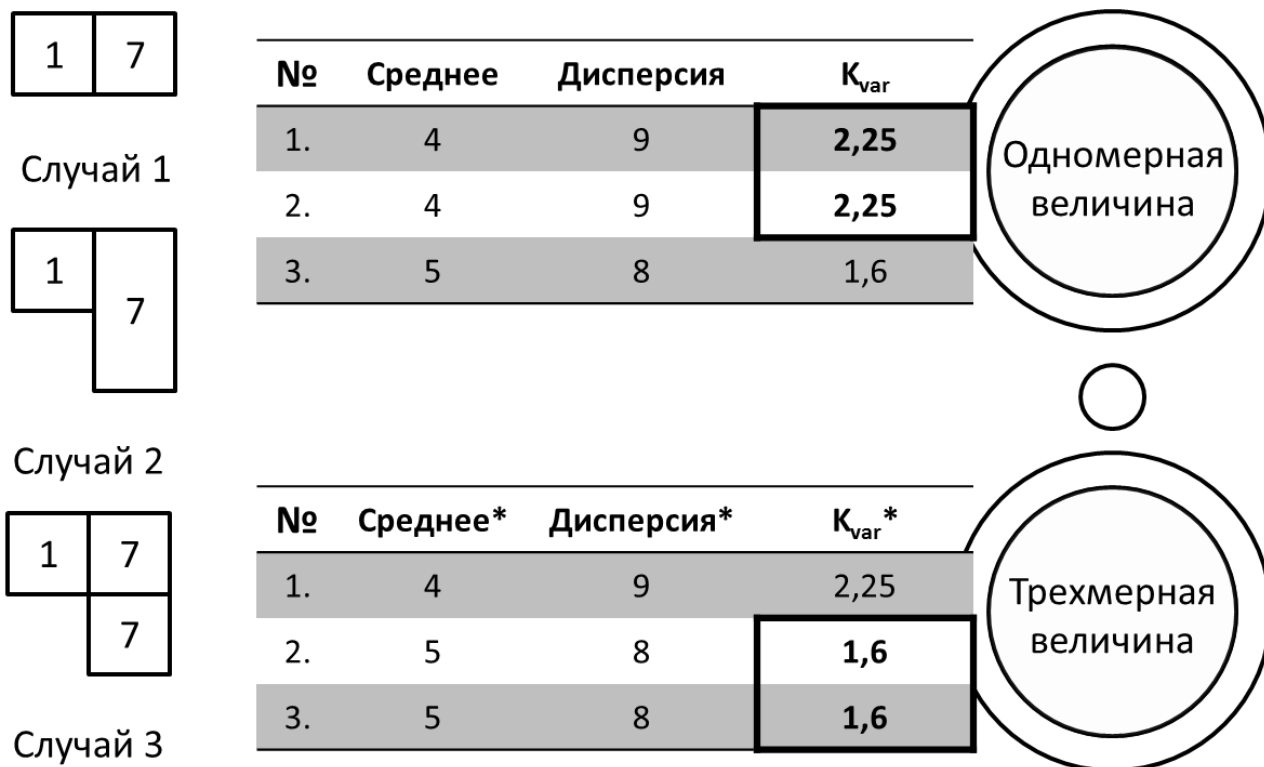
В частности, возможно посчитать оценки среднего и дисперсии. Однако, так как кроме данной выборки у нас нет иных данных, то оценки моментов имеет смысл вычислять только для определенных совокупностей индексов. При этом сами индексы для нас эквивалентны, следовательно, составление совокупностей может быть произведено только на основе дополнительной

информации. Так как индексы относятся к территориям, то логично группировать их по территориальному признаку. Проблемой является то, что возможных группировок может быть довольно большое число (например, при размере группы равном 10 регионам, для 83 субъектов РФ их будет, примерно, $9 \cdot 10^{61}$). Даже если группировка будет проводиться на основе некоторых объективных факторов (климат, рельеф, население и т.д.), то возникнет проблема включения границ, когда естественное разделение двух групп проходит по некоторой территории и не понятно, в какую из них ее следует включить. Наиболее непротиворечивым структурированием будет группировка в соответствие с административным делением. Однако, очевидно, что административное деление сложилось исторически, редко меняется и поэтому не соответствует локализации реальных явлений (например, при оценке неоднородности территорий по федеральным округам, она не будет постоянной для всего округа, но будет возрастать и убывать в его отдельных частях).

Дополнительной проблемой являются различные физические размеры округов (например, Северо-Кавказский ФО меньше Дальневосточного в 36 раз). При расчете дисперсии учитывается лишь размер выборки – число субъектов в округе. Поэтому, к примеру, имеет смысл переносить отношение таких дисперсий между ЮФО и ЦФО на отношение неоднородностей между этими округами – при сопоставимых размерах большая сумма отклонений ЦФО скорректируется на большее число субъектов. Однако можно усомниться в правомерности аналогичного переноса между СКФО и ДФО.

Неоднородность – по сути, пространственная характеристика. Адекватно в одномерном виде она может быть представлена разве только в форме своего среднего. Но это среднее должно быть обратно пропорционально площади изучаемой территории. Поэтому, если мы все-таки хотим получить численные оценки моментов для имеющихся данных (а также основанных на них мер, типа неоднородности), то *необходимо будет постулировать их независимость от разбиения всего пространства на зоны оценки*. Использование представления

одномерной величины такой независимостью не обладает (см. Рисунок 1). Действительно, для случаев 1 и 2 мы получаем одинаковые результаты, хотя очевидно, что среднее всей территории в случае 2 больше. Тогда как для случая 2 и случая 3 – идентичных по оценкам, но отличающихся разбивкой территорий – результаты различны.



* - при условии, что площадь квадратных территорий = 1ед.

Рисунок 1 - Расчет меры неоднородности при различном рассмотрении исходных данных

Для того чтобы требование независимости выполнялось, будет необходимо при расчете моментов использовать поправки на площади территорий. А это приводит нас к рассмотрению исходных данных в виде трехмерной величины (каковой она изначально и являлась). В таком представлении, случай 2 и случай 3 будут оценены правильно: их неоднородность одинакова, но меньше, чем в первом случае (Стоит заметить, что первое представление тоже может давать правильный результат, но только в случае равенства площадей всех территорий).

Однако, устранив ошибки в оценке, возникающие из-за различной площади территорий, мы остаемся с проблемой локализации неоднородности. Она может быть решена посредством расчета пространственной дисперсии. Представив наши (по сути, точечные) данные как непрерывные для своих территорий, мы можем нанести их на карту, и для каждой точки карты рассчитать дисперсию ее окрестности определенного радиуса, которую снова сопоставить исходной точке. Пример подобной локализации неоднородности представлен ниже (Рисунок 2). Для расчетов использовались ArcGIS и Map Viewer, разрешение растеризации - 87 тыс. точек. Информация, представленная в таком виде, позволяет не только сравнить неоднородность групп территорий (здесь – федеральных округов), но и показать конкретные места и силу ее проявления.

Такой подход, располагая значительными плюсами, также имеет (возможно, устранимые) недостатки. Во-первых, существует проблема объективного определения радиуса расчета. При чрезмерном радиусе мы потеряем то, что искали, то есть локальность представления. При слишком малом радиусе, неоднородность сведется к сравнению смежных территорий и будет, по существу, отражать скачки, происходящие на их границах. Возможно, стоило бы варьировать радиус в зависимости от плотности информации (количества попадающих в окрестность субъектов), увеличивая его для восточной части страны и уменьшая – для западной.

Во-вторых, остается открытым вопрос о внутренней территории субъекта. В соответствие с имеющимися данными для нас она представляется однородной (более точное рассмотрение возможно только при увеличении глубины статистического охвата). Однако, очевидно, что это лишь приближение. Тогда как наиболее адекватно это можно отобразить?

Ответы на эти вопросы помогут получить более адекватные оценки существующих проблем и, следовательно, более эффективные их решения.

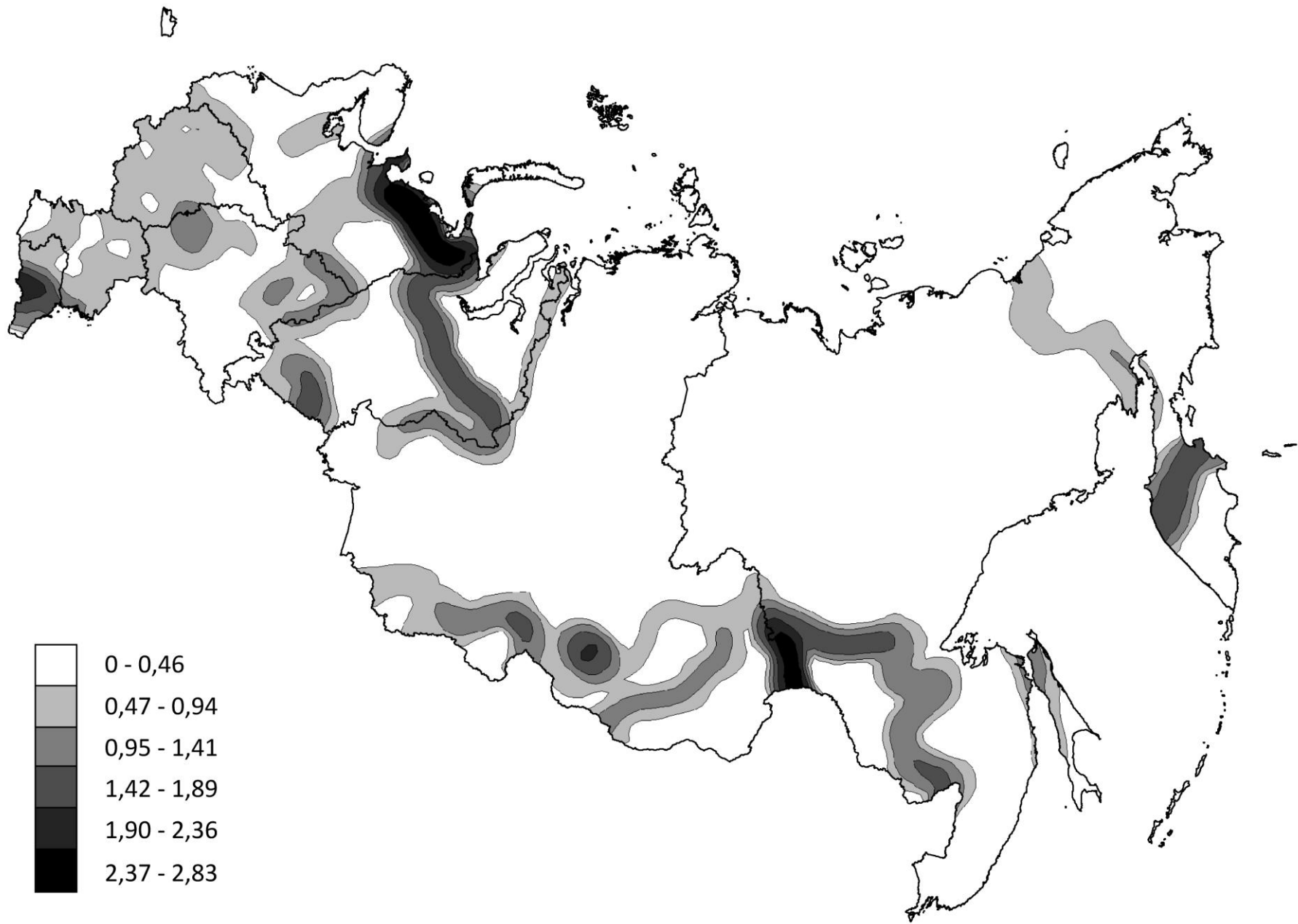


Рисунок 2 – Зоны неоднородности на карте России (пространственный коэффициент вариации на основе данных по развитию здравоохранения в 2007 г.)