

**XVI Всероссийский конкурс научных работ молодежи
«Экономический рост России»**

**Изучение пороговых совокупностей методом
декомпозиции смесей вероятностных распределений**

**Research of the threshold aggregates by decomposition of
mixtures distributions**

Автор: Зайков К.А.

студент Новосибирского государственного университета экономики и управления «НИНХ», Институт Экономики, курс 4, Бакалавр статистики.

Author: Zaykov K.A.

student of Novosibirsk State University of Economics and Management, Institute of Economics, 4 year, Bachelor of statistics.

Наименование ВУЗа: ФГБОУ ВПО «Новосибирский государственный университет экономики и управления «НИНХ».

University's name: FGBEI HPE "The Novosibirsk State University of Economics and Management.

Научный руководитель: Серга Людмила Константиновна, к.э.н., доцент, доцент кафедры статистика

Scientific adviser: Serga Lyudmila Konstantinovna, Cand. of economic science, Associate professor of department statistics.

Оглавление

Введение.....	3
1. Пороговые совокупности как объект статистического исследования	6
2. Методические аспекты исследования пороговых совокупностей.....	9
3. Применение методов разделения смеси распределений в изучении пороговых совокупностей	15
Заключение	20
Библиографический список	21

АННОТАЦИЯ

Экономический рост России зависит от множества факторов, одним из которых является эффективность управления. Принятие оптимальной стратегии органами власти и бизнесом во многом детерминировано точностью и достоверностью статистической информации. Качество информации, в свою очередь, определяется четкостью границ объекта исследования – статистической совокупности. В работе предлагается методика ограничения таких сложных социально-экономических явлений, как пороговые совокупности, для целей обследования и управления ими.

ANNOTATION

Russia's economic growth depends on many factors, one of which is the efficiency of the management. The optimal strategy-making of the state and business is determined by accuracy and reliability of statistical information. The quality of information, by-turn, is determined by the definition of the boundaries of the research object - statistical aggregates. Research formulates a method of limitation of complex socio-economic phenomena, as the threshold aggregates, for the purposes of the survey and management.

Ключевые слова: теория совокупностей, пороговая совокупность, порог, смесь распределений, декомпозиция (разделение) смесей.

Key-words: the theory of sets, threshold aggregates, threshold, mixtures distributions., decomposition of mixtures distributions.

Введение

Быстрый и скачкообразный темп развития современного общества порождает политические, экономические и социальные изменения и сдвиги. В этой связи, органы исполнительной и законодательной власти должны своевременно получать статистическую информацию о демографических, социальных и экономических явлениях. От полноты и достоверности полученной информации будет зависеть анализ эффективности деятельности органов управления и создания стратегии дальнейшего развития, что во многом детерминирует экономический рост России

Постоянное движение и вызванная им нестабильность, особо остро ставит вопрос о необходимости совершенствования основ теории статистического исследования, прежде всего, теории совокупностей. Весомый вклад в становление теории совокупностей внесли исследования Б.Г. Плошко, И.И. Елисейевой, Л.В. Некраша.

Ряд ученых подчеркивают необходимость в исследовании дополнительных классификационных разрезов реальных совокупностей, таких как динамический и пороговый. Центральное место в становлении и развитии новых подходов к изучению совокупности занимают работы экономистов-статистиков Новосибирского университета экономики и управления «НИНХ» Глинского В.В. и Серга Л.К.

Одним из подходов к изучению пороговых совокупностей является использование широко известных методов разделения смесей вероятностных распределений: метод максимального правдоподобия; группа методов, основанных на использовании мер близости; группирующие функции; методы, основанные на дискриминантном анализе; EM-алгоритм; метод моментов; k-means и др.

Целью исследования является практическое применение методов разделения смесей вероятностных распределений для изучения пороговых совокупностей.

Объектом исследования выступила совокупность работников Российской Федерации.

Предмет исследования – распределение работников Российской Федерации по размеру начисленной заработной платы.

В ходе исследования были решены следующие задачи:

1. Выполнен анализ основ теории пороговых совокупностей и различных методов разделения конечных смесей вероятностных распределений;
2. Изучены существующие методики разделения смесей распределений;
3. Проведено исследование распределения численности работников РФ по размеру начисленной заработной платы как смеси распределения.

Научная новизна работы заключается в применении методов разделения смесей распределения – основ теории распознавания образов, используемой в вычислительной технике, информатике и физике, для экономического исследования.

В качестве информационной базы использовалась официальная статистическая информация Федеральной службы государственной статистики: контент официального веб-сайта Росстата (<http://www.gks.ru/>) и статистический сборник «Труд и занятость в России» за 2011 год.

Для вычислений использовались пакеты программ: Microsoft Office Excel, StatSoft STATISTICA, SPECSTAT*, Wolfram Mathematica.

* Пакет программ для статистического анализа, разработанный профессором кафедры статистики Новосибирского государственного университета экономики и управления «НИНХ» Иониным В.Г.

1. Пороговые совокупности как объект статистического исследования

Статистическая наука изучает количественную сторону массовых общественных явлений в неразрывной связи с их качественной стороной, исследует количественное выражение закономерностей общественного развития в конкретных условиях места и времени.

Для изучения статистика использует свои категории, к которым относится и статистическая совокупность. Статистическая совокупность – это множество единиц изучаемого явления, объединенных единой качественной основой, общей связью, но отличающихся друг от друга отдельными признаками. [11, с. 11]

В теории совокупностей используются следующие классификации: по охвату единиц (генеральная, выборочная, основной массив, моно-совокупность); по степени однородности (общие, частные); по характеристике качества (сущности) единиц (население, учащиеся, микропредприятия и проч.) (рис. 1.). Однако данные классификации не могут в полной степени отвечать современным требованиям и решать весь спектр поставленных перед статистикой задач. В этой связи, учеными-статистиками В.В.Глинским и Л.К.Серга была предложены дополнительные классификационные разрезы для совокупностей: динамический и пороговый. [3, 4, 6]



Рис. 1. Классификация статистических совокупностей.

Пороговые совокупности – это такие реальные множества, отнесение единиц к которым происходит на базе статистических критериев, вводимых искусственным путем. Критерий выбирается по различным причинам: статистический, для решения задач анализа развития объектов; управленческий, для управления развитием явления и др.

Кроме того, благодаря установлению пороговых значений приводят к разбиению изначальной качественно неоднородной совокупности на однородные частные совокупности, благодаря чему, решаются задачи типологии данных.

В качестве примера пороговых совокупностей можно привести совокупность наемных работников по размеру начисленной заработной платы. Как известно, уровень заработной платы зависит от установленного в законодательном порядке минимального размера оплаты труда. Минимальный размер оплаты труда – гарантируемый государством размер месячной заработной платы за труд неквалифицированного работника, полностью

отработавшего норму рабочего времени при выполнении простых работ в нормальных условиях труда; минимальная заработная плата. [8,ст. 129]

Минимальный размер оплаты труда и будет своеобразным «порогом» для всей совокупности, введенным в целях искусственного регулирования государством оплаты труда населения. То есть, изменяя размер МРОТ, руководство страны может задать вектор развития общества, что уменьшает степень неопределенности для стратегического планирования в условиях современных темпов развития.

Созданные в результате деления генеральной совокупности множества, типы подчиняются собственным законам развития, присущим только однородным единицам данной частной совокупности. Однако на практике, статистический учет единиц генеральной совокупности происходит одинаково, без учета особенностей каждой частной пороговой совокупности, что приводит к уменьшению качества, достоверности и точности статистической информации.

Учитывая сказанное выше, и то, что установление границ-порогов совокупности носит субъективный характер (не всегда имеет место содержательное обоснование его величины), перед статистикой стоит задача определения объективных правил и критериев нахождения пороговых значений, чтобы они четко обозначали переходы одного качественного состояния явления в другой. Поставленную задачу, по нашему мнению, можно решить методами разбиения реальной совокупности на частные подсовкупности (типы), рассмотрев исходную как конечную смесь вероятностных распределений. В результате деления смеси на составляющие ее компоненты получаем объективный порог. Пороговое значение будет соответствовать величине явления на пересечении кривых вероятностных распределений.

2. Методические аспекты исследования пороговых совокупностей

В процессе ограничения пороговой совокупности широко применяются статистические методы типологии данных. Одним из таких методов классификации выступает разделение смеси вероятностных распределений.

Смеси вероятностных распределений играют важную роль в статистике и теории вероятностей, прежде всего, в области применения этих наук. Используя нормальное приближение, исследователь недооценивает большие значения результирующего распределения, считая, что их вероятности малы. Однако, на практике, как правило, результирующее распределение имеет более тяжелые «хвосты» и острую вершину, по сравнению с нормальным распределением. Поэтому, в сферах, где важна высокая точность, например, теории надежности и страховании, используют модели конечных смесей распределения.

Возможным объяснением отклонений распределения эмпирических данных от нормального может служить то, что на разные эксперименты влияет разное число случайных факторов. В этом случае центральная предельная теорема не применима, и необходимо рассматривать суммы числа случайных величин, предельным распределением которой выступают смеси вероятностных распределений.

Конечной смесью распределения называют распределение вероятностей, представляющее собой линейную функцию некоторого числа компонент распределений вероятностей. Такие распределения используются для моделирования совокупностей, которые предположительно содержат отдельные группы наблюдений. Первоначальным примером применения такого распределения является распределение, смешанное из двух нормальных и применённое Пирсоном в 1894 году:

$$f(x) = pN(x_1, \sigma_1) + (1 - p)N(x_2, \sigma_2) \quad (1)$$

где p - доля первой группы в совокупности;

\bar{x}_1, σ_1 – соответственно средняя и среднеквадратическое отклонение переменной в первой группе;

\bar{x}_2, σ_2 – соответствующие значения во второй группе.

Плотность распределения первого компонента будет иметь вид:

$$f_1(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\bar{x}_1)^2}{2\sigma_1^2}} \quad (2)$$

Плотность распределения второго компонента:

$$f_2(x) = \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x-\bar{x}_2)^2}{2\sigma_2^2}} \quad (3)$$

Тогда, плотность смеси распределения, предложенного Пирсоном имеет вид:

$$f(x) = p \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{(x-\bar{x}_1)^2}{2\sigma_1^2}} + \frac{(1-p)1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(x-\bar{x}_2)^2}{2\sigma_2^2}} \quad (4)$$

Решить задачу распределения смеси значит по имеющейся выборке классифицируемых наблюдений, извлеченной из генеральной совокупности, являющейся смесью частных одномодальных совокупностей известного параметрического вида, построить статистические оценки для числа компонентов смеси, их удельных весов и параметров, их определяющих.

В теоретическом варианте задача расщепления смеси заключается в восстановлении компонентов смеси и смешивающей функции (удельных весов) по заданному распределению всей (т.е. смешанной) генеральной совокупности и называется задачей идентификации компонентов смеси.

Существует множество методик разделения смесей распределения. Каждый метод имеет свои достоинства и недостатки, связанные со сложностью реализации вычислений, а так же точностью полученных результатов.

Статистический анализ смесей распределений проводится обычно в рамках одной из двух логических схем. В первой из них реализуется логика «от оценивания параметров смеси к классификации» (EM-алгоритмы, основанные на методе максимального правдоподобия, методе моментов и т.д.). Во второй, напротив, идут «от классификации к оцениванию», затем, имея оценки параметров распределений внутри классов, уточняют классификацию и т.д.

Метод моментов

Первым методом решения одномерной задачи разделения смесей был метод моментов, предложенный Пирсоном в 1897 году. Этот метод заключался в том, что к моментам смеси (выборочным или заданным) приравниваются теоретические моменты смеси, выраженные через теоретические моменты составляющих совокупностей и их веса.

$$\mu_k = M_k - C_k^1 M_{k-1} + C_k^2 M_{k-2} (M_1)^2 - C_k^3 M_{k-3} (M_1)^3 + \dots + (M_1)^k \quad (5)$$

Данный метод сравнительно прост в теории и популярен, однако «узкими местами» подхода являются его вычислительная реализация (особенно в случае высоких размерностей анализируемых распределений и большого числа смешиваемых классов) и относительно невысокое качество статистических свойств получаемых при этом оценок.

Использование мер близости

Применение любого метода группировки объектов требует введения критерия, в соответствии с которым разделение на подсовокупности будет оптимальным. Чаще всего таким критерием считают условие максимума или

минимума какой-либо функции от задаваемой тем или иным образом взаимной «близости» между элементами смеси.

Существует три основных способа количественного выражения близости (сходства) элементов — с помощью коэффициентов подобия, коэффициентов корреляции и различных метрических показателей расстояния (в том числе с использованием потенциальных функций). Необходимо отметить, что коэффициенты подобия используются в основном при решении задач иерархической классификации (построения так называемых дендрограмм), а коэффициенты корреляции — в задачах факторного анализа.

Выбор меры близости между элементами в значительной степени произволен, однако он некоторым образом связан с выбором критерия разделения. Последний задает условие, по которому будет оцениваться группирование элементов заданной совокупности. Правильное разбиение будет считаться выполненным, если достигнуто экстремальное значение критерия. Предполагается, что существует единственное разбиение, которое доставляет экстремум выбранному критерию, и результирующее разбиение должно совпадать с объективно существующим если границы последнего четко определены. При этом критерий должен быть таким, чтобы было возможно построение эффективного алгоритма для нахождения оптимального разбиения.

Метод максимального правдоподобия. EM-алгоритм

Группа методов максимального правдоподобия основана на принципе выполнения различных итерационных схем, приводящих к максимуму функцию правдоподобия.

Алгоритм вычисляет по некоторым значениям параметров новые, более правдоподобные значения. Главной положительной чертой итерационных схем является то, что независимо от начальных значений параметров, алгоритм обеспечивает сходимость к оценкам максимального правдоподобия, что было

неоднократно подтверждено в экспериментальных проверках для случая большого числа неизвестных параметров.

Основными «узкими местами» подхода, основанного на методе максимального правдоподобия статистического оценивания параметров смеси распределений, являются (помимо необходимости «угадать» общий параметрический вид распределения, задающего каждый из классов) требование ограниченности анализируемой функции правдоподобия, высокая сложность и трудоемкость процесса вычислительной реализации соответствующих процедур и медленная сходимость порождаемых ими итерационных алгоритмов.

В основе идеи EM-алгоритма лежит метод получения последовательностей оценок параметра, которые при постоянных условиях стремятся к оценке максимального правдоподобия.

Основное предположение EM-алгоритма заключается в том, что исследуемое множество данных может быть смоделировано с помощью линейной комбинации многомерных распределений, а целью является оценка параметров распределения, которые максимизируют логарифмическую функцию правдоподобия:

$$Q(\theta, \theta^m) = \int_{\mathbb{R}^m} [\log[f_{\theta}(x, y)]] f_{\theta^m}(x|y) \mu_Y(dy) \rightarrow \max \quad (6)$$

Иными словами, предполагается, что данные в каждом кластере подчиняются определенному закону распределения. С учетом этого предположения можно определить параметры - математическое ожидание и дисперсию, которые соответствуют закону распределения элементов в кластере, наилучшим образом «подходящему» к наблюдаемым данным.

Каждая итерация алгоритма состоит из двух шагов, известных как E или шаг ожидания (от Expectation step), и M- шаг максимизации (от Maximization). На первом шаге ожидаемое значение логарифмической функции правдоподобия основано на наблюдаемых данных, и находятся текущие оценки параметров. На

М-шаге эта функция максимизируется для получения улучшенных оценок параметров, которые увеличивают правдоподобие. Шаги чередуются до достижения сходимости. Данный алгоритм может в некоторых случаях сходиться очень медленно.

В нашем исследовании был применен EM-алгоритм, используемый в математической статистике для нахождения оценок максимального правдоподобия параметров вероятностных моделей, в случае, когда модель зависит от некоторых скрытых переменных.

3. Применение методов разделения смеси распределений в изучении пороговых совокупностей

Применение методов разделения смеси распределений осуществлялось на примере совокупности наемных работников, попавших в выборочное обследование, проводимое Федеральной службой государственной статистики (Росстат) в апреле каждого года, по размеру оплаты труда.

Совокупность наемных работников по размеру оплаты труда, как уже говорилось ранее, является пороговой совокупностью, так как размер начисленной заработной платы зависит от установленного Федеральным Собранием РФ минимального размера оплаты труда. Показатели МРОТ по РФ за 2000-2011 годы приведены в таблице 1.

Таблица 1

Размер МРОТ РФ в 2000-2013 гг.*

Дата введения	Сумма (рублей в месяц)
01.07.2000	132
01.01.2001	200
01.07.2001	300
01.05.2002	450
01.10.2003	600
01.01.2005	720
01.09.2005	800
01.05.2006	1100
01.09.2007	2300
01.01.2009	4330
01.06.2011	4611
01.01.2013	5205

*Источник: Федеральный закон №82-ФЗ «О минимальном размере оплаты труда»

Для исследования были использованы официальные данные Росстата о распределении численности работников по размерам начисленной заработной платы за 2000-2011 гг. Все данные представлены в таблице 2.

Таблица 2

Распределение численности работников по размерам начисленной заработной платы за 2000-2011 гг., в%

Заработная плата, руб.	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
до 1000,0	34,1	28,4	14,7	11,5	6,9	4,0	2,9	1,0	-	-	-	-
1000,1-1800,0	23,7	21,3	16,8	14,3	10,7	7,8	5,2	2,9	1,3	-	-	-
1800,1-2600,0	15,1	15,3	15,5	13,5	11,1	8,8	6,8	4,4	2,4	-	-	-
2600,1-3400,0	9,3	10,2	12,8	11,7	10,7	9,1	7,2	5,1	2,9	-	-	-
3400,1-4200,0	5,5	6,6	9,2	9,2	9,1	8,4	6,9	5,7	3,6	2,6	1,8	1,5
4200,1-5000,0	3,5	4,6	7,0	7,6	8,1	7,8	6,8	5,7	4,2	6,9	4,8	4,7
5000,1-5800,0	2,3	3,2	5,2	6,0	7,0	7,0	6,3	5,5	4,4	4,4	4,2	3,3
5800,1-7400,0	2,6	3,8	6,6	8,3	10,3	11,2	11,6	10,9	8,8	8,7	7,8	6,9
7400,1-9000,0	1,5	2,2	3,9	5,3	7,3	8,8	9,4	9,4	8,3	8,2	7,5	6,6
9000,1-10600,0	0,8	1,3	2,4	3,5	4,9	6,5	7,5	8,2	8,0	7,8	7,1	6,6
10600,1-13800,0	0,8	1,3	2,5	3,8	5,7	8,0	10,2	12,8	14,1	13,9	13,3	12,4
13800,1-17000,0	0,4	0,7	1,3	2,0	3,0	4,5	6,4	8,4	10,3	11,2	11,2	11,1
17000,1-20200,0	0,4	0,7	0,7	1,1	1,7	2,6	4,0	5,7	7,8	8,5	9,1	9,3
20200,1-25000,0	-		0,6	0,9	1,4	2,1	3,4	5,2	7,8	8,8	9,9	10,5
25000,1-35000,0	-	0,3	0,6	1,1	1,2	1,9	3,0	4,9	8,1	9,6	11,3	12,6
35000,1-50000,0	-				0,6	0,8	1,4	2,4	4,5	5,3	6,7	7,9
50000,1-75000,0	-	0,1	0,2	0,1	0,2	0,4	0,6	1,1	2,2	2,6	3,3	4,1
свыше 75000,0	-	-	-	0,1	0,1	0,2	0,4	0,7	1,3	1,5	2,0	2,5

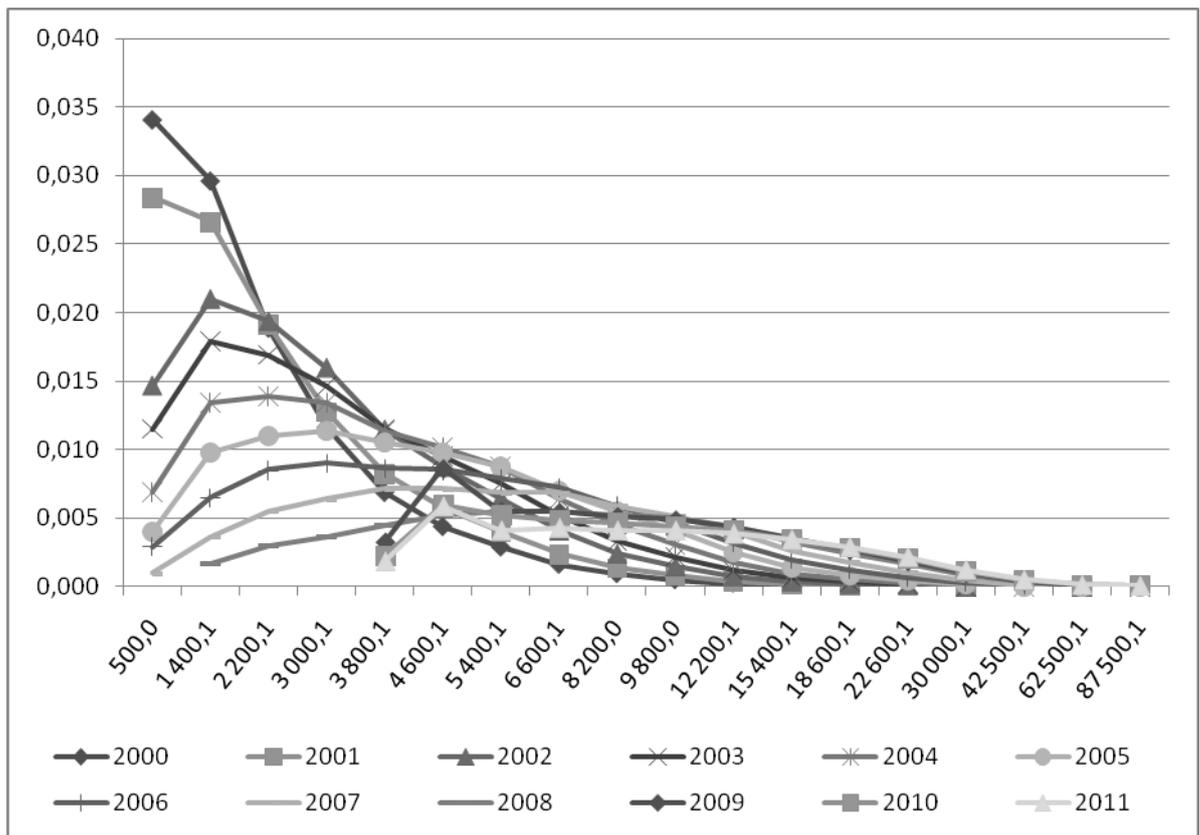


Рис.2. Эмпирические плотности распределения работников РФ по начисленной заработной плате в 2000-2011 гг.

В результате вычислений было выявлено, вся совокупность-смесь подчиняется логнормальному распределению, что соответствует общепринятым суждениям о законе распределения заработной плате населения. Вероятность принятия нулевой гипотезы о логнормальном законе распределения в среднем равна 84%.

Однако, при последующем разбиении смеси на составляющие совокупности видно, что оптимальным является декомпозиция смеси на два компонента с нормальным и логнормальным законами распределения.

Данная ситуация наиболее характерна для 2002, 2004, 2009, 2010 и 2011 годов.

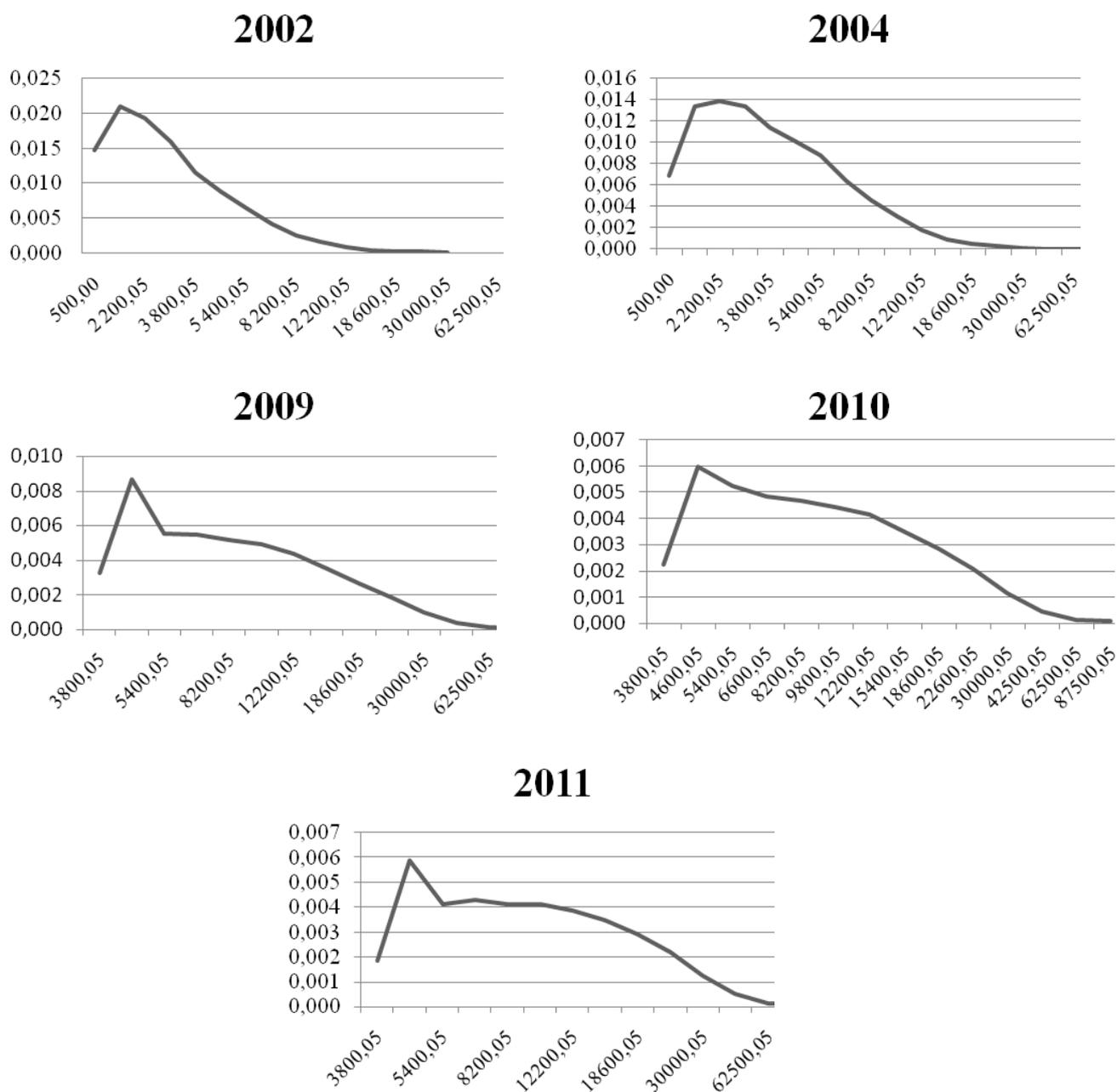


Рис. 3. Эмпирические плотности распределения работников РФ по начисленной заработной плате за 2002-2011 гг.

В результате декомпозиции образуются два распределения-компонента, первый – «левый» компонент аппроксимируется нормальным законом распределения. Второй – «правый» аппроксимируется логарифмически нормальным законом распределения. Причем вероятности принятия нулевых гипотез о данных распределениях выше, чем вероятность принятия гипотезы о логнормальном распределении всей совокупности (89 и 91% соответственно).

Что касается причин возникновения данной ситуации, то необходимо сделать вывод о наличии систематической ошибки наблюдения, возникшей в результате сокрытия респондентом размеров своей заработной платы.

Таким образом, для выведения хотя бы части реальной заработной платы из теневого сектора в целях повышения точности статистических данных и придания стимула для развития экономики и общества России, необходимо, по нашему мнению, установить минимальный размер оплаты труда в значении, равном границе между двумя совокупностями-компонентами, то есть своеобразному «естественному» МРОТ. Размеры предполагаемого и реально существующего МРОТ для сравнения занесены в таблицу 3.

Таблица 3

Естественный и официальный МРОТ РФ в 2002, 2004, 2009-2011 гг

Год	Официальный МРОТ, руб.	«Естественный» МРОТ, руб.	Отставание, руб.
2002*	400	3400	-3000
2004	600	5800	-5200
2009	4330	7400	-3070
2010	4330	8200	-3870
2011*	4493	8200	-3707

* Расчетные значения за год.

Заключение

Данное исследование показало обоснованность введения дополнительных классификационных разрезов совокупностей, таких как пороговые совокупности и их дальнейшего изучения в целях улучшения точности и качества полученных статистических данных, используемых для оперативного управления и стратегического планирования.

Пороговые совокупности – это такие реальные множества, отнесение единиц к которым происходит на базе статистических критериев, вводимых искусственным путем.

В ходе работы были апробированы методы декомпозиции смесей вероятностных распределений для изучения пороговой совокупности работников по начисленной заработной плате.

В результате исходное множество может быть разбито на две группы:

Первая – частная совокупность, сформированная под влиянием порога – МРОТ и аппроксимированная нормальным законом распределения. Здесь наблюдается систематическая ошибка наблюдения.

Вторая – частная совокупность, подчиняющаяся логнормальному закону.

Граница между двумя частными совокупностями должна быть принята как минимальный размер оплаты труда. То есть в 2010 и 2011 годах МРОТ должен был равен 8200 рублей, что больше официального уровня МРОТ почти на 4000 рублей.

В заключении можно отметить, что проведенное исследование подтвердило возможность применения методов разделения вероятностных смесей (композиций) распределения для изучения пороговых совокупностей и установления их оптимальных порогов.

Библиографический список

1. Венецкий И.Г., Венецкая В.И. Основные математико-статистические понятия и формулы в экономическом анализе. М.: Статистика, 1974.
2. Вентцель Е.С., Овчаров Л.А. Теория вероятности и ее инженерные приложения. Учеб. Пособие для втузов – 2-е изд. - М.: Высшая школа, 2000. - 480 с.
3. Глинский В.В. Статистические методы поддержки управленческих решений: Монография. - Новосибирск: Издательство НГУЭУ- 2008. - С. 108-118.
4. Глинский В.В., Серга Л.К. Нестабильные совокупности: концептуальные основы методологии статистического исследования // Вестник НГУЭУ. - 2009. - № 2. С. - 137–142.
5. Ионин В.Г., Ярославцева Л.П. Стратификация населения Новосибирской области по уровню денежных доходов // Вестник НГУЭУ. - 2010. - № 1. - С. 116 –130.
6. Серга Л.К. Об одном подходе к определению пороговых значений в решении задач классификации // Вестник НГУЭУ. 2012. №1. С. 54-60.
7. Труд и занятость в России. 2011: Стат.сб.Росстат. М., 2011. С. 426.
8. Трудовой кодекс Российской Федерации.
9. Ту Дж., Гонсалес С. Принципы распознавания образов: Монография. – М: Изд-во Мир, 1978. - С. 233-259.
10. Федеральный закон №82-ФЗ «О минимальном размере оплаты труда»
11. Харченко Л.П., Ионин В.Г. и др. Статистика. Учеб. Пособие. 3-е изд. М.: «ИНФРА-М», 2008. 445 с.
12. Эверитт Б. С. Большой словарь по статистике / науч. ред. перевода И.И. Елисеева. 3-е изд. М.: Проспект, 2010.
13. Официальный сайт Федеральной службы государственной статистики. URL: <http://www.gks.ru>

The bibliography

1. Venetskiy I.G., Venetskaya VI. The basic of the mathematical-statistical concepts and formulas in the economic analysis. M.: Statistics, 1974.
2. Ventsyl E.S., Ovcharov L.A. Theory of probability and its engineering applications. Ucheb. The manual for the system of technical colleges - 2-e Izd. - M.: The higher school, 2000. - 480 p.
3. Glinskiy V.V. Statistical methods of management decision support: a Monograph. - Novosibirsk: Publishing house of NSUEM - 2008. - With. 108-118.
4. Glinskiy V.V., Serg L.K.. Unstable population: the conceptual basis of the methodology of statistical research // Vestnik of NSUEM. - 2009. - № 2. With. - 137-142.
5. Jonas V.G., Yaroslavtseva L.P.. Stratification of the population of Novosibirsk region on the level of cash income // Vestnik NSUEM. - 2010. - № 1. - With. 116-130.
6. Serga L.K.. About one approach to determination of threshold values in solving classification problems // Vestnik NSUEM. 2012. №1. With. 54-60.
7. Labor and employment in Russia. 2011: Stat.sb.Rosstat. M., 2011. With. 426.
8. The labour code of the Russian Federation.
9. The J., Gonzalez With. Principles of pattern recognition: a Monograph. - Moscow: Publishing house of the World, 1978. - With. 233-259.
10. Federal law №82-FZ «On the minimum wage»
11. Kharchenko L.P., Jonas V. and others. The statistics. Ucheb. The allowance. 3-e изд. М.: «INFRA-M», 2008. 445.
12. Everitt B.C. Cambridge dictionary of Statistics / scientific Ed. translation I.I. Eliseeva. 3-e Izd. M.: Prospectus, 2010.
13. Official site of the Federal service of state statistics. URL: <http://www.gks.ru>