

Санкт-Петербургский государственный университет
Высшая школа менеджмента

НАУЧНЫЕ ДОКЛАДЫ

И.Ю. Чуракова

**НАПРАВЛЕНИЯ ИСПОЛЬЗОВАНИЯ
МЕТОДИК ВЫЯВЛЕНИЯ
АНОМАЛЬНЫХ НАБЛЮДЕНИЙ
ПРИ РЕШЕНИИ ЗАДАЧ
ОПЕРАЦИОННОГО МЕНЕДЖМЕНТА**

№ 13 (R)–2010

Санкт-Петербург

2010

И.Я. Чуракова. Направления использования методик выявления аномальных наблюдений при решении задач операционного менеджмента. Научный доклад № 13 (R)–2010. СПб.: ВШМ СПбГУ, 2010.

Ключевые слова и фразы: аномальные наблюдения (выбросы), процедуры выявления выбросов, статистический отчет для менеджмента, статистический процесс, принятие решений.

Работа представляет собой исследовательский материал, классифицирующий и систематизирующий выбросы и процедуры их выявления по основным характеристикам: численность совокупности, к которой применяется алгоритм, характер ее распределения, особенности функционирования алгоритма и работа с выходными данными. Для практического использования процедур выявления выбросов и дальнейшего применения результатов исследования в практике принятия решений автором предложен алгоритм статистического анализа данных, содержащих выбросы. Сделан обзор зарубежной литературы по практике принятия решений на основе данных, содержащих выбросы, проанализированы особенности статистического отчета для менеджера как основного документа, позволяющего повысить эффективность взаимодействия руководителя и специалиста и сделать статистический анализ мощным и действенным инструментом менеджера для принятия обоснованного решения.

Чуракова Ийя Юрьевна — кандидат экономических наук, старший преподаватель кафедры операционного менеджмента Высшей школы менеджмента Санкт-Петербургского государственного университета.

e-mail: churakova@gsom.pu.ru

© Чуракова И.Ю., 2010

© Высшая школа менеджмента СПбГУ, 2010

St. Petersburg State University
Graduate School of Management

WORKING PAPER

Iya Churakova

**THE WAYS OF USING
OUTLIERS DETECTING TECHNIQUES
IN COMPANIES' OPERATIONAL MANAGEMENT
TASKS**

13 (R)–2010

Saint Petersburg
2010

I.Yu.Churakova. The Ways of Using Outliers Detecting Techniques in Companies' Operational Management Tasks. Working Paper # 13 (R)–2010. Graduate School of Management, St. Petersburg State University: SPb, 2010.

Keywords and phrases: Outliers (extreme values), outliers detecting procedures, statistical report, statistical study, decision making.

This work represents the research material, classifying and systematizing commonly used outliers detecting procedures for their main characteristics: population or sample size, the nature of its distribution, work with the output data and procedures essential characteristics. For the practical use of detecting outliers (extreme values) procedures and further application of research results in practice, the author has proposed an algorithm of the statistical analysis of data containing outliers. The review of foreign literature on decision-making practice on the basis of data containing outliers is also presented in the work. The statistical report for the manager proposed to be the main special statistical instrument providing more effective business communications between managers and specialists and allows the statistical analysis become a powerful and effective tool for managers to make informed decisions. Executive summary is available at pp. 24.

Iya Yu. Churakova — Candidate of Sciences in Economics (First Doctoral Degree), Senior Lecturer, Graduate School of Management, St. Petersburg State University

e-mail : churakova@gsom.pu.ru

Содержание

Введение.....	6
Выбросы и их типы.....	7
Примеры использования алгоритмов выявления выбросов при принятии решений в области операционного менеджмента	18
Статистический отчет для менеджмента и отражение в нем результатов тестирования массива данных на присутствие выбросов.....	21
Выводы	22
Список литературы.....	23
Executive summary	24

Введение

Окружающая среда современного бизнеса предъявляет к управленцам все больше требований: борьба за владение ограниченными ресурсами, конкуренция в сфере сбыта, найма рабочей силы и прочие задачи заставляют принимать решения все быстрее, превращая то, что еще десять лет назад было просто неудачным («мелочь» в просторечии), в серьезную проблему, которая в режиме «здесь и сейчас» во многом может определить как вектор и динамику последующего развития фирмы, так и поставить вопрос о самом факте выживания бизнеса (дела) в современных реалиях функционирования мировой экономики. Глобальный экономический кризис добавляет неопределенности в хорошо отработанные алгоритмы, значимость привычных индикаторов изменяется под воздействием внешних и внутренних факторов. Сегодня на первый план выходит задача повышения компетентности самого руководителя, так как наработка собственного аналитического опыта, расширение поля возможностей за счет дополнительной информации, получаемой из доступных источников, но не лежащей на поверхности (требующей минимально-грамотной статистической обработки), позволяет менеджеру чувствовать себя гораздо увереннее в неопределенных ситуациях. Поэтому важной частью процесса принятия обоснованных управленческих решений наряду с интуицией и опытом менеджера сегодня становятся статистические методы обработки данных, как традиционные (анализ динамики, вариации, корреляционно-регрессионные методы), так и современные, требующие серьезной подготовки и применения специальных программных продуктов (интеллектуальный анализ данных, углубление данных и экспериментальный дизайн). Качество информации, элементарные умения и навыки по структурированию и отбору значимых данных, а также грамотно подобранные статистические алгоритмы превращают несложный анализ данных в мощный инструмент практикующего менеджера.

Массовые явления и процессы, составляющие предмет статистического анализа, обладают свойством изменчивости в пределах одной совокупности, или вариацией. Разные единицы одного и того же множества проявляют одно и то же изучаемое свойство в разной степени. Иногда различия столь велики, что обобщать и анализировать информацию становится трудно, а подчас и невозможно. Поэтому изучение характера изменчивости и управление вариацией представляют собой актуальную задачу статистической науки и практики. Именно контроль над вариацией и лежит в основе многих широко известных статистических процедур, применяющихся в качестве коли-

чественной основы принятия решений, например, в статистическом контроле качества используется информация о нормальных размерах вариации для выявления случаев, требующих принятия тех или иных решений. Одним из моментов, приводящих к искажению результатов статистического исследования и, соответственно, к ошибочному решению, является присутствие в совокупности наблюдений экстремально больших или малых значений — выбросов. Поэтому изучение методов выявления и, при необходимости, нивелирование значимости выбросов, несомненно, является важным для принятия управленческих решений.

Таким образом, целью работы является изучение и обобщение накопленного опыта по применению основных процедур выявления выбросов разных типов при решении задач операционного менеджмента.

Выбросы и их типы

Что же понимают под выбросами? Аномальными наблюдениями (выбросами, англ. Outliers, Extreme values) называют элементы совокупности (выборки), значительно отличающиеся от остальных по изучаемому признаку. Считается, что поиск отличающихся единиц начался с работы Сэра Френсиса Бэкона, великого философа и экономиста, точнее, с его работы «Новый Одеон», в которой он высказал предположение о том, что именно в различиях кроется движущая сила развития, предусмотренная самой Природой. Методология работы с аномальными значениями развивалась в работах Лежандра, Тьюкея, Дженкинса, Груббса и других.

Существует несколько типов выбросов, выделенных на основании статистических свойств исследуемой совокупности, схематически они представлены на рис. 1.

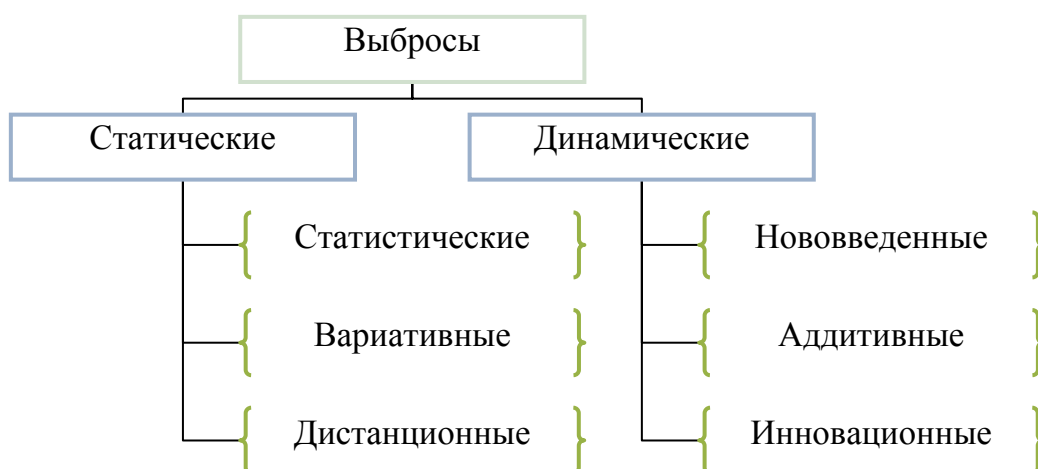


Рис. 1. Типы выбросов, выделенных по статистическим свойствам совокупностей

Теперь рассмотрим особенности каждого типа выбросов. Статистические выбросы, присутствуют в совокупности с известным распределением (например, нормальным или равномерным) как не укладывающиеся в нормальные границы совокупности. Это может быть бракованное изделие в серии, базовые характеристики которого не соответствуют норме и существуют в единственном экземпляре, или завышенная цена на продукт в одном магазине, резко отличающаяся от средних цен на рынке.

Вариативные выбросы принадлежат совокупности, но имеют резко отличающиеся значения по изучаемым характеристикам. Их необходимо не удалять, а, напротив, включать в анализ, так как присутствие вариативных выбросов означает не ошибки наблюдения или брак в работе, но может фиксировать наличие элементов другой совокупности в выборке или даже качественные новообразования в процессе, например, технологию производства автомобильных двигателей стандартного объема и мощности с экстремально малыми значениями показателей загрязнения окружающей среды.

Дистанционные выбросы, считаются таковыми, пока выборка или совокупность имеют существенную вариацию. Если объем выборки возрастет, она станет более однородной за счет появления единиц с промежуточными значениями, такие выбросы могут переместиться в нормальные границы совокупности. Таким образом, выделение описанных выше трех типов аномальных наблюдений тесно связано с характером вариации совокупности и с типом распределения массива данных. Причем в одном и том же массиве могут содержаться выбросы сразу нескольких типов, что необходимо учитывать при обработке данных.

В динамике термин «выброс» относится уже не к единице совокупности, а к уровню временного ряда. Различают три основных типа динамических выбросов: аддитивные, нововведенные и инновационные, отличные по источнику влияния и степени распространения. Аддитивные динамические выбросы — это уровни динамического ряда, подвергшиеся нерегулярной интервенции, оказавшей влияние только на данные уровни. Разница между теоретическим уровнем (по результатам сглаживания или аналитического выравнивания) и аддитивным выбросом будет представлена в трендовой модели как случайная ошибка.

Нововведенные динамические выбросы представляют собой нестандартные значения, причины появления которых (экстраординарный шок) воздействуют не только на текущие, но и на последующие уровни ряда. Например, введение запрета на вывоз пшеницы окажет влияние как на текущее значение объема продаж компаний-

экспортеров, так и на последующие, а также на объемы производства зерна, даже в том случае, если запрет будет отменен, ведь потерю эффективности экономической деятельности компаний компенсировать за один период вряд ли удастся.

Инновационные динамические выбросы представляют собой внутреннее изменение в системе, например, сдвиг уровней в связи с внутренними причинами, такими, как переход компании на новую технологию, или случайное однократное эндогенное изменение параметров системы.

С точки зрения причин, вызвавших появление выбросов, различают ошибочно аномальные значения (например, ошибки регистрации при обследовании, недостаточно качественные лабораторные образцы, смещенные выборки и т.п.) и отличающиеся данные. Разумеется, статистические свойства совокупностей и природа аномальных наблюдений влияют на выбор алгоритмов их обнаружения, а также на подготовку основы для принятия решений в случае, когда полученные данные содержат выбросы. На практике выявление выбросов не ограничивается выбором только одной стандартной процедуры и получением результатов с ее применением. Тестировать выборку на присутствие аномальных наблюдений приходится несколько раз в зависимости от сложности задачи, таким образом, для одного массива данных могут быть применены несколько процедур на разных этапах исследования: обработка и сводка первичных данных, изучение характера распределения после удаления или корректировки ошибок наблюдения, тестирование гипотез и получение аналитических выводов.

Зачем выявлять выбросы? Аномально высокие или низкие значения искажают показатели статистических свойств совокупностей — среднюю, дисперсию, коэффициент вариации, затрудняют анализ характера распределения совокупности. Иногда присутствие множественных выбросов в хвостах распределения может означать присутствие элементов частных совокупностей, обладающих специфическими свойствами. Такие распределения требуют моделирования по смешанному типу, используя несколько аналитических функций для разных частных совокупностей. В многомерных данных выбросы по одной переменной могут быть плохо заметны на первый взгляд, ведь во многих исследованиях число переменных измеряется десятками. Поэтому их принято обнаруживать при помощи специальных методов. Ведь все процедуры статистического вывода (планирование эксперимента, тестирование гипотез на основе выборок) по наборам данных, содержащим выбросы, могут дать решение, статистически незначимое, ценность которого на практике сомнительна.

Итак, предположим, массив данных содержит выбросы. Что делать в этом случае? Ошибки наблюдения можно удалить или заменить дополнительно собранными данными, на этот случай при организации выборочного наблюдения обычно обследуется дополнительно некоторое количество единиц. В случае стратифицированной выборки замена должна производиться только из единиц той же страты, а в случае кластерной выборки замене подлежит весь кластер.

Когда ошибки наблюдения удалены, но выбросы еще присутствуют в выборке, пользуются процедурами математического преобразования данных: логарифмированием, извлечением квадратного корня или стандартизацией. Зачастую после преобразований совокупность принимает характер какого-либо стандартного распределения без выбросов. Однако математические результаты не должны входить в противоречие с целью исследования: чрезмерные преобразования маскируют выбросы, а недостаточные зачислят в выбросы те единицы, которые находятся в нормальных границах. Практики решают это вопрос в соответствии с принципом парсимонии: упрощать, но не упрощаться. Иными словами, здесь большое значение имеет опыт предыдущих исследований по сходным массивам данных. Если же после преобразований аномальные значения остаются, то это и есть выбросы — отличающиеся данные, т.е. не характерные для совокупности в целом, но, тем не менее, присутствующие в ней.

Сколько выбросов можно считать нормой? Очевидно, что для небольших выборок два-три выброса могут составлять почти треть распределения и качество выборки нельзя считать приемлемым, тогда как для выборок более 30 единиц такое же число выбросов вполне допустимо. В литературе часто можно встретить приблизительную оценку допустимого числа выбросов — 20% от общей численности выбросов [Burke, 1998], но в каждом конкретном случае решение о допустимой численности выбросов должно приниматься в зависимости от характера распределения и интенсивности вариации выборочных данных, а также значимости выбросов в контексте принятия решений.

Отличающиеся данные при известном характере распределения совокупности являются тем материалом, который интересен руководителю при принятии решения. Например, поставщик с аномально низкой ценой и приемлемыми базовыми характеристиками товара или технология, с аномально высокой производительностью при том же объеме выпуска и качестве продукции может стать конкурентным преимуществом и обеспечить компании лидирующее положение на рынке. Но в процессе математической и статистической обработки данных как основы для принятия решений выявление выбросов про-

исходит с целью их удаления, т.к. значительная часть статистических и математических процедур чувствительна к выбросам. Еще Лежандр в 19 веке рекомендовал при расчете наименьших квадратов удалять чрезмерно большие наблюдения с целью уменьшить ошибки моделирования [Hadi, Rahmatullah, Werner, 2009]. Если присутствие выбросов все же необходимо при анализе ситуации, то можно провести моделирование без аномальных значений, а затем с ними, и сравнить результаты, либо выбирать робастные процедуры, т.е. устойчивые по отношению к наличию выбросов. Оба подхода применяются при подготовке статистических отчетов для менеджмента.

Анализ выбросов в динамических рядах представляет собой отдельную задачу, так как уровни динамических рядов расположены в порядке, в котором их значения появлялись, т.е. хронологическом. Кроме того, автокорреляция в рядах динамики может распространять влияние выброса на последующие наблюдения и простое исключение выброса из системы, хотя бы на время построения модели, не содержащей выбросы, здесь не всегда подходит. Зачастую выбросы появляются множественно, сразу на нескольких уровнях, появляется так называемый маскировочный эффект, скрывающий выбросы. Поэтому к анализу параметров уравнения тренда следует подключать периодизацию динамики, включение лаговых переменных, интерполяцию для пропущенных значений и другие процедуры, характерные для анализа временных рядов.

Алгоритмы обнаружения выбросов в совокупности или выборке будут различаться в зависимости от типа аномальных наблюдений. Рассмотрим наиболее часто использующиеся на практике алгоритмы, границы их применения и возможные направления использования для решения задач операционного управления.

1. Выявление выбросов на основе анализа стебельчато-листового графика (Steam-and-leaf plot). Стебельчато-листовой график позволяет сохранить индивидуальные значения признака и, вместе с тем, обладает основным свойством гистограммы — отражает характер распределения совокупности:

- массив данных упорядочивается по возрастанию значений;
- значение признака делится на «стебель» — повторяющуюся часть у нескольких единиц, и «листья» — остальные элементы;
- анализируется расположение «листьев» на стебле. Одиноко расположенные «листья», отдаленные по значениям «стебля» и есть выбросы.

Данный алгоритм является весьма субъективным с точки зрения выделения «стебля», однако он позволяет быстро оценить потенциальный риск попадания статистических, вариативных и даже дистанционных выбросов в массив данных. Подходит для массивов необработанных данных, широко используется в первичном анализе данных (Initial Data Analysis), в том числе для выявления ошибок наблюдения.

2. Выявление выбросов на основе межквартильного расстояния (Inter-Quartile Range, IQR) и построенных на его основе диаграммы «Ящик с усами» (двумерная диаграмма размаха, Box-and-Whisker plot) и гистограммы (Histogram):

- строится вариационный ряд, определяются пять основных характеристик (пятичисловая сводка, Five-number summary), включающих в себя минимальное значение признака, первое квартильное, медианное, третье квартильное и максимальное значения;
- рассчитывается межквартильное расстояние и на его основе определяются верхняя (Upper fence) и нижняя (Lower fence) границы совокупности;
- строится диаграмма и определяются наблюдения, являющиеся для данной совокупности в данный момент времени выбросами.

Используется в первичном анализе данных и в исследовательском анализе данных (Exploratory Data Analysis) для выявления структуры совокупности в нормальных границах и наблюдений за ее пределами. Основан на структурных характеристиках массива данных и подходит для выявления как статистических, дистанционных так и вариативных выбросов.

3. Выявление выбросов на основе теста Диксона, учитывающего не распределение всех единиц совокупности, а только максимальное значение признака и значения, окружающие «подозрительное» значение, т.е. предположительно выброс:

- определяется максимальное значение в выборке или совокупности;
- для всех элементов выборки (иногда только для подозрительных) рассчитывается тестовая статистика;
- фактическое значение критерия сравнивается с табличной величиной и принимается решение об отнесении значения к категории выбросов.

Часто тесты Диксон-типа применяются для порядковых переменных, т.е. рангов. Достоинством теста Диксона является возможность

его применения для экстремально малых выборок (до 7 единиц), а также нечувствительность к характеру распределения исследуемой совокупности. Недостатком является выявление единственного выброса за одну процедуру.

4. **Выявление выбросов на основе усеченных средних (Trimmed means):**

- строится диаграмма «Ящик с усами» по имеющейся выборке;
- из выборочной совокупности для получения α -усеченной средней отбрасывается определенная доля элементов ($\alpha/2$) с начала и конца выборочного распределения;
- рассчитывается средняя из оставшихся элементов.

Данный способ является робастной оценкой генеральной средней в случае, когда высок риск появления статистических или вариативных выбросов. Для подтверждения наличия выбросов следует сравнить обычную и усеченную среднюю — существенность различий средних дает уверенность в присутствии выбросов.

5. **Выявление выбросов на основе экстремального студентизованного отклонения** применяется для поиска единственного выброса в совокупностях, имеющих нормальное распределение:

- рассчитывается максимальное студентизованное отклонение от выборочной средней;
- полученное значение сравнивается с табличной величиной. Если максимальное отклонение больше табличной величины, то данное наблюдение считается выбросом и удаляется из выборки, после чего процедура повторяется.

Для применения данного метода необходимо предположение о нормальном распределении совокупности, подтвердить его можно, например, с помощью формального тест Андерсена-Дарлинга или эмпирического правила при объеме выборки более 30 единиц.

6. **Выявление выбросов на основе экстремального стандартизованного отклонения** применяются для выявления множественных выбросов в совокупностях, имеющих нормальное распределение:

- для выборки из n наблюдений, в которой предположительно имеется k выбросов, рассчитывается тестовая статистика (наибольшее стандартизованное отклонение (ESD), стандартизованный размах (STR), эксцесс (KUR));
- выборочные значения любой из этих статистик сравниваются с соответствующим критическим значением для $i = k$.

Если статистики, рассчитанные по выборке, превышают табличное значение, считается, что в выборке присутствует k выбросов. Если выборочная статистика меньше, чем табличная величина, то ее рассчитывают еще раз для $k-1$ выбросов. Процедура повторяется, пока не будет либо получен вывод о наличии выбросов, либо об отсутствии.

Данная процедура позволяет выделить множественные аномальные наблюдения на основании стандартной процедуры.

7. Выявление выбросов для двух количественных признаков, между которыми имеется статистическая связь, на основании показателей корреляции:

- строится поле корреляции (диаграмма рассеяния, Scatterplot), делится на квадранты (центральные линии соответствуют значениям средних по каждому из признаков), визуально обнаруживаются единицы или группы единиц, отстоящие от основного скопления точек;
- строится график остатков с выбросами и без них, визуально анализируются расхождения в графике;
- рассчитывается значение теста Кука (Cook's distance measure), на основании которого выделяются потенциальные выбросы;
- рассчитываются коэффициенты детерминации (R^2) и среднеквадратические ошибки для моделей с выбросами и без. Если статистическая значимость различий доказана, то наблюдения следует считать выбросами.

Выявление выбросов в данном случае крайне важно, так как коэффициент детерминации очень чувствителен к наличию выбросов. Смещенная в сторону выбросов дисперсия может показать наличие тесной связи только из-за одновременного наличия выбросов по значениям факторного и результативного признаков. Кроме того, искаженными оказываются коэффициент наклона линии регрессии и, в конечном счете, значимость уравнения корреляционной связи.

8. Выявление выбросов при многомерной группировке:

- визуально с помощью дендрограмм, многомерных графиков рассеяния;
- с помощью тестов Хаусманн-типа, состоящих в проверке существенности различий регрессионных моделей, построенных обычным методом наименьших квадратов (чувствительным к наличию выбросов) и любым робастным методом (например, регрессия наименьших абсолютных откло-

нений) выявляются наблюдения, дающие наибольшие расхождения между обычными и робастными оценками. Такие значения и есть наиболее вероятные выбросы.

Такие методы позволяют особенно эффективно выявлять статистические, вариативные и дистанционные выбросы. Сложность ситуации с выбросами в многомерном массиве состоит в том, что выбросы могут представлять собой отклоняющиеся значения по каждой переменной (многомерные выбросы), а могут отклоняться только по одной переменной, во всех остальных вполне укладываться в границы нормальной совокупности (такие единицы обычно маркируют по обозначению переменной, например, выбросы X-типа).

9. Выявление выбросов в рядах динамики имеет свои особенности, хотя, определение выброса остается стандартным. Известны два наиболее распространенных на практике подхода:

- **диагностика**, выбросы в этом случае оцениваются как уровни, имеющие наибольшее расхождение при обычном МНК и робастной процедуре оценивания параметров динамической модели;
- **робастный подход**, в этом случае изменяется сама процедура поиска параметров модели так, чтобы она была нечувствительна к наличию выбросов. Соответственно, максимальные отклонения фактических и выровненных по такой модели значений и есть выбросы.

Такие способы обнаружения выбросов подходят для аддитивных или нововведенных выбросов, когда порядок появления единиц в выборке задан, например, в динамических рядах. Кроме того, аномальные уровни рядов динамики могут представлять собой как однократные выбросы, так и предшествовать сдвигу уровней на более высокие значения.

Нельзя сказать, что в работе рассмотрены все существующие способы выявления выбросов, это, наверное, непосильная задача для одного автора. Однако, вновь разрабатываемые процедуры чаще всего возможно соотносить с уже функционирующими алгоритмами, чтобы понять, где именно их лучше использовать. В современной статистической практике нередки случаи, когда новым процедурам выявления выбросов, по сути являющимся модификацией того или иного алгоритма, присваивается название, общее для целого класса, например, «процедура Диксон-типа» или «тест Хаусман-типа». Этот факт позволяет быстро определить, пользуясь классификацией, какого типа вы-

бросы можно выявить на основе такой процедуры и для какой совокупности применять.

Обобщая приведенные в работе типы выбросов и способов их обнаружения, можно определить наиболее важные для практического использования моменты: характеристика совокупности, к которой применяется алгоритм, по численности элементов и характеру распределения, особенности функционирования алгоритма и работа с выходными данными. В таблице 1 представлен результат такого обобщения, то есть, дана краткая характеристика способов выявления выбросов, наиболее часто применяющихся на практике для анализа данных, в том числе и для подготовки статистической основы для принятия решений. Там же приведены особенности массива данных, к которым целесообразно применять тот или иной алгоритм, а также дан ответ на вопрос «как идентифицировать выброс», что важно знать для практического использования процедур.

Таблица 1

Способы выявления выбросов и особенности их применения для данных разных типов

Наименование способа	Характеристика совокупности	Особенности	Выявление выбросов
Стебельчато-листовой график (первичный анализ данных)	Необработанные массивы первичных данных	Отражает характер распределения совокупности, но сохраняет значения признака по единицам	Визуальное
Структурные средние и межквартильное расстояние (исследовательский анализ данных)	Необработанные массивы первичных данных, данные после удаления ошибок наблюдения	Выявляет характер распределения совокупности	Визуальное (гистограмма, диаграмма «Ящик с усами»), расчет нормальных границ совокупности (UF, LF)
Тест Диксона	Экстремально малые выборки (до 7 единиц)	Может рассчитываться по ранжированному переменным, выявляет единственный выброс за одну процедуру	Сравнение фактического значения критерия с табличной величиной

Продолжение Таблицы 1

Усеченные средние	Необработанные массивы первичных данных, данные после удаления ошибок наблюдения	Является робастной оценкой средней	Оценка существенности расхождения обычной и усеченной средней
Экстремальное студентизованное отклонение	Нормально распределенные совокупности	Выявляет единственный выброс за одну процедуру	Сравнение отклонения с табличной величиной
Экстремальное стандартизованное отклонение	Нормально распределенные совокупности	Выявляет множественные выбросы	Сравнение тестовой статистики с табличной величиной
Корреляционно-регрессионный анализ	Распределение единиц по значениям двух количественных переменных	Выявляет множественные выбросы, в том числе X-выбросы, Y-выбросы и выбросы по значениям обеих переменных	Визуально (график рассеяния, график остатков), сравнение тестовой статистики с табличной величиной, оценка существенности различий коэффициента детерминации, полученного по совокупности с выбросами и без
Многомерная группировка и множественная регрессия	Распределение единиц по значениям нескольких количественных переменных	Выявляет множественные выбросы, по значениям нескольких переменных	Визуально (дендрограмма) сравнение тестовой статистики с табличной величиной, оценка существенности различий коэффициента детерминации, полученного по совокупности с выбросами и без
Моделирование рядов динамики	Динамические ряды	Выявляет динамические выбросы	Оценка максимальной разницы теоретических значений, полученных по моделям, построенным обычным МНК и робастной процедуре

Таким образом, среди множества активных алгоритмов выявления выбросов выделяют два подхода: преобразование данных с тем, чтобы сделать выбросы более заметными при стандартных процедурах и методы, основанные на расчете расстояний между точками с помощью процедур, устойчивых к присутствию аномальных наблюдений [Hadi, Rahmatullah, Werner, 2009].

Примеры использования алгоритмов выявления выбросов при принятии решений в области операционного менеджмента

Рассмотрим теперь практические ситуации принятия решений в области операционной деятельности компаний, в случаях, когда изучаемый массив данных содержал выбросы.

Задача выбора промышленного робота по 20 базовым характеристикам [Booth, Khojja, Нц, 1991] была рассмотрена американскими учеными почти 20 лет назад. Выявление выбросов (т.е. роботов с нереальным или непривлекательным сочетанием характеристик) осуществлялось на основе сопоставления результатов многомерного моделирования обычным МНК и применения робастных процедур (расстояние Махаланобиса, метод главных компонент), выявленные выбросы представляли собой многомерные значения, аномальные сразу по нескольким переменным.

Авторы мечтали, что принятие решений о выборе лучшего варианта без выбросов в будущем можно будет переложить полностью на обладателя искусственного интеллекта. А в момент исследования применение статистических процедур выявления аномальных наблюдений избавили бы руководителей от подробного изучения технических характеристик и особенностей каждого робота, за ними осталось определение важнейших требований к объекту исследования и принятие решения после результатов исследований из наиболее реалистичных вариантов.

Задача выбора экстремально больших выбросов при проведении строительных контрактных аукционов решена исследователями из Австралии и Гонконга [Skitmore M, Lo H, 2002]. Особенностью исследования стало выявление вариативных выбросов (строительных проектов) с неправдоподобными техническими характеристиками (так как в случае аукциона возможно сознательное занижение стоимости или сроков строительства, которое впоследствии не оправдывается) в условиях экстремально малой выборки. Авторы создали выборку предложений по данным о проведенных ранее в мире аукционов аналогичного типа в заданном промежутке времени и построили обучающую функцию распределения относительной разницы между наибольшим и следующим за ним предложением. Для моделирования

они использовали стандартные функции: равномерного, нормального и лог-нормального распределения. Сопоставление относительных различий вариации по фактической и обучающей выборке и дало информацию о наличии несоответствующих конкурсу предложений. Предложенный подход может быть использован не только в строительной отрасли, а также при проведении государственных закупок зерна, используемых, в том числе, и в качестве целевых интервенций для «выравнивания» конъюнктуры рынка, плюс решение иных задач, базирующихся на конкурсном выборе предложений по нескольким параметрам.

В задаче построения статистической модели контроля качества (расширенной диаграммы контроля CUSUM) также были выявлены вариативные и статистические выбросы [Dror, 2009]. Автор для получения матрицы данных, формирующих вход, получил три значения по каждому параметру и использовал в качестве оценки средней медиану, чтобы избежать влияния аномальных значений, так как она служит оценкой в случае смещения средней, а выбранный метод статистической обработки данных (ANOVA) чувствителен к выбросам. Выбросы в данном случае оценивались как наблюдения за пределами нормальных границ совокупности.

Несколькими годами ранее, описывая особенности применения метода статистического контроля качества (SPC) авторы другого исследования заметили, что выявление и элиминирование выбросов с последующим пересчетом значений контрольных границ обязательно, так как R-карты и S-карты при учете выбросов дают смещенные оценки параметров процесса [Atienza, Ang, Tang, 1997]. В исследовании также предлагалось строить комплексные модели, объединяющие статистический контроль качества и прогнозирование по трендовым моделям. В этом случае аддитивные или инновационные выбросы также учитывался, так как их эффект может сказываться на величине контрольных лимитов.

Задачу выявления динамических выбросов в маркетинговых исследованиях решал профессор маркетинга из Франции — автор исследовал возможности построения регрессионной модели для динамического прогнозирования [Caruana, 2001]. Он использовал метод многократной множественной регрессии, сначала для выявления выброса на основании анализа графика остатков, а затем, после удаления выброса, для построения аддитивной трендовой модели с учетом сезонности, так как основной статистикой в данном случае является коэффициент детерминации, чувствительный к выбросам и требующий нормального распределения остатков.

Рассмотренные примеры представляют собой типичные случаи использования алгоритмов выявления выбросов разных типов для подготовки обоснованного управленческого решения. В практическом плане, как результат проведенного исследования, считаем возможным предложить для решения задач менеджмента алгоритм подготовки данных на основе статистического процесса с применением процедур, выявляющих выбросы:

- A. Постановка задачи исследования, определение совокупности и изучаемых признаков.
- B. Разработка основных гипотез исследования.
- C. Формирование выборки или подготовка специального исследования.
- D. Первичный анализ данных (Initial Data Analysis) и тестирование совокупности на выбросы, выявление ошибок наблюдения и их корректировка (Data Cleaning).
- E. Исследовательский анализ данных (Exploratory Data Analysis), изучение структуры и характера распределения выборочных данных. Выявление выбросов, представляющих собой отличающиеся данные по каждому значению признака.
- F. Тестирование гипотез с использованием специальных критериев и формирование статистического вывода. Выявление аномальных наблюдений на основании критерия или сопоставления робастной и обычной процедуры.
- G. Составление статистического отчета по результатам проведенного исследования. Представление результатов по выборке, содержащей выбросы и выборке с удаленными выбросами.

Далее следует разработка сценариев решений и выбор наиболее рационального решения. Конечно, можно выбрать вариант решения и без анализа выбросов, да и без статистического отчета вообще. Однако наличие в совокупности аномальных наблюдений может привести к использованию ложных предпосылок при принятии решения. Например, в случае проведения строительных аукционов было бы выбрано предложение только на основании нескольких критериев, самый важный из которых — цена и срок окупаемости проекта. Вполне вероятно, что таким предложением оказался бы совершенно нереалистичный проект, отнесенный в фактическом исследовании к выбросам, что на практике и происходит достаточно часто. Результат? Объект не построен, репутации компании-устроителя нанесен урон. Цена ошибки в результате некомпетентного или недостаточно подготовленного решения бывает порой довольно высока.

Статистический отчет для менеджмента и отражение в нем результатов тестирования массива данных на присутствие выбросов

Работа современного менеджера представляет собой комплекс сложных взаимодействий с насущной необходимостью принятия решений в предельно сжатые сроки. Статистический отчет призван повысить эффективность коммуникаций в деловой сфере, ускорить и упростить взаимодействие специалиста и руководителя.

Статистический отчет представляет собой систематизированную информацию о проведенном статистическом исследовании, основных гипотезах и полученных результатах, способ статистической поддержки компетентного менеджерского решения.

Цели, задачи и структура статистического отчета не регламентируется нормативными документами, но его внешний вид, ключевые параметры и структура сложились в процессе практического применения статистических методов в бизнесе и сейчас широко описаны в основном в зарубежной литературе [Higham, 1998].

Традиционно отчет содержит следующие разделы: аннотация, постановка задачи исследования и основные гипотезы, используемый метод исследования, результаты, выводы и рекомендации специалистам. Дополнительно приводятся основные этапы статистической обработки данных и ссылки на исследования аналогичного типа. Основной задачей специалиста, составляющего отчет, является подача важных результатов исследования сжато, понятно, увлекательно.

Отчет для руководителя должен содержать информацию по статистическому процессу, включая данные о выбросах, так как наличие или отсутствие аномальных наблюдений может существенным образом изменить подход к задаче, например, перевести ее из класса хорошо структурированных в класс слабоструктурированных. Соответственно, выбор метода принятия решений существенно изменится. Обычно рекомендуется делать два сценария: анализ проблемы, содержащий в массиве данных выбросы и анализ с исключенными выбросами.

Оформление отчета должно соотноситься с классом решаемых управленческих задач и категорией пользователей: для ежедневных отчетов по работе отдела продаж или производства необходим краткий свод показателей в простой форме и минимум выводов, а для привлечения новых клиентов потребуется красочная визуализация и понятная интерпретация основных показателей, наглядно демонстрирующая преимущества решения.

Выводы

Выявление аномальных наблюдений статистическими методами является весьма полезным при подготовке результатов статистического процесса, предназначенного для принятия решений. Присутствие выбросов различных типов в статических системах может искажать базовые характеристики совокупности или выборки, в частности, среднюю, размах и коэффициент вариации, дисперсию и коэффициент детерминации для множественных и парных уравнений регрессии. Выявление выбросов на разных этапах обработки данных позволяет разделить единицы в нормальных границах и за их пределами, выявить ошибки наблюдения и даже фиктивные данные, как, например, в случае с организацией аукционов. В динамике выбросы представляют собой уровни, подвергшиеся экстраординарному шоку снаружи системы или изнутри.

Быстро сориентироваться в разнообразных процедурах выявления аномальных наблюдений позволит представленная в работе классификационная таблица (Таблица 1), в которой автором составлена краткая характеристика по ключевым параметрам наиболее известных и используемых в мировой практике алгоритмов.

В результатах статистических исследований обычно мы видим уже обработанные массивы данных: даже несгруппированные выборки уже очищены от ошибок наблюдения или дистанционных выбросов. Сталкиваясь с необходимостью проведения исследования и обработкой результатов немудрено растеряться — полученная информация так сильно отличается от красивых таблиц в статьях и монографиях! Поэтому в контексте поставленной задачи в работе предложен подробный алгоритм практической работы с массивами данных, содержащих выбросы, подкрепленный практическими примерами решения аналогичных задач учеными и практиками из разных стран.

Разумеется, руководители используют не только собственные расчеты при разработке альтернатив решений, но и расчеты специалистов компании, консалтинговых агентств, независимых экспертов. Для повышения эффективности коммуникаций в деловой сфере, ускорения и упрощения взаимодействия специалиста и руководителя в мировой практике появился специальный стандарт результатов статистического процесса — статистический отчет. В отечественной практике он не распространен, а под термином понимается в основном отчет компании в органы государственной статистики. По мнению автора, время для изменения ситуации настало. Можно, должно и нужно сделать элементарные статистические исследования максимально полезными для основной деятельности менеджера — принятия компетентных и обоснованных решений.

Список литературы

Atienza O., Ang B., Tang L. Statistical process control and forecasting. *International Journal of Quality Science*, Vol.2 N1, 1997, p.37–51.

Barnett V., Lewis T. *Outliers in Statistical Data*, Wiley, 3 edition, 1998.

Booth D., Khouja M., Hu M. A Robust Multivariate Statistical Procedure for Evaluation and Selection of Industrial Robots. *International Journal of Operations and Production Management*, Vol.12, N2 1992, p.15–24.

Burke S. Missing values, Outliers, Robust statistics and Non Parametric methods, *VAM Bulletin*, 19 (1998), Autumn, 22–27.

Caruana A. Steps in forecasting with seasonal regression: a case study from the carbonated soft drink market. *Journal of product & brand management*, vol.10 N2, 2001, p.94–102.

Denon C., Gassner M., Verardi V. A New Hausmann Type Test to Detect the Presence of Influential Outliers. *ECARES Working paper*, 2008_006.

Dror S. A methodology for realignment of quality cost elements. *Journal of Modelling in Management*, vol.5N2, 2010.

Hadi A., Rahmatullah I., Werner M. Detection of outliers. *John Wiley & Sons, Inc. WIREs Comp Stat 2009*, Volume 1, July/August 2009.

Higham Nicholas J., *Handbook of Writing for the Mathematical Sciences*, 2nd edition, SIAM, 1998.

Skitmore M, Lo H. A method for identifying high outliers in construction contract auctions. *Blackwell Science Ltd., Engineering, Construction and Architectural Management*, 9/2–2002, p.90–130.

Сигел, Эндрю. *Практическая бизнес-статистика.: Пер.с англ.* — М.: Издательский дом «Вильямс», 2008.

Эверитт Б.С. *Большой словарь по статистике / науч.ред.перевода И.И. Елисеева.* 3-е издание. — М.: Проспект, 2010.

Executive Summary

The environment of modern business imposes increasing demands on managers: the struggle for the possession of limited resources, competition in sales, human resources and other problems to make decisions faster. The global economic crisis adds uncertainty to the well-developed algorithms, the significance of regular indicators changes under the influence of external and internal factors. Today the problem of increasing competence of the manager stands at the forefront. Therefore, an important part of informed management decisions, along with intuition and experience of managers today are statistical data-processing techniques, both traditional (time series, variation, correlation and regression techniques), and modern, require special knowledge and statistical software use (data mining, data dredging and experimental design). The nature of variability and variation control studies represents the actual problem of statistical science and practice. One of the things that lead to the distortion of statistical process and making wrong decision is the presence in samples (population) observations with extremely large or small values — outliers. Therefore, the study of outliers detection methods and, if necessary, leveling the significance of outliers is important for competent decision making.

The aim is to study and generalize on the application of the basic procedures for the identification of outliers of different types to solve problems of operational management.

Outliers are the elements of the population (sample) significantly different from most of the observations. It is believed that the search for different units began with the work of Sir Francis Bacon, the great philosopher and economist, rather, with his work “The new Odeon” in which he suggested that the difference lies precisely in the driving force for development provided by nature. Methodology of work with abnormal values was developed in Legendre, Tukey, Jenkins, Grubbs works and many others.

There are several types of outliers, allocated on the basis of the statistical properties of the studied population: statistical, deviation-based and distance-based. Outlier’s types are closely related to the nature of variation together with the type of data distribution. And in the same data set several types of outliers may be presented, this fact must be taken into account when processing data. In time series outliers are specific levels in series. There are three basic types of dynamic outliers: additive outliers, intervention events and innovative outliers. Additive outliers distort only current level (value in one time point), intervention events spreads their influence over the next few levels. Innovative outliers often anticipate the emergence of the level shift.

Outliers may occur in the population for several reasons: by chance (error in the measurement of a variable, sampling errors, etc.) or represent unusual value but common within a population. Abnormally high or low values distort the performance of statistical properties of sets — the mean, variance, coefficient of variation; make it difficult to identify the shape of a distribution. After all, statistical inference procedures (experimental design, hypothesis testing using sampling data) based on data set with outliers may provide a solution not statistically significant, and the results cannot be used in decision making practice.

Observation error can be removed or replaced with additionally collected data. When the observation errors are removed, but the outliers are still present in the sample, usually mathematical transformation of data is used: the logarithm, the square root or standardization. Often, after the transformation data set takes the form of any standard distribution. How many outliers can be considered the norm? It is usually assumed that the presence of 20% of the total number of units outside the normal boundaries is permitted, but the number of extreme values also depends on the distribution and spread of the sample data, as well as the importance of outliers in the context of decision-making.

Usually outliers detecting is used to remove them from data set. If the presence of outliers is still necessary to analyze the situation, it is possible to model without extreme values, and then with them, and compare the results, or choose a robust procedure. Robust statistical methods minimize the effect of an outlier observation. Both approaches are applied during statistical process and reporting the results for management.

There are various approaches to outlier detection depending on the application and number of observations in the data set:

1. analysis of Steam-and-Leaf plot, usually is used for raw data;
2. Inter-Quartile Range and normal boundaries of the sample (Lower Fence, Upper Fence) calculation, Boxplots and Box-and-Whisker plot developing;
3. Dyxon-type tests are based on the ratio of the ranges and allow to identify a single outlier in extremely small data sets (10 observations and less);
4. trimmed means (alpha-trimmed means are usually used instead of simple average when outliers may present in data sets, they represent robust measure of central tendency);
5. extreme studentized deviations are used to identify a single outlier in data set with normal distribution (normal samples). Maximum deviation from the mean is compared with a tabled value;
6. extreme standardized deviation (using different statistics, for example, KUR, ESD, STR) for normal samples;

7. for two variables using correlation and regression techniques;
8. from a multi-dimensional grouping visually using dendrograms, multidimensional plots, as well as using Hausmann-type tests;
9. in time series by use diagnostic and robust approach.

Summing up the above in the types of outliers and detecting procedures we can identify the most important points for practical use: characteristics of the population to which the algorithm is applied, the number of elements and shape of the distribution, specific features of the algorithm and work with the output data. Thus, speaking about extreme values detecting procedures we can identify two main ideas: data transformation in order to make the outliers more visible under standard procedures and methods based on calculating the distances between points using procedures that are resistant to the presence of outliers.

We now consider the practical situation of decision-making in the field of operating companies in cases where the studied data set contains outliers.

The problem of evaluating and selecting an industrial robot for 20 baseline characteristics has been examined by American scientists nearly 20 years ago [Booth, Khouja, Hu, 1991]. Identification of outliers (robots with unrealistic or unattractive combination of characteristics) was based on comparing the results of multivariate modeling ordinary least squares and use of robust procedures (Mahalanobis distance, principal component analysis) identified outliers is a multidimensional value, extraordinary by values of several variables.

The task of choosing extremely large outliers during construction contract auctions is solved by researchers from Australia and Hong Kong [Skitmore M, Lo H, 2002]. The feature of the study was to identify variative outliers (construction projects) to the improbable characteristics (as in the case of the auction may deliberately understating the cost or construction schedule, which would never be justified) using extremely small samples. For modeling the standard functions of the actual and the training sample were used.

In the problem of developing statistical model of quality control (enhanced control chart, CUSUM) were also identified variative and statistical types of outliers [Dror, 2009]. Outliers in this case were estimated as observations outside the normal boundaries of sample.

The task of identifying outliers in time series used in marketing research was solved by a marketing professor from France — the author investigated the possibility of constructing a regression model for prediction purposes [Caruana, 2001]. He used the method of multiple regression for the total number of units and then for the data set without extreme values.

In practical terms, as a result of our investigation, we consider it possible to offer for solving management algorithm training data based on statistical process with the procedures that detect outliers:

- A. Formulation of research problems, definition of population and the studied traits.
- B. Development of the main hypotheses of the study.
- C. Data collection and sampling.
- D. Initial Data Analysis and testing set for emissions, identification of errors of observation and correction (Data Cleaning).
- E. Research Data Analysis (Exploratory Data Analysis), to study the structure and distribution of sample data. Identification of emissions, which are characterized by the breakdown of each feature value.
- F. Hypothesis testing using specific criteria and the formation of statistical inference. Identification of outliers on the basis of a criterion or comparison of robust and normal procedure.
- G. Compilation of statistical reports on the results of the study. Presentation of results from a sample containing emissions and retrieval from remote emissions.

To increase the effectiveness of communication in business, accelerate and simplify the interaction between the expert and leader in the world is now a special standard for the results of the statistical process — a statistical report. The statistical report is a systematized information set about the statistical study, the main hypotheses and results, the method of statistical support for a competent manager's decision. The report contains the following sections: abstract, statement of research objectives and the main hypotheses, used method of investigation, findings, conclusions and recommendations to specialists. The report to the chief executive should contain the information on the statistical process including data on extreme values, since the choice of decision-making may change. In domestic practice, the statistical report is not common, and the term is understood in the main report of the company to state statistics committee. The author believes that the time has come to change the situation. You can, should and need to do basic statistical research as useful for the core business manager — take informed and sound decisions.

Опубликованные научные доклады

№ 1(R)–2005	А. В. Бухвалов Д. Л. Волков	Фундаментальная ценность собственного капитала: использование в управлении компанией
№ 2(R)–2005	В. М. Полтерович О. Ю. Старков	Создание массовой ипотеки в России: проблема трансплантации
№1(E)–2006	I. S. Merkuryeva	The Structure and Determinants of Informal Employment in Russia: Evidence From NOBUS Data
№ 2(R)–2006	Т. Е. Андреева В. А. Чайка	Динамические способности фирмы: что необходимо, чтобы они были динамическими?
№ 3(R)–2006	Д. Л. Волков И. В. Березинец	Управление ценностью: анализ основанных на бухгалтерских показателях моделей оценки
№ 4(R)–2006	С. А. Вавилов К. Ю. Ермоленко	Управление инвестиционным портфелем на финансовых рынках в рамках подхода, альтернативного стратегии самофинансирования
№ 5(R)–2006	Г. В. Широкова	Стратегии российских компаний на разных стадиях жизненного цикла: попытка эмпирического анализа
№ 6(R)–2006	Д. В. Овсянко В. А. Чайка	Особенности организации процесса непрерывного улучшения качества в российских компаниях и его связь с процессами стратегического поведения
№ 7(R)–2006	А. Н. Козырев	Экономика интеллектуального капитала
№ 8(R)–2006	Н. А. Зенкевич, Л. А. Петросян	Проблема временной состоятельности кооперативных решений
№ 9(R)–2006	Е. А. Дорофеев, О. А. Лапшина	Облигации с переменным купоном: принципы ценообразования
№ 10(E)–2006	Т. Е. Andreeva V. A. Chaika	Dynamic Capabilities: what they need to be dynamic?
№11(E)–2006	G. V. Shirokova	Strategies of Russian Companies at Different Stages of Organizational Life Cycle: an Attempt of Empirical Analysis
№12(R)–2006	А. Е. Лукьянова, Т. Г. Тумарова	Хеджевые фонды как инструменты снижения рисков и роста ценности компании
№13(R)–2006	Л. Н. Богомолова	Применение этнографических методов для изучения процессов принятия потребительских решений

№14(R)–2006	Е. К. Завьялова	Особенности профессионально-личностного потенциала и развития карьеры линейных менеджеров отечественных производственных предприятий
№15(R)–2006	С. В. Кошелева	Удовлетворенность трудом как комплексный диагностический показатель организационных проблем в управлении персоналом
№16(R)–2006	А. А. Румянцев, Ю. В. Федотов	Экономико-статистический анализ результатов инновационной деятельности в промышленности Санкт-Петербурга
№17(R)–2006	Е. К. Завьялова	Взаимосвязь организационной культуры и систем мотивации и стимулирования персонала
№18(R)–2006	А. Д. Чанько	Алгебра и гармония HR-менеджмента. Эффективность обучения персонала и диагностика организационной культуры
№19(E)–2006	T. E. Andreeva	Organizational change in Russian companies: findings from research project
№20(E)–2006	N. E. Zenkevich, L. A. Petrosjan	Time-consistency of Cooperative Solutions
№21(R)–2006	Т. Е. Андреева	Организационные изменения в российских компаниях: результаты эмпирического исследования
№22(R)–2006	Д. Л. Волков, Т. А. Гаранина	Оценивание интеллектуального капитала российских компаний
№23(R)–2006	А. В. Бухвалов, Ю. Б. Ильина, О. В. Бандалюк	Электронное корпоративное управление и проблемы раскрытия информации: сравнительное пилотное исследование
№24(R)–2006	С. В. Кошелева	Особенности командно-ролевого взаимодействия менеджеров среднего и высшего звена международной и российских компаний
№25(R)–2006	Ю. В. Федотов, Н. В. Хованов	Методы построения сводных оценок эффективности деятельности сложных производственных систем
#26(E)–2006	S. Kouchtch, M. Smirnova, K. Krotov, A. Starkov	Managing Relationships in Russian Companies: Results of an Empirical Study
№27(R)–2006	А. Н. Андреева	Портфельный подход к управлению люксовыми брендами в фэшн-бизнесе: базовые концепции, ретроспектива и возможные сценарии

№28(R)–2006	Н. В. Хованов, Ю. В. Федотов	Модели учета неопределенности при построении сводных показателей эффективности деятельности сложных производственных систем
№29(R)–2006	Е. В. Соколова, Ю. В. Федотов, Н. В. Хованов.	Построение сводной оценки эффективности комплексов мероприятий по повышению надежности функционирования объектов электроэнергетики
#30(E)–2006	M. Smirnova	Managing Buyer-Seller Relationships in Industrial Markets: A Value Creation Perspective
№31(R)–2006	С. П. Куш, М. М. Смирнова	Управление взаимоотношениями в российских компаниях: разработка концептуальной модели исследования
№32(R)–2006	М. О. Латуха, В. А. Чайка, А. И. Шаталов	Влияние «жестких» и «мягких» факторов на успешность внедрения системы менеджмента качества: опыт российских компаний
№33(R)–2006	А. К. Казанцев, Л. С. Серова, Е. Г. Серова, Е. А. Руденко	Индикаторы мониторинга информационно-технологических ресурсов регионов России
№34(R)–2006	Т. Е. Андреева, Е. Е. Юртайкин, Т. А. Солтицкая	Практики развития персонала как инструмент привлечения, мотивации и удержания интеллектуальных работников
#35(E)–2006	T.Andreeva, E.Yurtaikin, T.Soltitskaya	Human resources development practices as a key tool to attract, motivate and retain knowledge workers
№36(R)–2006	А. В. Бухвалов, В. Л. Окулов.	Классические модели ценообразования на капитальные активы и российский финансовый рынок. Часть 1. Эмпирическая проверка модели CAPM. Часть 2. Возможность применения вариантов модели CAPM
№37(R)–2006	Е. Л. Шекова	Развитие корпоративной социальной ответственности в России: позиция бизнеса (на примере благотворительной деятельности компаний Северо-Западного региона)
№38(R)–2006	Н. А. Зенкевич, Л. А. Петросян	Дифференциальные игры в менеджменте

№39(R)–2006	В. Г. Беляков, О. Р. Верховская, В. К. Дерманов, М. Н. Румянцева	Глобальный мониторинг предпринимательской активности Россия: итоги 2006 года
№40(R)–2006	В. А. Чайка, А. В. Куликов	Динамические способности компании: введение в проблему
№41(R)–2006	Ю. Е. Благов	Институционализация менеджмента заинтересованных сторон в российских компаниях: проблемы и перспективы использования модели «Арктурус»
№42(R)–2006	И. С. Меркурьева, Е. Н. Парамонова, Ю. М. Битина, В. Л. Гильченко	Экономический анализ на основе связанных данных по занятым и работодателям: методология сбора и использования данных
#43(E)–2006	I.Merkuryeva, E. Paramonova, J. Bitina, V. Gilchenok	Economic Analysis Based on Matched Employer-Employee Data: Methodology of Data Collection and Research
№44(R)–2006	Н. П. Дроздова	Российская «артельность» — мифологема или реальность' (Артельные формы хозяйства в России в XIX — начале XX в.: историко-институциональный анализ)
№1(R)–2007	Е. В. Соколова	Бенчмаркинг в инфраструктурных отраслях: анализ методологии и практики применения (на примере электроэнергетики)
№2(R)–2007	С. П. Куш, М. М. Смирнова	Управление поставками в российских компаниях: стратегия или тактика
№3(R)–2007	Т. М. Скляр	Проблема ленивой монополии в российском здравоохранении
№4(R)–2007	Т. Е. Андреева	Индивидуальные предпочтения работников к созданию и обмену знаниями: первые результаты исследования
№5(R)–2007	А. А. Голубева	Оценка порталов органов государственного управления на основе концепции общественной ценности
№6(R)–2007	С. П. Куш, М. М. Смирнова	Механизм координации процессов управления взаимоотношениями компании с партнерами
#7(E)–2007	D.Volkov, I.Berezinets	Accounting-based valuations and market prices of equity: case of Russian market

№8(R)–2007	М. Н.Барышников	Баланс интересов в структуре собственности и управления российской фирмы в XIX – начале XX века
#9(E)–2007	D.Volkov, T.Garanina	Intellectual capital valuation: case of Russian companies
№10(R)–2007	К. В. Кротов	Управление цепями поставок: изучение концепции в контексте теории стратегического управления и маркетинга.
№11(R)–2007	Г. В. Широкова, А. И. Шаталов	Характеристики компаний на ранних стадиях жизненного цикла: анализ факторов, влияющих на показатели результативности их деятельности
№12(R)–2007	А. Е. Иванов	Размещение государственного заказа как задача разработки и принятия управленческого решения
№ 13(R)–2007	О. М. Удовиченко	Понятие, классификация, измерение и оценка нематериальных активов (объектов) компании: подходы к проблеме
№14(R)–2007	Г. В. Широкова, Д. М. Кнатько	Влияние основателя на развитие организации: сравнительный анализ компаний управляемых основателями и наемными менеджерами
#15(E)–2007	G.Shirokova, A.Shatalov	Characteristics of companies at the early stages of the lifecycle: analysis of factors influencing new venture performance in Russia
#16(E)–2007	N.Drozdova	Russian “Artel’nost” — Myth or Reality' Artel’ as an Organizational Form in the XIX — Early XX Century Russian Economy: Comparative and Historical Institutional Analysis
#1(E)–2008	S.Commander, J. Svejnar, K. Tinn	Explaining the Performance of Firms and Countries: What Does the Business Environment Play'
№1(R)–2008	Г. В.Широкова, В. А. Сарычева, Е. Ю. Благов, А. В. Куликов	Внутрифирменное предпринимательство: подходы к изучению вопроса
№1A(R)–2008	Г. В. Широкова, А. И. Шаталов, Д. М. Кнатько	Факторы, влияющие на принятие решения основателем компании о передаче полномочий профессиональному менеджеру: опыт стран СНГ и Центральной и Восточной Европы

№ 2(R)–2008	Г. В. Широкова, А. И. Шаталов	Факторы роста российских предпринимательских фирм: результаты эмпирического анализа
№ 1(R)–2009	Н.А. Зенкевич	Моделирование устойчивого совместного предприятия
№ 2 (R)–2009	Г. В. Широкова, И. В. Березинец, А. И. Шаталов	Влияние организационных изменений на рост фирмы
№ 3 (R)–2009	Г.В. Широкова, М.Ю. Молодцова, М.А. Арепьева	Влияние социальных сетей на разных этапах развития предпринимательской фирмы: результаты анализа данных Глобального мониторинга предпринимательства в России
# 4 (E)–2009	N. Drozdova	Russian Artel Revisited through the Lens of the New Institutional Economics
№ 5 (R)–2009	Л.Е. Шепелёв	Проблемы организации нефтяного производства в дореволюционной России
№ 6 (R)–2009	Е.В. Соколова	Влияние государственной политики на инновационность рынков: постановка проблемы
№ 7 (R)–2009	А.А. Голубева, Е.В. Соколова	Инновации в общественном секторе: введение в проблему
# 8 (E)–2009	A. Damodaran	Climate Financing Approaches and Systems: An Emerging Country Perspective
№ 1 (R)–2010	И.Н. Баранов	Конкуренция в сфере здравоохранения
№ 2 (R)–2010	Т.А. Пустовалова	Построение модели оценки кредитного риска кредитного портфеля коммерческого банка (на основе методологии VAR)
№ 3 (R)–2010	Ю.В.Лаптев	Влияние кризиса на стратегии развития российских МНК
№ 4 (R)–2010	А.В. Куликов, Г.В. Широкова	Внутрифирменные ориентации и их влияние на рост: опыт российских малых и средних предприятий
# 5 (E)–2010	M. Storchevoy	A General Theory of the Firm: From Knight to Relationship Marketing
№ 6 (R)–2010	А.А.Семенов	Появление систем научного менеджмента в России
# 7 (E)–2010	D. Ivanov	An optimal-control based integrated model of supply chain scheduling
№ 8 (R)–2010	Н.П. Дроздова, И.Г. Кормилицына	Экономическая политика государства и формирование инвестиционного климата: опыт России конца XIX — начала XX вв.

№ 9 (R)–2010	Д.В. Овсянко	Направления применения компонентов менеджмента качества в стратегическом управлении компаниями
# 10 (E)–2010	V. Cherenkov	Toward the General Theory of Marketing: The State of the Art and One More Approach
№ 11 (R)–2010	В.Н.Тишков	Экономические реформы и деловая среда: опыт Китая
№ 12 (R)–2010	Т.Н. Клёмина	Исследовательские школы в организационной теории: факторы формирования и развития