Учреждение Российской академии наук
ЦЕНТРАЛЬНЫЙ ЭКОНОМИКО–МАТЕМАТИЧЕСКИЙ ИНСТИТУТ
CENTRAL ECONOMICS AND MATHEMATICS INSTITUTE

РОССИЙСКАЯ АКАДЕМИЯ НАУК

RUSSIAN ACADEMY OF SCIENCES

АНАЛИЗ И МОДЕЛИРОВАНИЕ ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ

Сборник статей

Выпуск 8

Москва 2011

Анализ и моделирование экономических процессов. Сборник статей под ред. В.З.Беленького, выпуск 8. – М.: ЦЭМИ РАН, 2010. – 131 с. (рус).

Коллектив авторов: **О.А.Андрюшкевич, В.З.Беленький, Т.А.Белкина, М.А.Гогулин, И.М.Денисова, Л.Я.Клеппер, Е.С.Паламарчук, В.А.Разумовская, С.А.Смоляк, Н.А.Трофимова, В.М.Четвериков**.

Восьмой выпуск ежегодного Сборника продолжает серию, начатую в 2004 году. Как и прежде, выделено четыре тематических раздела: "Анализ реальных экономических процессов", "Модели финансовых и рыночных механизмов", "Динамические модели", "Дискуссии, заметки и письма". Всего представлено 8 статей.

Analysis and modelling of economic processes. The Collection of Articles, ed. V.Z.Belenky, issue 8. – Moscow: CEMI RAS, 2011. – 131 p. (Rus).

The eighth issue of annual Collection of articles prolongs the series beginning in 2004. Four sections are share out: "Analysis of of actual economic processes", "Modelling of financial and market mechanisms", "Dynamic models", "Discussions, Notes and Letters". As a whole eight articles are presented.

Ответственный редактор – доктор физ.-мат. наук, проф. В.З.Беленький

Рецензенты:

доктор экономических наук, проф. О.Б.Брагинский
доктор физ.-мат. наук, проф. В.М.Хаметов

Редактор И.А.Левина

ISBN 978-5-8211-0583-7

Содержание

От редактора

Сборник стал традиционным и фактически обрел статус Ежегодника, издаваемого в ЦЭМИ РАН. Намечается в будущем издавать Сборник не только в печатном виде, но и открыть страницу на сайте ЦЭМИ, в которой Сборник будет представлен в электронной форме.

Как обычно, благодарю всех авторов за участие в Сборнике и желаю всем нам успехов в дальнейшей работе.

В.З.Беленький

О.А. Андрюшкевич, И.М. Денисова

ОСОБЕННОСТИ СОВРЕМЕННОГО РАЗВИТИЯ АУТСОРСИНГА В РОССИИ[1]

В последние десятилетия аутсорсинг, как одна из форм взаимодействия субъектов крупного, среднего и малого бизнеса приобретает все большее значение в мировой практике. Его применение позволяет корпорациям в значительной мере повысить свою конкурентоспособность за счет снижения издержек, рационализации производственной и управленческой деятельности, фокусирования на инновациях. Развитие современной экономической и инновационно-технологической системы нашей страны происходит в русле общемировых тенденций, поэтому аутсорсинг, являясь наиболее эффективной формой управления бизнесом, активно завоевывает российский рынок. Этим объясняется актуальность изучения основных тенденций и особенностей современного развития российского аутсорсинга.

В мировой практике применение аутсорсинга началось в сфере информационных технологий, в дальнейшем охватив такие отрасли экономики, как финансовая сфера, юридические услуги, управление инвестициями, здравоохранение, фармацевтика, биотехнологии, страхование и т.д. [1]. Большое распространение получил офшорный аутсорсинг – это форма аутсорсинга, предполагающая передачу некритичных для бизнеса процессов компаниям, находящимся в географическом удалении. При таком сотрудничестве компаний физическое расположение их офисов не имеет значения, а наиболее важным остается экономия за счет разного уровня оплаты труда. Разновидностями офшорного аутсорсинга являются:

- аутсорсинг информационных технологий (ITO – Information technology outsourcing), в частности разработка программного обеспечения (ПО) и вынос второстепенных служб поддержки инфраструктуры (Infrastructure technology outsourcing);
- аутсорсинг бизнес-процессов (BPO) - вынос некритичных для бизнеса процессов, требующих большого объема относительно неквалифицированного труда.

Основными причинами использования офшорных компаний считаются традиционное снижение издержек экономии на внутреннем штате или

[1] Работа выполнена при финансовой поддержке Российского Гуманитарного научного фонда, проект 10-02-00271.

на локальных подрядчиках (оно может составлять от 20 до 40%), а также такие факторы как временные ограничения и необходимые навыки. Ежегодный оборот мирового рынка офшорного аутсорсинга составляет приблизительно (в млрд. долл.) 315 (InfrastructureTO - 160, BPO - 140, ПО – 15) [2]. Наиболее сильную позицию на рынке офшорного аутсорсинга занимают Индия (аутсорсинг услуг) и Китай (аутсорсинг промышленного производства). Кроме того, сильные позиции занимает Россия (ПО), Пакистан (ПО, аутсорсинг услуг), Бангладеш (ITO, разработка услуг), Болгария (ПО), Украина (ПО), Румыния (ПО), Филлипины (аутсорсинг услуг), Египет (аутсорсинг услуг, ПО).

Мировой рынок аутсорсинга развивался поэтапно. Первой стадией был тактический и безотлагательный аутсорсинг вспомогательных функций как эффективный инструмент реструктуризации предприятий, при котором отношения заказчика и аутсорсинговой компании строились по схеме «покупатель-продавец». Второй фазой стал стратегический аутсорсинг, который характеризуется расширением спектра услуг, внедрением новых технологий, при этом участники осознают необходимость партнерских отношений для эффективности работы данной модели аутсорсинга. Современный этап развития аутсорсинга в мире характеризуется переходом к трансформационному аутсорсингу, при котором основное внимание уделяется инновациям, а также масштабу услуг, предоставляемых заказчику для активного роста предприятия. Российский рынок аутсорсинга несколько отстает в своем развитии от мирового - отечественные сервисные компании совсем недавно приступили к внедрению стратегического аутсорсинга. В данной работе основное внимание акцентируется на изучении проблем и особенностей развития российского рынка аутсорсинговых услуг – IT-аутсорсинга и BPO.

1. Аутсорсинг информационных технологий

Как и на Западе, локомотивом аутсорсинга в России выступает сфера информационных технологий. Своим ростом этот рынок обязан двум основным движущим факторам:

- расширение общего проникновения IT в бизнес-процессы компаний, механизмы организации государства и повседневную жизнь людей,
- тенденция к аутсорсингу внутренних функций IT и IT-инфраструктуры.

Развитие отечественных корпоративных информационных систем привело к возможности применения IT- аутсорсинга. Если в западных странах распространен комплексный ITO, в рамках которого подрядчику передается вся IT-инфраструктура заказчика, то в России такой подход используется крайне редко - компании предпочитают использовать его ограниченные варианты, такие как хостинг приложений или поддержка корпоративных сайтов. Многие современные российские компании пользуются услугами аутсорсинга, рассматривая его как сервис, организованный определенной компанией, где несколько услуг предоставляются комплексно для полного охвата потребностей клиента, но обычно акцент ставится на одну конкретную услугу.

1.1.Основные виды ITO. На практике в России наиболее распространенными являются следующие услуги ITO.

1.1.1. Абонентское обслуживание компьютеров. Заказчику предлагается комплексный набор услуг, позволяющий ему обойтись без собственного системного администратора или же значительно снизить его загрузку. Как правило, это услуги по настройке аппаратной части оборудования, настройке и обновлению программного обеспечения, зашита против взломов и проникновения в сеть, оперативный ремонт и замена оборудования, профилактические мероприятия по предотвращению поломок и износа оборудования. На практике те же фирмы часто оказывают сопутствующие услуги – поддержка работоспособности оргтехники, модернизация компьютерного парка, прокладка локальных сетей, IP–телефония и настройка АТС, IT-аудит и консалтинг.

*1.1.2. Центры обработки данных (ЦОД***)***. Многие компании сталкиваются с необходимостью пользоваться услугами ЦОД. Строительство собственного ЦОД требует от компании привлечения финансирования, проведения общестроительных работ, решения проблем с энергоснабжением, закупки серверного оборудования, организации службы эксплуатации ЦОД и обеспечения безопасности. ЦОДы периодически требуют текущей и капитальной модернизации. Многие компании предпочитают вместо создания корпоративного ЦОДа заключить договор с коммерческим ЦОДом и получить услуги ЦОДа как сервис.

1.1.3. Модель Software as a Service, (SaaS) – программное обеспечение как услуга. С появлением глобальной сети открылись новые возможности для развития IT-аутсорсинга и примером является модель Software as a Service, (SaaS). Первые компании, предлагающие программное обеспечение в аренду, появились в западных странах в конце 90-х г.[3]. SaaS-аутсорсер не только предоставляет программное обеспечение и физическое оборудование для размещения информационных систем, но и обеспечивает их установку, поддержку и обновление. В рамках модели SaaS заказчики платят не за владение программным обеспечением, а за его аренду и при этом, в отличие от классической схемы лицензирования ПО, заказчик несет сравнительно небольшие периодические затраты и ему не требуется инвестировать существенные средства в приобретение системы. Схема предполагает, что, если необходимость в программном обеспечении временно отсутствует, заказчик может приостановить выплаты.

1.1.4.Офшорное программирование. Неотъемлемой составляющей IT-стратегии большинства компаний на Западе являются офшорные услуги, а наиболее популярным в мировой практике направлением офшорных аутсорсинговых услуг является офшорное программирование – разработка ПО на заказ. Учитывая уровень развития отрасли ПО и разрабатываемых приложений, сложность аутсорсинговых заказов постоянно растет. Наблюдается рост заказов на разработку экспертных систем (автоматизированные службы помощи), обеспечение для call-центров и др. Этот рынок растет на 70-80% в год [4].

Офшорное программирование является одним из самых прибыльных видов аутсорсинга. Тем не менее, по оценкам экспертов, на фоне мирового аутсорсинга, российский сектор офшорного программирования пока еще остается достаточно слабым - значимые объемы экспорт ПО из России принял лищь в 2003 г., достигнув 546 млн. долл. Рост по отношению к 2002 г. тогда составил почти 50% [5]. Большинство Российских IT-компаний ориентировалось на европейский рынок, но лишь 14% из них занималось офшорным программированием, остальные оказывали и IT- услуги. Объемы рынка офшорного программирования несопоставимы с мировыми и фактически не видны на фоне IT-отрасли. Объем офшорного сегмента крупной компании не превышает нескольких миллионов долларов, мелкие фирмы довольствуются сотнями тысяч долларов в год. Большая часть заказчиков идет из США (примерно 85-90% всех заказов), оставшиеся – из Западной Европы [6]. Только фирмы, переориентировавшиеся с офшорного на производство своего ПО, имеют в качестве заказчиков крупные российские корпорации и государственный сектор. В России в настоящее время несколько тысяч офшорных компаний, расположенных чаще всего в крупных городах. Среди них множество некрупных фирм или просто команд программистов. Трудные ресурсоемкие проекты такие фирмы выполнять не могут, но небольшие задачи с дальнейшей поддержкой им по силам. В среднем российские программисты работают сегодня за 16-25 долл. в час (для сравнения украинцы просят за ту же работу – 7-10 долл., индусы – 25-50 долл., американцы – 70-140 долл. в час [6]).

В России аутсорсинг программных услуг (ПО) наиболее востребован. Изучением этого рынка занималась компания Elashkin Research [7]. Для этого был проведен опрос IT-директоров средних и крупных российских компаний. При этом были представлены почти все секторы экономики – производственный, нефтегазовый, энергетический, финансовый, транспортный и др., а также государственный. Свыше половины предприятий имели более 500 работников (их можно отнести к крупному по отечественным меркам бизнесу), 35% предприятий составляли средний бизнес. Подавляющее большинство опрошенных являлись коммерческими предприятиями, доля государственных структур составляла около 5%. Годовой бюджет предприятий распределился практически поровну в следующих категориях: до 50 тыс. долл., 50-100 тыс. долл., 100-500 тыс. долл. и более 500 тыс. долл. При этом преобладали предприятия с бюджетом свыше полумиллиона долларов, а небольшие компании с бюджетом до 50 тыс. долл. составили около 18%. Важной характеристикой, позволяющей оценить серьезность намерений предприятий в отношении аутсорсинга, являлось наличие выделенной в бюджете строки «аутсорсинг». На момент обследования 47 компаний такой строки не имели. Исследования показали, что услуги аутсорсинга рынком востребованы – это доказывает тот факт, что 24% компаний выделяют на него 10% бюджета. Доля внешних поставщиков при создании программных систем составляет 30%. Суммарная доля компаний, использующих услуги программного аутсорсинга, составляет 49%, а доля относящихся к аутсорсингу отрицательно – 12%. Достаточно большая группа компаний еще не определила свое отношение к этому виду услуг (24%), 5% планируют попробовать, а 10% имеют единичный опыт применения. Подобная картина – группа активных пользователей, небольшая группа противников и много колеблющихся – характерна для начального этапа распространения новых технологий или подходов. Это говорит о том, что в этой колеблющейся группе, в сумме составляющей 39%, находятся основные ресурсы роста этого рынка услуг. При этом используются все основные модели взаимодействия между разработчиками и заказчиком: аутсорсинг ресурсов (аренда специалистов, управление которыми осуществляет заказчик) – 34%, аутсорсинг отдельных задач (передача проектов, ответственность за которые лежит на исполнителе) – 42% и аутсорсинг процесса (передача процесса разработки) – 24%. Таким образом, наиболее популярен аутсорсинг отдельных задач, наименее – полный аутсорсинг [7].

1.2. Основные тенденции развития рынка ITO. Остановимся на анализе специфики развития в России наиболее распространенных видов услуг ITO.

Российский рынок аутсорсинговых услуг в сфере информационных технологий развивался очень быстрыми темпами: начиная с 2000-х годов он демонстрировал стабильный рост, прирост выручки в сегменте ITO превышал темпы роста IT-индустрии в целом, а в период с 2005 по 2006 гг. объем экспорта услуг увеличился на 80% (с 1 млрд. долл. в 2005 г. до 1.8 млрд. долл. в 2006 г.) [2]. Исследованием особенностей современного российского рынка IT-аутсорсинга - изучением изменений его структуры и основных тенденций развития занимается немало крупных аналитических компаний, таких как IDC, Gather, Астерос, in4media и др.

Компания Астерос представила результаты своих исследований российского рынка IT-аутсорсинга за 2010 г., проведенных по ее заказу in4media. По данным проведенных исследований удалось выявить специфику развития рынка IT-аутсорсинга в кризисные для отечественной экономики годы и сделать следующие выводы:

1) Большинству заказчиков, воспользовавшихся услугами ITO, все-таки удалось достичь поставленных целей: оптимизировать расходы на IT и обеспечить стабильное качество приобретаемых услуг. Более того, 51% компаний намерены в ближайшее время расширить объемы приобретаемых услуг, хотя основным фактором выбора поставщика ITO по-прежнему остается стоимость услуг [8].

2) Крупный российский бизнес оценил преимущества использования модели аутсорсинга в отношении IT-услуг, в особенности таких как разработка ПО, обслуживание информационных сетей и телекоммуникаций. В ближайшие два года большинство предприятий намерено сохранить или даже расширить объемы потребляемых IT-услуг. Почти три четверти (73%) опрошенных предприятий уже привлекают внешних поставщиков IT-услуг, еще более 13% планируют начать использовать аутсорсинг в ближайшее время.

3) Наиболее популярными услугами на крупнейших российских предприятиях остаются аутсорсинг ПО (71%) и аутсорсинг сетевых/телекоммуникационных услуг (51%). Более 40% опрошенных уже сегодня используют аутсорсинг IT-инфраструктуры, а 35% - аутсорсинг приложений.

4) На российских предприятиях большинство предприятий сотрудничают одновременно с несколькими поставщиками IT-услуг, а половина использующих ITO заказчиков работают по относительно небольшим аутсорсинговым контрактам (до 1 млн. долл.). И лишь 9% опрошенных компаний используют ITO в более крупных масштабах с суммами контрактов более 10 млн. долл.

Оценивая опыт аутсорсинга, абсолютное большинство респондентов (91%) признало свой опыт аутсорсинга IT-услуг успешным и утверждало, что аутсорсинг помог им достичь поставленных целей. Лишь для 3% компаний опыт по его использованию оказался неудачным. Основными причинами неудач они считали срыв сроков, отсутствие грамотного управления проектом со стороны поставщика, завышенное ожидание от проектов со стороны высшего руководства и акционеров предприятий–заказчиков. При этом выбор поставщика услуг определялся ими в первую очередь ценой услуг (83% опрошенных), и в меньшей степени известностью и репутацией поставщика IT-услуг.

Таким образом, результаты исследований позволили сделать весьма оптимистичные выводы относительно перспектив развития рынка ITO в России: в условиях кризиса, когда бизнес был поставлен в очень жесткие условия, многие крупные предприятия были вынуждены пользоваться услугами ITO и на своем опыте убедились в его преимуществах. И это, несомненно, должно способствовать тому, что они еще охотнее будут расширять свои аутсорсинговые контракты.

Растет значимость России на мировом рынке аутсорсинга. Международная ассоциация профессионалов аутсорсинга (IAOP) опубликовала рейтинг 2008 г. The Global Outsourcing 100, учитывающий деятельность 100 лучших аутсорсеров во всех отраслях. Рейтинг формировался на основе заявок, которые обрабатывает независимое жюри. Рассматривались такие показатели, как величина компании, рост ее оборота и клиентской базы, набор компетенций, а также потенциал в области менеджмента. В рейтинг 2008 г. были включены сразу 6 IT-компаний, которые являются российскими или изначально имеют в своем составе центры разработок в России. Этими компаниями являлись Auriga, DataArt, EPAM Systems, IBA, Luxoft и Mera Networks. Еще в 2007 г. в рейтинге The Global Outsourcing 100 фигурировали всего 4 российские компании (DataArt, EPAM Systems, Luxoft и StarSoft) [9].

Все включенные в рейтинг 2008 г. The Global Outsourcing 100 компании входят в состав ассоциации РУССОФТ, представляющей собой некоммерческое партнерство разработчиков программного обеспечения. В настоящее время РУССОФТ объединяет крупнейшие и наиболее влиятельные компнии-разработчики программного обеспечения России, Украины и Белоруссии и объединяет около 80 компаний, работающих в области информационных технологий со штатом более 17000 высококвалифицированных сотрудников [10]. Одна из таких компаний - компания Luxoft, включенная в рейтингThe Global Outsourcing 100, является признанным лидером на российском рынке производства и экспорта программного обеспечения. Ее офисы находятся в России, США, Канаде, Великобритании и Украине. Она по праву считается крупнейшим отечественным разработчиком программного обеспечения на заказ. Luxoff стала первой в мире компанией, система управления качеством которой прошла сертификацию на соответствие требованиям моделей SEI SW-CMM и SEI CMMI SW/SE 5 уровня одновременно. Компания владеет центрами разработок в Москве, Дубне, Омске, Санкт-Петербурге, Одессе, Днепропетровске и Ванкувере. Клиентами компании являются Deutsche Bank, IBM, UBS, T-Mobile, Ping Identity и др. Количество сотрудников – 4000, доход 2009 г. – 149 млн. долл.[10, 11]

Таким образом, мы убедились, что на фоне мирового финансово-экономического кризиса произошло повышение значимости IT-аутсорсинга в IT-стратегиях российских компаний-заказчиков. В отличие от большинства западных стран, где основной движущей силой развития рынка IT- не заниматься непрофильным бизнесом и тем самым снять с компании связанных с ним заботы, в России, по-прежнему, в основном прибегают к аутсорсингу из-за дефицита ресурсов и с целью экономии. При этом параметры технологичности и качества могут оставаться без внимания. Это способствует появлению на рынке компаний, оказывающих дешевые, но некачественные услуги. На российском рынке наиболее востребованными является гибкость и мобильность IT-услуг – поэтому в настоящее время преобладают не фиксированные контракты на 5 и более лет, а гибкие, учитывающие изменения в потребностях клиентов. Российскому рынку аутсорсинга свойственны непрозрачность и закрытость. Отсутствие знаний об успешных проектах приводит к тому, что многие заказчики считают, что выводить IT-сервисы на аутсорсинг небезопасно. Однако законы рыночного выживания, заставляющие компании добиваться эффективности и конкурентоспособности за счет снижения издержек, повышения качества продукции и услуг, доступа к передовым технологиям, неизбежно ведут к трансформации подходов и изменению рынка IT-аутсорсинга в сторону потребления стандартизованных, типовых сервисов, а в дальнейшем и в сторону стратегического аутсорсинга.

2. Аутсорсинг бизнес-процессов

Аутсорсинг бизнес-процессов подразумевает передачу другой организации каких-либо определенных процессов, которые не являются для компании ведущими в основной деятельности. На аутсорсинг могут быть переданы управление персоналом, бухгалтерский учет, логистика, маркетинг, реклама и др. За рубежом, в развитых странах, большое распространение получила такая форма правового обеспечения бизнеса, как юридический аутсорсинг. В российской экономике такая форма обслуживания бизнеса появилась относительно недавно, и рынок юридических услуг в стране неуклонно растет. Новыми, нетрадиционными для российского рынка аутсорсинговых услуг, стали архивный аутсорсинг, аутсорсинг печатных работ, появились аутсорсинговые компании в области экологии. Наиболее освоенными и востребованными направлениями рынка аутсорсинга бизнес-процессов являются производственный или промышленный аутсорсинг, финансовый и логистический. Ниже мы проанализируем основные тенденции развития и особенности для России этих форм аутсорсинга.

2.1. Производственный (промышленный) аутсорсинг. Промышленный аутсорсинг можно условно разделить на два вида – аутсорсинг заготовок и комплектующих и аутсорсинг функций и операций по обеспечению производства продукции (уборка производственных помещений, внутризаводская логистика, ремонт оборудования).

Для изучения тенденций развития промышленного аутсорсинга в России, полезно проследить историю его развития в мировой практике. В мире этот вид аутсорсинга гораздо более развит, чем в России. По данным исследования Week Census on Manufacturing – 54,9% американских компаний используют аутсорсинг в производстве, и 43,8% в обслуживании оборудования [12]. Основной мотив перехода на аутсорсинг на Западе – необходимость сконцентрировать ограниченные ресурсы на основной деятельности и достичь в ней преимуществ перед конкурентами за счет более низких издержек и/или более эффективного производства. По данным опроса 29 директоров по производству крупных американских, австралийских и европейских компаний, проведенному исследовательской компанией Plant Maintenance Resource Center, переход на аутсорсинг ими обосновывается стремлением увеличить производительность труда, уменьшить затраты, сфокусироваться на основной деятельности. Основные выгоды от перехода - быстрые сроки выполнения работ и доступность необходимого оборудования.

В России условия для перехода на аутсорсинг отличаются в силу ряда причин. Во-первых, бизнес на Западе изначально развивался исходя из рыночных соображений, которым соответствуют и принципы организации производства. Во-вторых, правовая среда позволяет с первых контактов устанавливать доверительные отношения. И, в-третьих, уровень развития поставщиков такой, что в большинстве случаев есть гарантия обеспечения требуемого качества услуг.

Традиции развития производства в России иные – практически все процессы, отвечающие за производство продукции, включены в структуру компаний. Кратко охарактеризовать современное состояние отечественной промышленности можно следующим образом:

1) типичный принцип построения российской производственной компании - « натуральное хозяйство»: собственное производство заготовок (литейный и кузнечные цеха), свои ремонтные и транспортные цеха, и т.д.;

2) усталость производственных фондов (производственные мощности изношены, низкий уровень загрузки, с трудом достигается требуемое качество);

3) ограниченные ресурсы (требуются затраты на переоборудование, отсутствуют возможности инвестирования в полный комплекс, поэтому возникает необходимость выбора сфер инвестирования);

4) производственные компании готовы выделять заготовительные и вспомогательные цеха и оставлять в структуре компаний только ключевые процессы и бизнесы.

Опыт успешного использования аутсорсинга российскими компаниями позволяет выделить основные побуждающие и сдерживающие факторы перехода на промышленный аутсорсинг в России.

Побуждающие факторы:

1) *Наличие периодически выполняемых работ* (ремонтные, техническое обслуживание зданий и оборудования). Основной мотив перехода - возможность избавится от чрезмерных накладных расходов.

2) *Изготовление простейших заготовок (*литые заготовки, неосновные элементы изделий). Пример - загруженные на 10% мощности по производству простейших изделий на Волгоградском тракторном заводе с приходом новой команды менеджмента были проданы, как убыточные.

3*) Наличие заготовок или операций, требующих специального оборудования*.

4) *Сезонные колебания спроса*. Пример - сезонность спроса на мороженое. Летом среднемесячные продажи поднимаются до 10 раз, поэтому в этот период крупные современные производители переводят на аутсорсинг производство части продукции.

5) *Подстраховка на случай поломки оборудования*. Некоторые компании используют аутсорсинг в небольших объемах, даже если имеются собственные производственные мощности, чтобы в случае поломки собственного оборудования можно было не останавливать отгрузки товара, а переключиться на производство на стороне. Аутсорсинг такого типа используется в случае, если собственное оборудование сильно изношено.

6) *Аутсорсинг как бизнес-модель*. Некоторые компании изначально определяют, какое производство будут осуществлять сторонние организации. Пример - компания «Август», производящая женскую одежду под маркой “OGGI”, самостоятельно производит разработку дизайна, конструкторскую подготовку, закупку материала и продажу одежды. Производство одежды для них осуществляют около 30 компаний-субконтрактников[12].

Сдерживающие факторы:

1*) Неспособность посчитать полную себестоимость*. Зачастую решения принимаются в отсутствии экономической проработки и не все факторы учитываются при анализе ситуации, поскольку необходимо учитывать полный набор затрат не только на производство, но и затраты на ремонт и обслуживание оборудования, обогрев, электроэнергию и уборку мусора и т.д.

2*) Отсутствие надежных поставщиков*. От компаний, которые используют аутсорсинг, требуется мобилизация управленческих, организаторских и переговорных ресурсов, чтобы добиться требуемого уровня качества и стабильности поставок.

3) *Снижение оперативности ниже требуемого уровня.* Один из аргументов за собственное производство заключается в опасении, что процесс подготовки к производству на стороне очень долог, особенно для сложных товаров.

4) *Высокие барьеры переключения или потенциальная монополия со стороны поставщика.* Объективные барьеры, делающие аутсорсинг экономически невыгодным – это отсутствие производителей подобных изделий в регионе, крупные габариты изделий, что делает невозможным их перевозку, а также риск возникновения монопольного положения поставщика.

5) *Усложнение управления.* Для активного использования аутсорсинга, помимо наличия экономической выгоды, необходимо, чтобы работники компании и система управления была способна эффективно управлять контрактными отношениями.

6*) Отсутствие гибкости поставщиков.* Основные претензии крупных производителей к поставщикам деталей и узлов состоят в том, что обычно это мелкие предприниматели, не принимающие во внимание проблемы долгосрочного развития и инвестиций и стремящиеся из данных им производственных активов «выжать» максимум возможного на данный момент. Компании, успешно использующие аутсорсинг, стремятся преодолеть эти факторы и обезопасить выполнение своих договорных обязательств.

Практика использования аутсорсинга показала, что переход к стратегическому аутсорсингу в России встречает много преград. В частности, активному аутсорсингу в российской промышленности в основном мешает отсутствие доверия в бизнесе: руководители боятся вступать в деловые отношения со сторонними компаниями. Для внедрения стратегического аутсорсинга в промышленность требуется изменить мышление не только предпринимателей, но и сотрудников, которые должны оценить выгоды аутсорсинга. Разработка и реализация стратегического аутсорсинга доступна только крупной профессиональной компании, обладающей, в первую очередь, значительными кадровыми и финансовыми резервами и опытом работы на рынке, поскольку стратегическое партнерство предполагает создание глубоко интегрированных схем, направленных на качественное развитие бизнеса. Совместное развитие и взаимная интеграция с аутсорсинговой компанией способно не только обеспечить развитие профильных компетенций у компании-заказчика услуг, но и создать новые, которые аутсорсинговая компания, в свою очередь, сможет освоить и предложить новые услуги.

Следует особо отметить, что для успешного перехода на аутсорсинг руководство компании должно быть очень осторожным. Уже принимая решение на выделение в аутсорсинг непрофильных активов, необходимо правильно определить, какие функции являются для компании непрофильными. Имеются примеры, когда переданные функции считались второстепенными, а со временем они становились стратегическими для бизнеса, в то время как компетенции компании в этой области были утеряны. Существует ряд требований, которым должна удовлетворять профессиональная сервисная компания. Одно из них – гибкость и быстрота принятия решений. Здесь следует обратить внимание на опыт работы аутсорсера на рынке, клиентскую базу в сфере деятельности предприятия-заказчика, важным является прозрачность действий компании. Когда аутсорсинговая компания приступает к выполнению своих обязательств по стратегическому партнерству, отношения компаний переходят на новый уровень развития, при которых обе стороны будут осуществлять мониторинг производительности, оценивать результаты и решать проблемы. Стратегическое партнерство становится тем позитивным динамическим фактором, который существенно ускоряет развитие, помогая при этом снизить операционные издержки и обеспечить разделение рисков между компаниями. Крупные отечественные аутсорсеры успешно вступают в соревнование со своими иностранными конкурентами в области качества предоставляемых услуг.

2.2. Логистический аутсорсинг. В процессе дальнейшего развития тенденций снижения затрат компании стали использовать аутсорсинг в других, помимо производства, бизнес- процессах, в частности в области логистических услуг. Для крупных компаний логистический аутсорсинг может стать необходимым элементом управления сложными алгоритмами поставок. Перед руководителем компании встает вопрос о целесообразности применения аутсорсинговых схем, и, в случае его положительного решения, выбрать, какие логистические функции отдать подрядчику, а какие оставить за внутренними отделами логистики.

Следует отметить, что развитие собственной внутренней логистики имеет ряд преимуществ. Во-первых, это то, что помимо наработанного опыта и независимости от сторонних посредников, она обеспечивает компании персонифицированный сервис, в ряде случаев значительную экономию средств, а также возможность в дальнейшем провести диверсификацию бизнеса. Использование аутсорсинга подразумевает качественный, но стандартизированный сервис, а также сокращение финансовых рисков, поскольку логистический посредник берет на себя ответственность за исполнение логистической операции. Помимо этого, существует возможность сокращения расходов на логистическую часть бизнеса за счет сокращения транспортных расходов,затрат на приобретение оборудования и содержание штата специалистов. Кроме того, при наличии конкуренции на рынке логистических услуг, аутсорсинговые компании будут стремиться сделать для своих клиентов выгодные предложения в части оптимального соотношения цены и качества услуг.

На российском рынке востребованные логистические услуги условно можно объединить в 5 групп [13].

1) Транспортно-экспедиционные услуги. Существуют отдельно транспортные услуги и экспедиционные. Как правило, транспортную услугу заказывают те компании, в которых уже существует отдел экспедирования. В этом случае транспортная компания производит только перевозку груза. Чаще всего оказывается востребованной транспортно-экспедиционная услуга, которая включает полную координацию грузоперевозки.

2) Услуги по таможенному оформлению. Таможенное оформление – достаточно специфическая сфера логистической деятельности, подразумевающая работу непосредственно с таможенными органами в регламентированном правовом поле. Это обуславливает наличие в компании аттестованных специалистов по таможенному оформлению и уплаты таможенных платежей. Поэтому внутри компании такую работу организовать трудно. Помимо этого, при возникновении разногласий с таможенными органами в части декларирования товаров вся ответственность, как юридическая, так и финансовая в виде штрафов, ложится непосредственно на компанию.

3) Услуги склада. Организация складского комплекса внутри компании достаточно дорогостоящее мероприятие, подразумевающее большие финансовые инвестиции как в покупку либо аренду помещения, так и в оборудование склада в соответствии с требуемыми нормами.

4) Координация процесса закупок, упаковка/переупаковка товаров. Данный вид услуг востребован чаще компаниями, имеющими в своей организационной структуре достаточно разветвленные филиальные (торговые) сети. В данном случае заказ такого рода логистических услуг целесообразен и оправдан.

5) Комплексные услуги (комплексный аутсорсинг). Здесь в роли компании-подрядчика выступает единый логистический оператор. Как правило, это крупные логистические компании, имеющие в своем распоряжении все ресурсы, необходимые для реализации логистических бизнес-процессов. Данная комплексная услуга востребована компаниями, алгоритм поставок которых достаточно сложен и имеет несколько промежуточных этапов. По сути, это услуга по товародвижению от производителю к потребителю

Таким образом, при решении вопроса о том, какие логистические функции доверить посреднику, необходимо ориентироваться на возможности компании в части организации собственной логистики, корпоративную стратегию в части дальнейшего развития бизнеса (концентрацию или диверсификацию), частоту и сложность поставок и т.д. На данный момент логистический рынок в России достаточно развит, и компания имеет возможность найти для себя оптимальное решение.

Можно выделить три уровня отечественных посредников, осуществляющих свой бизнес на логистическом рынке.

1) Логистические компании уровня 1. Эти компании осуществляют свою деятельность в рамках одной услуги. Примером могут служить таможенные брокеры, специализирующиеся исключительно на таможенном оформлении грузов.

2) Логистические компании уровня 2. Это компании – традиционные логистические посредники, исполняющие традиционные логистические функции.

3) Логистические компании уровня 3. Такие компании исполняют несколько или все логистические функции и наиболее приближены к единому логистическому оператору.

Сегодня в России растет объем товарооборота, который проходит через кооперацию в области логистики. Тем не менее, на аутсорсинг логистическим операторам в рамках контрактной логистики передано лишь 27% товарооборота, а 73% собственных грузопотоков торговые и промышленные предприятия обслуживают самостоятельно [14]. Среднее значение роста абсолютных значений годового оборота логистических операторов за 2006-2007 г.г. составило около 8%, для транспортно-экспедиторских компаний – 23%, а складских операторов – около 2%. Это показывает, что конкуренция на рынке еще недостаточно высока. Среднее значение рентабельности по отрасли в 2007 г. составило примерно 11%, причем диапазон изменения значений рентабельности исследуемых компаний – от 0,5% до 48% [14]. Таким образом, снижение рентабельности в перспективе будет стимулировать более агрессивное поведение компаний и рост конкуренции на рынке. Высокие темпы роста объема рынка привлекут новых конкурентов и смогут в кратко– и среднесрочной перспективе улучшить ситуацию на этом рынке.

В настоящее время на рынок логистических услуг большое влияние оказывает растущая глобализация экономической активности и выход на российский рынок более мощных иностранных компаний, требующих от российских логистических компаний не только повышения качества услуг и гибкости, но и развития комплексной логистической услуги. Это обуславливает на только развитие отдельных логистических операторов в крупные логистические провайдеры, но и объединение отдельных логистических операторов в союзы как с другими логистическими операторами, так и с партнерами в различных сферах бизнеса. В конечном итоге, это ведет к повышению качества логистических услуг в целом и переходу от комплексной логистики к логистике интегрированной, которая позволяет более эффективно реализовывать цели бизнеса.

2.3. Аутсорсинг в финансовой сфере. В последнее время за рубежом и в России все большую популярность приобретает финансовый аутсорсинг, в состав которого входят бухгалтерский аутсорсинг, аутсорсинг заработной платы и др.

Бухгалтерский аутсорсинг – один из вариантов бухгалтерского обеспечения функций учета и отчетности на предприятии, за которое несет ответственность, в соответствии с законодательством, руководитель предприятия. Аутсорсинговая компания используется для обеспечения бухгалтерского учета и предоставления необходимой отчетности в налоговые органы и внебюджетные фонды. Формы сотрудничества могут быть весьма разнообразны. Теоретически и практически бухгалтерская компания может полностью выполнять функции бухгалтерии предприятия вплоть до исполнения банковских платежей и выставления первичных документов контрагентам. Однако это не исключает возможности построения достаточно сложных и гибких бизнес-процессов с использованием внешних компаний и собственных сотрудников для достижения максимальной эффективности.

Аутсорсинг заработной платы подразумевает именно расчет заработной платы, премиальных, командировочных, пособий, компенсаций и т.д., а не осуществление выплат. Данный бизнес отдают на аутсорсинг во многих странах мира. Так, в США, стране родоначальнице самого понятия аутсорсинга, 80% компаний осуществляют расчет заработной платы через аутсорсинг. В Европе эта услуга используется повсеместно. В России на данный момент только 7-10% [15] компаний в той или иной степени используют эту услугу, в странах СНГ – их количество еще меньше. Есть два возможных пути решения для реализации этой услуги. Это IT-модули по расчету заработной платы – когда клиент покупает некий продукт и сам в нем реализует необходимые расчеты. Второй вариант – полный аутсорсинг функции – когда клиент покупает комплексную услугу, и весь процесс расчета остается в зоне ответственности аутсорсера. В том и другом случае разработки и IT- решения компаний-поставщиков очень важны.

Финансовый аутсорсинг, помимо свойственных аутсорсингу общих преимуществ, имеет преимущества, присущие только ему. Главным из них является переложение ответственности за организацию учета и его правильность на аутсорсинговую компанию, причем возмещение убытков, связанных с неправильным расчетом налогов и несвоевременным предоставлением отчетности, происходит за счет поставщика услуг или по страховому договору. Кроме того, приобретая услуги по бухучету, клиент получает доступ к юридическим и налоговым ресурсам аудиторских компаний, способных найти нетрадиционные решения возникающих проблем и отстоять их в суде. Таким образом, снижается риск принятия ошибочных решений и возрастают шансы на успех в дискуссиях с налоговыми органами при защите выгодных для клиента решений по тем спорным вопросам, по которым существует двоякое толкование законодательства.

Однако у компаний-заказчиков существуют основания избегать аутсорсинговых услуг. Во-первых, многие из них просто не знакомы с принципами аутсорсинга бизнес-процессов. Кроме того, компании боятся потерять контроль над ситуацией и избегают доверять посторонним коммерческие тайны, а основания для таких опасений есть как в России, так и за рубежом. Одних заказчиков останавливает дополнительные расходы сил и средств на этапе перехода и чисто психологический барьер, связанный с передачей «своей» бухгалтерии в «чужие» руки. У других возникают опасения за сохранность принадлежащей им финансовой информации, за правильность ведения и предоставления налоговой отчетности в регулирующие органы. Кроме того, хотя крупные аутсорсинговые компании обладают соответствующей компетенцией и несут ответственность перед заказчиком за правильность и своевременность предоставления информации, тем не менее, именно финансовый отдел и бухгалтерия заказчика отвечают перед государством за правильность и своевременность отчетности. В России одной из особенностей ведения бизнеса является то, что многие компании не просто заботятся о конфиденциальности информации, но и не заинтересованы в какой бы то ни было открытости, так как зачастую бухгалтерский учет носит полулегальный характер. Как следствие, многие крупные заказчики неохотно идут на аутсорсинг, предпочитая вести бухгалтерию самостоятельно.

Сейчас в России наметилась тенденция отказа от «серых» схем ведения бизнеса и деления бухгалтерии на теневую и официальную. Для компании, готовящейся к выходу на международный рынок капитала, использование независимой внешней бухгалтерии служит своего рода декларацией о намерениях построить прозрачную финансовую систему, и это, несомненно, является дополнительным аргументом в пользу компании в глазах иностранных инвесторов и партнеров.

2.4. Исходящий телемаркетинг и обработка телефонных вызовов. Аутсорсинг обработки телефонных вызовов и передача на сторону колл-центров - еще один распространенный пример аутсорсинга бизнес-процессов в области управления взаимоотношениями с клиентами. Многие западные компании используют именно этот тип аутсорсинга, размещая колл-центры в местах с дешевой рабочей силой, зачастую в других странах. Лидирующим поставщиком по аутсорсингу колл-центров является Индия, способная за счет дешевой рабочей силы предоставить западным компаниям приемлемое качество по низким ценам. Достаточно часто колл-центры размещаются в странах Восточной Европы, а также некоторых арабских государствах (например, в Египте). Аутсорсинг колл-центров в России также достаточно распространен, однако имеет несколько другую форму. Требования к владению русским языком практически исключают размещение колл-центров в дальнем зарубежье и даже в некоторых странах СНГ. Поэтому многие компании предпочитают переносить центры обслуживания в российские провинциальные города со сравнительно дешевой рабочей силой.

2.5. Аутсорсинг сервисных функций. Рынок аутсорсинговых услуг в области юридической поддержки, сервисов в области документов и экологии появился в России сравнительно недавно и продолжает расти. Мы рассмотрим новые и наиболее востребованные в настоящее время виды этих сервисных аутсорсинговых услуг.

2.5.1. Юридический аутсорсинг - это правовое абонентское обслуживание предприятий. В современном мире все руководители компаний сталкиваются с проблемой оптимизации работы своего бизнеса, эффективного сопровождения сделок, увеличения качественных показателей в работе. В этих условиях выгодным становится передать правовое управление бизнесом юридической аутсорсинговой компании, осуществляющей мониторинг изменений в законодательстве и разработку механизмов защиты собственности клиента, чем осуществлять затраты на собственный квалифицированный юридический отдел.

Юридический аутсорсинг имеет следующие преимущества:

1) Экономия средств. Пользование услугами юридического аутсорсинга дает предприятию экономию денежных средств на административных расходах, расходах на офис и уплате налогов с заработной платы сотрудников.

2) Рост конкурентоспособности. Штатный юрист не может обладать глубокими знаниями по всем отраслям права, а поэтому не сможет быстро разрешать все юридические вопросы, возникающие в ходе деятельности предприятия. Заключая договор правового абонентского обслуживания, предприятие получает автоматически для осуществления своей деятельности целый штат опытных юристов, позволяющий решать вопросы, связанные с трудовым, налоговым, гражданским, административным, а иногда и с уголовным правом, обеспечить надежную защиту интересов работодателя и оперативно справиться с разнообразным потоком правовых вопросов. Это намного дешевле, чем затраты на организацию собственного юридического отдела.

К минусам юридического аутсорсинга относятся:

1) Отсутствие комплексного видения правовой ситуации, что впоследствии приводит к юридическому анализу, построенному с учетом не полного количества факторов. Договор аутсорсинга заключается, как правило, уже при наличии давних крупных юридических проблем, причем в момент их обострения. Юрист-аутсорсер, привлеченный для решения возникших острых проблем, концентрируется на решении конкретно поставленной юридической задачи, без учета глубины, степени и всей тяжести юридической проблемы. Штатный юрист, работающий длительное время, может и должен предпринять все юридические действия, направленные на предотвращение и недопущение подобных юридических проблем в будущем.

2) Дороговизна услуги юридического аутсорсинга. Услуги юридического аутсорсинга гораздо более дороги, чем привлечение юриста по договору об оказании услуг.

В России для развития юридического аутсорсинга существует немало преград:

- низкая правовая культура предпринимательской деятельности, в том числе недисциплинированность в плане исполнения договорных обязанностей;

- отсутствие достаточного количества юристов высокой квалификации;

- отсутствие практики заключения сложных комплексных договоров;

- предпринимательский риск, обусловленный увеличением числа договорных связей.

Юридический аутсорсинг обеспечивает эффективную работу бизнеса таким важным фактором как регулярное юридическое сопровождение в режиме реального времени и в соответствии с последними изменениями действующего законодательства. Причем в его рамках могут предлагаться консультационные, правовые и юридические услуги, представительство в интересах клиента в различных организациях, судах, участие в ведении переговоров и т. д.

2.5.2. *Аутсорсинг в сфере управления персоналом* - передача аутсорсинговой компании решения задач, связанных с управлением персоналом – наймом сотрудников, расчетом компенсаций и налоговой отчетностью. Любая компания вынуждена решать подобные задачи, но профильной подобная деятельность является лишь для кадровых агентств. Именно поэтому практика аутсорсинга управления персоналом приобрела значительную популярность. Чаще передаются процессы, связанные с подбором и поиском сотрудников (а также лизингом персонала и аутстаффингом), реже – процессы кадрового администрирования. В России практика аутсорсинга процессов управления персоналом распространена достаточно широко. Примером компаний, предоставляющих услуги по набору и найму персонала, являются многочисленные кадровые агентства. Значительно меньше имеется поставщиков комплексных услуг в сфере управления персоналом и кадрового администрирования.

2.5.3. Аутсорсинг издательской деятельности - передача стороннему исполнителю (издательству) функций перевода, печати и постпечатной обработки документов, издание корпоративных газет и журналов. Издательство имеет штат редакторов, корректоров, отдел верстки, внештатных авторов. Для наиболее ресурсоемких и дорогих направлений деятельности, таких как перевод материалов, печать изданий и их распространение, издательство может привлекать другие организации.

2.5.4. Аутсорсинг печатных работ является для российского рынка новым. По договору аутсорсинга организация либо передает специализированной фирме в управление офисную печатную технику с дальнейшей установкой дополнительного оборудования уже аутсорсинговой компанией - исполнителем работ, либо заказчик отказывается от приобретения собственной техники и все функции по обеспечению офисной печатной техникой, бумагой, расходными материалами, а также сервисным обслуживанием ложатся на аутсорсинговую фирму. Таким образом, заказчик освобождается от затрат на приобретение техники, снижает текущие эксплуатационные затраты на расходные материалы, бумагу, сервис, получая возможность заниматься непосредственно своим бизнесом, не отвлекаясь на офисную поддержку.

2.5.5. Архивный аутсорсинг - сравнительно новый вид аутсорсинговых услуг, заключающийся в передаче для внеофисного хранения архивных документов специализированной организации. Как известно, в любых организациях создается собственный архив. Для этого необходимо специальное помещение, а также наличие в штате должности архивариуса. Для организации правильного хранения документов необходимы еще обязательные системы пожарной и охранной сигнализации, поддержание постоянной температуры и влажности в помещении архива а также дополнительные расходы на архивные стеллажи, компьютеры, рабочие места для архивистов и др. Профессиональная фирма имеет специализированное архивное хранилище, штат профессиональных архивистов, опыт хранения и работы с документами организаций и предпринимателей различного профиля деятельности. Архивный аутсорсинг выгоден практически для любых предпринимательских структур, так как они получают соответствующий и дистанционно-управляемый во многих случаях архив, в котором документы безопасно хранятся в установленном порядке, до истечения сроков их хранения. При этом документы защищены от несанкционированного доступа к ним неуполномоченных работников организации и третьих лиц. При организации хранения, как правило, используются автоматизированные процессы, компьютеризация. По договору архивного аутсорсинга хранитель несет полную ответственность за утрату или порчу документов, взятых на хранение. Подобные аутсорсинговые компании работают не только в Москве и Санкт-Петербурге, но и во многих других городах России.

2.5.6. Экологический аутсорсинг также появился на российском рынке сравнительно недавно. В соответствии с договором об аутсорсинге профессиональные компании в соответствии с федеральным законодательством производят коммерческое экологическое сопровождение деятельности предприятий, особенно занимающихся производственной деятельностью. Эти компании берут на себя и обязанности по защите интересов заказчика со стороны контролирующих органов, разработку необходимой организации экологической документации, и т.д.

Таким образом, по мере развития корпораций в России, появляются все новые виды аутсорсинга. Хотя применение на Западе аутсорсинга носит массовый характер, для российских компаний существует ряд проблемных факторов системного характера, объективно препятствующих развитию рынка аутсорсинга в России, основными из которых являются низкая правовая культура предпринимательской деятельности, недостаток информации и отсутствие достаточного количества квалифицированных специалистов, отсутствие страхования рисков деятельности, проводимой на условиях аутсорсинга, отсутствие индустриальных стандартов.

Заключение

Изменившиеся экономические условия в России, а также глобальная тенденция сокращения издержек стимулировали новый всплеск интереса к выведению части функций предприятия вовне – аутсорсингу. К настоящему моменту в России для этого сформировалось достаточно возможностей как на технологическом, так и на предпринимательском уровне. Многие российские компании уже прибегли к этой практике. Как и в любом деле, в выделении на аутсорсинг процессов компании есть как положительные, так и отрицательные моменты.

Опыт удачного применения аутсорсинга показывает, что детально описанные и проанализированные бизнес-процессы организации позволяют оптимально распределять задачи между внутренними и внешними исполнителями, грамотно принимать решения относительно вынесения той или иной функции на аутсорсинг.

Но для развития отечественного аутсорсинга существуют определенные, свойственные только России, проблемы. В основном это связано с наличием ряда факторов системного характера, препятствующих его развитию, такие как проблемы сохранения коммерческой тайны, отсутствие индустриальных стандартов, низкая правовая культура предпринимательской деятельности, недостаток информации и отсутствие достаточного количества квалифицированных специалистов, отсутствие страхования рисков деятельности и т.д. Но основное препятствие для использования аутсорсинговых услуг в России кроется в специфике российского менталитета – отсутствии доверия к партнерам по бизнесу и желании сконцентрировать всю деятельность в единой компании. Но менталитет российского бизнесмена постепенно меняется и уже сегодня растет число компаний, оценивших преимущества аутсорсинга и предполагающих в своей деятельности его активное использование.

Сегодня аутсорсинг является базовым трендом развития мировой экономики в условиях глобализации, потому что современный технологический уровень развития средств коммуникации позволяет использовать наиболее подходящие и наиболее качественные ресурсы, вне зависимости от того, где они расположены. Современные тенденции развития аутсорсинга свидетельствуют о его высокой роли как инструмента современного межфирменного соперничества. Они создали благоприятные предпосылки для национальной индустрии, чтобы начать продвижение России и ее компаний в качестве надежных партнеров на мировом рынке аутсорсинга, в первую очередь, США и Западной Европы – ключевых для российского аутсорсинга регионах. Максимально оправданное и своевременное использование всех возможностей аутсорсинга, несомненно, будет способствовать совершенствованию российской инновационно-технологической среды.

Литература

1. Андрюшкевич О.А., Денисова И.М. Аутсорсинг как новая форма управления бизнесом и современная модель предпринимательства. //Анализ и моделирование экономических процессов, вып.5. ЦЭМИ РАН, 2008 г.
2. Офшорный аутсорсинг. http://ru.wikipedia.org/wiki/%D0%9E%D1%84%D1%84%D1%88%D0%BE%D1%80%D0.
3. Публикации . SaaS как новое направление ИТ-аутсорсинга 08.10. 2009. http://www.outsourcing.ru/content/rus/291/2915-article.asp.
4. IT-аутсорсинг в мире и в России. http://www.itspecial.ru/post/10086/.
5. Приглашение к аутсорсингу. http://www.iks-media.ru/articles/23469.html.
6. Благодаря индийским программистам через 5 лет Россия станет лидером ИТ-аутсосринга. 18.08.05. http://www.klerk.ru/soft/articles/29659/.
7. Аутсорсинг программных услуг в России. .№4 (22),22 декабря 2003 u/http://www.crn.ru/numbers/spec-umbers/detail_print.php?ID=12043&print=Y.
8. Астерос исследовал российский рынокИТ-аутсорсинга. http://ict-online.ru/news/n73901/.
9. Шесть российских ИТ- аутсорсеров вошли в мировой ТОП-100 http://www.cnews.ru/news/top/index.shtml?2008/03/26/293788.
10. Ассоциация разработчиков программного обеспечения РУССОФТ. Как начать свое дело. http://start.fbd.spb.ru/09.prochie-otraslevye-obedineniya/.
11. Информация о компании. http://www.physcareer.ru/company/42.html.
12. Развитие промышленного аутсорсинга в России, - где крупный бизнес может заработать на малом. http://www.subcontract.ru/Docum/DocumShow-DocumID-10.html
13. Аутсорсинг в логистике. Статьи.16.04.2007. http://www.custom.ru/articles_6.htm.
14. Аутсорсинг – что это такое? http://www.lobanov-logost.ru/index.php?newsid=4306] 15/03/ 2011.
15. Будущее аутсорсинга бизнес-процессов. http://future.erpnews.ru/doc14.htm.

Н.А. Трофимова, В.А. Разумовская

Гравитационная модель трудовой миграции

В статье рассматриваются основные проблемы, связанные с прогнозированием трудовой миграции в России, исследуются основные направления моделирования процессов трудовой миграции, обосновывается возможность применения класса гравитационных моделей, предлагается модифицированная гравитационная модель, а также приведены результаты прогнозных расчетов для миграционных потоков из стран СНГ в ЦФО России.

1. Проблема трудовой миграции

Одной из характерных черт развития современной мировой экономики является миграция населения, и в первую очередь, трудовая миграция -- перемещение людей, связанное с изменением места жительства и работы. Миграция играет важную роль в современном мире, в миграционный процесс вовлечены миллионы человек, миграция формирует новый рынок труда и новые финансовые и социальные отношения.

Миграция населения и трудовых ресурсов возникает при наличии значительного контраста в уровнях экономического и социального развития и темпах естественного демографического прироста стран, принимающих и отдающих рабочую силу. Главной производительной силой в экономике являются трудовые ресурсы. С учетом старения кадров, возрастания смертности, уменьшения численности населения возникает вопрос привлечения трудовых ресурсов из других стран, т.е. мигрантов.

В 2010 г. Россия заняла второе место по числу прибывших мигрантов, после Америки. (Публикация статистического доклада на 2011 г. Всемирного банка (ВБ), посвященный трудовой миграции, денежным трансграничным переводам мигрантов и влиянию этих процессов на национальные экономики). Данный процесс затрагивает политические, социальные и, главным образом, экономические аспекты. В период кризиса в связи с ухудшением финансового состояния предприятий, снижением заработной платы, работодателям потребовалась дешевая рабочая сила. И тут как нельзя, кстати, пришелся большой поток трудовых мигрантов из стран СНГ и ближнего зарубежья.

В данной работе будет рассматриваться трудовая миграция в Россию из стран бывшего СССР. Отслеживание и прогнозирование количества трудовых мигрантов необходимо, чтобы избежать увеличения безработицы для трудоспособного населения регионов России.

Целью исследования является разработка модели, позволяющей прогнозировать количество мигрантов, которые могут прибыть в поисках работы из стран СНГ в ту или иную область России. Практическая значимость работы состоит в том, что результаты моделирования процесса трудовой миграции дают возможность региональным властям и миграционным службам проводить мониторинг трудовых ресурсов, прибывших из других стран и эффективнее управлять трудовыми потоками.

2. Основные направления моделирования процессов трудовой миграции

Основной задачей моделирования процессов трудовой миграции является прогноз притока и выбытия трудовых ресурсов для более эффективного управления социально-экономическим развитием страны или региона. Неслучайно разработка моделей этого направления началась еще в 20-х гг. прошлого столетия. Рассмотрим основные направления моделирования процессов трудовой миграции.

2.1. ***Экстраполяционные методы***. Модели, основанные на применении экстраполяционных методов наиболее распространены при прогнозировании трудовой миграции. В этих моделях используются линейные и экспоненциальные функции, при этом делается предположение о неизменности абсолютных среднегодовых приростов или о среднегодовых темпах роста численности мигрантов за определенный период. Такое предположение может быть относительно верным только для очень кратких периодов времени (не более 5 лет), поэтому в долгосрочных прогнозных расчетах эти модели дают не всегда корректные результаты.

2.2. ***Имитационные вероятностные модели.*** Такие модели базируются на методе Монте-Карло, согласно которому для каждого потенциального субъекта миграции вероятность совершения/несовершения акта миграции определяется с помощью датчика случайных чисел. Основным препятствием для широкого применения таких моделей служит отсутствие необходимой информации, а также громоздкость и трудоемкость вычислений.

2.3. ***Формирование миграционных потоков.*** Существует несколько подходов к формированию миграционных потоков.

Первый подход рассматривает поток как перемещение населения между исследуемым районом и всей окружающей территорией. При этом для прогнозирования миграции используются марковские цепи, задаваемые распределением переходных вероятностей. В этих моделях учитываются коэффициенты интенсивностей миграции всех выделенных групп населения, время адаптации в новом районе, а также миграционная подвижность коренного населения.

Второй подход рассматривает маятниковые потоки населения, интенсивность которых зависит от количества мигрантов на границе области зарождения миграционного потока, и сил сопротивления, препятствующих оседанию мигрантов в населенном пункте. Маятниковая миграция — условное название регулярных (обычно ежедневных) поездок населения из одного населенного пункта (места жительства) в другой — на работу или учебу и обратно. Маятниковая миграция является результатом несоответствия размещения производства и расселения людей; это одна из частей миграции, не отражающая полную картину.

Третий подход – это матричные модели миграционных потоков. Обычно их рассматривают, как частный вариант социально- демографического баланса, в котором не учитываются родившиеся и умершие.

2.4. ***Регрессионные модели.*** Используются два вида регрессионных моделей: аддитивные и мультипликативные. Для прогноза интенсивностей миграционных потоков чаще всего используются мультипликативные модели, а для прогноза величины миграционных потоков – аддитивные.

2.5. ***Факторные модели***. В этих моделях миграционные потоки между двумя регионами увязываются с социальными, экономическими и демографическими факторами, инициирующими акт миграции. К этому классу относят *гравитационные модели*, в которых величина потоков между двумя районами прямо пропорциональна расстоянию между ними, взятому в некоторой степени. Преимущество этих моделей в том, что они основаны на реальных статистических данных и отражают притягивающие и отталкивающие факторы миграции; к числу последних часто относят расстояние между регионами.

Классическая гравитационная модель имеет следующий вид:

$$Mij = \frac{Pi * Pj}{Dij^2} * A, \tag{1}$$

где

Mij -- прогнозируемое число мигрантов из региона i в регион j
Pi -- численность населения региона i
Pj -- численность населения региона j
Dij -- расстояние между регионами i и j
A -- коэффициент соответствия.

В настоящей работе использована гравитационная модель, в которой отталкивающий фактор – расстояние между регионами -- модифицирован и, кроме того, вместо коэффициента соответствия введен показатель привлекательности.

3. Модифицированная гравитационная модель

Модифицируя формулу (1) классической гравитационной модели, примем следующие посылки.

1. Регионами i (источниками миграции) выступают страны СНГ. Вместо показателя численности Pi используется не все население , а только количество Bi ***потенциальных мигрантов,*** каковыми считаются люди с уровнем дохода не большим прожиточного минимума.

2. Регионами j (стоками миграционных потоков) выступают области России. При расчете показателя численности населения Pj берется только активное население ; это связано с тем, что целью работы является анализ и прогноз трудовой миграции.

3. Модель (1) не учитывает ряд важных факторов, влияющих на миграционные потоки. Так, при современных информационных технологиях, облегчающих поиск работы, и современных способах перемещения, расстояние в чистом виде уже не является значимым фактором. Чаще всего люди планируют переезд, исходя из стоимости переезда, привлекательности региона, в который переезжает мигрант и ряда других факторов. Поэтому в нашей модели вместо расстояния Dij между страной СНГ и областью России используется показатель стоимости проезда Cij , причем цена билета рассчитывается от столицы страны-источника до областного центра.

4. Вместо коэффициента соответствия А в формулу (1) введены *коэффициенты привлекательности* миграции, рассчитываемые по формуле:

$$K_{ij} = \frac{D_j}{D_i} * F_j , \qquad (2)$$

Где

Di -- заработная плата в стране-источнике i

Dj -- заработная плата в области-стоке j

Fj – доля области в общей квоте ЦФО.

Первоначально было сделано предположение, что на перемещение мигрантов влияет соотношение валового внутреннего продукта (ВВП) страны СНГ и ВВП областью России, Но проведенные расчеты не подтвердили данное предположение. В результате был сделан вывод, что мотивами людей, которые едут в поисках работы в другие регионы и страны являются в основном разница в заработных платах и в условиях труда. Мигранты едут в места с более высокими заработками и возможностью трудоустроиться; это и отражено в формуле (2).

Таким образом, модифицированная гравитационная модель трудовой миграции будет иметь следующий вид:

$$M_{ij} = \frac{A_j * B_i}{C_{ij}^{\;2}} * K_{ij} \qquad (3)$$

где

Mij -- потенциальное число мигрантов

Aj -- численность активного населения области России

Bi -- потенциальное количество мигрантов в стране-источнике

C_{ij} -- стоимость проезда от источника до стока

Kij -- коэффициент миграционной привлекательности.

С помощью модифицированной гравитационной модели был построен прогноз на 2010г.

4. Экспериментальные расчеты и анализ полученных результатов

Модифицированная гравитационная модель предназначена для краткосрочного прогнозирования (именно -- на один год) притока мигрантов из стран СНГ в регионы России,. Расчеты проводились для десяти стран СНГ: Азербайджан, Армения, Беларусь, Казахстан, Кыргызская Республика, Молдова, Таджикистан, Туркменистан, Узбекистан и Украина.

Разработанная модель проверялась для всех регионов России, но анализ полученных результатов показал, что данная модель может применяться только для регионов ЦФО РФ. Для других округов Российской Федерации необходимо вводить дополнительные переменные, отражающие особенности миграционных потоков, например, климатические условия, наличие национальных диаспор, возможность добираться домой, затратив меньшее количество времени и т.д.

Данную модель также необходимо усовершенствовать при расчете прогнозов для Украины, т.к. в этом случае наблюдается высокий уровень скрытой миграции из Украины в Россию, которая объясняется невысокой стоимостью проезда, знанием языка и большого числа украинских граждан, проживающих постоянно на территории РФ. На Украине также велика численность людей с уровнем дохода не большим прожиточного минимума (именно они рассматриваются в нашей модели как потенциальные мигранты).

Для расчета модифицированной гравитационной модели был использован коэффициент миграционной привлекательности, который вычислялся на основе произведения соотношения заработных плат в регионе, куда направляется миграционный поток, и откуда он движется на долю квоты данного региона в общей квоте округа.

При прогнозировании численности мигрантов для Московской области коэффициент миграционной привлекательности рассчитывался не по формуле (2), а по формуле

$$K_{ij} = \frac{D_j - D_I}{D_i} * F_j \qquad (4)$$

Это связано с особенностями социально-экономического развития этого региона. Так, этот регион имеет самую большую долю квоты в общей квоте ЦФО и самую высокую заработную плату, а также наиболее удобные условия проезда до места назначения.

4.1. ***Анализ исходных данных.*** Расчеты проводились на основе опубликованных статистических данных 2009 г. Для расчетов использовались официальные данные использовались официальные опубликованные данные Межгосударственного статистического комитета Содружества независимых государств: численность населения регионов России, численность населения стран СНГ с уровнем дохода равным или меньше прожиточного минимума, средняя заработная плата в государствах СНГ и России. Там же были взяты данные на прогнозный 2010 г.: стоимость проезда и размер квот.

Для стран СНГ для расчета была выбрана среднемесячная номинальная заработная плата, которая переводилась в рубли по курсу на декабрь 2009 г. Эти данные представлены на Рис. 1.

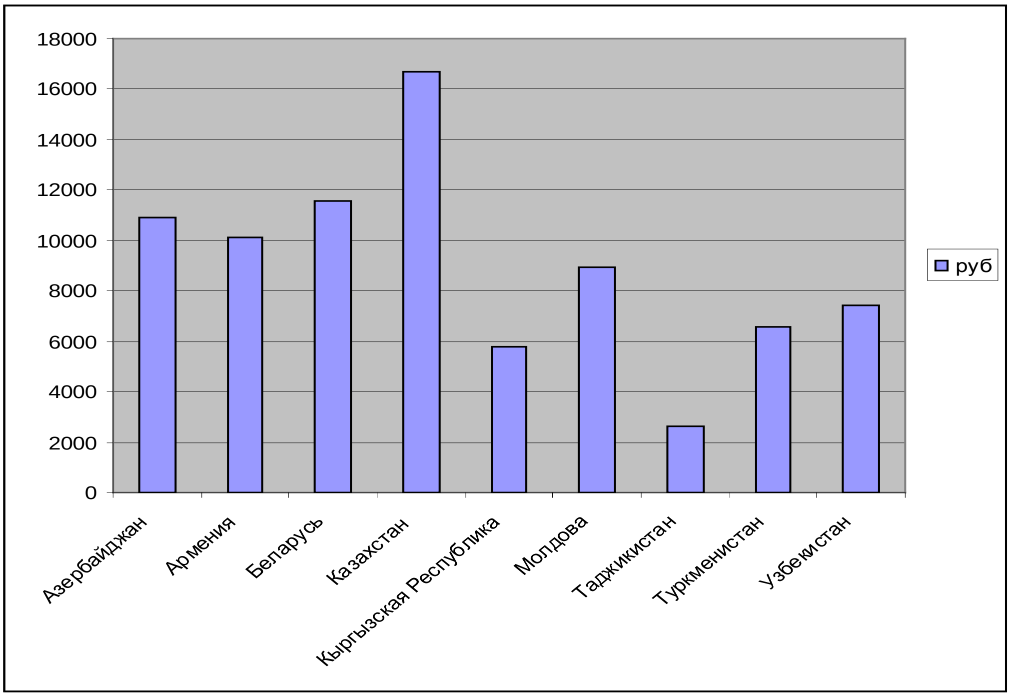

Рис. 1. Средняя заработная плата в странах СНГ за 2009 г.

Как видно из рисунка, самая низкая средняя заработная плата наблюдается в Таджикистане, Кыргызской Республике, Туркменистане и Узбекистане. Можно предположить, что самый большой приток мигрантов в Россию будет именно из этих стран. Самые высокие зарплаты зафиксированы в Беларуси и Казахстане, поэтому поток из этих стран будет предположительно ниже, чем из других.

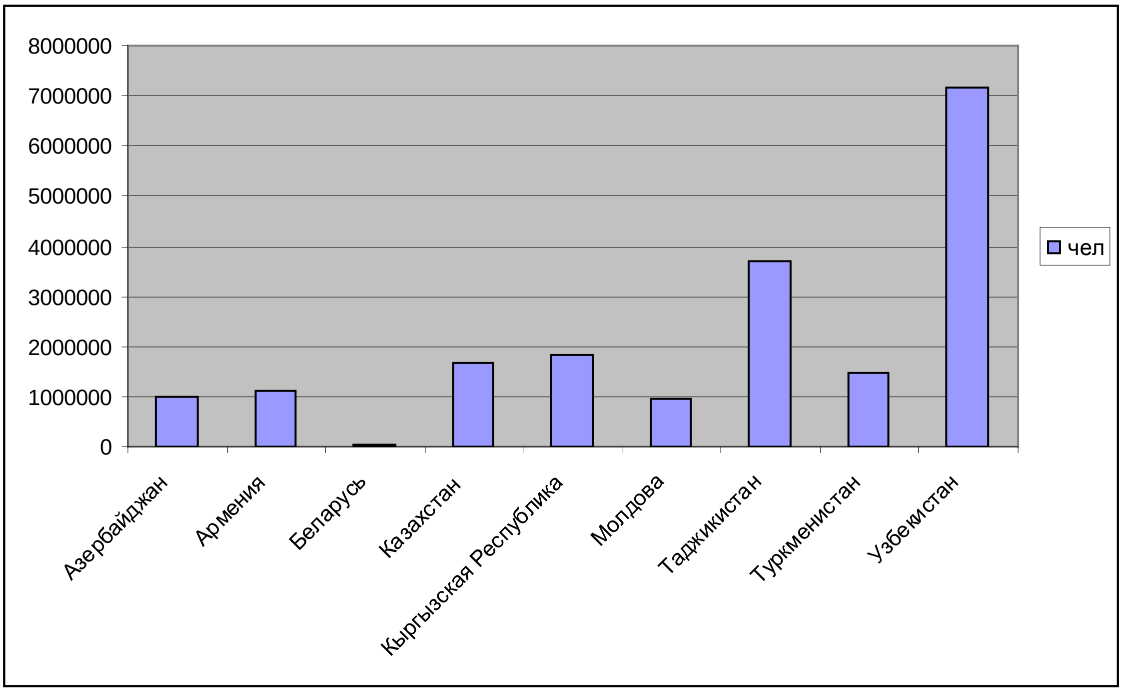

Рис. 2. Численность потенциальных мигрантов из стран СНГ.

Информация о численности людей с уровнем дохода не большим прожиточного минимума (т.е. потенциальных мигрантов), была найдена на различных сайтах; результаты представлены на Рис. 2. Из рисунка следует, что самое большое число потенциальных мигрантов наблюдается в Узбекистане, Таджикистане и Кыргызской Республике, тогда как в Беларуси, Молдове и Азербайджане число людей с низким уровнем дохода значительно ниже.

Для расчета стоимости проезда был использован интернет ресурс, в котором приводится совокупная стоимость проезда от страны СНГ до региона России.

Источником данных о региональном распределении квот на иностранную рабочую силу на 2010 г. стала публикация газеты «КоммерсантЪ» на сайте www.retailer.ru. Квота на осуществление трудовой деятельности для иностранцев была введена в 2002 г. федеральным законом «О правовом положении иностранных граждан». Общее количество квот представляет собой приглашения на работу для приезжих из визовых стран, и квоты для выходцев из СНГ. Для работников из стран СНГ порядок получения квот отличается от приезжих из визовых стран. Количество квот ежегодно формируется на основе запроса исполнительных органов регионов РФ. Потребность региона в иностранной силе утверждает специальная межведомственная комиссия. Квоты публикуются для всех граждан и «визовых» и из стран СНГ, но вопрос о том, сколько граждан приедет, конкретно из каждой страны остается открыт. Именно это положение делает

Таблица 1

Плановое географическое распределение квот в 2010 г.

Доля квот	Выделено квот (тыс. чел)
Области ЦФО	480.0
Белгородская	7.2
Брянская	3.1
Владимирская	12.2
Воронежская	9.8
Ивановская	0.7
Калужская	29.4
Костромская	3.9
Курская	1.6
Липецкая	6.6
Московская	335.6
Орловская	2.2
Рязанская	17.4
Смоленская	6.0
Тамбовская	3.5
Тверская	13.3
Тульская	19.7
Ярославская	7.6

прогноз количества мигрантов актуальным. Плановое географическое распределение квот в 2010 г. представлено в табл. 1.

Как следует из таблицы, самая большая потребность в иностранной рабочей силе в Московской, Калужской, Тульской, Рязанской, Тверской областях. Предположительно именно в эти регионы и будут направлены потоки мигрантов, т.к. именно там есть большая возможность официально трудоустроиться

Источником данных о средней заработной плате и численности населения регионов России стал опубликованный на сайте Федеральной службы государственной статистики ежегодник «Регионы России. Социально-экономические показатели за 2010 г.». Из ежегодника были взяты данные о номинальной заработной плате и численности населения за 2009 г.

На Рис. 3 изображено распределение средней номинальной заработной платы (руб./месяц) по областям ЦФО. Именно этот фактор является главным притягивающим для мигрантов, ищущих работу.

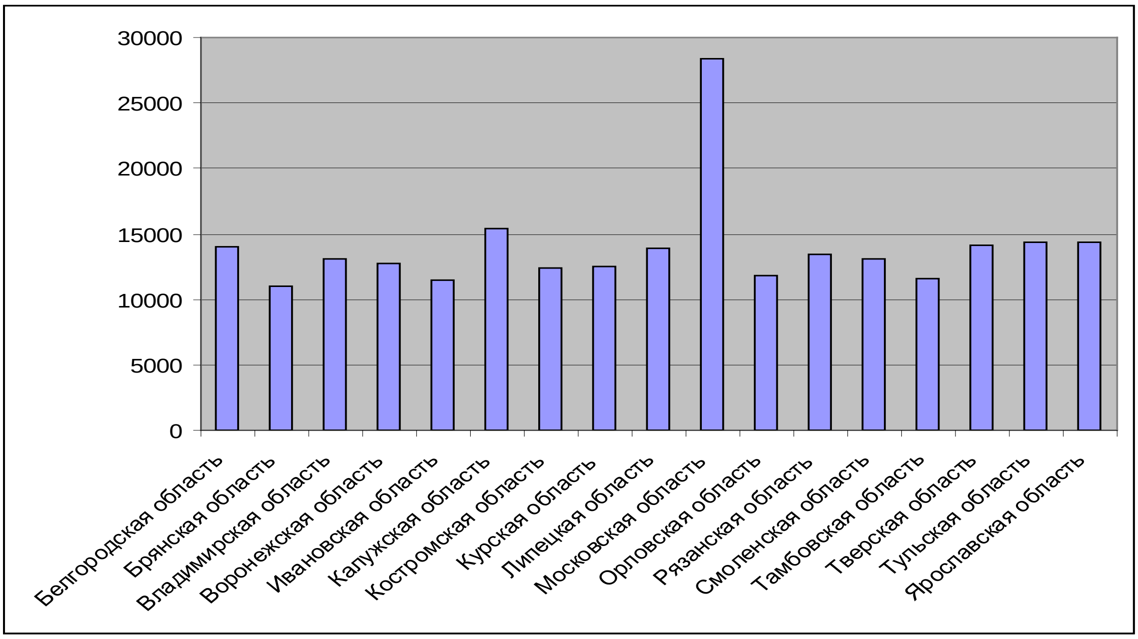

Рис. 3. Средняя номинальная заработная плата в областях ЦФО.

На основе анализа Рис. 3 можно сделать вывод, что самые крупные потоки миграции будут направлены в Московскую, Калужскую, Тульскую, Ярославскую, Белгородскую и Тверскую области. Меньшее число иностранных рабочих будет прибывать в Брянскую, Ивановскую, Орловскую и Тамбовскую области.

4.2. ***Результаты расчетов.*** В результате расчетов с помощью формул (2) – (4), было получено прогнозируемое число мигрантов из каждой рассматриваемой страны СНГ, которые могут прибыть в конкретную область ЦФО. В качестве примера в табл. 2 приведены результаты расчетов для Московской и Тульской областей.

Таблица 2

Количество прогнозируемых мигрантов

в Московскую и Тульскую области (тыс. чел.)

	Московская область	**Тульская область**
Население	17222	1553
Зарплата (тыс. руб.)	28.3	14.3
Доля региональной квоты в общей квоте ЦФО (%)	70	4
Азербайджан	252.1	1.7
Армения	142.3	0.6
Беларусь	41.1	0.2
Казахстан	213.7	1.6
Кыргызская Республика	337.0	2.6
Молдова	409.8	4.0
Таджикистан	295.8	8.1
Туркменистан	50.4	1.7
Узбекистан	297.7	8.0

Сводные результаты расчетов распределения количества прогнозируемых мигрантов по областям России представлены в табл. 3.

Таблица 3

Распределение количества прогнозируемых мигрантов

по областям ЦФО России (тыс. чел.)

Области ЦФО	Количество мигрантов
Белгородская область	9.2
Брянская область	4.1
Владимирская область	13.1
Воронежская область	19.2
Ивановская область	0.7
Калужская область	38.9
Костромская область	2.4
Курская область	1.4
Липецкая область	6.1
Московская область	2039.9
Орловская область	1.6
Рязанская область	16.6
Смоленская область	6.2
Тамбовская область	2.9
Тверская область	13.4
Тульская область	28.6
Ярославская область	12.2

Видно, что Московская область занимает особое положение (это понятно); численность мигрантов в остальные области можно интерпретировать как характеристику их привлекательности, Данные табл. 3, построенные в виде диаграммы показаны на Рис. 4. Как видно из рисунка наиболее привлекательными областями для мигрантов из стран СНГ, помимо Московской, являются Калужская, Тульская, Воронежская и Рязанская области. Это объясняется высоким уровнем заработной платы и потребностью региона в иностранной рабочей силе. Но несомненным лидером является Московская область. На нее приходится самый большой поток миграции. Это связано не только с потребностью в рабочей силе, но и с желанием работодателей сэкономить, т.к. рабочие из стран СНГ, готовые работать на любых условиях становятся наиболее дешевым вариантом заполнения свободных мест.

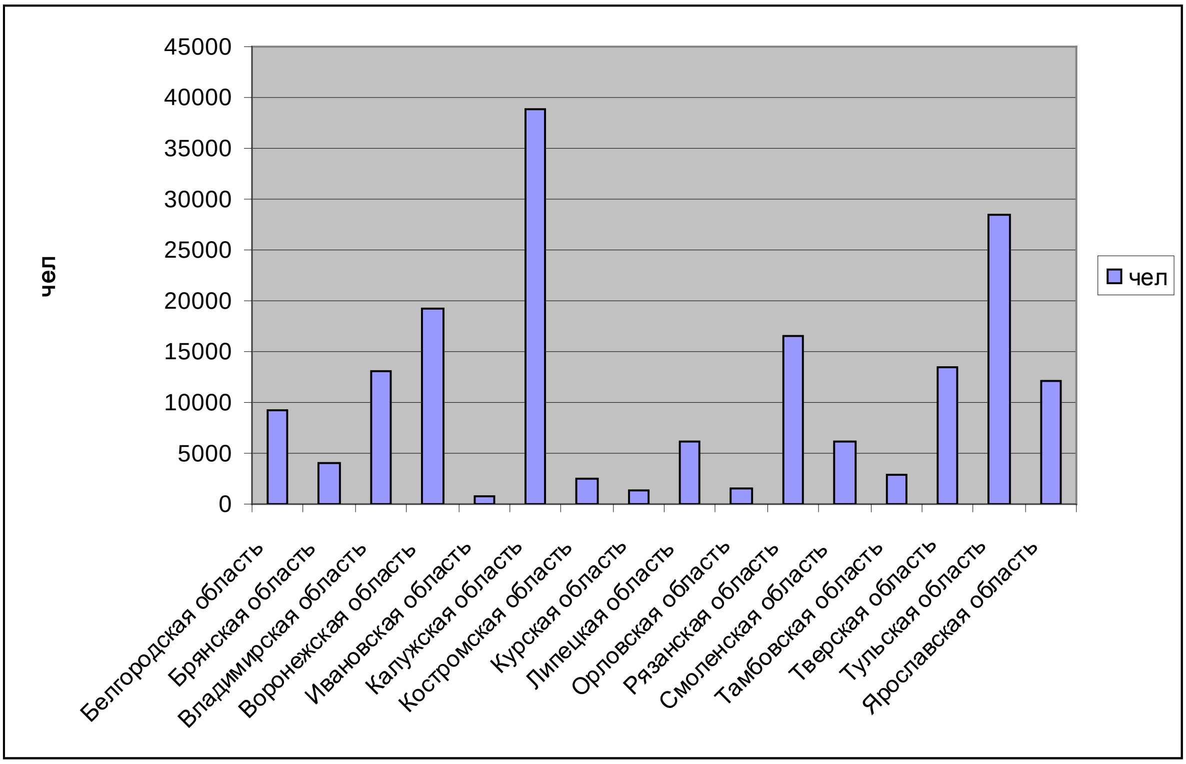

Рис. 4. Диаграмма привлекательности регионов ЦФО.

На Рис. 5 показано общее прогнозируемое количество мигрантов в ЦФО. Из рисунка следует, что самый большой поток мигрантов в ЦФО идет из Молдовы. Далее идут Узбекистан, Таджикистан и Кыргызская Республика. За ними следуют Азербайджан, Армения, Туркменистан и Беларусь.

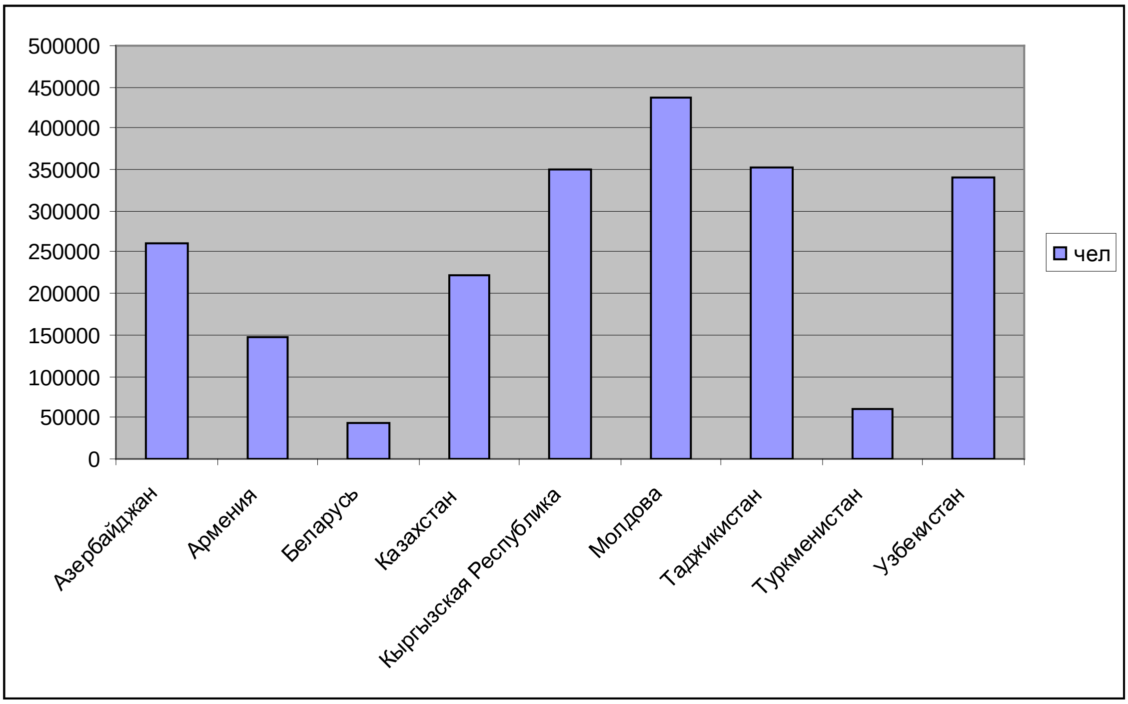

Рис. 5. Количество мигрантов в ЦФО из стран СНГ .

В основном миграционный поток из Молдовы приходится на Московскую область. Как и в случае с Украиной это явление связано с наличием скрытой миграции. Знание языка помогает быстрее адаптироваться. В отличие от мигрантов из других стран СНГ, молдаване заняты не только на стройках, рынках и в сфере оказании услуг населению. Они, также как и украинские граждане, работают в различных фирмах и организациях, занимаясь офисной работой.

Большое число мигрантов прибывает из Таджикистана, Киргизии и Узбекистана. В большинстве случаев граждане этих стран не владеют русским языком и резко выделяются на фоне жителей российских регионов. Но такой большой поток мигрантов именно в Россию объясняется давно сложившейся системой поставки рабочей силы из этих стран на территорию нашего государства.

Относительно небольшой приток мигрантов из Беларуси объясняется достаточно малой численностью в этой стране людей с уровнем дохода не большим прожиточного минимума. Поэтому в основном населению не имеет смысла искать работу в России.

Проверка полученных результатов проводилась экспертным путем на основе анализа публикаций в СМИ, на основе исследования региональных сайтов и на основе косвенных статистических данных. Это связано с тем фактом, что нет источников, дающих полную достоверную картину миграционной ситуации. Комбинировать или сравнивать имеющиеся данные крайне сложно, т.к. методики сбора первичной информации и, тем более, ее обработки, не согласованы. В рамках отдельных систем определяется либо число мигрантов какой-то определенной категории (например, лиц, ищущих убежища), либо фиксируется число событий, а не людей (как это делает пограничная статистика в отношении фактов пересечения границы). В настоящее время ведется работа по объединению всех источников информации об иностранных мигрантах в рамках Центрального банка данных иностранцев. К сожалению, большинство административных источников закрыто для общественности, даже в тех случаях, когда для этого нет объективных оснований. Многие мигранты, въезжая на территорию России, вообще не попадают в поле статистического наблюдения. Несмотря на проблему с получением исходных данных, ошибка прогнозных расчетов не превысила 8%. Это говорит о том, что разработанная модифицированная гравитационная модель прогноза трудовой миграции может успешно использоваться для прогнозирования миграционных процессов.

Заключение

Миграционный поток из стран СНГ в Россию довольно велик. И в последние годы он стал практически не регулируемый. При составлении квот региональные власти опираются на общую цифру требуемых мигрантов, не обращая внимания из каких стран они едут и в каком количестве. Есть страны, такие как Таджикистан и Узбекистан, жители которых не знают ни русского языка, ни законов страны, в которую въезжают, чем и пользуются недобросовестные работодатели.

С помощью модифицированной гравитационной модели были сделаны расчеты, результатами которых стали цифры, отражающие потенциальное число мигрантов, которые могут въехать на территорию регионов России в поисках работы. Прогноз по этой модели получается краткосрочный на 1 год и отражает реальную ситуацию.

Данная модель работает при определенных ограничениях. Она может быть использована только для Центрального федерального округа, так как для других регионов недостаточно тех факторов, которые использовались при расчетах. Для проведения расчетов по другим регионам надо ввести дополнительные переменные, которые будут отражать влияние климатических условий регионов России, наличие диаспор, возможности добираться домой, затратив меньшее количество времени и т.д.

Модифицированная гравитационная модель (2) – (4) не может быть применена для расчетов потока мигрантов из Украины из-за слишком большого уровня скрытой миграции. Знание языка и большое количество украинских граждан проживающих на территории России постоянно, помогают мигрантам быстрее адаптироваться, а невысокая стоимость проезда дает возможность часто совершать поездки домой. В случае с Украиной речь скорее идет о маятниковой миграции.

Главной особенностью модифицированной гравитационной модели является изменение акцентов в анализе влияния отталкивающих и притягивающих факторов. Так главным отталкивающим фактором является стоимость проезда, а не просто расстояние. Это связано с изменением характера транспортных технологий и с увеличением роли стоимостных расходов. Главным притягивающим фактором в этой модели стал коэффициент миграционной привлекательности, который отражает индивидуальность каждого региона России. Это позволяет оценить численность потенциальных мигрантов для каждой области России. Данную оценку можно будет использовать для составления страновых квот для любого региона и проведения более эффективной миграционной политики.

При определенных условиях модифицированную гравитационную модель можно использовать для расчета числа нелегальных мигрантов в регионах России.

М.А.Гогулин, В.М.Четвериков

Модель ценовой конкуренции на рынке однородного товара

Известно, что классическая модель ценовой конкуренции Бертрана [1] при гомогенной олигополии приводит к тому, что фирмы (продавцы) вынуждены опустить цены до уровня предельных издержек (если они одинаковы). Другими словами, в модели заложено предпочтение всех покупателей приобретать товар только по минимальной цене.

В модели, предлагаемой в данной статье, предпочтения покупателей не столь жестко привязаны к минимальной цене, как в модели Бертрана, а общий объем рынка определяется максимальной ценой предложения. Мы ограничились подробным рассмотрением конкуренции двух фирм (дуополии), хотя основные формулы написаны для общего случая конкуренции n фирм.

Если объем рынка в целом не изменяется при снижении конкурирующими фирмами цен, исходя из желания получить большую долю фиксированного рынка, то модель приводит к существованию точки ценового равновесия по Нэшу.

Однако если объем рынка растет при уменьшении максимальной цены предложения, то, начиная с определенной «скорости» этого роста, вместо равновесной цены мы получаем непрерывное множество равновесных цен. При определенном соотношении параметров модели это множество может содержать даже квазимонопольную цену, т.е. цену, которая является выгодной для всех фирм в целом при единой ценовой политике.

Несмотря на то, что при исследовании предлагаемой модели находятся цены, равновесные по Нэшу, модель не является моделью теории игр, поскольку в ней не заложен механизм игры – правил, по которым участники игры могут принимать решения. Наш подход имеет простую аналогию с щироко используемым во многих работах поиском таких конфигураций взаимодействующих физических подсистем, которые имеют минимальную энергию. Однако, в отличие от этих работ, в которых механизм и скорость достижения этих конфигураций не рассматривается, мы описываем и динамику процесса.

Принято считать [2], что цены влияют на объемы продаж сильнее, чем все остальные инструменты маркетинга. Однако ценовая конкуренция может дать положительный выигрыш лишь при определенных условиях.

Рассмотрим n фирм, продающих на рынке один и тот же товар (или товары, схожие функционально, но несколько отличающиеся по своим характеристикам) по разным ценам. Предположим, что предельные издержки каждой i -ой фирмы постоянны и равны c_i, а объем продаж q_i определяется как ценой продаж $p-d_i$ i-ой фирмы, так и ценами всех конкурентов $p-d_k$, $k \neq i$. Используемая запись цены в виде разности двух положительных величин позволяет интерпретировать цену как разность некоторой базовой цены p $(p > c_i \quad \forall i)$, одинаковой для всех фирм и торговых скидок $d_i \geq 0$, которую фирмы определяют из конкурентных соображений.

Задачей фирмы является определение такой торговой скидки, которая приводит к максимуму прибыли

$$\pi_i = (p - c_i - d_i) \cdot q_i.$$

1. **Динамическая модель установления рынка**

В этом разделе приводится модель рынка, из которой будут получены объемы продаж q_i в условиях ценовой конкуренции.

Динамика продаж q_i определяется теми «усилиями» $\mu_i \geq 0$, которые прилагают фирмы для привлечения покупателей. Это могут быть рекламные затраты, дисконтные программы, различные бонусы и другие мероприятия по продвижению товара. В конечном итоге именно дифференциация продукта в сознании покупателя определяет предпочтение покупателя одного продукта другому [3, 4].

Кроме того, у каждой фирмы есть (помимо «усилий» μ_i) и какая-то своя доля рынка κ_i (конечно $\kappa_i \geq 0,\ \sum_{i=1}^{n} \kappa_i = 1$), – определяемая симпатиями покупателей или удобным месторасположением торговых точек; эти доли считаются постоянными.

Для описания динамики рынка мы воспользуемся моделью конкурентного маркетинга [5,6], имеющей большое сходство с моделью Ланчестера (Lanchester) для описания военных действий [7,8].

$$\frac{d}{dt}q_i = \sum_{j=1}^{n} a_{ij} q_j + Q \cdot \nu \cdot \kappa_i, \tag{1.1}$$

где

$$a_{ij} = \begin{cases} -(\nu + \mu - \mu_i) & при \quad j = i, \\ \mu_i & при \quad j \neq i \end{cases} \tag{1.2}$$

$$\mu = \sum_{i=1}^{n} \mu_i, \qquad Q > 0, \quad \nu > 0. \tag{1.3}$$

Как будет видно далее, параметр Q (предполагаемый постоянным) – это объем рынка в предельном установившемся режиме, а параметр ν имеет смысл обратного времени установления равновесия в системе (1.1) при «включении» постоянных μ_i.

Матрица a_{ij} является матрицей простой структуры и имеет n линейно независимых собственных векторов u_k $(k = 1, 2, ..., n)$, несмотря на существование лишь двух собственных значений

$$\begin{aligned} &\lambda_1 = -\nu \quad кратности \quad 1, \\ &\lambda_2 = -(\mu + \nu) \quad кратности \quad n - 1. \end{aligned} \tag{1.4}$$

Собственный вектор u_1, соответствующий λ_1, имеет компоненты $u_{1i} = \mu_i$, а собственные векторы u_k (k=2,…, n), соответствующие кратному собственному значению λ_2, могут быть выбраны, например, в виде

$$u_{ki} = \begin{cases} 1 \;\; для \;\; i = k - 1 \\ -1 \;\; для \;\; i = n \\ \quad 0 \;\; для \; всех \; остальных \; i \end{cases} \qquad\qquad k = 2, 3, ..., n \quad .$$

Простота структуры матрицы a_{ij} обеспечивает возможность представления общего решения задачи Коши для системы (1.1) в виде

$$q_i(t) = \overline{q}_i + \sum_{k=2}^{n} A_k \cdot u_{ki} \cdot \exp(\lambda_2 \cdot t) + A_1 \cdot u_{1i} \cdot \exp(\lambda_1 \cdot t), \tag{1.5}$$

где константы A_k определяются начальными условиями $q_i(0) = q_i^0$, а стационарное решение:

$$\overline{q}_i = Q \cdot (\mu_i + \kappa_i \cdot \nu) \cdot (\mu + \nu)^{-1}. \tag{1.6}$$

Заметим, что, как и следовало ожидать из вида системы (1.1), суммарный объем продаж $Q(t)$ подчиняется уравнению

$$\frac{d}{dt}Q(t) = -\nu \cdot Q(t) + Q \cdot \nu, \quad Q(t) = \sum_{i=1}^{n} q_i(t) \quad . \tag{1.7}$$

Таким образом, система (1.1) описывает такое поведение потребителей, которое после «включения» постоянных «усилий» μ_i со стороны фирм приводит к стационарному распределению (1.6) объемов продаж по фирмам за характерное время $\approx \nu^{-1}$. Установление почти равновесного соотношения между объемами продаж фирм произойдет быстрее, поскольку согласно формулам (1.4)-(1.5), из двух экспонент быстрее затухает $\exp(\lambda_2 \cdot t)$ и асимптотика решения системы (1.1) на больших временах будет иметь вид:

$$q_i = \frac{Q}{\mu + \nu} \cdot \{\mu_i \cdot [1 + B \cdot \exp(-\nu \cdot t)] + \kappa_i \cdot \nu\}, \textit{ где } B = \frac{\mu + \nu}{\mu} \cdot \sum_{j=1}^{n} (q_j^0 - \overline{q}_j).$$

Чаще всего в подобных моделях характерные «усилия» фирм представляются в виде

$$\mu_i = \beta_i \cdot x_i^{\alpha_i}, \tag{1.8}$$

где x_i – финансовые затраты фирмы, на которые реагируют потребители, β_i и α_i – положительные эмпирические коэффициенты. При этом стационарное распределение объемов продаж (1.6) будет выпуклым вверх по x_i при $0 \le \alpha_i \le 1 \quad \forall i$ или иметь S-образную форму при $\alpha_i > 1 \quad \forall i$.

В данной работе мы будем считать $\alpha_i = 1 \quad \forall i$, а в качестве финансовых затрат, на которые реагируют потребители, будет фигурировать торговая скидка (дисконт) $x_i = d_i$. Кроме того, мы будем предполагать, что переходные процессы установления объемов продаж протекают гораздо быстрее возможных изменений дисконтов фирмами. Поэтому в качестве q_i будем брать стационарное распределение (1.6).

Наши вычисления существенно упростятся, если считать

$$\kappa_i \cdot \nu << \beta_i \cdot d_i \tag{1.9}$$

и пренебречь членами, содержащими k_i. Такое пренебрежение может привести к необходимости пересмотра стратегий фирм при малых дисконтах; этот случай в данной работе не рассматривается, поэтому мы в пространстве дисконтов d малую окрестность начала координат.

Общий объем продаж Q зависит не только от начальной цены p, но и от величин торговых скидок d_i. Т.к. мы уже предположили, что в распределении объемов продаж между фирмами отражается соотношение между скидками различных фирм, то на величину общего объема продаж должна действовать некоторая интегральная величина, характеризующие все скидки. Поскольку потребители не склонны производить серьезный количественный анализ цен на рынке (вряд ли им доступен анализ средневзвешенной по объемам продаж цены, о чем упоминается в Приложении 2), то такой интегральной величиной может служить, например, минимальная величина d_i, которая характеризует максимальную цену предложения среди конкурирующих фирм.

Будем предполагать, что уменьшение максимальной цены сопровождается увеличением общего объема продаж (подробнее см. Приложение 1). Другими словами, введем в нашу модель предположение о возможной связи дисконтов с общим объемом продаж следующего вида:

$$Q = Q_0 \cdot (1 + \gamma \cdot \min_j d_j), \qquad \gamma \geq 0, \qquad Q_0 > 0. \qquad (1.10)$$

Таким образом, общая модель ценовой конкуренции ссостоит в следующем: фирмы стремятся установить такие торговые скидки для своих покупателей, которые, с учетом аналогичных действий конкурентов, максимизируют получаемую прибыль

$$\pi_i = (p - c_i - d_i) \cdot q_i, \quad p > 0, \quad c_i > 0, \quad d_i > 0, \quad i = 1,2,...,n \qquad \text{(A)}$$

$$q_i = Q_0 \cdot (1 + \gamma \cdot \min_j d_j) \cdot (\beta_i \cdot d_i) \cdot (\sum_{j=1}^{n} \beta_j \cdot d_j)^{-1}, \gamma \geq 0, Q_0 > 0 \qquad \text{(B)}$$

2. **Конкуренция двух одинаковых фирм**

Рассмотрим вначале простой симметрический случай: две фирмы имеют одинаковые издержки $c_1 = c_2 = c$ и одинаковое отношение потребителей $\beta_1 = \beta_2$. В этом случае, согласно формулам (A) и (B) прибыль i-ой фирмы в зависимости от торговых скидок d_1, d_2 имеет вид

$$\pi_i = (p - c - d_i) \cdot d_i \cdot (d_1 + d_2)^{-1} \cdot Q_0 \cdot (1 + \gamma \cdot \min(d_1, d_2)), \quad i = 1, 2 \quad (2.1)$$

В плоскости (d_1, d_2) кривая f_1 наилучшего отклика первой фирмы на величину скидки второй определяется вне диагонали $d_1 = d_2$, где производная $\frac{\partial \pi_1}{\partial d_1}$ непрерывна, условием $\frac{\partial \pi_1}{\partial d_1} = 0$. В этих областях кривая описывается двумя отдельными непрерывными функциями,

В области $d_2 < d_1$ (ниже диагонали)

$$d_2 = d_1^2 \cdot (p - c - 2 \cdot d_1)^{-1} \tag{2.2}$$

при $d_1 \in (0, d_1^*]$, где

$$d_1^* = 3^{-1} \cdot (p - c), \tag{2.3}$$

что определяется пересечением кривой (2.2) с диагональю.

В области $d_2 > d_1$ (выше диагонали) кривая наилучшего отклика описывается формулой

$$d_2 = d_1^2 \cdot (1 - \gamma \cdot (p - c) - 2 \cdot \gamma \cdot d_1) \cdot [p - c - 2 \cdot d_1 \cdot (1 - \gamma \cdot (p - c)) - 3 \cdot \gamma \cdot d_1^2]^{-1} \tag{2.4}$$

при $d_1 \in [d_1^{**}, \breve{d}_1)$, где число $\breve{d}_1$ является положительным корнем квадратной скобки в формуле (2.4), а $d_1^{**} > 0$ определяется точкой пересечения кривой (2.4) с диагональю, т.е. является положительным корнем уравнения:

$$5 \cdot \gamma \cdot d^2 + 3 \cdot d \cdot (1 - \gamma \cdot (p - c)) - (p - c) = 0, \tag{2.5}$$

который равен

$$d^{**} = (p - c) \cdot \{\sqrt{9 \cdot (1 - \kappa)^2 + 20 \cdot \kappa} - 3 \cdot (1 - \kappa)\} \cdot (10 \cdot \kappa)^{-1}, \quad \gamma = \kappa \cdot (p - c)^{-1}\ . \tag{2.6}$$

Заметим, что если величину γ считать изменяемым параметром, то

$$\lim_{\gamma \to 0} d_1^{**} = d_1^*, \quad \lim_{\gamma \to \infty} d_1^{**} = \frac{3}{5} \cdot (p - c)$$

$$\lim_{\gamma \to 0} \breve{d}_1 = \frac{p - c}{2}, \quad \lim_{\gamma \to \infty} \breve{d}_1 = \frac{2}{3} \cdot (p - c)\ .$$

В области $d_1 \in (d_1^*, d_1^{**})$ производная $\frac{\partial \pi_1}{\partial d_1}$ терпит скачок на диагонали, причем знак этой производной выше и ниже диагонали определяется равенством $sign \frac{\partial \pi_1}{\partial d_1} = sign(d_2 - d_1)$, как это и показано на Рис. 1 стрелками.

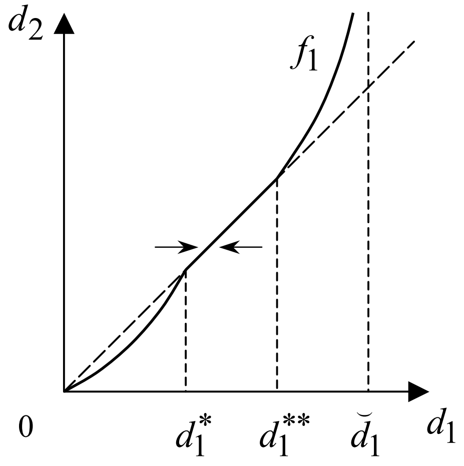

Рис. 1. Кривая наилучшего отклика d_1 первой фирмы на торговые скидки d_2 второй фирмы.

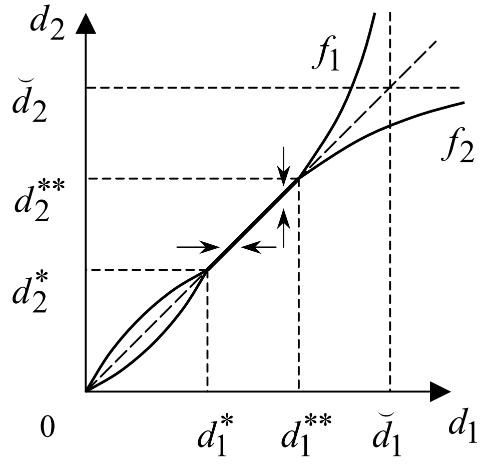

Рис. 2. Кривые наилучшего отклика для первой и второй фирм в симметричном случае.

Если рассмотреть вторую фирму, то в силу симметрии кривая ее наилучшего отклика будет симметрична относительно биссектрисы кривой наилучшего отклика первой фирмы; поэтому объединенная картина двух этих кривых будет выглядеть так, как показано на Рис. 2.

Отрезок диагонали между точками (d_1^*, d_2^*) и (d_1^{**}, d_2^{**}) является общим для этих кривых и представляет собой непрерывное множество точек равновесия Нэша.

Мы отнюдь не предполагаем (по крайней мере в рамках данной модели), что существует действенный механизм, заставляющий фирмы обязательно стремиться к одной из точек равновесия Нэша. По-видимому, в реальности, существование подобной *притягивающей* точки равновесия делает более вероятным выполнение договоренностей (в рамках коалиционной игры), если они касаются поддержания торговых скидок в размерах, определяемых этой точкой. И можно сказать, что в рассматриваемой модели множество торговых скидок, «доступных для переговоров», шире, чем в случае одной точки равновесия. Для приведенного симметричного случая нетрудно определить (из условия максимума прибыли) притягивающую точку из найденного непрерывного множества. Эта возможность становится не столь очевидной для несимметричного случая.

Ответим еще на один вопрос. Возможна ли ситуация, при которой непрерывное множество точек равновесия Нэша в пространстве торговых скидок содержит такую точку (*квазимонополия*), в которой суммарная прибыль двух фирм будет иметь максимальную величину? Для рассматриваемого симметричного случая положительный ответ на этот вопрос следует из следующих простых вычислений.

Как показано в Приложении 1, в общем случае это возможно лишь при $\gamma > (\overline{p-c})^{-1}$. В симметричном случае $\overline{p-c} = p-c$ и согласно (П.5)

$$\tilde{d} = 2^{-1} \cdot (p-c) \cdot (1-k^{-1}), \quad \gamma = k \cdot (p-c)^{-1} \quad . \tag{2.7}$$

При положительном ответе на поставленный вопрос должно существовать такое значение $k > 0$, при котором

$$d^* \le \tilde{d} < d^{**}, \tag{2.8}$$

где d^* определяется формулой (2.3), а d^{**} – формулой (2.6). Нетрудно показать, что неравенство (2.8) выполняются при $k \ge 3$.

3. **Конкуренция двух различных фирм**

Если $c_1 \ne c_2$ и $\beta_1 \ne \beta_2$, то рассмотренная в разделе 2 симметрия нарушается. Для удобства дальнейшего описания введем обозначения

$$a = (p-c_2) \cdot (p-c_1)^{-1}, \quad \beta = \beta_2 \cdot \beta_1^{-1} \tag{3.1}$$

В этих обозначениях формулы (А), (В) для прибыли принимают вид:

$$\begin{aligned} \pi_1 &= (p-c_1-d_1) \cdot \frac{d_1}{d_1 + \beta \cdot d_2} \cdot Q_0 \cdot [1 + \gamma \cdot \min(d_1, d_2)], \\ \pi_2 &= ((p-c_1) \cdot a - d_2) \cdot \frac{\beta \cdot d_2}{d_1 + \beta \cdot d_2} \cdot Q_0 \cdot [1 + \gamma \cdot \min(d_1, d_2)]. \end{aligned} \tag{3.2}$$

Пусть при $\gamma = 0$ (в отсутствии роста рынка в целом), кривые наилучшего первой (f_1) и второй (f_2) фирм пересекаются в области $d_2 < d_1$ в точке (d_1^0, d_2^0) (см. Рис. 3).

При $\gamma = 0$ кривая f_1 определяется условием

$$\frac{\partial \pi_1}{\partial d_1} = 0 \Rightarrow \quad d_2 = \frac{1}{\beta} \cdot \frac{d_1^2}{p-c_1-2 \cdot d_1} \quad , \tag{3.3}$$

а кривая f_2 – условием

$$\frac{\partial \pi_2}{\partial d_2} = 0 \quad \Rightarrow \quad d_1 = \beta \cdot \frac{d_2^2}{(p - c_1) \cdot a - 2 \cdot d_2} \qquad . \tag{3.4}$$

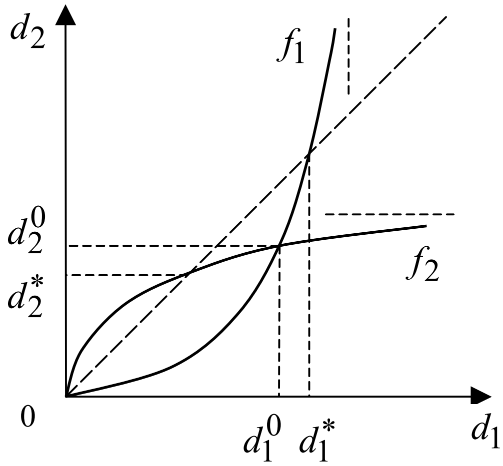

Рис. 3. Кривые наилучшего отклика для двух фирм при $\gamma = 0$ в несимметричном случае.

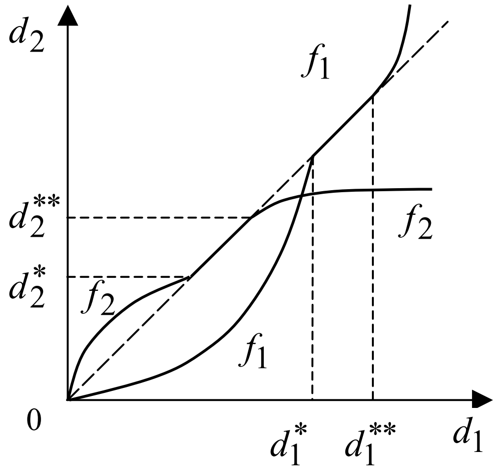

Рис. 4. Кривые наилучшего отклика для двух фирм при малых положительных γ.

Точка пересечения (d_1^0, d_2^0) кривых f_1 и f_2 определяется совместным решением уравнений (3.3) и (3.4). Задача может быть сведена к решению кубического уравнения:

$$3 \cdot \beta \cdot (d_2^0)^3 + 2 \cdot (2 - a \cdot \beta) \cdot (p - c_1) \cdot (d_2^0)^2 - 4 \cdot (p - c_1)^2 \cdot a \cdot d_2^0 + (p - c_1)^3 \cdot a^2 = 0,$$

которое, путем замены переменных

$$d_2^0 = (p - c_1) \cdot a \cdot x, \quad \omega = (a \cdot \beta)^{-1}, \tag{3.5}$$

приводится к виду

$$3 \cdot x^3 + 2 \cdot (2 \cdot \omega - 1) \cdot x^2 - 4 \cdot \omega \cdot x + \omega = 0. \tag{3.6}$$

Нетрудно заметить, что по определению величины a и β положительны и, следовательно, значение левой части равенства (3.6) в точке $x = 0.5$ отрицательно. Отсюда следует, что из двух положительных корней уравнения (3.6) только один не превышает 0.5 (что очевидно из равенства (3.4)) и является функцией произведения параметров $a \cdot \beta$.

Можно показать, что искомый корень уравнения (3.6) будет монотонно возрастающей функцией от ω, стремящейся к значению 0.5 при $\omega \to \infty$ и ведущей себя как $\sqrt{\omega/2}$ при $\omega \to 0$.

Из равенства (3.4) следует, что

$$d_1^0 = (p - c_1) \cdot \omega^{-1} \cdot x^2 \cdot (1 - 2 \cdot x)^{-1} \tag{3.7}$$

Таким образом, при $\gamma = 0$ равенства (3.5) и (3.6) вместе с единственным корнем уравнения (3.6), лежащим в интервале $x \in (0,\ 0.5)$, определяют равновесную точку Нэша (d_1^0, d_2^0).

Точки пересечения кривых f_1 и f_2 с биссектрисой в плоскости (d_1, d_2) определяются выражениями:

$$d_1^* = (p - c_1) \cdot \beta \cdot (2\beta + 1)^{-1}, \quad d_2^* = (p - c_1) \cdot a \cdot (2 + \beta)^{-1} \quad , \tag{3.8}$$

и условие того, что точка пересечения (d_1^0, d_2^0), являющаяся точкой равновесия Нэша, лежит ниже биссектрисы, эквивалентна неравенству

$$d_2^* < d_1^* \quad \Rightarrow \quad \beta > \beta_0(a) = \sqrt{(1-a)^2 + a} - (1-a) \quad . \tag{3.9}$$

При $\gamma > 0$ кривые f_1 и f_2 перестают быть гладкими во всем первом квадранте плоскости (d_1, d_2), как было показано в симметричном случае. При малых γ кривые f_1 и f_2 имеют вид, показанный на Рис. 4.

Начиная с $\gamma = \gamma_1 > 0$ одно из пересечений кривых f_1 и f_2 будет осуществляться на биссектрисе в точке $d_1^* = d_2^{**}$ (см. Рис. 5). Для определения величин d_1^{**} и d_2^{**}, характеризующих точки продолжения гладких участков кривых f_1 и f_2 соответственно выше и ниже биссектрисы, необходимо написать для каждой из них уравнение в соответствующей области и условие пересечение с биссектрисой.

Для кривой f_1 выше биссектрисы ($d_2 > d_1$):

$$\frac{\partial \pi_1}{\partial d_1} = 0 \quad \Rightarrow \quad d_2 = \frac{1}{\beta} \cdot \frac{d_1^2 \cdot (1 - \gamma \cdot (p - c_1) + 2 \cdot \gamma \cdot d_1)}{(p - c_1) - 2 \cdot d_1 \cdot [1 - \gamma \cdot (p - c_1) + 1{,}5 \cdot \gamma \cdot d_1]} \tag{3.10}$$

и пересечение с биссектрисой осуществляется в точке $d_1 = d_2 = d_1^{**}$, что приводит к уравнению для определения d_1^{**}:

$$\gamma \cdot (3 + \frac{2}{\beta}) \cdot (d_1^{**})^2 + (1 - \gamma \cdot (p - c_1)) \cdot (2 + \frac{1}{\beta}) \cdot d_1^{**} - (p - c_1) = 0\,. \tag{3.11}$$

Положительный корень этого уравнения имеет пределы:

$$d_1^{**} \underset{\gamma \to 0}{\to} d_1^* = (p - c_1) \cdot \beta \cdot (2 \cdot \beta + 1)^{-1}, \quad d_1^{**} \underset{\gamma \to \infty}{\to} (p - c_1) \cdot (2 \cdot \beta + 1) \cdot (3 \cdot \beta + 2)^{-1}.$$

Для кривой f_2 ниже биссектрисы ($d_2 < d_1$):

$$\frac{\partial \pi_2}{\partial d_2} = 0 \quad \Rightarrow \quad d_1 = \beta \cdot \frac{d_2^2 \cdot (1 - \gamma \cdot (p - c_1) \cdot a + 2 \cdot \gamma \cdot d_2)}{(p - c_1) \cdot a - 2 \cdot d_2 \cdot [1 - \gamma \cdot (p - c_1) \cdot a + 1{,}5 \cdot \gamma \cdot d_2]} \quad , \tag{3.12}$$

и точка пересечения ее с биссектрисой $d_1 = d_2 = d_2^{**}$ удовлетворяет уравнению

$$\gamma \cdot (3 + 2 \cdot \beta) \cdot (d_2^{**})^2 + (1 - \gamma \cdot (p - c_1) \cdot a) \cdot (2 + \beta) \cdot d_2^{**} - (p - c_1) \cdot a = 0 \quad (3.13)$$

положительный корень которого имеет пределы:

$$d_2^{**} \underset{\gamma \to 0}{\to} d_2^* = (p - c_1) \cdot \frac{a}{2 + \beta}, \quad d_2^{**} \underset{\gamma \to \infty}{\to} (p - c_1) \cdot a \cdot \frac{2 + \beta}{3 + 2 \cdot \beta} < \frac{1}{2} \cdot (p - c_1) \cdot a .$$

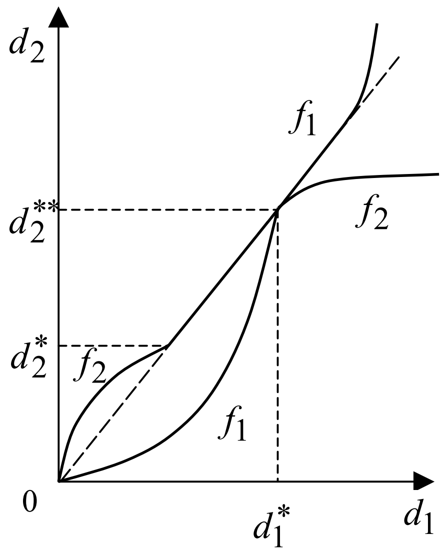

Рис. 5. Кривые наилучшего отклика для двух фирм при $\gamma = \gamma_1$.

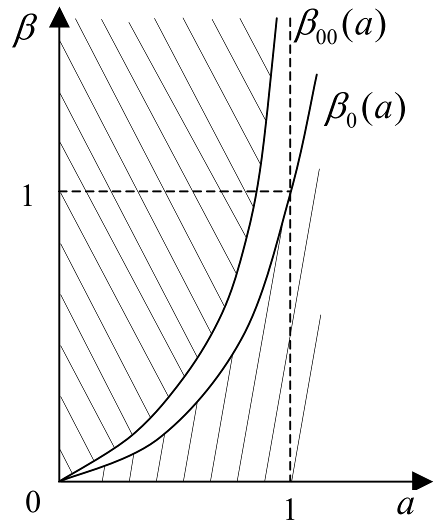

Рис. 6. В не заштрихованной области параметров a, β существует непрерывное множество точек равновесия Нэша при $\gamma > \gamma_1$.

Очевидно, что для того, чтобы существовало такое значение $\gamma_1 > 0$, для которого выполняется условие $d_2^{**}(\gamma_1) = d_1^*$ необходимо и достаточно, чтобы выполнялось неравенство

$$d_1^* < \lim_{\gamma \to \infty} d_2^{**}(\gamma) . \quad (3.14)$$

Это неравенство справедливо, если между положительными параметрами a и β выполняется соотношение:

$$\begin{aligned} &\beta < \beta_{00}(a) = \frac{5a - 3 + \sqrt{(5a-3)^2 + 16a \cdot (1-a)}}{4 \cdot (1-a)} \quad && \text{\textit{при}} \quad 0 < a < 1, \\ &\beta < \infty && \text{\textit{при}} \quad a > 1 \end{aligned} \quad . \quad (3.15)$$

Таким образом, объединение двух условий (3.9) и (3.15) приводит к разрешенной области значений параметров a и β, представленной на Рис. 6 незаштрихованной областью.

Эта область характеризуется двумя условиями: тем, что точка пересечения кривых наилучшего отклика f_1 и f_2 при $\gamma = 0$ лежит ниже биссектрисы (условие (3.9)), и существует такое значение $\gamma_1 > 0$, при котором кривые f_1 и f_2 имеют хотя бы одну общую точку на биссектрисе в плоскости (d_1, d_2) (условие (3.15)).

4. **Выгоды продавцов и покупателей**

От ценовой конкуренции продавцов выгоды (consumers' surplus) получают, прежде всего, покупатели. Это очевидно не только на растущем из-за торговых скидок рынке (когда $\gamma > 0$), но и в ситуации $\gamma = 0$, когда рынок в целом не растет, но происходит перераспределение объемов продаж между двумя фирмами. Источником этого перераспределения являются предлагаемые фирмами ценовые скидки d_1, d_2 и различия в реакции на них со стороны покупателей. Эти различия в нашей модели описываются коэффициентами β_1, β_2. Предположим, что установились скидки (d_1^0, d_2^0), которые соответствуют точке равновесия Нэша при $\gamma = 0$.

Насколько существенно зависят прибыли фирм и выгода покупателей от включенных в модель параметров c_i, β_i? Для ответа на этот вопрос запишем исходную задачу в такой форме, чтобы результаты рассмотрения зависели от одного безразмерного параметра $\omega = (a \cdot \beta)^{-1}$, где величины a и β определены формулами (3.1).

Введем безразмерные величины

$$\delta_i = d_i^0 \cdot (p - c_i)^{-1} < 1,\ i = 1, 2, \tag{4.1}$$

которые определяют равновесную точку Нэша. Экономический смысл величин (4.1) очевиден: это величины равновесных скидок, нормированных на величину прибыли за единицу товара без скидок. Согласно формулам (3.3) – (3.7) величины δ_i определяются единственными, не превышающими значение 0.5 положительными корнями двух кубических уравнений:

$$3 \cdot \delta_2^3 + 2 \cdot (2 \cdot \omega - 1) \cdot \delta_2^2 - 4 \cdot \omega \cdot \delta_2 + \omega = 0 \tag{4.2}$$

$$3 \cdot \omega \cdot \delta_1^3 + 2 \cdot (2 - \omega) \cdot \delta_1^2 - 4 \cdot \delta_1 + 1 = 0 \quad . \tag{4.3}$$

Оба эти корня являются функциями параметра $\omega = (a \cdot \beta)^{-1}$, причем из сравнения (3.6) и (4.3) следует, что если x удовлетворяет уравнению (3.6), то

$$\delta_2(\omega) = x, \quad \delta_1(\omega) = \omega^{-1} \cdot x^2 \cdot (1 - 2 \cdot x)^{-1} \quad . \tag{4.4}$$

В новых обозначениях формулы (3.2) прибыли фирм имеют вид:

$$\begin{gathered} \pi_i = (p - c_i - d_i) \cdot \beta_i \cdot d_i \cdot (\beta_1 \cdot d_1 + \beta_2 \cdot d_2)^{-1} \cdot Q_0 = (p - c_i) \cdot F_i \cdot Q_0, \\ F_i = (\delta_i^{-1} - 1) \cdot f_i, \quad f_1 = \omega \cdot \delta_1^2 \cdot (\delta_2 + \omega \cdot \delta_1)^{-1}, \quad f_2 = \delta_2^2 \cdot (\delta_2 + \omega \cdot \delta_1)^{-1}. \end{gathered} \tag{4.5}$$

Величины F_i можно интерпретировать как условную «выгоду» i-ой фирмы, определяющую как бы приходящуюся на нее долю рынка при условии продажи без скидок. Общие выгоды покупателей могут быть также представлены через уже введенные величины f_i – условные выгоды покупателей:

$$CS = (p - c_1) \cdot f_1 \cdot Q_0 + (p - c_2) \cdot f_2 \cdot Q_0 \quad . \tag{4.6}$$

Заметим, что согласно определению (4.5) величин F_i, f_i

$$F_1 + f_1 + F_2 + f_2 = 1,$$

т.е. сумма условных выгод продавцов и покупателей одинакова при любых значениях параметров.

Воспользовавшись численным решением уравнений (4.2), (4.3) (результаты которого приведены в табл. 1), можно представить графически (см. Рис. 7) характер изменений условных выгод F_i и f_i при изменении $\delta_1 \cdot \delta_2^{-1}$ отношения нормированных скидок (4.1).

Введенные выше условные выгоды имеют то преимущество перед реальными, что не зависят явно от величин $p - c_i$, что весьма удобно для анализа. Суммарная условная выгода потребителей достигает максимального значения при $\omega = \delta_1 \cdot \delta_2^{-1} = 1$. Отметим, что это значение параметра ω соответствует не только абсолютно симметричному случаю $a = 1, \beta = 1$, но тем не менее приводит к равенству нормированных скидок δ_1 и δ_2.

Максимальное значение нормированных скидок в приведенных вычислениях $\delta_2 = 0.445$ наблюдалось при $\omega = 16$, а минимальное $\delta_2 = 0.125$ при $\omega = 0.05$. Другими словами, в численном эксперименте нормированные скидки менялись в 2.5 раза.

Таблица 1

	omega	delta1/delta2	F1	f1	F2	f2	f1+f2	F1+F2
	0,050	3,600	0,084	0,069	0,742	0,106	0,175	0,825
	0,100	2,730	0,118	0,096	0,656	0,130	0,226	0,774
	0,250	1,760	0,180	0,125	0,533	0,162	0,287	0,713
	0,500	1,330	0,250	0,150	0,431	0,169	0,319	0,681
	1,000	1,000	0,334	0,167	0,334	0,167	0,333	0,667
	2,000	0,760	0,431	0,172	0,248	0,149	0,321	0,679
	4,000	0,570	0,533	0,164	0,179	0,125	0,288	0,712
	8,000	0,440	0,624	0,156	0,121	0,099	0,255	0,745
	16,000	0,310	0,715	0,118	0,091	0,076	0,194	0,806
max	16,000	3,600	0,715	0,172	0,742	0,169	0,333	0,825
min	0,050	0,310	0,084	0,069	0,091	0,076	0,175	0,667
Разность	15,950	3,290	0,631	0,104	0,651	0,093	0,158	0,158
Отношение	320,000	11,613	8,520	2,510	8,155	2,231	1,907	1,238

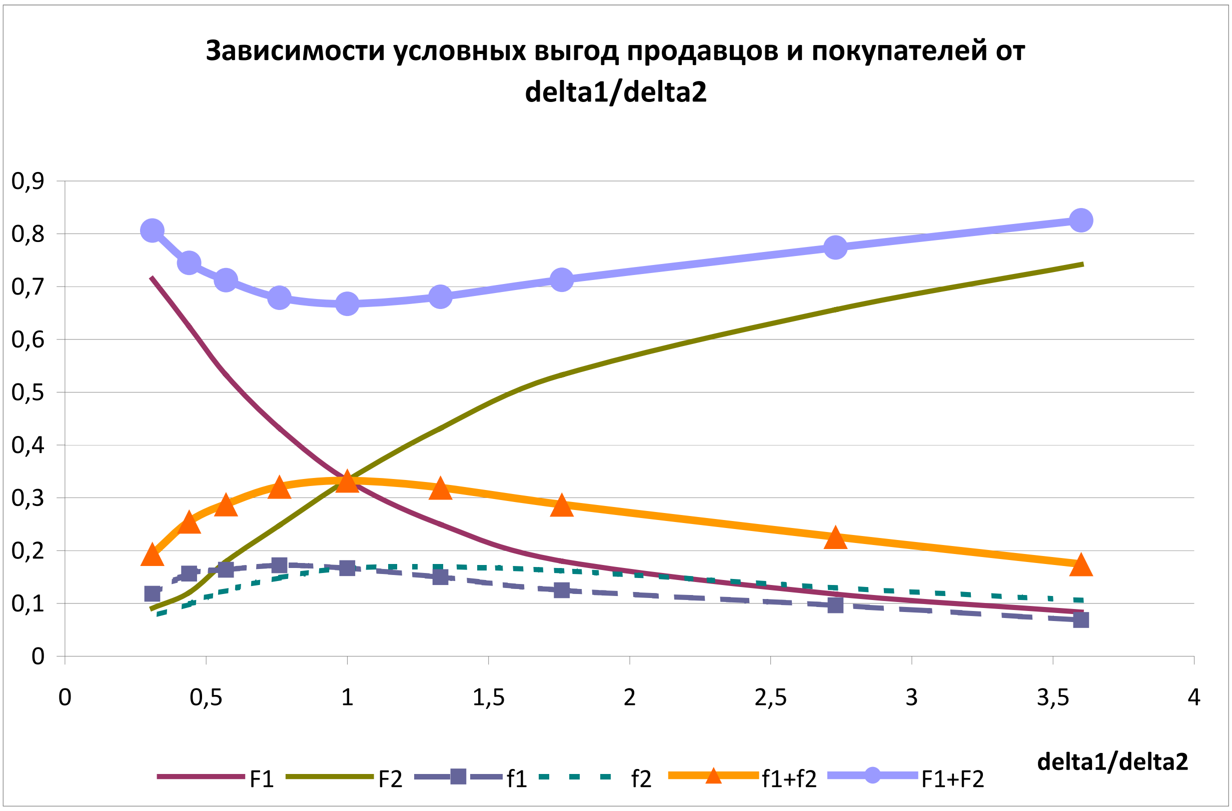

Рис. 7. Жирные линии – сумма условных выгод продавцов (верхняя линия F1+F2) и покупателей (f1+f2). Круто растущая линия – условная выгода второй фирмы (F2), круто убывающая линия – условная выгода первой фирмы (F1). Пунктирными линиями представлены условные выгоды потребителей, покупающих товар у соответствующей фирмы.

Бросается в глаза существенное различие в масштабах изменений условных выгод отдельных фирм и отдельных групп потребителей (отношение максимума к минимуму для различных показателей в табл. 1). Другими словами в нашей модели ценовой конкуренции перераспределение условных выгод между фирмами выражено гораздо ярче, чем перераспределение условных выгод групп потребителей.

5. **Краткие выводы**

Снижение цен торговыми организациями на отдельные группы товаров являются весьма распространенным маркетинговым инструментом. При этом преследуются различные цели. Рассмотренная в данной статье модель ценовой конкуренции обращает внимание на одну из них: увеличение объемов продаж за счет уменьшения их у конкурентов.

Если общий объем рынка при введении торговых скидок не увеличивается, то для данной модели существует такое распределение скидок между продавцами, которое соответствует равновесию Нэша. Оно характерно тем, что отклонение от него для любого участника является не выгодным, если остальные участники сохраняют прежние значения скидок. Численные расчеты для двух конкурирующих фирм показывают, что при изменении предельных издержек и относительного предпочтения покупателей в широком диапазоне равновесные скидки могут изменяться в 2.5 раза, а прибыль фирм может изменяться в восемь раз.

Если общий объем рынка при введении торговых скидок растет, то на примере двух фирм показано, что при не очень жестких предположениях относительно значений параметров возможна ситуация, когда вместо уединенной точки равновесия Нэша реализуется непрерывное множество, обладающее теми же свойствами. При более жестких условиях на параметры модели это множество может даже содержать такую точку, которая соответствует максимальной величине общей прибыли для обоих участников, если бы они принимали скоординированные одинаковые решения относительно величины скидок (ситуация квазимонополии или сговора).

Если исходить из математической структуры рассмотренной модели, то некоторые ситуации, возникающие в теории управления организационными системами [9 – 11], могут приводить к похожим математическим конструкциям.

В заключение авторы приносят благодарность редактору сборника В.З.Беленькому за полезные замечания.

Приложение 1

Рост рынка и квазимонопольная цена

Предположим, что связь между объемом Q продаж и ценой P линейна:

$$Q = \gamma_0 \cdot (P_{\max} - P) \quad . \tag{П.1}$$

Представим цену P в виде

$$P = p - d\,, \tag{П.2}$$

где p – цена, соответствующая объему продаж Q_0, а d – величина торговой скидки. Тогда из (П.1) и (П.2) следует, что объем продаж при наличии скидки

$$Q = Q_0 \cdot (1 + \gamma \cdot d), \qquad \gamma = \gamma_0 \cdot Q_0^{-1} = (P_{\max} - p)^{-1} \tag{П.3}$$

Если цена p определена как монопольная при постоянных предельных издержках c, то тогда

$$p = P_{mon} = 0{,}5 \cdot (P_{\max} + c), \quad Q_0 = Q_{mon} = 0{,}5 \cdot \gamma_0 \cdot (P_{\max} - c), \quad \gamma = \gamma_{mon} = 2 \cdot (P_{\max} - c)^{-1}$$

Однако если цена p без скидок выше монопольной цены P_{mon}, то $Q_0 < Q_{mon}$ (что вполне возможно для нового товара), то $\gamma = \gamma_0 \cdot Q_0^{-1} > \gamma_{mon}$, и для установления P_{mon} и Q_{mon} требуется торговая скидка $d > 0$.

Если рассматривать все фирмы как одну квазимонополию, продающую товар по одной цене $p - d$, то суммарная прибыль всех фирм будет определяться выражением

$$\Pi(d) = \sum_{i=1}^{n} \pi_i = Q_0 \cdot [\overline{p-c} - d] \cdot (1 + \gamma \cdot d), \quad \overline{p-c} = \sum_{i=1}^{n} (p - c_i) \cdot \beta_i \cdot (\sum_{j=1}^{n} \beta_j)^{-1}. \tag{П.4}$$

Определим величину

$$\tilde{d} = 0{,}5 \cdot (\overline{p-c} - \gamma^{-1}) \quad . \tag{П.5}$$

Очевидно, что если $\tilde{d} < 0$, то, согласно (П.4) $\max\limits_{d \ge 0} \Pi(d) = \Pi(0) = Q_0 \cdot \overline{p-c}$.

Это означает, что базовая цена p и является квазимонопольнй. Если же $\tilde{d} > 0$, то

$$\max_{d \ge 0} \Pi(d) = \Pi(\tilde{d}) = 0{,}25 \cdot Q_0 \cdot \overline{p-c} \cdot (2 + \overline{p-c} \cdot \gamma + (\overline{p-c} \cdot \gamma)^{-1}) \ge \Pi(0), \tag{П.6}$$

и, следовательно, существует такая величина торговой скидки $\tilde{d} > 0$, обеспечивающей максимум прибыли для квазимонополии. Очевидно, что это возможно (согласно (П.5) и (П.6)) лишь при $\gamma > (\overline{p-c})^{-1}$, что вполне соответствует предыдущим рассуждениям данного приложения о случае, когда базовая цена p была выше монопольной.

Приложение 2

О модели Шерера и Росса

В книге Шерера и Росса [12] рассматривается следующая модель ценовой конкуренции при дуополии.

Общий объем продаж

$$Q = \gamma_0 \cdot (P_{\max} - P)$$

определяется средневзвешенной ценой

$$P = P_1 \cdot s_1 + P_2 \cdot s_2,$$

где P_i цены продаж i-ой фирмы, а s_i – доля рынка i-ой фирмы, которая определяется по следующим формулам:

в области $0,4 \cdot P_{\max} < P_2 < P_1 < P_{\max}$

$$s_1 = \frac{1}{2} - \frac{1}{3} \cdot \frac{P_1 - P_2}{P_2}, \quad s_2 = \frac{1}{2} + \frac{1}{3} \cdot \frac{P_1 - P_2}{P_2};$$

в области $0,4 \cdot P_{\max} < P_1 < P_2 < P_{\max}$

$$s_1 = \frac{1}{2} + \frac{1}{3} \cdot \frac{P_2 - P_1}{P_1}, \quad s_2 = \frac{1}{2} - \frac{1}{3} \cdot \frac{P_2 - P_1}{P_1} \ .$$

Множитель $1/3$ появляется из предположения, что цена не может опуститься ниже $0,4 \cdot P_{\max}$. Поэтому максимально возможное отношение цен равно $10/4$ и при таком соотношении цен доля рынка одной из фирм равна 1. Полные издержки фирмы описываются квадратичной зависимостью

$$TC_i = a_0 + a_1 \cdot s_i \cdot Q + a_2 \cdot (s_i \cdot Q)^2, \quad a_k > 0, \quad k = 0, 1, 2.$$

В этой модели, к сожалению, не просматривается возможность обобщения на случай ценовой конкуренции многих фирм с учетом предпочтений покупателей, поскольку определение долей рынка очень специфическое.

Литература

1. Байе М.Р. Управленческая экономика и стратегия бизнеса. – М.: ЮНИТИ-ДАНА, 1999. – 743 с.
2. Липсиц И.В. Ценообразование (управление ценообразованием в организации). – М.: Экономист, 2004, – 448 с.
3. Четвериков В.М. Эффективность рекламных затрат для различных рынков. Труды международной конференции «Управление экономическим потенциалом региона в условиях международной интеграции», ноябрь 2004 г., Москва-Гомель, с. 45-52.
4. Четвериков В.М. Оптимальные затраты на рекламу в моделях рынков различных типов. Сб. «Анализ и моделирование экономических процессов», вып.1. – М.: ЦЭМИ РАН 2004, с 45-62.
5. Kimball G.E. Some industrial applications of military operations research methods. Operation Research, 1957, №5, pp. 201-204.
6. Schmalensee R. A model of advertising and product quality. Journal of political economy. 1978, 86, pp. 485-503.
7. Little John D.C. Aggregate advertising models: the state of the art. Operation Research, 1979, v. 27, №4, pp. 629-667.
8. Горшунова С.Н. Оптимизация затрат на рекламу: модели и методы (реферативный обзор). Сб. «Моделирование механизмов функционирования экономики России на современном этапе», вып. 4. – М.: ЦЭМИ РАН, 2000, с. 85-106.
9. Новиков Д.А. Теория управления организационными системами. – М.: Физматлит, 2007. – 584 с.
10. Четвериков В.М. Влияние конкуренции при распределении корпоративного заказа. Сб. «Анализ и моделирование экономических процессов», вып. 6. – М.: ЦЭМИ РАН, 2009, с. 69-92.
11. Четвериков В.М. Декларируемая эффективность и конкуренция при распределении корпоративного заказа. Сборник научных статей «Теоретические и практические проблемы формирования инновационной экономики» по материалам чтений, посвященных 100-летию со дня рождения П.Друкера.
    Беларусь, Гомель, ноябрь 2009 г., Гомель, ЦИИР, 2009, с. 47-51.
12. Шерер Ф.М., Росс Д. Структура отраслевых рынков. – М.: Инфра-М, 1997, 698 с.

Т.А.Белкина

Теоремы достаточности для вероятности неразорения в динамических моделях страхования с учетом инвестиций [1]

Одно из центральных мест в работах, посвященных изучению моделей динамики капитала страховых компаний, участвующих на финансовых рынках, является оценивание вероятности разорения, соответствующей той или иной стратегии инвестиций, анализ ее асимптотического поведения при больших значениях первоначального капитала (см., например, [1]-[10]). Капитал компании в указанных работах представляется, как правило, однородным (возможно, управляемым) марковским процессом, описываемым стохастическим дифференциальным уравнением (будем называть его в дальнейшем процессом риска). С помощью аппарата производящих операторов выписываются дифференциальные, интегро-дифференциальные (и др., в зависимости от вида оператора) уравнения для вероятности неразорения. Дальнейший анализ моделей основан на исследовании решений этих уравнений, в частности, изучении их асимптотических свойств. При этом авторы сталкиваются с некоторыми трудностями, состоящими в том, что, с одной стороны, вывод уравнений и получение асимптотик решений основывается на некоторых недоказанных в той или иной из рассматриваемых ситуаций предположениях (в частности, предположении существовании производных соответствующего порядка у вероятности неразорения), а с другой стороны, изучение асимптотики решения предполагает обоснование краевого условия, которому удовлетворяет вероятность неразорения. Если в моделях страхования без инвестиций, например, в классической модели Крамера-Лундберга [11] и ее модификации со стохастическими премиями [12] дифференцируемость вероятности неразорения доказывается с помощью несложных преобразований, а краевые условия обосновываются с помощью оценки Лундберга (и подобных ей оценок, легко получаемых, если удается построить функцию от процесса риска, являющегося мартингалом), то в ситуации

[1]Работа поддержана Российским фондом фундаментальных исследований, проекты № 10-01-00767 и 11-01-00219

погружения в финансовый рынок проблема оказывается намного сложнее. Так, в работе [5] получение асимптотического представления вероятности неразорения при вложении всех средств в рисковый актив сопряжено с доказательством ее верхних и нижних оценок (см. также [4]), полученных путем достаточно сложных построений, которые, видимо не удалось воспроизвести автору [8] для видоизмененной модели со стохастическими премиями при такой же стратегии инвестиций, вследствие чего соответствующая асимптотика содержит неопределенную аддитивную константу, определяющую вероятность неразорения при бесконечно большом начальном капитале.

В данной работе на примере моделей, рассмотренных в [5], [8], будет показано, что таких проблем, как доказательство дифференцируемости функции вероятности неразорения, получение для нее оценок типа Лундберга для обоснования асимптотических представлений, можно избежать, если корректно поставить задачу определения искомой вероятности на всей положительной полуоси и доказать существование ее решения. Для этой цели используется подход, связанный с доказательством утверждений типа так называемых проверочных теорем в задачах оптимального стохастического управления.

1.Описание моделей и интегро-дифференциальные уравнения для вероятности разорения

Рассмотрим сначала классический процесс риска в модели Крамера-Лундберга:

$$R_t = u + ct - \sum_{k=1}^{N_t} Z_k, \tag{1}$$

где R_t – величина капитала страховой компании в момент времени t, $t \geq 0$, u - величина начального капитала, c - скорость поступления страховых взносов (премий), N_t, $t \geq 0$, - пуассоновский процесс с параметром λ, определяющий для каждого t число предъявленных исков за временной промежуток $(0, t]$; $Z_1, Z_2, \ldots$ - последовательность независимых одинаково распределенных случайных величин (н.о.р.с.в.) с некоторой функцией распределения $F(z)$ ($F(0) = 0$, $\mathbf{E}Z_1 = m < \infty$), представляющих собой величины последовательных страховых выплат, которые, кроме того, не зависят от процесса N_t. Предполагается также выполненным условие положительности нагрузки безопасности, т.е. положительности математического ожидания (м.о.) разности собранных премий и суммарных исков в единицу времени: $c > \lambda m$.

В модификации модели Крамера-Лундберга – модели со стохастическими премиями, исследуемой в [8], [12], детерминированный процесс поступления премий заменяется случайным процессом, а именно рассматривается процесс риска R_t вида

$$R_t = u + \sum_{i=1}^{N_1(t)} C_i - \sum_{j=1}^{N(t)} Z_j, \tag{2}$$

где первая сумма в правой части - совокупный страховой взнос к моменту времени t, $N_1(t)$ — пуассоновский процесс с параметром λ_1, определяющий для любого $t > 0$ число премий, внесенных клиентами страховой компании за временной промежуток $(0, t]$, C_i — н.о.р.с.в. с функцией распределения $G(y)$ ($G(0) = 0$, $\mathbf{E}C_1 = n < \infty$), которые определяют размеры премий и предполагаются независимыми от процесса $N_1(t)$; все остальные переменные определяются так же, как и в классической модели. Кроме того, предполагается, что процессы суммарных премий и суммарных страховых выплат являются независимыми. Естественно считать, что $\lambda_1 > \lambda$, т.е. премии поступают чаще, чем предъявляются иски. Наличие положительной нагрузки безопасности в данном случае имеет вид $n\lambda_1 > m\lambda$.

Пусть страховая компания с процессом риска (1) или (2) непрерывно инвестирует весь свой капитал в акции, изменение цены которых описывается уравнением

$$dS_t = S_t(\mu\, dt + \sigma\, dw_t),$$

где S_t - цена акции в момент t, μ - ожидаемая доходность акции, $\sigma > 0$ - волатильность, $\{w_t\}$ - стандартный винеровский процесс.

В этом случае динамика капитала (результирующий процесс риска) описывается стохастическим дифференциальным уравнением

$$dX_t = \mu\, X_t dt + \sigma\, X_t dw_t + dR_t, \qquad X_0 = u, \tag{3}$$

где процесс R_t определяется в (1) или (2) (в дальнейшем соответствующие модели будем называть "модель I" или "модель II").

Замечание. Уравнение типа (3) может описывать ситуацию, когда в акции вкладывается не весь капитал, а только фиксированная доля α, при этом оставшаяся доля инвестируется в безрисковый актив (банковский счет) при процентной ставке r, $0 \le r < \mu$, эволюция которого описывается уравнением

$$dB_t = rB_t\,dt,$$

где B_t - величина банковского счета. Тогда уравнение (3) при замене параметров μ, σ на

$$a = \alpha\mu + (1-\alpha)r, \quad b = \alpha\sigma \tag{4}$$

соответственно можно рассматривать как уравнение динамики капитала, полностью инвестируемого в акции с ожидаемой доходностью a и волатильностью b. ■

В моделях динамики капитала, описываемых уравнением (3), мы не будем предполагать наличия положительной нагрузки безопасности для исходного процесса риска $\{R_t\}$, так как положительный снос, необходимый для увеличения резерва, возникает при инвестировании в ценные бумаги. Предположим, что соотношение

$$\lim_{h\downarrow 0} \frac{1}{h}\mathbf{E}[(f(X_h) - f(X_0))|X_0 = u] = (\mathcal{A}f)(u) \quad \forall u \in \mathbb{R} \tag{5}$$

выполнено для всех функций f из некоторого подкласса в пространстве $\mathcal{C}^2(\mathbb{R})$ дважды непрерывно дифференцируемых функций на $\mathbb{R}$; оператор $\mathcal{A}$ - *инфинитезимальный оператор* однородного марковского процесса (3), который определяется одним из равенств

$$(\mathcal{A}f)(u) = \lambda \int_0^u f(u-z)\,dF(z) - \lambda f(u) + f'(u)[\mu u + c] + \frac{1}{2}\sigma^2 u^2 f''(u),$$

$$(\mathcal{A}f)(u) = \lambda \int_0^u f(u-z)\,dF(z) + \lambda_1 \int_0^\infty f(u+y)\,dG(y) - (\lambda+\lambda_1)f(u) +$$
$$+ f'(u)\mu u + \frac{1}{2}\sigma^2 u^2 f''(u),$$

в модели I или II соответственно.

Обозначим $\tau = \tau(u) = \inf\{t : X_t < 0 | X_0 = u\}$ - момент разорения процесса $\{X_t\}$, тогда $\mathbf{P}(\tau < \infty)$ - вероятность разорения в течение бесконечного интервала времени. Известно, что для марковского процесса (3) вероятность неразорения как функция начального капитала

$$\varphi(u) = \mathbf{P}(\tau = \infty) \tag{6}$$

при соответствующих предположениях должна удовлетворять следующему уравнению:

$$(\mathcal{A}\varphi)(u) = 0, \quad u > 0, \tag{7}$$

т.е.

$$(\sigma^2/2)u^2\varphi''(u) + (\mu u + c)\varphi'(u) - \lambda\varphi(u) + \lambda\int_0^u \varphi(u-x)dF(x) = 0 \tag{8}$$

$$0 < u < \infty,$$

(см. [5], [10]) или

$$\begin{aligned}(\sigma^2/2)u^2\varphi''(u) + \mu u\varphi'(u) - (\lambda+\lambda_1)\varphi(u) + \lambda\int_0^u \varphi(u-x)\,dF(x) + \\ +\lambda_1\int_0^\infty \varphi(u+y)\,dG(y) = 0\end{aligned} \tag{9}$$

$$0 < u < \infty,$$

(см. [8]) в модели I или II соответственно. Оба уравнения являются линейными интегро-дифференциальными уравнениями (ИДУ). Для того, чтобы поставить задачу нахождения вероятности неразорения $\varphi(u)$ в соответствующей модели, необходимо добавить краевое условие - предельное значение искомой вероятности при стремлении начального капитала к бесконечности. Этому посвящен следующий раздел.

2. Обоснование краевого условия и теорема достаточности для вероятности неразорения

Для вероятности неразорения, удовлетворяющей (8), в работе [5] была доказана следующая

Теорема 1. *Пусть* $F(x) = 1 - e^{-x/m}$, *$m > 0$, $\rho := 2\mu/\sigma^2$. Тогда:*

1) если $\rho > 1$, то

$$\varphi(u) = 1 - Ku^{1-\rho}(1 + o(1)), \quad u \to \infty, \tag{10}$$

для некоторого $K > 0$;

2) если $\rho < 1$ то $\varphi(u) = 0$ для любого u.

В работе [8] было доказано аналогичное утверждение для случая экспоненциально распределенных требований и премий, т.е. в случае $F(x) = 1 - e^{-x/m}$, $G(y) = 1 - e^{-x/n}$, однако асимптотическое представление типа (10) содержит в этом утверждении не одну, а две неопределенные константы. Точнее, в [8] было показано в частности, что при $\rho > 1$

$$\varphi(u) = K_1 - K_2 u^{1-\rho}(1 + o(1)), \quad u \to \infty, \tag{11}$$

для некоторых K_1, $K_2 > 0$.

Решение ИДУ (8), обладающее асимптотическим представлением (10) (если оно существует), очевидно, удовлетворяет краевому условию

$$\lim_{u\to\infty} \varphi(u) = 1. \tag{12}$$

Это условие, которому должна удовлетворять вероятность неразорения по крайней мере при определенных параметрах, вообще говоря, требует обоснования с использованием вероятностных методов. Так, для модели I условие (12) следует из верхней оценки для вероятности разорения, полученной в [4] в рамках более общей модели (см. также [5], где эти оценки были уточнены). Для модели II аналогичные оценки не были получены, поэтому представление (11) содержит неизвестную константу K_1. Что касается второй неизвестной константы, K_2, а также константы K в представлении (10), они не могут быть определены методом локального анализа, точнее, при исследовании асимптотического поведения решений уравнения на бесконечности. Вероятностные же оценки, полученные в [4] для модели I, точнее, нижняя оценка для вероятности неразорения, также не дает представления о значении соответствующей константы, но позволяет утверждать, что эта константа положительна.

Ниже мы покажем, каким образом можно избежать как получения вероятностных оценок для обоснования (12), так и доказательства дважды непрерывной дифференцируемости функции вероятности неразорения - достаточно доказать существование решения корректно сформулированной сингулярной проблемы для ИДУ (8) или (9) с краевым условием (12). Для этого нам понадобится следующая теорема (мы ее называем теоремой достаточности для вероятности неразорения).

Теорема 2. *Пусть* $\mu > \sigma^2/2$ *и уравнение* (7) *(в модели* I *или* II*) имеет на* $[0,\infty)$ *решение* $\varphi(u)$, *являющееся дважды непрерывно дифференцируемой на* $(0,\infty)$ *функцией, удовлетворяющей условию* (12) *и ограничению*

$$0 \le \varphi(u) \le 1. \tag{13}$$

Тогда это решение определяет вероятность неразорения процесса, заданного в (3), *т.е.* $\varphi(u) = \boldsymbol{P}(\tau = \infty)$, *где* $\tau = \tau(u) = \inf\{t : X_t < 0\}$ *- момент разорения процесса* $\{X_t\}$.

Доказательство. Для модели I покажем сначала, что для всех $t > 0$ и любого начального капитала $u \ge 0$ [2]

$$\varphi(u) = \mathbf{E}\,\varphi(X_{t\wedge\tau(u)}). \tag{14}$$

Обозначим $\tau_n = \tau_n(u) = \inf\{\, t > 0 \mid X_t \notin [0,n]\,\}$, $n = 1, 2, \ldots$ Воспользуемся теперь обобщенной формулой Ито. Имеем $\varphi\colon R^1 \to R^1$, $\varphi \in C^2(R^1)$, при этом

$$X_{t\wedge\tau_n} = u + \int_0^{t\wedge\tau_n} (\mu X_{s-} + c)\,ds + \int_0^{t\wedge\tau_n} X_{s-}\sigma\,dw_s - \sum_{i=1}^{N_{t\wedge\tau_n}} Z_i.$$

Тогда

$$\begin{aligned}\varphi(X_{t\wedge\tau_n}) =& \varphi(u) + \int_0^{t\wedge\tau_n} [\varphi^{'}(X_{s-})(\mu X_{s-} + c) + \frac{1}{2}\sigma^2 X_{s-}^2 \varphi^{''}(X_{s-})]\,ds + \\ &+ \int_0^{t\wedge\tau_n} \varphi^{'}(X_{s-})\sigma X_{s-}\,dw_s + \sum_{i=1}^{N_{t\wedge\tau_n}} [\varphi(X_{\theta_i}) - \varphi(X_{\theta_i -})],\end{aligned}$$

[2]Здесь $a \wedge b = min(a,b)$

где θ_i, $i = 1, 2, ...$ - случайные моменты скачков процесса $\{N_t\}$. Последнюю сумму можно переписать в виде $\int_0^{t\wedge\tau_n}[\varphi(X_s) - \varphi(X_{s-})]\, d\pi_s$, где $d\pi_s$ - случайная мера, соответствующая пуассоновскому процессу с параметром λ. Тогда перепишем последнее равенство в следующем виде:

$$\begin{aligned}
\varphi(X_{t\wedge\tau_n}) =& \varphi(u+\varepsilon) + \int_0^{t\wedge\tau_n} [\lambda\mathbf{E_s}[\varphi(X_{s-} - Z) - \varphi(X_{s-})] + \\
&+ \varphi^{'}(X_{s-})\{\mu X_{s-} + c\} + \frac{1}{2}\sigma^2 X_{s-}^2 \varphi^{''}(X_{s-})]\, ds + \\
&+ \int_0^{t\wedge\tau_n} \varphi^{'}(X_{s-})\sigma X_{s-}\, dw_s - \int_0^{t\wedge\tau_n} \varphi(X_{s-})\, d(\pi_s - \lambda s) - \\
&- \int_0^{t\wedge\tau_n} \lambda\mathbf{E_s}\left[\varphi(X_{s-} - Z)\right] ds + \int_0^{t\wedge\tau_n} \varphi(X_s)\, d\pi_s
\end{aligned}$$

(символом $\mathbf{E_s}$ здесь обозначено условное математическое ожидание $\mathbf{E_s}(\cdot) = \mathbf{E}(\cdot\,|\,\mathbf{X_{s-}})$, т.е. усреднение берется по распределению случайной величины Z). В соответствии с уравнением Беллмана первый интеграл почти наверное равен нулю. Второй интеграл является стохастическим интегралом Ито от предсказуемого процесса (напомним, что по построению процесс X_t в пределах интегрирования ограничен). Математическое ожидание этого интеграла равно нулю. То же верно и для следующего интеграла от предсказуемого процесса по мартингальной мере (детерминированный процесс λt является компенсатором пуассоновского случайного процесса с параметром λ, см. [13]). Покажем, что математическое ожидание суммы двух оставшихся интегралов также равно нулю. Действительно,

$$\begin{aligned}
\mathbf{E}\left\{\int_0^{t\wedge\tau_n} \varphi(X_s)\, d\pi_s - \int_0^{t\wedge\tau_n} \lambda\mathbf{E_s}\left[\varphi(X_{s-} - Z)\right] ds\right\} = \\
= \mathbf{E}\int_0^{t\wedge\tau_n} \mathbf{E_s}\left[\varphi(X_{s-} - Z)\right] d(\pi(s) - \lambda s) = 0.
\end{aligned}$$

Следовательно, $\varphi(u) = \mathbf{E}\,\varphi(X_{t\wedge\tau_n(u)})$, и, устремляя n к бесконечности и применяя теорему о монотонной сходимости, получаем, что верно (14).

Обозначим $X_t(u)$, $t \geq 0$, - процесс, определенный (3), подчеркивая, что его начальное состояние равно u. Тогда покажем, что

$$X_t(u+\varepsilon) \to \infty \qquad \text{на множестве} \qquad \{\tau = \infty\}, \tag{15}$$

где $\tau = \tau(u)$. Действительно, процесс, удовлетворяющий (3), в модели I представляется в виде

$$X_t = X_t(u) = \exp\{H_t\}(u + c\int_0^t \exp\{-H_\tau\}\,d\tau - \sum_{i=1}^{N_t} Z_i \exp\{-H_{\theta_i}\}),$$

где $H_t = (\mu - \frac{\sigma^2}{2})t + \sigma w_t$, откуда

$$X_t(u+\varepsilon) - X_t(u) = \varepsilon \exp\{H_t\}. \tag{16}$$

При выполнении условия $\mu > \sigma^2/2$ имеет место соотношение $\exp\{H_t\} \to \infty$ п.н. (см., например, [14]), откуда и следует (15). Для модели II соотношения (14), (15) и (16) доказываются аналогично. Очевидным следствием (16) является соотношение

$$\{\tau(u) = \infty\} \subseteq \{\tau(u+\varepsilon) = \infty\}. \tag{17}$$

Осталось доказать теперь, что $\varphi(u)$ - вероятность неразорения процесса $\{X_t\}_{t\geq 0}$, определяемого в (3) (как для модели I, так и для модели II). В силу (14) при начальном состоянии $u+\varepsilon$ и любом $t > 0$

$$\varphi(u+\varepsilon) = \mathbf{E}\,\varphi(X_{t\wedge\tau(u+\varepsilon)}(u+\varepsilon)).$$

Устремляя в этом равенстве t к бесконечности и используя свойство (15), условие (12), неотрицательность функции φ и соотношение (17), получим:

$$\varphi(u+\varepsilon) = \lim_{t\to\infty} \mathbf{E}\,\varphi(X_{t\wedge\tau(u+\varepsilon)}(u+\varepsilon)) \geq \lim_{t\to\infty} \mathbf{E}\,\varphi(X_{t\wedge\tau(u+\varepsilon)}(u+\varepsilon))\mathbf{I}_{\{\tau(u+\varepsilon)=\infty\}} \geq$$

$$\geq \lim_{t\to\infty} \mathbf{E}\,\varphi(X_{t\wedge\tau(u+\varepsilon)}(u+\varepsilon))\mathbf{I}_{\{\tau(u)=\infty\}} = \mathbf{E}\,I_{\{\tau(u)=\infty\}} = \mathbf{P}\,\{\tau(u) = \infty\}.$$

Устремляя теперь ε к нулю, получаем, что

$$\varphi(u) \geq \mathbf{P}\,\{\tau(u) = \infty\}. \tag{18}$$

С другой стороны, учитывая ограничение (13), условие (12) и полагая $\varphi(u) = 0$ для $u < 0$, имеем

$$\varphi(u) = \lim_{t\to\infty} \mathbf{E}\,\varphi(X_{t\wedge\tau(u)}(u)) = \mathbf{E}\,\varphi(X_{\tau(u)}(u)) = \mathbf{E}\,\varphi(X_{\tau(u)}(u))\mathbf{I}_{\{\tau(u)=\infty\}} \leq$$

$$\leq \mathbf{E}\,\mathbf{I}_{\{\tau(u)=\infty\}} = \mathbf{P}\,\{\tau(u) = \infty\},$$

т.е. $\varphi(u) \leq \mathbf{P}\,\{\tau(u) = \infty\}$, откуда с учетом (18) получаем $\varphi(u) = \mathbf{P}\,\{\tau(u) = \infty\}$, что и требовалось доказать.

3. Теоремы существования в случае экспоненциальных распределений размеров требований и премий

Здесь мы приведем сначала полученный ранее (совместно с Н.Б.Конюховой, А.О.Куркиной в [10]) результат, связанный с постановкой сингулярной задачи для ИДУ (8) в модели I для случая экспоненциального распределения требований, утверждением существования и единственности ее решения, а также асимптотическими представлениями этого решения в нуле и на бесконечности.

Далее приведем аналогичный результат, полученный в настоящее время совместно с Н.Б.Конюховой, С.В.Курочкиным [15] в модели II и касающийся ИДУ (9) для случая экспоненциальных распределений требований и премий.

Из этих результатов, в частности, будет следовать выполнение условия теоремы 2 (теоремы достаточности для вероятности неразорения) относительно существования решений ИДУ, обладающих нужными свойствами для того, чтобы эти решения определяли вероятность неразорения в соответствующих моделях.

Модель I. Итак, пусть $F(x) = 1 - e^{-x/m}$. В этом случае из (8) получаем следующее ИДУ, определенное на $\mathbb{R}_+$:

$$(\sigma^2/2)u^2\varphi''(u)+(\mu u+c)\varphi'(u)-\lambda\varphi(u)+\frac{\lambda}{m}\int_0^u \varphi(u-x)\exp\left(-x/m\right)dx = 0. \tag{19}$$

Для этого ИДУ с вольтерровым интегральным оператором и сильными особенностями в нуле и на бесконечности, в [10] была поставлена следующая сингулярная задача на $\mathbb{R}_+$:

$$(\sigma^2/2)u^2\varphi''(u)+(\mu u+c)\varphi'(u)-\lambda\varphi(u)+\lambda(J_m\varphi)(u) = 0, \quad 0 < u < \infty, \tag{20}$$

$$\lim_{u\to+0}\varphi(u) \quad \text{и} \quad \lim_{u\to+0}\varphi'(u) \quad \text{существуют и конечны,} \tag{21}$$

$$\lim_{u\to+0}\left[c\varphi'(u) - \lambda\varphi(u)\right] = 0, \tag{22}$$

$$0 \le \varphi(u) \le 1, \qquad u \in \mathbb{R}_+, \tag{23}$$

$$\lim_{u\to\infty}\varphi(u) = 1, \qquad \lim_{u\to\infty}\varphi'(u) = 0. \tag{24}$$

Здесь введено обозначение

$$(J_m\varphi)(u) = \frac{1}{m}\int_0^u \varphi(u-x)\exp(-x/m)dx, \tag{25}$$

где J_m – вольтерров интегральный оператор, $J_m : C[0,\infty) \to C[0,\infty)$, $C[0,\infty)$ – пространство непрерывных ограниченных на $\mathbb{R}_+$ функций. В [10] была доказана следующая

Теорема 3. *Пусть в ИДУ* (20) *все параметры* σ^2, c, λ, m, μ *– фиксированные положительные числа, и пусть выполняется условие* $\mu > \sigma^2/2$ *("надежности акций"). Тогда:*

1) существует, и притом единственное, решение $\varphi(u)$ *сингулярной линейной задачи* (20)–(22), (24)*;*

2) это решение удовлетворяет требованиям (23) *и является бесконечно дифференцируемой монотонно возрастающей на* $\mathbb{R}_+$ *функцией; если выполнено условие* $m(\mu-\lambda)+c > 0$, *то решение* $\varphi(u)$ *– вогнутая на* $\mathbb{R}_+$ *функция, а если* $m(\mu-\lambda)+c \le 0$, *то* $\varphi(u)$ *выпукла на некотором отрезке* $[0,\widehat{u}]$, *где* $0 < \widehat{u}$ *– точка перегиба;*

3) при малых и справедливо асимптотическое представление

$$\varphi(u) \sim C_0\left[1 + \frac{\lambda}{c}\left(u + \sum_{k=2}^{\infty} C_k u^k\right)\right], \qquad u \sim +0; \tag{26}$$

здесь $C_0 :\ 0 < C_0 < 1$, $C_k = D_k/k$, $k = 2,3,\ldots$, *где* D_k *определяются по следующим формулам:*

$$D_2 = -[(\mu-\lambda)/c + 1/m], \tag{27}$$

$$D_3 = -[D_2(b^2 + 2\mu - \lambda + c/m) + \mu/m]/(2c), \tag{28}$$

$$D_k = -\{D_{k-1}[(k-1)(k-2)b^2/2 + (k-1)\mu - \lambda + c/m]+$$
$$+D_{k-2}[(k-3)b^2/2 + \mu]/m\}/[c(k-1)], \quad k = 4,5,\ldots. \tag{29}$$

4) при больших и справедливо представление

$$\varphi(u) = 1 - Ku^{1-2\mu/b^2}[1 + o(1)], \qquad u \to \infty, \tag{30}$$

где $K > 0$.

Модель II. Пусть теперь $F(x) = 1 - e^{-x/m}, \quad G(y) = 1 - e^{-y/n}$. В этом случае в [15] была поставлена следующая краевая сингулярная задача для ИДУ (9) с вольтерровым и невольтерровым интегральными операторами, определенного на $\mathbb{R}_+$:

$$(\sigma^2/2)u^2\varphi''(u) + \mu u\varphi'(u) - (\lambda + \lambda_1)\varphi(u) + \lambda(J_m\varphi)(u) + \lambda_1(J_{1,n}\varphi)(u) = 0, \tag{31}$$

$$|\lim_{u\to+0}\varphi(u)| < \infty, \qquad \lim_{u\to+0}[u\varphi'(u)] = 0, \tag{32}$$

$$(\lambda + \lambda_1)\lim_{u\to+0}\varphi(u) = \frac{\lambda_1}{n}\int_0^\infty \varphi(y)\exp(-y/n)\,dy, \tag{33}$$

$$0 \le \varphi(u) \le 1 \qquad \forall u \in \mathbb{R}_+, \tag{34}$$

$$\lim_{u\to+\infty}\varphi(u) = 1, \qquad \lim_{u\to+\infty}\varphi'(u) = 0. \tag{35}$$

ИДУ (31) содержит интегральный оператор J_m, определенный в (25), и невольтерров интегральный оператор $J_{1,n} : C[0,\infty) \to C[0,\infty)$,

$$(J_{1,n}\varphi)(u) = \frac{1}{n}\int_0^\infty \varphi(u+y)\exp(-y/n)\,dy, \tag{36}$$

где $C[0,\infty)$ – линейное пространство непрерывных ограниченных на $\mathbb{R}_+$ функций.

Без подробного описания результата, скажем здесь лишь, что в [15] при условии $\mu > \sigma^2/2$ доказаны существование и единственность решения поставленной сингулярной задачи и получено его асимптотическое представление (11) при $K_1 = 1$. Следовательно, с учетом теоремы 2, мы получаем обоснование этого представления для вероятности неразорения при больших значениях начального капитала.

Автор выражает благодарность С.В.Курочкину, обратившему внимание на необходимость доказательства теорем достаточности для вероятности неразорения в рассматриваемых моделях при постановке сингулярных задач для ее нахождения.

Список литературы

[1] Paulsen J. Risk theory in a stochastic environment// Stoch. Proc. and Appl. 1993, v. 21, p. 327-361.

[2] Norberg R. Ruin problems with assets and liabilities of diffusion type// Stoch. Proc. and Appl., 1999, v. 81, p. 255-269.

[3] Hipp C., Plum M. Optimal investment for insurers// Insurance: Mathematics and Economics, 2000, v. 27, 2, p. 215-228.

[4] Kalashnikov V., Norberg R. Power tailed ruin probabilities in the presence of risky investments// Stoch. Proc. and Appl., 2002, v. 98, p. 211-228.

[5] Frolova A., Kabanov Yu., Pergamenshchikov S. In the Insurance business risky investments are dangerous// Finance and Stochastics, 2002, v. 6, 2, p.227-235.

[6] Hipp C., Plum M. Optimal investment for investors with state dependent income, and for insurers// Finance and Stochastics, 2003, v. 7, 3, p. 299-321.

[7] Gaier J., Grandits P., Schachermayer W. Asymptotic ruin probabilities and optimal investment// Ann. Appl. Probab., 2003, v.13, 3, p.1054-1076.

[8] Бойков А.В. Стохастические модели капитала страховой компании и оценивание вероятности неразорения// Дисс. ... канд. физ.-мат. наук, 2003, М.: МИ РАН.

[9] Белкина Т.А., Конюхова Н.Б., Куркина А.О. Оптимальное управление инвестициями в динамических моделях страхования: I. Инвестиционные стратегии и вероятность разорения// Обозрение прикладной и промышленной математики, 2009, т. 16, вып.6, с. 961–981.

[10] Белкина Т.А., Конюхова Н.Б., Куркина А.О. Оптимальное управление инвестициями в динамических моделях страхования: II. Модель Крамера–Лундберга при экспоненциальном распределении размера требований// Обозрение прикладной и промышленной математики, 2010, т.17, вып.1, с. 3–24.

[11] Grandell I. Aspects of risk theory. Springer, Berlin, 1991.

[12] Бойков А.В. Модель Крамера-Лундберга со стохастическими премиями// Теория вероятностей и ее применения, 2002, т. 47, вып. 3, с. 549-553.

[13] Bremaud P. Point processes and queues. Springer-Verlag, New York, 1981.

[14] Оксендаль Б. Стохастические дифференциальные уравнения.М.: Мир, 2003.

[15] Белкина Т.А., Конюхова Н.Б., Курочкин С.В. Сингулярная краевая задача для линейного интегродифференциального уравнения, возникающего в моделях страховой математики: анализ и численное решение// Ж. вычисл. матем. и матем. физ. (готовится к печати).

Е.С.Паламарчук

Управление динамикой равновесной цены в экономике с мультипликативной неопределенностью[1]

Ценообразование в результате взаимодействия спроса и предложения является одним из основных рыночных механизмов. Получающаяся таким образом равновесная цена может быть рассмотрена в динамике, что позволяет отразить влияние различных факторов на ее формирование. Такой динамический подход является актуальным для ряда рынков (энергетического, агро-рынка, рынка ресурсов) [1, 2, 3], где модель ценообразования, основанная только лишь на понятии избыточного спроса не отражает специфики происходящих процессов, в т.ч. мультипликативный эффект влияния неопределенности.

В анализе поведения производителей и потребителей можно выделить две важных составляющих.

Во-первых, учет экономическими агентами информации о некоторой *эталонной* ценовой траектории P_t^0, которая выступает как "желаемая" , и естественным является стремление приблизиться к ней [4, 5]. В этом случае потери, возникающие из-за отклонения от эталонной траектории, традиционно учитываются в квадратичном виде $(P_t - P_t^0)^2$ [4]. С другой стороны, может возникнуть ситуация, когда траекторию P_t^0, напротив, рассматривают как нижнюю границу, падение ниже которой было бы нежелательно для рыночной цены (из соображений безубыточности и т.д.). При этом превышение P_t над P_t^0 не обладает таким эффектом, т.е. упомянутоый выше подход, основанный на квадратичном виде потерь, не соотвествует анализируемому поведенческому аспекту. Вводится другая мера учета возникающих издержек - в виде логарифма отношения P_t^0 и текущей равновесной цены P_t на рынке [6]. Экономические агенты могут влиять на динамику равновесной цены, изменяя характеристики функций спроса и предложения, неизбежно сталкиваясь при этом с дополнительными расходами.

Второй важной составляющей модели является включение в анализ временны́х предпочтений субъектов рынка, что выражается в дисконтировании ими своих будущих денежных потоков.

[1]Работа поддержана Российским фондом фундаментальных исследований, проект 10-01-00767.

В данной работе рассматривается задача управления равновесной ценой на рынке с мультипликативной неопределенностью при стремлении горизонта планирования к бесконечности. При этом интегральный целевой функционал учитывает потери из-за приближения цены к заданной траектории P_t^0 как логарифм их отношения и включает издержки по изменению цены за плановый период при наличии у агентов временных предпочтений. Ставится вопрос поиска оптимальных в среднем управлений на бесконечном интервале времени и изучение их вероятностных свойств.

1. Описание модели и постановка задачи

Рассматривается рынок, где равновесная цена устанавливается в каждый момент времени из условия равенства спроса и предложения. Специфика некоторых рынков такова, что излишки спроса и предложения не допускаются (например, если товар нельзя хранить). Примером такой ситуации является рынок электроэнергии [1]. Предположим, что функции зависимости спроса D и предложения S от цены P - изоэластичны и имеют вид [1, 7]

$$D = M \cdot P^{-\beta} \tag{1}$$

$$S = N \cdot P^{\alpha}\,, \tag{2}$$

где константы $\alpha, \beta \geq 0$ и $\alpha + \beta > 0$. Случайные величины M и N в соответствующих уравнениях характеризуют мультипликативное влияние неопределенности, подробнее о них будет сказано ниже. Случайная равновесная цена P определяется из условия $D = S$, откуда

$$P = (M/N)^{1/\gamma}\,, \quad \gamma = \alpha + \beta\,. \tag{3}$$

Функции спроса со случайным воздействием вида (1) рассматривались в [8, 9], вместе с тем во многих ситуациях функции предложения также подвержены влиянию "шума" (производство электроэнергии, сельскохозяйственной продукции и т.д.). В качестве мультипликативной неопределенности будем рассматривать случайные процессы $\{M_t\}_{t=0}^{\infty}$ и $\{N_t\}_{t=0}^{\infty}$ заданные на вероятностном пространстве $\{\Omega, \mathcal{F}, \mathbf{P}\}$ с фильтрацией $\mathbf{F} = (\mathcal{F}_t)_{t\geq 0}$, и описываемые уравнениями [1]

$$dM_t = (b_M - a_M \ln M_t)M_t dt + \sigma_M M_t dW_t\,, \tag{4}$$

$$dN_t = (b_N - a_N \ln N_t)N_t dt + \sigma_N N_t dW_t\,, \tag{5}$$

где $\{W_t\}$ - одномерный стандартный винеровский процесс и все параметры $b_M, b_N, \sigma_M, \sigma_N,\ a_M, a_N > 0$ - константы. Помимо этого считается, что известно распределение случайной цены $P_0 = p_0$ в начальный момент времени.

По формуле Ито из (3)-(5) находим уравнение для динамики случайной равновесной цены P_t

$$dP_t = (\gamma b_P + K - \gamma(a_M \ln M_t - a_N \ln N_t))P_t dt + \sigma_P P_t dW_t\,, \quad P_0 = p_0\,, \quad (6)$$

где $b_P = b_M - b_N$, $K = \dfrac{\gamma}{2}((\gamma+1)\sigma_N^2 + (\gamma-1)\sigma_M^2 - 2\gamma\sigma_M\sigma_N)$, а в качестве σ_P обозначено $\gamma(\sigma_M - \sigma_N)$.

Далее, предположим, что процессы (4), (5) обладают одинаковой степенью средневозвратности, т.е. $a_N = a_M = a$, тогда уравнение (6) преобразуется к виду

$$dP_t = (\gamma b_P + K - a \ln P_t)P_t dt + \sigma_P P_t dW_t\,, \quad P_0 = p_0. \quad (7)$$

Пусть P_t^0 - детерминированная функция, задающая ценовую траекторию. Производитель и потребитель учитывают ее динамику в процессе принятия своих решений (например, для производителя таким индикатором может выступать цена, обеспечивающая безубыточность). Свойства, которым должна удовлетворять P_t^0, будут указаны ниже. Далее, b_P будем считать управляющим параметром и рассматривать (7) как управляемый случайный процесс.

Перейдем к описанию системы управления. На вероятностном пространстве с фильтрацией $\{\Omega, \mathcal{F}, \mathbf{P}, \mathbf{F}\}$ определим процесс эволюции равновесной цены $\{P_t\}_{t=0}^{\infty}$ как решение следующего стохастического дифференциального уравнения:

$$dP_t = (\gamma \bar{U}_t + K - a \ln P_t)P_t dt + \sigma_P P_t dW_t, \quad (8)$$

$\bar{U}_t$ - неупреждающее управление. В качестве допустимых управлений будем рассматривать процессы $\{\bar{U}_t\}_{t=0}^{\infty}$, согласованные с фильтрацией $(\mathcal{F}_t)_{t\geq 0}$, $\mathcal{F}_t = \sigma\{W_s,\ s \leq t\}$, такие что уравнение (8) имеет решение. Множество допустимых управлений обозначим как $\mathcal{U}$. Введем [6] меру отклонения цены P_t от P_t^0 в виде $\mathcal{Z}_t := \ln \dfrac{P_t^0}{P_t}$. Значение этой меры велико, если цена P_t мала по сравнению с P_t^0. Издержки по управлению ценой будем учитывать в виде $\bar{U}_t^2$. Таким образом, совокупные "потери" в момент времени t равны $\ln\dfrac{P_t^0}{P_t}+\bar{U}_t^2$.

Временные предпочтения агентов отражаются в виде дисконтирующей функции f_t, с помощью которой они оценивают издержки, относящиеся к разным моментам времени. Предполагается, что

1) $f_t > 0$ для $t \geq 0$, $f_0 = 1$;

2) f_t - невозрастающая и дифференцируемая на $[0,\infty)$;

3) $\lim\limits_{t\to\infty}\dfrac{f_t'}{f_t} = const \leq 0$.

Функционал потерь за плановый период $[0,T]$ имеет вид

$$J_T^P(\bar{U}^T) = \int_0^T f_t(\ln\frac{P_t^0}{P_t}+\bar{U}_t^2)\,dt\ , \tag{9}$$

где T - конечный момент времени, $\bar{U}^T = \{\bar{U}_t\}_{t\leq T} \in \mathcal{U}^{\mathcal{T}}$ - управление на интервале $[0,T]$. На горизонте планирования длины T стандартная задача состоит в минимизации совокупных ожидаемых потерь, т.е.

$$EJ_T^P(\bar{U}^T) \to \inf_{\bar{U}^T\in\mathcal{U}^{\mathcal{T}}}\ . \tag{10}$$

При неограниченном возрастании горизонта планирования T, т.е. в случае, когда $T\to\infty$, будем искать управление $\bar{U}^*$, *оптимальное в среднем на бесконечном интервале времени.*

Определение 1. Управление $\bar{U}^* \in \mathcal{U}$ будем называть *оптимальным в среднем на бесконечном интервале времени*, если оно является решением задачи

$$\phi := \limsup_{T\to\infty}\mathcal{N}(T)EJ_T^P(\bar{U}) \to \inf_{\bar{U}\in\mathcal{U}} = \phi^*\,, \qquad \mathcal{N}(T) := \frac{1}{\int\limits_0^T f_t\,dt} = \ . \tag{11}$$

($\mathcal{N}(T)$ - *осредняющий множитель*). ■

Наша задача - найти это управление и исследовать его вероятностные свойства.

2. Сведение к стандартной задаче стохастического линейного регулятора

Заметим, что $\ln \dfrac{P_t^0}{P_t} = \ln P_t^0 - \ln P_t$, т.е. выражение в (9) преобразуется:

$$J_T^P(\bar{U}^T) = \int_0^T f_t(-\ln P_t + \bar{U}_t^2)\, dt + \int_0^T f_t \ln P_t^0\, dt\,.$$

Так как второе слагаемое не зависит от управления, то в задаче минимизации будем рассматривать функционал

$$J_T(\bar{U}^T) = \int_0^T f_t(-\ln P_t + \bar{U}_t^2)\, dt\ , \tag{12}$$

и тогда на траекторию P_t^0 накладывается следующее ограничение: существует константа $C > 0$, такая что для любого $T > 0$

$$\mathcal{N}_T \left| \int\limits_0^T f_t \ln P_t^0\, dt \right| < C. \tag{13}$$

На основании (8) по формуле Ито запишем уравнение для динамики процесса $\ln P_t$

$$d(\ln P_t) = (-a(\ln P_t)dt + \gamma \bar{U}_t)dt + (K - \frac{\sigma_P^2}{2})dt + \sigma_P dW_t\,, \quad P_0 = p_0. \tag{14}$$

Теперь можно заметить, что (12),(14) представляет собой линейную стохастическую систему управления с квадратичным критерием качества. Для ее сведения к стандартному линейно-квадратическому регулятору произведем замену переменных

$$G_t := \sqrt{f_t}\,, \quad X_t := G_t \ln P_t\,, \quad U_t := G_t \bar{U}_t. \tag{15}$$

В новых обозначениях динамика управляемого процесса будет описываться уравнением

$$dX_t = (-a + \frac{G'_t}{2G_t})X_t dt + \gamma U_t dt + (K - \frac{\sigma_P^2}{2})G_t dt + \sigma_P G_t dW_t\,, \quad X_0 = \ln p_0\,. \tag{16}$$

а целевой функционал примет вид

$$J_T(U^T) = \int_0^T (-G_t X_t + U_t^2)\, dt\,. \tag{17}$$

3. Задача на бесконечном интервале времени

3.1 Решение задачи

Введем в рассмотрение функцию

$$r_t = -\frac{1}{2}\int\limits_t^\infty \Phi(s,t) G_t\, dt, \quad \text{где } \Phi(t,s) = \frac{G_t}{G_s} e^{-a(t-s)} \tag{18}$$

Из условия 2) невозрастания дисконтирующей функции f_t получаем экспоненциальную оценку для $\Phi(t,s)$ вида

$$\Phi(t,s) \le e^{-a(t-s)} \tag{19}$$

и поэтому r_t - ограниченная функция. Нетрудно проверить, что r_t - решение линейного неоднородного дифференциального уравнения

$$dr_t = (a - \frac{G'_t}{2G_t}) r_t + \frac{1}{2} G_t \tag{20}$$

с начальным условием $r_0 = -\frac{1}{2}\int\limits_0^\infty \Phi(t,0) G_t\, dt$.

Теорема 1 . Управление $\bar{U}_t^*$, имеющее вид

$$\bar{U}_t^* = -\gamma \frac{r_t}{G_t}\,, \tag{21}$$

где r_t удовлетворяет уравнению (20), является решением задачи (11).

Доказательство. Доказательство разбивается на два этапа: сначала показывается, что $\bar{U}^*$ обеспечивает экстремум функционала ϕ, а затем проверяется, что значение ϕ^* конечно.

Будем рассматривать управляемоую систему (16)-(17). В качестве управления, оптимальность которого будем доказывать, возьмем $U_t^* = G_t \bar{U}_t^* = -\gamma r_t$. Зафиксируем произвольное допустимое управление U и соотвествующий ему процесс X_t по уравнению (16). В [10] для разности функционалов $J_T(U^*) - J_T(U)$ было получено сооотношение, которое имеет место при оценке (19)

$$J_T(U^*) \leq J_T(U) + 2c_0 r_T^2 \,, \quad c_0 > 0 \text{ - некоторая константа.} \tag{22}$$

Умножая обе части 22 на $\mathcal{N}(T)$, получим

$$\mathcal{N}(T) J_T(U^*) \leq 2c_0 \mathcal{N}(T) r_T^2 + \mathcal{N}(T) J_T(U) \,. \tag{23}$$

Очевидно, что в случае $\mathcal{N}(\infty) = 0$ ввиду ограниченности функции r_T^2 ясно, что

$$\limsup_{T\to\infty} \mathcal{N}(T) E J_T(U^*) \leq \limsup_{T\to\infty} \mathcal{N}(T) E J_T(U) \tag{24}$$

(не исключено, что обе части этого неравенства обращаются в нуль).

Переходя к ситуации, когда $\mathcal{N}(\infty) > 0$, рассмотрим предел $\lim\limits_{T\to\infty} r_T$. Заметим, что из (18),(19) следует

$$|r_t| \leq \frac{1}{2} \int\limits_t^\infty e^{-a(s-t)} G_s \, ds \,.$$

Правая часть имеет предел, который можно найти по правилу Лопиталя:

$$\lim_{t\to\infty} \frac{\int\limits_t^\infty e^{-as} f_s \, ds}{e^{-at}} = \frac{1}{a} \lim_{t\to\infty} f_t =: \tilde{f} \,.$$

Отсюда и из (18) следует, что существует предел $\tilde{r} := \lim\limits_{t\to\infty} r_t$, причем $\tilde{r} = 0$ при $\tilde{f} = 0$ и некоторой константе $\tilde{r} = const$ при $\tilde{f} > 0$.

Поэтому (24) будет выполняться и для $\mathcal{N}(\infty) > 0$, т.е. во всех случаях управление $\bar{U}^*$ доставляет минимум в задаче (11).

Осталось проверить конечность значения функционала ϕ^*. Из условия (13) следует, что достаточно рассмотреть поведение $\mathcal{N}(T) J_T(U^*)$. По формуле Ито из (16),(20) записываем соотношение

$$d(r_t X_t^*) = r_t U_t^* dt + (K - \frac{\sigma^2}{2}) G_t r_t dt + \sigma_P G_t r_t dW_t + \frac{1}{2} G_t X_t^* dt,$$

с учетом того, что $U_t^* = -\gamma r_t$, можно получить выражение для $J_T(U^*)$ в виде

$$J_T(U^*) = 2X_0 r_0 - 2X_T^* r_T + 2(K - \frac{\sigma_P^2}{2}) \int\limits_0^T G_t r_t \, dt + 2 \int\limits_0^T \sigma_P G_t r_t dW_t - \gamma^2 \int\limits_0^T r_t^2 \, dt \tag{25}$$

и для среднего значения

$$EJ_T(U^*) = 2EX_0 r_0 - 2EX_T^* r_T + 2(K - \frac{\sigma^2}{2}) \int\limits_0^T G_t r_t \, dt - \gamma^2 \int\limits_0^T r_t^2 \, dt \, . \tag{26}$$

Выпишем решение уравнения (16) при $U_t = U_t^* = -\gamma r_t$

$$X_T^* = \Phi(T,0)X_0 - \gamma^2 \int\limits_0^T \Phi(T,t) r_t \, dt + (K - \frac{\sigma^2}{2}) \int\limits_0^T \Phi(T,t) G_t \, dt + \sigma_P \int\limits_0^T \Phi(T,t) G_t \, dW_t . \tag{27}$$

Экспоненциальная оценка (19), ограниченность функций r_t и f_t, приводят к тому, что $|EX_T^*| \leq C_0$, т.е. среднее значение процесса X_T^* ограничено некоторой константой $C_0 > 0$.

Воспользовавшись уравнением (20) для r_t, нетрудно найти выражение для третьего слагаемого в (25)

$$\int\limits_0^T G_t r_t \, dt = \frac{r_T - \frac{1}{2} \int_0^T f_t \, dt}{a}.$$

Аналогичным образом по уравнению динамики получаем оценку для интеграла

$$\int\limits_0^T r_t^2 \, dt \leq \frac{r_T^2}{2a} + \frac{\int_0^T G_t r_t \, dt}{2a} \, .$$

Учитывая все эти соотношения, можно привести оценки для математического ожидания (26) в виде двойного неравенства

$$C_0(r_T+r_0)-\gamma^2\frac{r_T^2}{2a}+(2(K-\frac{\sigma^2}{2})/a-\gamma^2/(4a^2))r_T+(\gamma^2/(4a^2)-(K-\frac{\sigma^2}{2})/a)\int\limits_0^T f_t\,dt$$

$$\leq EJ_T(U^*)\leq 2/a(K-\frac{\sigma^2}{2})r_T-1/a(K-\frac{\sigma^2}{2})\int\limits_0^T f_t\,dt. \qquad (28)$$

Отсюда ввиду ограниченности величин r_T,$\mathcal{N}(T)$ следует конечность функционала ϕ^*. Теорема 1 доказана. ■

3.2. Вероятностные свойства оптимального в среднем управления

Известно [10], что более сильным (в вероятностном смысле) видом оптимальности является так называемая g-оптимальность почти наверное.

Определение 2. Пусть g_T - положительная невозрастающая функция. Управление $\bar{U}^*\in\mathcal{U}$ называется *g-оптимальным почти наверное*, если $\limsup\limits_{T\to\infty} g_T(J_T^P(\bar{U}^*)-J_T^P(U))=0$ с вероятностью единица при любом $U\in\mathcal{U}$. ■

В [10] для задачи стандартного стохастического линейно-квадратического регулятора была получена оценка $g_T=o(\frac{1}{\ln T})$. Если в такой системе функционал качества включает дисконтирование, то $g_T=\max\{o(1),o(\frac{1}{f_T\ln T})\}$ (см. [5]).

Перейдем к рассмотрению вопроса о поиске функции g_T, такой что управление $\bar{U}^*$ будет являться *g-оптимальным почти наверное.*

Теорема 2 . Управление $\bar{U}^*$, задаваемое (21), является *g-оптимальным почти наверное* для любой положительной функции с условием $g_T=o(1)$. Если $\lim\limits_{t\to\infty} f_t=0$, то в качестве g_T можно взять произвольную положительную константу.

Доказательство. Обратившись к оценке (22) и полученным соотношениям для предела функции r_t, будем иметь результат утверждения. Действительно, из ограниченности r_t следует, что для справедливости $\limsup_{T\to\infty} g_T(J_T(U^*) - J_T(U) \le 0$ достаточно взять любую $g_T \to 0$ при $T \to \infty$. Кроме того, было показано, что $\lim_{t\to\infty} f_t = 0$ влечет $\lim_{t\to\infty} r_t = 0$, т.е. в этом случае $\limsup_{T\to\infty}(J_T(U^*) - J_T(U) \le 2c_0 \cdot \lim_{T\to\infty} r_T^2 = 0$, т.е. g_T может быть константой. ■

В следующем утверждении дается ответ на вопрос, обеспечивает ли оптимальное в среднем управление $\bar{U}^*$ экстремум не только ожидаемого значения, но и самого функционала $\limsup_{T\to\infty} \mathcal{N}(T) J^P(\bar{U})$ в смысле почти наверное.

Теорема 3 . Управление $\bar{U}^*$, задаваемое (21), является решением задачи $\limsup_{T\to\infty} \mathcal{N}(T) J_T^P(\bar{U}) \to \inf_{\bar{U}\in\mathcal{U}}$ с вероятностью единица.

Доказательство . Соотношение $\limsup_{T\to\infty} \mathcal{N}(T) J_T(U^*) \le \limsup_{T\to\infty} \mathcal{N}(T) J_T(U)$ получается из (23) путем предельного перехода.

Теперь остановимся на традиционном вопросе конечности функционала $\limsup_{T\to\infty} \mathcal{N}(T) J_T(U^*)$ в смысле стремления его к некоторой случайной величине п.н. или же к константе. Из (25),(26) имеем

$$J_T(U^*) = EJ_T(U^*) - 2EX_0 r_0 + 2EX_T^* r_T + 2X_0 r_0 - 2X_T^* r_T + 2\int_0^T \sigma_P G_t r_t dW_t.$$

Для $EJ_T(U^*)$ была получена оценка (28), а для исследования поведения $\mathcal{N}(T)X_T^*$ сформулируем специальную Лемму.

Лемма . Для решения X_T^*, задаваемого (27), предел $\lim_{T\to\infty} \mathcal{N}(T) X_T^* = 0$ с вероятностью 1.

Доказательство . Применяя экспоненциальную оценку (19), исходя из (18), можно показать, что при $T \to \infty$ выражение $\Phi(T,0)X_0 \to 0$ почти наверное, а слагаемые $R_1(T) := -\int_0^T \Phi(T,t) r_t\, dt$ и $R_2(T) := \int_0^T \Phi(T,t) G_t\, dt$

имеют оценки

$$0 \le R_1(T) \le -\int\limits_0^T e^{-a(T-t)} r_t\, dt,$$

$$0 \le |K - \frac{\sigma^2}{2}| R_2(T) \le |K - \frac{\sigma^2}{2}| \int\limits_0^T e^{-a(T-t)} G_t\, dt.$$

Таким образом, если

$$\int\limits_0^\infty e^{-at} r_t\, dt < \infty, \ \int\limits_0^\infty e^{-at} \sqrt{f_t}\, dt < \infty\,,$$

то автоматически при $T \to \infty$ следуют сходимости $R_1(T) \to 0\,, R_2(T) \to 0$.

В остальных случаях из полученного ранее результата $\lim\limits_{T\to\infty} r_T = 0$ при $\lim\limits_{t\to\infty} f_t = 0$, с применением правила Лопиталя можно показать, что $R_1(T) \to 0\,, R_2(T) \to 0$, а при $\lim\limits_{t\to\infty} f_t = const > 0$ будем иметь стремление $R_1(T)$ и $R_2(T)$ к некоторым константам. В итоге, $\mathcal{N}(T)R_1(T) \to 0\,, \mathcal{N}(T)R_2(T) \to 0$.

Осталось показать, что для $\mathcal{N}(T)R_3(T)$, где $R_3(T) := \int\limits_0^T \Phi(T,t) G_t\, dW_t$ с вероятностью 1 выполняется аналогичное соотношение. Для этого воспользуемся определением функции $\Phi(t,s)$ и преобразуем это выражение, т.е. $R_3(T) = G_T e^{-aT} \int\limits_0^T e^{at}\, dW_t$. Затем, применим закон повторного логарифма для винеровского процесса, предварительно произведя замену времени в интеграле $\int\limits_0^T e^{at}\, dW_t$. Тогда существует момент времени $T_0(\omega)$, такой, что при $T > T_0(\omega)$ с вероятностью 1 справедливо неравенство

$$G_T e^{-aT} |W_{\int\limits_0^T e^{2at}\, dt}| \le c G_T e^{-aT} e^{aT} \sqrt{\ln T},$$

$c > 0$ - константа, т.е. $|R_3(T)| \leq cG_T\sqrt{\ln T}$. В [5] было показано, что $\mathcal{N}(T)f_T \ln_T \to 0$ при $\mathcal{N}(T) \to 0$. Аналогичное утверждение справедливо и в случае $N(\infty) < \infty$. Действительно, для $t_0 > 0$ запишем $f_T \ln T = \int\limits_{t_0}^{T} \frac{f_t}{t} dt + f_{t_0} \ln f_{t_0} + \int\limits_{t_0}^{T} f_t' \ln t \, dt$. Пользуясь условием конечности $\int\limits_{t_0}^{\infty} f_t \, dt$, нетрудно увидеть, что выражение $\int\limits_{t_0}^{T} \frac{f_t}{t} dt$ ограничено, неубывает и, следовательно, имеет предел, равный некоторой константе. Функция $\int\limits_{t_0}^{T} f_t' \ln t \, dt$, отрицательная и невозрастающая, она также имеет предел, что следует из условия $f_T \ln T \geq 0$. Таким образом, $\lim\limits_{T\to\infty} f_T \ln T = c_2 \geq 0$, c_2 - константа. Если $c_2 > 0$, то по определению предела неравенство $f_t \geq \frac{c_3}{\ln t}$ будет выполняться для $t > t_1$, где $c_3 > 0$ - некоторая константа. При интегрировании оказывается, что $\int\limits_{t_1}^{T} f_t \, dt \geq c_3 \int\limits_{t_1}^{T} \frac{1}{\ln t} dt \to \infty$, это противоречит условию $N(\infty) < \infty$. Следовательно, $\lim\limits_{T\to\infty} f_T \ln T = 0$. Тогда во всех случаях $\mathcal{N}(T)|R_3(T)| \leq c\mathcal{N}(T)G_T\sqrt{\ln T} \to 0$ и Лемма доказана. ■

Пусть $\mathcal{M}_T$ - мартингал, $< \mathcal{M}_T >$ - его квадратичная характеристика. В [10] было показано, что если $< \mathcal{M}_T > \to \infty$, то $\lim\limits_{T\to\infty} \frac{\mathcal{M}_T}{< \mathcal{M}_T >} = 0$ п.н. Положим $\mathcal{M}_T = \int\limits_{0}^{T} G_t r_t dW_t$, тогда $< \mathcal{M}_T > = \int\limits_{0}^{T} f_t \, r_t^2 \, dt$. Если $\mathcal{M}_T < \infty$, то $\mathcal{M}_T \to \mathcal{M}_\infty$, где $\mathcal{M}_\infty$ - некоторая случайная величина. В противном случае, применяя утверждение из [10], имеем $\frac{|\mathcal{M}_T|}{< \mathcal{M}_T >} \to 0$ почти наверное.

Заметим, что т.к. $r_T^2 \le c_1$ ($c_1 > 0$ - константа), то

$$< \mathcal{M}_T > = \int\limits_0^T f_t\, r_t^2\, dt \le c_1 \int\limits_0^T f_t\, dt \quad \text{и} \quad \frac{1}{< \mathcal{M}_T >} \ge \mathcal{N}(T),$$

из чего следует, что $\mathcal{N}_T \mathcal{M}_T \to 0$ с вероятностью единица. В совокупности с оценкой (28) и с применением Леммы это доказывает утверждение Теоремы. $\blacksquare$

4. Предмет дальнейших исследований

В работе была рассмотрена модель управления равновесной ценой на рынке, где спрос и предложение описываются изоэластичными функциями с мультипликативной неопределенностью. На практике коэффициенты эластичности также оказываются неизвестными параметрами, можно лишь делать предположения об их распределении. Для учета этого обстоятельства в дальнейшем можно рассмотреть ситуацию, когда эластичности спроса и предложения являются случайными величинами. Функционал качества (9) представляет совокупные потери за плановый период, однако, в ряде ситуаций необходимо отдельным образом учесть соотношение между равновесной ценой P_T и траекторией P_T^0 в терминальный момент времени, например, когда в конце горизонта планирования большая разница между ними нежелательна. Это можно сделать путем включения в критерий качества слагаемого $\left(\ln \dfrac{P_T^0}{P_T}\right)^2$ и исследовать систему управления уже с таким целевым функционалом.

Список литературы

[1] Koekebakker S., Sodal S. The value of an operating electricity production unit. Working Paper 14/2001. Agder University College, 2001.

[2] Wengler J. Managing Energy Risk: A Nontechnical Guide to Markets and Trading. Pennwell Books, 2001, 250 p.

[3] Methods to analyse agricultural commodity price volatility, edited by Isabelle Piot-Lepetit. Springer, 2011, 300 p.

[4] Sengupta J. K. Optimal Stabilization Policy with a Quadratic Criterion Function. The Review of Economic Studies, Vol. 37, No. 1, 1970

[5] Паламарчук Е. С. Управление процессом сходимости цены к равновесному значению при наличии случайных факторов. Сборник Анализ и моделирование экономических процессов. М: ЦЭМИ, 2010, Вып. 7

[6] Optimal pricing, inflation, and the cost of price adjustment, edited by Eytan Sheshinski,Yoram Weiss. MIT Press, 1993, 534 p.

[7] Turvey C. G., Stoke J. Market Structure and the Value of Agricultural Contingent Claims, Canadian Journal of Agricultural Economics, Vol. 56, Issue 1, 2008

[8] Monahan G. E, Petruzzi N. C., Wen Zhao. The Dynamic Pricing Problem from a Newsvendor's Perspective. Manufacturing & Service Operations Management, Vol. 6, Issue 1, 2004

[9] Talluri K. T, Van Ryzin G. Theory and practice of revenue management, Springer, 2005, 714 p.

[10] Белкина Т. А., Кабанов Ю. М, Пресман Э. Л. О стохастической оптимальности для линейно-квадратического регулятора. - Теория вероятностей и ее применения, 2003

Л.Я.Клеппер

ВЕРОЯТНОСТИ ОСЛОЖНЕНИЙ В ОРГАНАХ И ТКАНЯХ ПРИ ТЕРАПИИ С НЕОДНОРОДНЫМ ОБЛУЧЕНИЕМ

В радиационной онкологии лучевым терапевтам приходится иметь дело, в основном, с неоднородными распределениями дозы в облучаемых органах и тканях организма. Это затрудняет оценку планов лучевой терапии и выбор эффективного варианта лечения на множестве альтернатив. В настоящее время растет интерес к описанию неоднородных дозовых полей в виде дифференциальных и интегральных гистограмм доза-объем, которые, по сути дела, можно рассматривать как определенным образом упорядоченные векторные характеристики неоднородных дозовых распределений [1-7]. К сожалению, использовать эти характеристики дозового поля для оценки и сравнения альтернативных планов лучевой терапии для выбора эффективного варианта облучения можно только в отдельных случаях.

Один из возможных путей решения этой проблемы заключается в свертке векторных характеристик неоднородного дозового поля в адекватные дозы однородного облучения тканей, которые приводят к таким же значениям вероятностей лучевых осложнений, как и анализируемые неоднородные дозовые поля. Адекватные дозы являются скалярными оценками лучевого воздействия на ткани и дают возможность использовать математические модели, которые были разработаны для оценки однородных дозовых полей. Поэтому разработка процедур свертки векторных характеристик в скалярные представляет актуальную проблему современной онкорадиологии, решение которой позволит повысить эффективность планирования лучевой терапии и результатов лечения.

Большой вклад в развитие нового научного направления в радиационной биофизике, связанного с оценкой неоднородных лучевых воздействий на животных (и человека), внесли работы Кейрим-Маркуса и его учеников [1]. Они первыми предложили метод свертки неоднородных дозовых распределений в эквивалентные (по критерию вероятности кроветворной гибели животных) однородные дозовые распределения. Дозу однородного облучения организма, которая приводит к такой же вероятности гибели животного, как и рассматриваемое неоднородное распределение дозы, они предложили назвать равноценной дозой. В результате обсуждения вопросов терминологии с И.Б.Кейрим-Маркусом, в своих работах мы стали использовать термин "адекватная доза" [3-7].

Цель настоящей работы – дать систематическое описание и анализ основных наработанных к настоящему времени математических моделей для расчета вероятностей осложнений при однородном и неоднородном способах облучения органов и тканей человека.

1. Вероятности осложнений при однородном облучении

Для описания Вероятности Лучевых Осложнений (аббревиатура ВЛО) в тканях организма при однородном облучении, мы предложили [3,8] воспользоваться модифицированным распределением вероятностей Вейбулла [9]. Для фиксированного объема облученной ткани и схемы фракционирования дозы, функция распределения вероятности осложнений в зависимости от дозы имеет следующий вид:

$$P(D) = 1 - Q(D) \qquad Q(D) = \exp\left[-\left(\frac{D}{A_1}\right)^{A_2}\right] \quad , \tag{1}$$

где D – доза однородного облучения ткани; A_1, A_2 – параметры модели; отметим, что *доза* – это энергия облучения, приходящаяся на единицу объема ткани [1], единица измерения – "грей".

Определенная в (1) величина $Q(D)$ есть Вероятность Отсутствия Лучевого Осложнения (аббревиатура ВОЛО).

В соответствии с теорией надежности систем [10], *интенсивность отказов* $r(D)$ тканевой системы определяется выражением

$$r(D) = \frac{d}{dD}[-\ln(Q(D))] = \frac{A_2}{A_1} \cdot \left(\frac{D}{A_1}\right)^{A_2-1}$$

[1]Строго говоря – на единицу массы; мы рассматриваем однородную ткань и переходим от массы к объему. Объем может иметь произвольную размерность $n = 1, 2, 3$.

и должна быть **неубывающей функцией дозы** D; из этого условия следует, что параметр A_2 должен быть больше единицы. В действительности, как показали практические расчеты, для нормальных органов и тканей организма $A_2 > 2$. Поясним, что величина $r(D)dD$ есть вероятность "отказа" тканевой системы при ее однородном облучении дозой, лежащей в интервале $[D, D + dD]$.

1.1. *Толерантная доза и ее обобщение. Толерантная доза* – это доза однородного облучения органов и тканей, которая приводит к ВЛО, не превышающей 5%. Это понятие выработано коллективным опытом лучевых терапевтов. Рассматривая графики, которые описывают ВЛО в органах и тканях как функцию от дозы, можно понять причину, по которой радиологи стали выделять "толерантную дозу". Графики ВЛО имеют логистический вид; в области малых доз (до толерантного уровня) - они пологие, а затем достаточно резко возрастают, образуя "пороги". Эти пороги были установлены лучевыми терапевтами и названы толерантными дозами. В настоящее время толерантная доза (точнее, интервал, в котором она лежит) как функция от объема V установлена практически для всех нормальных органов и тканей. При планировании лучевой терапии толерантные дозы стали выступать как ограничения на терапевтические поля, позволяя решать основную задачу лучевого лечения: создать в очаге опухолевого заболевания терапевтическое дозовое поле, необходимое для его необратимой регрессии, без серьезных (необратимых) лучевых воздействий на нормальные органы и ткани организма.

В наших работах [3-8] введено обобщенное понятие "толерантная доза уровня" α : T_α. По определению, доза T_α находится из условия $P(D)|_{D=T_\alpha} = \alpha$; согласно (1),

$$T_\alpha = A_1[-\ln(1-\alpha)]^{1/A_2} \quad ; \qquad (2)$$

при α=0.05 это будет традиционная доза.

1.1. *Зависимость обобщенной толерантной дозы от объема облученной ткани.* Опыт лучевой терапии свидетельствует, что основным параметром, определяющим толерантную дозу, является объем V облучаемой ткани. В упомянутых наших работах принято следующее предположение, описывающее зависимость $T_\alpha(V)$.

Предположение 1. Толерантная доза представима в следующей форме:

$$T_\alpha(V) = C(\alpha)V^{-b} \qquad \forall\,(\alpha \in (0,1)\;,\; V > 0) \quad ; \tag{3}$$

здесь $b \in (0,1)$ – параметр, зависящий от вида ткани; $C(\alpha)$, $\alpha \in (0,1)$ – положительная, возрастающая функция. Это равносильно предположению, что в формуле (1) параметр $A_2 > 1$ не зависит от V, а параметр A_1 связан с объемом V соотношением:

$$A_1(V) = c1 \cdot V^{-b} \quad , \tag{4}$$

где $c1 > 0$ – константа; при этом

$$C(\alpha) = c1 \cdot [-\ln(1-\alpha)]^{1/A_2} \quad . \tag{5}$$

Параметры $c1, A_2, b$ мы называем *характеристическими* для облучаемых тканей. ■

2. Описание неоднородных дозовых полей с помощью гистограмм

При планировании объемной лучевой терапии трудности, связанные с визуальной оценкой и сопоставлением неоднородных объемных дозовых распределений, привели к их описанию в виде интегральных или дифференциальных гистограмм доза-объем (ИГ, ДГ) [1-7].

2.1. *Фрагментация и гистограммы.* Предположим, что облучаемая область ткани G объема V разделена на m равнообъемных элементарных областей (*фрагментов*) таким образом, что распределение дозы в каждом из фрагментов можно считать однородным (с заданной точностью). Упорядочим фрагменты индексом $j = 1, \ldots, m$; объем каждого фрагмента равен $v = V/m$, а значение дозы на фрагменте j обозначим d_j. Описанное разбиение назовем *фрагментацией.*

Пусть $D_{\text{мин}}, D_{\text{макс}}$ – минимальное и максимальное значения дозы в облучаемой области. Разделим дозовый интервал $[D_{\text{мин}}, D_{\text{макс}}]$ на N равных частей $i = 1, \ldots, N$ с шагом $\Delta D = (D_{\text{макс}} - D_{\text{мин}})/N$; для каждого i максимальным значением дозы будет $D_i = D_{\text{мин}} + i\Delta D$. Значению

D_i сопоставим область G_i, включающую в себя те фрагменты j, в которых доза облучения d_j не превосходит D_i; количество таких фрагментов обозначим $J(i)$, и тогда объем области G_i составит величину $V_i = J(i)v$. Множество пар $\{(D_i, V_i)|\ i = 1, \ldots, N\}$ образует *интегральную гистограмму* доза-объем (построенную на данной фрагментации).

Ясно, что расширяющиеся области $G_1, \ldots, G_m$ последовательно вложены друг в друга, причем $G_N = G$; очевидно также, что значение дозы на области G_i (обозначим его d^i, индекс **вверху**) возрастает по i. Таким образом, любое неоднородное дозовое поле преобразуется в дозовое поле с вложенными областями G_i с возрастающими значениями дозы d^i. В частности, если область G шаровая радиуса R, областями G_i могут быть вложенные шары.

Дифференциальная гистограмма получается из интегральной путем "разностного дифференцирования": полагаем

$$v_i = V_{i+1} - V_i \qquad i = 1, \ldots, (N-1) \quad .$$

Отметим, что значения v_i неотрицательны, причем некоторые из них могут быть нулевыми. Множество пар $\{(D_i, v_i)|\ i = 1, \ldots, (N-1)\}$ образует *дифференциальную гистограмму* доза-объем.

Зная ДГ, можно без труда восстановить ИГ (путем суммирования).

2.2. *Свойства гистограмм.* Рассмотрим вопрос о том, насколько характеристики дозового поля, получаемые на основе гистограмм, отличаются от характеристик реального дозового поля. Анализ показывает, что:

1. Гистограммы упорядочивают неоднородное дозовое поле, но при этом реальная пространственная структура поля нивелируется.

2. При описании дозового поля с помощью гистограмм исчезает информация о многоэкстремальности дозового поля, хотя она может играть важную роль при оценке дозовых полей (например, как это имеет место при внутритканевой лучевой терапии).

3. ИГ и ДГ можно взаимно и однозначно преобразовать одну в другую.

4. Процедура обработки реального дозового поля и замена его на гистограммы необратима. Невозможно, используя гистограммы, восстановить пространственную структуру реального дозового поля.

5. Актуальной проблемой клинической радиобиологии является обоснование утверждения о том, что однородные и неоднородные дозовые поля в одной и той же ткани, имеющие одинаковые гистограммы, имеют равные (или близкие) значения ВЛО, т.е. эквивалентны по ВЛО.

ИГ и ДГ являются упорядоченными векторными характеристиками дозового поля, которые дают представление о характере неоднородного распределения дозы в ткани. Однако, на наш взгляд, применение гистограмм для анализа, сопоставления и выбора, на множестве альтернативных планов, рационального плана облучения пациента неоднозначно и проблематично.

3. Расчет ВЛО при неоднородном облучении на основе понятия "адекватная доза"

Понятие адекватная доза связано с принятием следующего предположения, дополняющего принятое выше Предположение 1.

Предположение 2. При неоднородном облучении ткани ВЛО дается формулой (1), в которой параметры A_1, A_2 зависят от объема V в соответствии с Предположением 1, а значение $D = D_{\text{ад}}$ является, **по определению**, *адекватной дозой*, которая подходящим образом строится для заданного неоднородного дозового поля; характеристические параметры $c1, A_2, b$ зависят только от вида ткани (но не зависят от формы облучаемой области и распределения неоднородного дозового поля). ■

Таким образом, адекватная доза является эквидозиметрической скалярной величиной (Кейрим-Маркус [1]), играющей ключевую роль при вычислении ВЛО. Вопрос о том, как строить адекватную дозу, теоретически не решен до сих пор; на практике $D_{\text{ад}}$ строится обычно на основе фрагментации исходного дозового поля, описанной в п. 2.1, по формуле

$$D_{\text{ад}} = \left[\frac{1}{m}\sum_{j=1}^{m} d_j^{A_2}\right]^{1/A_2} \quad . \tag{6}$$

Очевидно, что если облучение однородно с дозой D, то $D_{\text{ад}} = D$.

Формула (6) соответствует следующему предположению.

Предположение 3. Пусть на области ткани G объема V задано неоднородное дозовое поле с плотностью облучения (на единицу объема размерности n, см. сноску 1) $\delta(x)$, $x \in G$; этому полю отвечает адекватная доза

$$D_{\text{ад}} = \left[\frac{1}{V} \int\limits_{x \in G} \delta^{A_2}(x) dx \right]^{1/A_2} \quad . \quad \blacksquare \tag{7}$$

4. Математические свойства предложенных моделей ВЛО и АД

Следующие утверждения характеризуют математические свойства моделей, отвечающих Предположениям 1-3.

4.1. *Оптимальность однородного поля.*

Утверждение 1. Пусть G – облучаемая область произвольной формы в ткани с заданными параметрами A_1, A_2. Для фиксированного значения средней дозы облучения $D_{\text{ср}}$ наименьшее значение ВЛО достигается при однородном облучении; любое неоднородное облучение приводит к возрастанию ВЛО.

Доказательство. Пусть область G облучается с некоторой дозовой плотностью $\delta(x)$, удовлетворяющей условию

$$\frac{1}{V} \int\limits_{x \in G} \delta(x) dx = D_{\text{ср}} \quad . \tag{8}$$

Т.к. $A_2 > 1$, функция $f(\delta) = \delta^{A_2}$ выпукла [2], поэтому

$$D_{\text{ад}}^{A_2} = \frac{1}{V} \int\limits_{x \in G} \delta^{A_2}(x) dx \geq D_{\text{ср}}^{A_2} \quad ,$$

причем равенство достигается только при однородном облучении. Согласно Предположениям 2,3 для ВЛО имеем $P(D_{\text{ад}}) \geq P(D_{\text{ср}})$, что и утверждается. $\blacksquare$

[2]По другой терминологии *выпукла вниз.*

4.2. *Ситуация локальной независимости*. Предположение 3 позволяет ввести для неоднородного дозового поля понятие *эффективная дозовая плотность*; ее удобно определить формулой

$$\gamma(x) = \left(\frac{\delta(x)}{c1}\right)^{A_2} \quad , \tag{9}$$

где $c1$ – характеристический параметр формулы (4). Согласно (7), адекватную эффективную дозу можно записать в виде

$$\Gamma = \frac{1}{V}\int_{x\in G} \gamma(x)dx \quad , \tag{10}$$

а ВОЛО Q в виде

$$Q = \exp(-\Gamma V^{A_2 b}) \quad . \tag{11}$$

Утверждение 2. Пусть параметры ткани b, A_2 связаны соотношением $bA_2 = 1$. Тогда локальные (в частях ткани бесконечно малого объема dx) лучевые осложнения являются независимыми случайными событиями с плотностью вероятности $\gamma(x)$.

Доказательство. При данном условии ВОЛО (11) принимает вид

$$Q = \exp(-\Gamma V) = \exp\left[-\int_{x\in G} \gamma(x)dx\right] \quad ; \tag{12}$$

эта формула и составляет смысл доказываемого утверждения. ∎

4.3. *Точка перегиба.*

Утверждение 3. График функции $P(D)$, описывающей ВОЛО согласно (1), имеет (единственную) точку перегиба

$$\hat{D} = A_1(A_2 - 1)^{1/A_2} > 0 \quad , \tag{13}$$

слева от которой функция P выпукла вниз, а справа – вверх.

Доказательство. Точка перегиба определяется условием $P''(D) = 0$ (штрих означает производную по D). Знак $P''(D)$ совпадает со знаком выражения

$$(A_2 - 1) - \left(\frac{D}{A_1}\right)^{A_2} \quad ,$$

которое слева от точки (13) положительно, а справа отрицательно, что соответствует доказываемому утверждению. ■

4.4. *Возрастание адекватной дозы по параметру* A_2. Следующее важное свойство адекватной дозы обнаружено автором и по его просьбе доказано М.Л.Бронштейном [12]; приводимое ниже доказательство следует оригиналу, несколько упрощая его.

Утверждение 4. Адекватная доза, определяемая формулой (7), возрастает по параметру A_2.

Доказательство. Достаточно показать, что $\ln D_{\text{ад}}$ возрастает по параметру. Обозначив для удобства параметр A_2 буквой a, рассмотрим, исходя из (7), функции

$$F(a) = \ln D_{\text{ад}} = \frac{1}{a}[L(a) - \ln V] \qquad L(a) = \ln\left[\int\limits_{x \in G} \delta^a(x) dx\right] \quad . \qquad (14)$$

Дифференцируя по a, получаем

$$F'(a) = -\frac{1}{a^2}\{[L(a) - \ln V] - aL'(a)\} \quad ;$$

надо показать, что $F'(a) \geq 0$, т.е.

$$f(a) = aL'(a) - [L(a) - \ln V] \geq 0 \quad .$$

Т.к. $f(0) = 0$, достаточно показать, что

$$f'(a) = aL''(a) \geq 0 \quad \sim \quad L''(a) \geq 0 \quad ; \qquad (15)$$

покажем это. Обозначив интеграл в (14) буквой I, находим

$$L'(a) = \frac{1}{I} \int\limits_{x \in G} \delta^a(x) \ln \delta(x) dx$$

$$L''(a) = \frac{1}{I^2}\left[-\left(\int\limits_{x\in G}\delta^a(x)\ln\delta(x)dx\right)^2 + I\cdot\int\limits_{x\in G}\delta^a(x)\ln^2\delta(x)dx\right] \quad . \quad (16)$$

Требуемая, согласно (15), неотрицательность квадратной скобки в (16) следует из интегрального неравенства Коши-Буняковского ([13, Глава 2, §1])

$$\left(\int\limits_{x\in G}u(x)v(x)dx\right)^2 \leq \int\limits_{x\in G}u^2(x)dx\cdot\int\limits_{x\in G}v^2(x)dx$$

применительно к функциям

$$u(x) = \delta^{a/2}(x) \quad , \quad v(x) = \delta^{a/2}(x)\ln\delta(x) \qquad x\in G \quad . \quad \blacksquare$$

5. Обсуждение и практические выводы

Описанные модели расчета ВЛО на основе понятия "адекватная доза" открывают замечательную возможность изучения важной для радиобиологии и радиационной медицины проблемы оптимизации дозовых полей. Изложенные результаты позволяют сделать некоторые выводы, полезные для практики лучевой терапии; эти выводы касаются как нормальных, так и опухолевых тканей.

5.1. *Нормальные ткани.* При лучевой терапии облучению подвергаются не только пораженная область (опухоль, мишень), но и, с неизбежностью, прилегающие нормальные ткани. Терапевт стремится добиться резкого снижения дозы за пределами мишени, чтобы минимизировать охваченную облучением область нормальной ткани. Из Утверждения 1 следует, что нужно стараться при этом, чтобы наружное (за пределами мишени) облучение было по возможности однородным.

Таким образом, для уменьшения лучевой нагрузки на нормальные органы и ткани организма необходимо стремиться к минимизации в них дозы облучения (как средней $D_{\text{ср}}$, так и адекватной $D_{\text{ад}}$) и к ее равномерному распределению.

5.2. *Опухолевые ткани*. Опыт показал, что формула (1) может быть использована для описания вероятности локального излечения опухолевого заболевания (ВЛИОЗ). Об этом свидетельствуют графики зависимости ВЛИОЗ от дозы, построенные в результате обработки систематизированной клинической информации. Они имеют логистический вид. Формулу (1) можно модифицировать и для расчета зависимости ВЛИОЗ от объема опухолевой ткани и дозы облучения. Представляют интерес теоретические выводы, которые можно сделать из Утверждения 1 применительно к опухолевым тканям. Стремление к формированию однородного дозового поля в очаге опухолевого заболевания продиктовано необходимостью избежать рецидива опухолевого заболевания в областях минимума дозы и образования лучевых некрозов в областях максимума дозы. В современной радиологии требования к однородности дозового поля в объеме мишени достаточно высоки. Согласно рекомендациям Международной комиссии по радиационным единицам [11] однородность дозового поля необходимо поддерживать в пределах [-5%,+7%].

5.3. *Контактные методы лучевой терапии*. В настоящее время существует контактная лучевая терапия, которая характеризуется чрезвычайно неоднородным дозовым распределением в опухоли. Опыт применения контактных методов свидетельствует, что неоднородное терапевтическое дозовое поле в опухоли может быть весьма эффективным. А это означает, что медико-биологические критерии оптимизации терапевтического дозового поля при лечении опухоли не разработаны в должной мере; проблема радиобиологии заключается в том, чтобы установить, каковы же должны быть эти критерии.

Из Утверждения 1 следует, что если опухоль можно рассматривать как ткань (со всеми присущими ткани свойствами), а не как простое, не связанное в тканевую систему множество клеток, тогда, при фиксированной интегральной дозе, эффективным должно быть неоднородное распределение дозы в опухоли, поскольку оно будет приводить к большей величине $D_{\text{ад}}$ и, следовательно, к большей вероятности локального излечения. Это справедливо в той мере, насколько опухоль проявляет тканевые свойства. Насколько нам известно, вопрос о мере связности опухолевых клеток в тканевую систему и ее влиянии на эффективность лучевого лечения в таком аспекте еще не рассматривался. Косвенным указанием на то, что опухоль можно рассматривать как организованное в ткань множество клеток, служит тот факт, что графики реальных кривых, которые описывают ВЛИОЗ в зависимости от дозы, являются достаточно пологими. Если бы опухоль являлась структурно не связанным множеством клеток, тогда ВЛИОЗ была бы круто (почти мгновенно) возрастающая функция дозы.

Заключение

В статье систематизированы исследования автора по разработке математических моделей, которые могут быть использованы для оценки вероятности лучевых осложнений при радиационных методах лечения опухолевых заболеваний. Значимость проблемы оценки ВЛО для корректного планирования лучевой терапии трудно переоценить. Дело заключается в том, что все наработанные в настоящее время методы оценки результатов лучевого воздействия на нормальные органы и ткани человека предполагают, что их облучения происходят однородными дозовыми полями, в то время как в действительности, в подавляющем большинстве случаев, поля неоднородны. Оценка параметров математической модели происходит, по сути дела, на клиническом материале, соответствующем неоднородным дозовым полям, а затем эти параметры используются для оценки лучевых воздействий на органы и ткани в предположении, что дозовые поля в них однородны. Поэтому используемые методы имеют лишь приближенный характер.

Ситуация могла бы коренным образом измениться, если бы оценка ВЛО на основе адекватной дозы соответствовала реальной действительности. Живой организм является чрезвычайно сложной системой и проблема перехода от неоднородных дозовых полей к адекватным дозам требует серьезных совместных усилий лучевых терапевтов, медицинских физиков и математиков.

****

Автор выражает искреннюю признательность редактору сборника В.З.Беленькому за полезные замечания, улучшившие изложение материала.

Литература

1. Кейрим-Маркус И.Б. Эквидозиметрия. – М.: Атомиздат, 1980.

2. Lyman J.T. Complication probability as assessed from dose volume histograms,. *Rad. Res.*, 1985, v. 104.

3. Клеппер Л.Я. Формирование дозовых полей радиоактивными источниками излучения. – М.: Энергоатомиздат, 1993.

4. Клеппер Л.Я. Неоднородные дозовые распределения в нормальных органах и тканях организма в лучевой терапии злокачественных опухолей. *Медицинская техника*, 1999, № 5.

5. Клеппер Л.Я. Проблема равноценной дозы в лучевой терапии злокачественных опухолей. *Медицинская радиология и радиационная безопасность*, 1999, № 6(44).

6. Клеппер Л.Я. Методы перехода от неоднородного распределения дозы в ткани к адекватному однородному облучению ткани. *Медицинская техника*, 2001, № 2.

7. Клеппер Л.Я. Дифференциальные гистограммы доза - объем, их "свертка" в Адекватные Дозы эквивалентного однородного облучения тканей и лучевая терапия злокачественных опухолей. *Медицинская техника*, 2008, № 4.

8. Клеппер Л.Я. Формирование дозовых полей дистанционными источниками излучения. – М.: Энергоатомиздат, 1986.

9. Veibull W. A statistical distribution function of wide applicability. *J. Appl. Mechanics*, 1951, v. 18.

10. Барлоу Р., Прошан Ф. Математическая теория надежности. – М.: Советское радио, 1969.

11. Назначение, протоколирование и отчет по фотонной лучевой терапии. 50 (62) доклад Международного комитета по радиационным единицам и измерениям (МКРЕ). Издание 1 сентября 1993, 7910 Woodmont Avenue Bethesda, Maryland 20814, USA.

12. Клеппер Л.Я., Паньшин Г.А., Сотников В.М, Бронштейн М.Л. О зависимости терапевтической дозы от объема опухоли. *Медицинская радиология и радиационная безопасность*, 2010, т. 55, № 2.

13. Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. – М.: Наука, 1972.

С.А. Смоляк

Об одном функциональном уравнении

При установлении общего вида критериев эффективности инвестиционных проектов в условиях вероятностной неопределенности понадобилось найти строго возрастающую функцию $f(x)$, удовлетворяющую функциональному уравнению:

$$f(x+y)=a(x)f(y)+a(y)f(x), \tag{1}$$

где $a(x)$ – некоторая функция.

Решение было впервые дано в статье [1] и затем, в более простой форме, изложено в [2, раздел 7.6]. В данной статье рассматривается n-мерное обобщение этой задачи и дается ее решение.

Задача

Пусть задано натуральное $n \geq 2$. Ищется такая (строго) монотонная (возрастающая или убывающая) функция $f(x)$ на вещественной оси, для которой тождественно по $(x_1, \dots, x_n)$ выполняется равенство

$$f(x_1+\ldots+x_n)=a_1(x_2,\ldots,x_n)f(x_1)+\ldots+a_n(x_1,\ldots,x_{n-1})f(x_n), \tag{2}$$

где $a_1, \dots, a_n$ - некоторые функции от n–1 переменных. ■

Решение

Пусть $(i_1, \ldots , i_n)$ – какая-то перестановка чисел $1, \dots, n$. Тогда в силу (2)

$$f\left(x_{i_1}+\ldots+x_{i_n}\right)=a_1\left(x_{i_2},\ldots,x_{i_n}\right)f\left(x_{i_1}\right)+\ldots+a_n\left(x_{i_1},\ldots,x_{i_{n-1}}\right)f\left(x_{i_n}\right).$$

Усреднив эти равенства по всем перестановкам $(i_1,\ldots,i_n)$, получим:

$$f(x_1+\ldots+x_n)=a(x_2,\ldots,x_n)f(x_1)+\ldots+a(x_1,\ldots,x_{n-1})f(x_n), \tag{3}$$

где $a(x_1,\ldots,x_{n-1})$ – среднее из $a_i\left(x_{i_1},\ldots,x_{i_{n-1}}\right)$ по всем i и по всем перестановкам $(i_1, \ldots , i_{n-1})$ чисел $1, \ldots , n$–1. Нетрудно убедиться, что функция a – *симметрична*, т.е. не изменяется при любых перестановках ее аргументов.

Таким образом, нам достаточно найти (строго) монотонную функцию $f(x)$ и симметричную функцию a, для которых выполняется равенство (3).

Положив в (3) $x_1=\ldots=x_n=0$, найдем: $f(0)=na(0,\ldots,0)f(0)$. Такое равенство возможно только если $f(0)=0$ или $f(0)\neq 0$, $a(0,\dots,0)=1/n$.

Рассмотрим обе эти ситуации последовательно.

А. $f(0)=0$. Здесь надо отдельно рассмотреть случаи $n=2$ и $n>2$.

А1. Для $n=2$ задача сводится к решению уравнения (1); решим ее.

Подставив в (1) $y=x$, найдем: $f(2x)=2a(x)f(x)$, откуда:

$$a(x)=\frac{f(2x)}{2f(x)}. \tag{4}$$

Далее, подставляя в (1) $y=0$, получаем:

$$f(x)=f(x+0)=a(x)f(0)+a(0)f(x)=a(0)f(x).$$

Такое равенство возможно только если $a(0)=1$.

Пусть d =1. Тогда в силу монотонности $f(d)\neq 0$; из (1) имеем:

$$f(x+d)=a(x)f(d)+a(d)f(x),$$

$$f(x+2d)=a(x)f(2d)+a(2d)f(x).$$

Исключив отсюда $a(x)$, получим:

$$f(d)f(x+2d)-f(2d)f(x+d)+\left[f(2d)a(d)-f(d)a(2d)\right]f(x)=0.$$

При $x=kd$ отсюда находим:

$$f(d)f((k+2)d)-f(2d)f((k+1)d)+\left[f(2d)a(d)-f(d)a(2d)\right]f(kd)=0.$$

Это значит, что величины $f(kd)$ удовлетворяют возвратному уравнению второго порядка с ненулевым старшим коэффициентом и граничным условием $f(0)=0$. Общее ненулевое решение такого уравнения имеет вид:

$$f(kd)=\begin{cases} c\left(\alpha^k-\beta^k\right), & \alpha\neq\beta, \alpha\neq 0, \beta\neq 0, \\ cdk\alpha^k, & \alpha=\beta\neq 0, \end{cases}$$

где α и β – корни соответствующего характеристического уравнения (возможно, комплексные), $c\neq 0$ – некоторая константа. При этом ни α, ни β не могут равняться нулю, иначе нулевыми окажутся все $f(kd)$.

Если обозначить $\lambda=\alpha^{1/d}$, $\mu=\beta^{1/d}$, то из полученного выражения для $f(kd)$ вытекает, что при всех x, кратных d, справедливо равенство:

$$f(x)=\begin{cases} c\left(\lambda^x-\mu^x\right), & \lambda\neq\mu; \\ cx\lambda^x, & \lambda=\mu. \end{cases} \tag{5}$$

Повторим теперь предыдущие рассуждения, приняв $d = 1/m$, где m – некоторое целое число. Мы получим, что значения $f(x)$ при всех x вида k/m, будут задаваться аналогичной формулой, возможно с другими значениями c, λ и μ. Однако при всех целых x обе формулы должны давать одинаковые значения $f(x)$, что возможно только тогда, когда c, λ и μ в обеих формулах совпадают. Поскольку m может быть выбрано произвольно, мы получаем, что формула (5) будет справедлива при всех рациональных значениях x. Но тогда она справедлива при всех x, поскольку функция $f(x)$ монотонная.

Заметим теперь, что функция (5) не будет монотонной, если величины λ и μ – комплексные или хотя бы одна из них – отрицательная. Не могут они оказаться и нулевыми. Отсюда следует, что λ и μ – положительны; будем считать, что $\lambda \geq \mu > 0$. Наконец, если обе величины λ и μ будут либо меньше 1, либо больше 1, то функция (5) не будет монотонной.

Легко проверить, что в противоположной ситуации $\lambda \geq 1 \geq \mu > 0$ она будет монотонной; при этом возможны два случая.

1) $\lambda = \mu = 1$. Здесь $f(x) = cx$, $c \neq 0$. Тогда из (4) находим $a(x) = 1$. Легко проверить, что эта пара функций удовлетворяет равенству (1).

2) $\lambda \geq 1 \geq \mu, \lambda \neq \mu$. Здесь $f(x) = c(\lambda^x - \mu^x)$ и $a(x) = \dfrac{f(2x)}{2f(x)} = \dfrac{\lambda^x + \mu^x}{2}$ в силу (4). Легко видеть, что эта пара функций удовлетворяет (1). ■ (A1)

A2. $n > 2$. Положив в (3) $x_3 = \ldots = x_n = 0$, находим:

$$f(x_1 + x_2) = a(x_2, 0, \ldots, 0) f(x_1) + a(x_1, 0, \ldots, 0) f(x_2).$$

Но это уравнение совпадает с (1), поэтому должно быть $f(x) = cx$ или $f(x) = c(\lambda^x - \mu^x)$. Проверим, что обе эти функции удовлетворяют (3). Если $f(x) = cx$, то (3) будет выполняться при $a(x_1, \ldots, x_{n-1}) = 1$. Если же $f(x) = c(\lambda^x - \mu^x)$, то имеет место равенство:

$$\lambda^{x_1 + \ldots + x_n} - \mu^{x_1 + \ldots + x_n} = (\lambda^{x_1} - \mu^{x_1})\lambda^{x_2 + \ldots + x_n} + (\lambda^{x_2} - \mu^{x_2})\lambda^{x_3 + \ldots + x_n}\mu^{x_1} + $$
$$+ (\lambda^{x_3} - \mu^{x_3})\lambda^{x_4 + \ldots + x_n}\mu^{x_1 + x_2} + \ldots + (\lambda^{x_n} - \mu^{x_n})\mu^{x_1 + \ldots + x_{n-1}}.$$

Отсюда следует, что

$$f(x_1 + \ldots + x_n) = f(x_1)\lambda^{x_2 + \ldots + x_n} + f(x_2)\lambda^{x_3 + \ldots + x_n}\mu^{x_1} + \ldots + f(x_n)\mu^{x_1 + \ldots + x_{n-1}}.$$

Это значит, что равенство (2) выполняется при

$$a_1(x_1, \ldots, x_{n-1}) = \lambda^{x_1 + \ldots + x_{n-1}}, \quad a_2(x_1, \ldots, x_{n-1}) = \lambda^{x_2 + \ldots + x_{n-1}}\mu^{x_1}, \ldots,$$
$$a_{n-1}(x_1, \ldots, x_{n-1}) = \lambda^{x_{n-1}}\mu^{x_1 + \ldots + x_{n-2}}, \quad a_n(x_1, \ldots, x_{n-1}) = \mu^{x_1 + \ldots + x_{n-1}}.$$

Как и в начале статьи, усредняя $a_i(x_{i_1}, \ldots, x_{i_{n-1}})$ по всем i и по всем перестановкам $(i_1, \ldots, i_{n-1})$ чисел $1, \ldots, n-1$, найдем соответствующую симметричную функцию a, при которой будет выполняться (3):

$$a(x_1, \ldots, x_{n-1}) = \frac{1}{n!}\sum_{k=0}^{n-1}\sum_{(i_1, \ldots, i_k)} \lambda^{x_{i_1} + \ldots + x_{i_k}} \mu^{x_1 + \ldots + x_{n-1} - (x_{i_1} + \ldots + x_{i_k})}.$$

Здесь внутренняя сумма распространяется на все наборы из k различных целых чисел от 1 до n–1 (при $k=0$ в ней будет только одно слагаемое).

Эту формулу можно записать и иначе:

$$a\left(x_1,...,x_{n-1}\right)=\frac{1}{n!}\sum_{k=0}^{n-1}\sum_{[i_1,...,i_k]}k!\left(n-1-k\right)!\lambda^{x_{i_1}+...+x_{i_k}}\mu^{x_1+...+x_{n-1}-\left(x_{i_1}+...+x_{i_k}\right)}. \quad (6)$$

Здесь внутренняя сумма распространяется уже на все *упорядоченные* наборы из k различных целых чисел от 1 до n (при $k=0$ в ней будет только одно слагаемое). Например, для $n=3$ эта формула дает:

$$a\left(x,y\right)=\frac{1}{6}\left[2\lambda^{x+y}+2\mu^{x+y}+\lambda^x\mu^y+\lambda^y\mu^x\right]. \blacksquare \quad \text{(A)}$$

Б. $f(0)\neq 0$, $a(0,\ldots,0)=1/n$.

Поскольку вместе с f решением задачи будет и cf при любом $c\neq 0$, то без ограничения общности можно считать, что $f(0)=1$.

Дальнейшие рассуждения проведем в несколько этапов.

Этап 1. Положив в (2) $x_1=x, x_2=...=x_n=0$, найдем:

$$f\left(x\right)=a\left(0,...,0\right)f\left(x\right)+\left(n-1\right)a\left(x,0,...,0\right)f\left(0\right)=\frac{1}{n}f\left(x\right)+\left(n-1\right)a\left(x,0,...,0\right).$$

Отсюда:

$$a\left(x,0,...,0\right)=\frac{1}{n}f\left(x\right). \quad (7)$$

В частности, при $n=2$ из (3) и (7) получим:

$$f\left(x_1+x_2\right)=\frac{1}{2}f\left(x_1\right)f\left(x_2\right)+\frac{1}{2}f\left(x_1\right)f\left(x_2\right)=f\left(x_1\right)f\left(x_2\right).$$

Любое монотонное решение этого уравнения при условии $f(0)=1$ будет экспонентой: $f\left(x\right)=\mu^x$, где $\mu>0$. Эта функция удовлетворяет (3), если в соответствии с (7) принять $a\left(x\right)=\frac{1}{2}\mu^x$. На следующих этапах будем рассматривать случай, когда $n>2$.

Этап 2. Положив в (3) $x_3=...=x_n=0$, с учетом (7) получим:

$$f\left(x_1+x_2\right)=a\left(x_2,0,...,0\right)f\left(x_1\right)+a\left(x_1,0,...,0\right)f\left(x_2\right)+$$

$$+\left(n-2\right)a\left(x_1,x_2,0,...,0\right)f\left(0\right)=\frac{2}{n}f\left(x_1\right)f\left(x_2\right)+\left(n-2\right)a\left(x_1,x_2,0,...,0\right).$$

Отсюда находим:

$$(n-2)a(x_1,x_2,0,...,0)=f(x_1+x_2)-\frac{2}{n}f(x_1)f(x_2)$$

или

$$a(x_1,x_2,0,...,0)=\frac{1}{n-2}f(x_1+x_2)-\frac{2}{n(n-2)}f(x_1)f(x_2). \qquad (8)$$

Этап 3. Положив в (3) $x_4=...=x_n=0$, с учетом (8) получим:

$$f(x_1+x_2+x_3)=a(x_2,x_3,0,...,0)f(x_1)+a(x_1,x_3,0,...,0)f(x_2)+$$
$$+a(x_1,x_2,0,...,0)f(x_3)+(n-3)a(x_1,x_2,x_3,0,...,0)f(0)=$$
$$=\frac{1}{n-2}\left[f(x_2+x_3)f(x_1)+f(x_1+x_3)f(x_2)+f(x_1+x_2)f(x_3)\right]-$$
$$-\frac{6}{n(n-2)}f(x_1)f(x_2)f(x_3)+(n-3)a(x_1,x_2,x_3,0,...,0).$$

Но тогда

$$(n-3)a(x_1,x_2,x_3,0,...,0)=f(x_1+x_2+x_3)-$$
$$-\frac{1}{n-2}\left[f(x_1+x_2)f(x_1)+f(x_1+x_3)f(x_2)+f(x_2+x_3)f(x_1)\right]+ \qquad (9)$$
$$+\frac{6}{n(n-2)}f(x_1)f(x_2)f(x_3).$$

Этап 4. На этом этапе мы получим окончательные результаты. Для этого нам понадобится ввести новые обозначения.

Зафиксируем множество m переменных $\{z_1,\ldots,z_m\}$ и обозначим через P_k ($0\le k<m$–1) некоторое разбиение этого множества на два подмножества, первое из которых содержит k элементов, а второе – все остальные: $\{z_1,...,z_m\}=\{z_{i_1},...,z_{i_k}\}\cup\{z_{i_{k+1}},...,z_{i_m}\}$. Поставим разбиению P_k в соответствие выражение

$$f(z_{i_1})...f(z_{i_k})f(z_{i_{k+1}}+...+z_{i_m}). \qquad (10)$$

Для суммы таких выражений по всем различным разбиениям P_k введем специальное обозначение:

$$f_k(z_1,...,z_m)=\sum_{P_k}f(z_{i_1})...f(z_{i_k})f(z_{i_{k+1}}+...+z_{i_m}).$$

В частности, для $k=0$ имеем: $f_0(z_1,...,z_m)=f(z_1+...+z_m)$.

Если формально применить введенное обозначение к случаю $k=m$–1, мы получим m различных разбиений P_{m-1}, каждому из которых будет поставлено в соответствие одно и то же выражение, так что окажется, что

$$f_{m-1}\left(z_1,...,z_m\right)=mf\left(z_1\right)...f\left(z_m\right).$$

Нам это неудобно, поэтому положим:

$$f_{m-1}\left(z_1,...,z_m\right)=f\left(z_1\right)...f\left(z_m\right).$$

Далее мы будем применять эти обозначения к разным множествам переменных (одновременно они будут аргументами функций f_k).

Заметим вначале, что имеют место следующие равенства:

$$f_k\left(z_2,...,z_m\right)f\left(z_1\right)+f_k\left(z_1,z_3,...,z_m\right)f\left(z_2\right)+...+f_k\left(z_1,...,z_{m-1}\right)f\left(z_m\right)=$$

$$=\begin{cases}(k+1)f_{k+1}\left(z_1,...,z_m\right), & k<m-2;\\ mf_{m-1}\left(z_1,...,z_m\right), & k=m-2.\end{cases} \qquad (11)$$

Действительно, при $k<m$-2 в левой части равенства каждое из выражений $f\left(z_{i_1}\right)...f\left(z_{i_{k+1}}\right)f\left(z_{i_{k+2}}+...+z_{i_m}\right)$, отвечающих некоторому разбиению P_k, встретится в левой части (11) ровно k+1 раз – в тех слагаемых, в которых аргументом второго множителем будет элемент первой группы выражения (10). Если же $k=m$-2, то в левой части (11) будет m одинаковых слагаемых $f\left(z_1\right)...f\left(z_m\right)$.

Рассмотрение формул (7), (8) и (9) позволяет предположить, что при всех $m \le n$ и некоторых значениях коэффициентов B_m^k имеет место равенство:

$$(n-m)a\left(x_1,...,x_m,0,...,0\right)=\sum_{k=0}^{m-1}B_m^k f_k\left(x_1,...,x_m\right) \ , \qquad (12)$$

при этом $B_1^0=\dfrac{n-1}{n}, B_2^0=1, B_2^1=-\dfrac{2}{n}; B_3^0=1, B_3^1=-\dfrac{1}{n-2}, B_3^2=\dfrac{6}{n(n-2)}$.

Равенство (12) мы докажем по индукции, одновременно определив и коэффициенты B_m^k. Допустим, что (12) справедливо при некотором m; докажем, что тогда оно верно и при $m=m+1$, т.е. что

$$(n-m-1)a\left(x_1,...,x_{m+1},0,...,0\right)=\sum_{k=0}^{m}B_{m+1}^k f_k\left(x_1,...,x_{m+1}\right).$$

Для этого, положив в (3) $x_{m+1}=...=x_n=0$, найдем:

$$f\left(x_1+...+x_{m+1}\right)=a\left(x_2,...,x_{m+1},0,...0\right)f\left(x_1\right)+...+a\left(x_1,...,x_m,0,...,0\right)f\left(x_{m+1}\right)+$$
$$+(n-m-1)a\left(x_1,...,x_{m+1},0,...,0\right)f(0).$$

Отсюда:

$$(n-m-1)a(x_1,...,x_{m+1},0,...,0)=f(x_1+...+x_{m+1})-$$
$$-a(x_2,...,x_{m+1},0,...)f(x_1)-...-a(x_1,...,x_m,0,...,)f(x_{m+1})=$$
$$=f_0(x_1,...,x_m)-\sum_{k=0}^{m-1}B_m^k\frac{f_k(x_2,...,x_{m+1})f(x_1)+...+f_k(x_1,...,x_m)f(x_{m+1})}{n-m}= \qquad (13)$$
$$=f_0(x_1,...,x_m)-\sum_{k=0}^{m-2}B_m^k\frac{(k+1)f_{k+1}(x_1,...,x_{m+1})}{n-m}-B_m^{m-1}\frac{(m+1)f_m(x_1,...,x_{m+1})}{n-m}.$$

Поэтому искомое равенство будет выполняться, если выполняются соотношения

$$B_{m+1}^0=1;\quad B_{m+1}^k=-\frac{k}{n-m}B_m^{k-1},(1\le k<m);\quad B_{m+1}^m=-\frac{m+1}{n-m}B_m^{m-1}\quad ,$$

из которых вытекают явные выражения для B_m^k:

- при $k<m$–1:

$$B_m^k=\left(-\frac{k}{n-m+1}\right)B_{m-1}^{k-1}=\left(-\frac{k}{n-m+1}\right)\left(-\frac{k-1}{n-m+2}\right)B_{m-2}^{k-2}=$$
$$...=\left(-\frac{k}{n-m+1}\right)\left(-\frac{k-1}{n-m+2}\right)...\left(-\frac{1}{n-m+k}\right)B_{m-k}^0=(-1)^k\frac{k!(n-m)!}{(n-m+k)!};$$

- при $k=m$–1:

$$B_m^{m-1}=\left(-\frac{m}{n-m+1}\right)B_{m-1}^{m-2}=\left(-\frac{m}{n-m+1}\right)\left(-\frac{m-1}{n-m+2}\right)B_{m-2}^{m-3}=$$
$$...=\left(-\frac{m}{n-m+1}\right)\left(-\frac{m-1}{n-m+2}\right)...\left(-\frac{3}{n-2}\right)B_2^1=(-1)^{m-1}\frac{m!(n-m)!}{n(n-2)!}.$$

Обратим теперь внимание на то, что равенство (13) будет справедливым и при $m=n$, если левую часть в нем считать равной нулю. Поскольку

$$B_n^k=(-1)^k,\quad (k<n-1);\quad B_n^{n-1}=(-1)^{n-1}(n-1),$$

то из (13) получаем:

$$0=\sum_{k=0}^{n-2}(-1)^k f_k(x_1,...,x_n)+(-1)^{n-1}(n-1)f_{n-1}(x_1,...,x_n). \qquad (14)$$

Положим здесь $x_1=x;\quad x_2=...=x_n=d$ и учтем, что $f_k(x_1,...,x_n)$ будет суммой величин $f(x),f(x+d),...,f(x+(n-k-1)d)$ с коэффициентами, зависящими от d. Тогда правая часть (14) окажется суммой величин $f(x),f(x+d),...,f(x+(n-1)d)$ с коэффициентами, зависящими от d.

Отсюда следует, что величины $f(md)$ удовлетворяют возвратному уравнению n-го порядка с постоянными коэффициентами. Но тогда для x, являющихся целыми кратными d, будет:

$$f(x) = \sum_i P_{r_i}(x)\mu_i^x, \tag{15}$$

где каждое $P_{r_i}(x)$ – многочлен от x степени r_i, а $\sum_i (r_i + 1) \le n$.

Повторяя рассуждения из п. AI, мы убедимся, что такое равенство будет справедливо и при всех x.

Докажем теперь, что все $r_i = 0$. Пусть это не так и, например, $r_1 > 0$. Если подставить (15) в $f_{n-1}(x_1,...,x_n) = f(x_1)...f(x_n)$, можно увидеть, что справа появится член $(x_1...x_n)^{r_1}\mu_1^{x_1+...+x_n}$ с некоторым ненулевым коэффициентом. Множитель перед $\mu_1^{x_1+...+x_n}$ при этом будет многочленом степени nr_1 от переменных x_i. Если же подставить (15) в $f_k(x_1,...,x_n)$ при $k < n-1$, раскрыть скобки и отобрать среди полученных слагаемых те, которые включают $\mu_1^{x_1+...+x_n}$, мы увидим, что стоящие перед ними множители будут многочленами от x_i меньшей степени. Например, в $f_k(x_1,...,x_n)$ входит слагаемое $f(x_1)...f(x_k)\,f(x_{k+1}+...+x_n)$. Если подставить сюда (15), то в полученном выражении найдется единственный подходящий член $(x_1...x_k)^{r_1}(x_{k+1}+...+x_n)^{r_1}\mu_1^{x_1+...+x_n}$ с некоторым ненулевым коэффициентом. Таким образом перед $\mu_1^{x_1+...+x_n}$ будет стоять многочлен от x_i степени $(k+1)r_1 < nr_1$. Отсюда следует, что в правую часть (14) входит член, пропорциональный $(x_1...x_n)^{r_1}\mu_1^{x_1+...+x_n}$ и не имеющий подобных. Но тогда правая часть (15) не может равняться нулю, что и требовалось доказать.

Итак, все $r_i = 0$, а тогда (15) примет вид:

$$f(x) = \sum_i P_i\mu_i^x, \tag{16}$$

где P_i – некоторые константы, количество которых не превосходит n. Более того, поскольку $f(0) = 1$, то сумма всех P_i равна 1.

Если ненулевое P_i только одно, то оно равно 1. При этом функция $f(x) = \mu^x$ при $\mu > 0, \mu \ne 1$ монотонна и удовлетворяет (3), если принять, что

$$a(x_1,...,x_{n-1}) = \frac{1}{n}\mu^{x_1+...+x_{n-1}}. \tag{17}$$

Рассмотрим случай, когда в (16) имеются два или более ненулевых P_i.

Пусть $P_1 \neq 0$, $P_2 \neq 0$. Подставим (16) в правую часть (14), раскроем все скобки и рассмотрим только члены, содержащие $\mu_1^{x_2+\ldots+x_n}\mu_2^{x_1}$. Найдем их.

В выражении $f_{n-1}(x_1,\ldots,x_n) = f(x_1)\ldots f(x_n)$ соответствующий член будет один: $\left(P_2\mu_2^{x_1}\right)\left(P_1\mu_1^{x_2}\right)\ldots\left(P_1\mu_1^{x_n}\right) = P_2P_1^{n-1}\mu_1^{x_2+\ldots+x_n}\mu_2^{x_1}$. Далее,

$$f_{n-2}(x_1,\ldots,x_n) = f(x_1+x_2)\,f(x_3)\ldots f(x_n)+\ldots+f(x_{n-1}+x_n)\,f(x_1)\ldots f(x_{n-2}).$$

Здесь $\mu_1^{x_2+\ldots+x_n}\mu_2^{x_1}$ появляется по одному разу в каждом из слагаемых, содержащих $f(x_1)$, при этом – с одним и тем же коэффициентом $P_2P_1^{n-2}$. Количество таких слагаемых равно C_{n-1}^2 – числу способов выбора двух элементов из набора $\{x_2,\ldots,x_n\}$.

Аналогично при $0<k<n-1$ найдем, что при раскрытии $f_k(x_1,\ldots,x_n)$ появятся C_{n-1}^{n-k} одинаковых членов $P_2P_1^k\mu_1^{x_2+\ldots+x_n}\mu_2^{x_1}$, а при раскрытии $f_0(x_1,\ldots,x_n)$ подобных членов не будет совсем.

Но сумма всех рассматриваемых членов равна 0, так что

$$0 = \sum_{k=1}^{n-2}(-1)^k C_{n-1}^{k-1}P_2P_1^k + (-1)^{n-1}(n-1)P_2P_1^{n-1} =$$

$$= -P_1P_2\left[\sum_{i=0}^{n-3}C_{n-1}^i(-P_1)^i + C_{n-1}^{n-2}(-P_1)^{n-2}\right] =$$

$$= -P_1P_2\left[(1-P_1)^{n-1} - C_{n-1}^{n-1}(-P_1)^{n-1}\right] = -P_1P_2\left[(1-P_1)^{n-1} - (-P_1)^{n-1}\right].$$

В таком случае $(1-P_1)^{n-1} = (-P_1)^{n-1}$, поскольку $P_1 \neq 0$, $P_2 \neq 0$. При *четном n* такое равенство невозможно вообще, а при *нечетном* – возможно только, когда $1-P_1=P_1$, т.е. при $P_1=1/2$. Аналогично можно убедиться, что все ненулевые $P_i=1/2$. Но сумма всех P_i равна 1, так что ненулевых P_i должно быть ровно два. В таком случае (16) принимает вид: $f(x) = \left(\lambda^x + \mu^x\right)/2$.

Ранее мы предположили, что $f(0)=1$. Если отказаться от этого предположения, то решением задачи в рассмотренной ситуации могут быть функции вида $f(x) = c\left(\lambda^x + \mu^x\right), c \neq 0$. Все они удовлетворяют (3). Действительно, при *нечетных n* справедливо равенство:

$$\lambda^{x_1+\ldots+x_n} + \mu^{x_1+\ldots+x_n} = \left(\lambda^{x_1}+\mu^{x_1}\right)\lambda^{x_2+\ldots+x_n} - \left(\lambda^{x_2}+\mu^{x_2}\right)\lambda^{x_3+\ldots+x_n}\mu^{x_1} +$$

$$+\left(\lambda^{x_3}-\mu^{x_3}\right)\lambda^{x_4+\ldots+x_n}\mu^{x_1+x_2} - \ldots + \left(\lambda^{x_n}-\mu^{x_n}\right)\mu^{x_1+\ldots+x_{n-1}}.$$

Оно означает, что функция $f(x)=c\left(\lambda^x+\mu^x\right)$ удовлетворяет (2) при

$$a_1(x_1,...,x_{n-1})=\lambda^{x_1+...+x_{n-1}},\ a_2(x_1,...,x_{n-1})=-\lambda^{x_2+...+x_{n-1}}\mu^{x_1},$$
$$a_3(x_1,...,x_{n-1})=\lambda^{x_3+...+x_{n-1}}\mu^{x_1+x_2},\ ...,\ a_n(x_1,...,x_{n-1})=\mu^{x_1+...+x_{n-1}}.$$

Усредняя эти функции по всем перестановкам чисел 1, 2, …, n, найдем отвечающую им симметричную функцию $a(x_1,...,x_{n-1})$, при которой выполняется равенство (3):

$$a(x_1,...,x_{n-1})=\frac{1}{n!}\sum_{k=0}^{n-1}(-1)^{n-1-k}\sum_{(i_1,...,i_k)}\lambda^{x_{i_1}+...+x_{i_k}}\mu^{x_1+...+x_{n-1}-\left(x_{i_1}+...+x_{i_k}\right)}.$$

Здесь внутренняя сумма распространяется на все наборы из k различных целых чисел от 1 до n–1 (при k = 0 в ней будет только одно слагаемое).

Это же выражение можно записать и иначе:

$$a(x_1,...,x_{n-1})=\sum_{k=0}^{n-1}(-1)^{n-1-k}\frac{k!(n-1-k)!}{n!}\sum_{[i_1,...,i_k]}\lambda^{x_{i_1}+...+x_{i_k}}\mu^{x_1+...+x_{n-1}-\left(x_{i_1}+...+x_{i_k}\right)}. \qquad (18)$$

Здесь внутренняя сумма распространяется уже на все *упорядоченные* наборы из k различных целых чисел от 1 до n–1 (при k = 0 в ней будет только одно слагаемое).

В частности, при $n=3$ найдем:

$$a(x,y)=\frac{2\left(\mu^{x+y}+\lambda^{x+y}\right)-\left(\lambda^x\mu^y+\lambda^y\mu^x\right)}{6}.$$

Остается заметить, что функция $f(x)=c\left(\lambda^x+\mu^x\right)$ будет монотонной только при $\lambda>\mu\ge 1$ или при $1\ge\lambda>\mu>0$. ■ (Б)

Полученное решение задачи сведено в следующую таблицу.

Ситуация	$f(x)$	Ограничения	$a(x_1,..,x_n)$
А. $f(0)=0$	cx	$c\ne 0$	1
	$c\left(\lambda^x-\mu^x\right)$	$\lambda\ge 1\ge\mu>0, \lambda\ne\mu, c\ne 0$	формула (6)
Б. $f(0)>0$	$c\mu^x$	$\mu>0, \mu\ne 1, c\ne 0$	$\frac{1}{n}\mu^{x_1+...+x_{n-1}}$
	$c\left(\lambda^x+\mu^x\right)$	n – нечетное, $c\ne 0$, $\lambda>\mu\ge 1$ или $1\ge\lambda>\mu>0$	формула (18)

Литература

1. Смоляк С. А. Об учете разброса эффекта при расчетах экономической эффективности в условиях неопределенности // Модели и методы стохастической оптимизации. - М.: ЦЭМИ АН СССР, 1983.

2. Смоляк С.А. Оценка эффективности инвестиционных проектов в условиях риска и неопределенности. - М.: Наука, 2002.

В.З.Беленький

Биномиальные функции

В данной заметке на базе элементарной математики введен в рассмотрение класс функций (названных "биномиальными"), который, на взгляд автора, может представить определенный научный интерес.

1. Универсальное биномиальное тождество

1.1. Альтернирующие биномиальные коэффициенты. При целом $m \geq 1$ детерминант

$$det \begin{vmatrix} 1 & 1 & \ldots & 1 \\ \lambda_0 & \lambda_1 & \ldots & \lambda_m \\ \lambda_0^2 & \lambda_1^2 & \ldots & \lambda_m^2 \\ \ldots & \ldots & \ldots & \ldots \\ \lambda_0^m & \lambda_1^m & \ldots & \lambda_m^m \end{vmatrix} = \prod_{i=0}^{m} \prod_{j=i+1}^{m} (\lambda_j - \lambda_i) \tag{1}$$

называется определителем Вандермонда $W_m(\Lambda)$, где $\Lambda = (\lambda_0, \lambda_1, \ldots, \lambda_m)$ – некоторый упорядоченный набор чисел; дополнительно полагаем $W_0(\lambda_0) = 1 \quad \forall\, \lambda_0$.

Для произвольного натурального n рассмотрим систему из $(n+1)$ линейных уравнений относительно $(n+1)$ неизвестных $(z_0, \ldots, z_n)$

$$\begin{array}{llllllllll} z_0 + & z_1 & + & z_2 & + & \ldots & + & z_n & = 0 \\ & 1 \cdot z_1 & + & 2 \cdot z_2 & + & \ldots & + & n \cdot z_n & = 0 \\ & 1^2 \cdot z_1 & + & 2^2 \cdot z_2 & + & \ldots & + & n^2 \cdot z_n & = 0 \\ & \ldots\ \ldots & \ldots & \ldots\ \ldots & \ldots & \ldots & \ldots & \ldots\ldots\ \ldots & \ldots \\ & 1^{n-1} \cdot z_1 & + & 2^{n-1} \cdot z_2 & + & \ldots & + & n^{n-1} \cdot z_n & = 0 \\ & 1^n \cdot z_1 & + & 2^n \cdot z_2 & + & \ldots & + & n^n \cdot z_n & = (-1)^n \cdot n! \end{array} \tag{2}$$

.

Детерминант этой системы есть определитель Вандермонда $W_n(0,1,\ldots,n)$

$$det\begin{vmatrix} 1 & 1 & \ldots & 1 \\ 0 & 1 & \ldots & n \\ 0^2 & 1^2 & \ldots & n^2 \\ \ldots & \ldots & \ldots & \ldots \\ 0^n & 1^n & \ldots & n^n \end{vmatrix} \quad , \qquad (3)$$

который, согласно (1), равен

$$W_n := \prod_{i=0}^{n}\prod_{j=i+1}^{n}(j-i) \quad . \qquad (4)$$

Найдем решение системы (2); будем вычислять его по правилу Крамера $z_k = \Delta^k/\Delta$, $k = 0,\ldots,n$, где $\Delta = W_n$, а Δ^k есть определитель матрицы, получающейся из (3) заменой k-го столбца на вектор-столбец правой части (2). Раскладывая этот определитель по столбцу k (в котором только один ненулевой элемент – (n,k)), находим

$$\Delta^k = (-1)^n n! \cdot (-1)^{n-k} W_{n-1} = (-1)^k \cdot \tilde{W}^k \quad ,$$

где $\tilde{W}^k = W_{n-1}(\{0,1,\ldots,n\}/k)$ (из списка аргументов удален элемент k). Вычисляя определитель $\tilde{W}^k$ по формуле (1), видим, что что новое произведение отличается от (4) тем, что в нем отсутствуют члены вида $(j-k)$, $j = k+1,\ldots,n$, и вида $(k-i)$, $i = 0,\ldots,k-1$; произведение первых есть$(n-k)!$, вторых – $k!$ (с учетом стандартного соглашения $0! = 1$).

Таким образом, решением системы (2) являются числа

$$Z_k^n := (-1)^k \cdot \frac{n!}{(n-k)!\cdot k!} = (-1)^k \cdot C_k^n \qquad k = 0,\ldots,n \quad , \qquad (5)$$

т.е. **альтернирующие** биномиальные коэффициенты.

В силу соотношений (2) имеют место равенства

$$\sum_{j=0}^{n} Z_j^n j^i = \begin{cases} 0 & i = 0,\ldots,n-1 \\ (-1)^n \cdot n! & i = n \end{cases} \quad . \qquad (6)$$

1.2. Универсальное биномиальное тождество. Придадим полученному результату двойственную форму. Согласно (6), система однородных уравнений относительно $(n+1)$ неизвестных $\mathbf{X} = (x_0, \ldots, x_n)$

$$\sum_{j=0}^{n} Z_j^n x_j = 0 \tag{7}$$

имеет фундаментальную систему из n решений

$$\mathbf{X}^{(i)} := (0^i, 1^i, \ldots, n^i) \qquad i = 0, 1, \ldots, n-1 \quad ; \tag{8}$$

и любое решение $\mathbf{A} = (a_0, \ldots, a_n)$ системы (7) имеет вид

$$\mathbf{A} = \sum_{i=0}^{n-1} \lambda_i \mathbf{X}^{(i)} \quad ,$$

где λ_i – произвольные числа. Введя в рассмотрение многочлен

$$p(x) := \sum_{i=0}^{n-1} \lambda_i x^i \quad ,$$

имеем $a_j = p(j)$, $k = 0, \ldots, n$.

Подведем итог в следующей форме.

Теорема 1.

1) Если числа $a_0, a_1, \ldots, a_n$ удовлетворяют равенству

$$\sum_{j=0}^{n} Z_j^n a_j = 0 \quad , \tag{9}$$

то существует и единствен многочлен $p(x)$ степени не выше $(n-1)$ такой, что $a_j = p(j)$, $k = 0, \ldots, n$.

2) Обратно, для любого многочлена $p(x)$ степени не выше $(n-1)$ имеет место равенство

$$\sum_{j=0}^{n} Z_j^n p(j) = 0 \quad . \quad \blacksquare \tag{10}$$

Содержательный смысл Теоремы 1 в том, что всякое равенство вида (9) может быть представлено, при подходящем выборе многочлена $p(x)$, в **универсальной** форме (10), К примеру, известные равенства (см., например, [1, стр. 12])

$$\sum_{j=k}^{n} (-1)^j C_j^n C_k^j = 0 \qquad 0 \le k \le n-1 \tag{11}$$

получаются[1] из (10) при $p = \pi_k$, где система многочленов $\boldsymbol{\pi} := \{\pi_k\ ,\ k = 0, 1, \ldots\}$ дается формулами

$$\pi_0(x) \equiv 1 \quad , \quad \pi_k(x) = (-1)^k \frac{1}{k!} \cdot \prod_{i=0}^{k-1} (x - i) \qquad k \geq 1 \quad , \tag{12}$$

поскольку, очевидно,

$$\pi_k(j) = \begin{cases} 0 & j \leq k-1 \\ Z_k^j & j \geq k \end{cases} = \begin{cases} 0 & j \leq k-1 \\ (-1)^k C_k^j & j \geq k \end{cases} \quad . \tag{13}$$

Примечание. Конечно, множитель $(-1)^k$ в определении многочленов (12) можно убрать; он введен потому, что именно в такой форме эти многочлены используются в дальнейшем, см. Раздел 3. ■

В теоретическом плане систему соотношений (6) удобно представить в следующей форме.

Теорема 2. Если $p(x)$ – многочлен степени n со старшим коэффициентом λ, то имеет место тождество

$$\sum_{j=0}^{n} Z_k^n p(x-k) \overset{x}{\equiv} \lambda \cdot n! \quad . \tag{14}$$

Доказательство. При любом k многочлен $p(x-k)$ можно записать в виде

$$p(x-k) \overset{x}{\equiv} \lambda \cdot (-1)^n k^n + \sum_{i=0}^{n-1} q_i(x) \cdot k^i \quad ,$$

где в сумме справа коэффициенты q_i (являющиеся многочленами от x, что сейчас несущественно) не зависят от k (именно это важно). Суммируя по $k = 0, \ldots, n$ эти тождества с коэффициентами Z_k^n и меняя в двойной сумме порядок суммирования, приходим, с учетом (6), к тождеству (14). ■

Примечание. Теорема 2 сохраняет силу и при $n = 0$. ■

Соотношения (10),(14) представляют собой две формы *универсального биномиального тождества.*

[1] В цитированном источнике опечатка: допускается $r = n$, что неверно.

2. Сопряженные последовательности

Обозначим через $\mathbf{N} = \{0, 1, \ldots\}$ множество натуральных чисел. Через L обозначим линейное пространство, элементами которого являются последовательности вещественных чисел. Каждую последовательность можно рассматривать как функцию на $\mathbf{N}$. Стандартным базисом в L является набор последовательностей, каждая из которых отлична от нулевой последовательности только единицей на одном из мест, и тогда члены последовательности выступают как ее координаты в стандартном базисе.

2.1 Биномиальный базис. Покажем, что набор элементов

$$\mathbf{u} := \{u_k := (Z_k^n\ ,\ n \in \mathbf{N}),\ k \in \mathbf{N}\}$$

также является базисом, т.е. каждый элемент $\mathsf{g} \in \mathsf{L}$ может быть представлен в виде ряда

$$\mathsf{g} = \sum_{k=0}^{\infty} b_k u_k \quad . \tag{15}$$

Введя бесконечную матрицу $\mathbf{U}$, строками которой являются элементы u_k, запишем (15) в виде $\mathsf{g} = \mathsf{b}\mathbf{U}\ ,\ \mathsf{b} = (b_k)$. Компоненты b_k элемента b естественно считать координатами элемента g в биномиальном базисе.

Для наглядности приведем начальные строки матрицы $\mathbf{U}$

$$\begin{array}{l} u_0 = (\ 1 \quad 1 \quad 1 \quad 1 \quad 1 \ \ldots) \\ u_1 = (\ 0 \ -1 \ -2 \ -3 \ -4 \ \ldots) \\ u_2 = (\ 0 \quad 0 \quad 1 \quad 3 \quad 6 \ \ldots) \\ u_3 = (\ 0 \quad 0 \quad 0 \ -1 \ -4 \ \ldots) \end{array}$$

Видим, что $\mathbf{U}$ есть верхнетреугольная матрица с ненулевыми элементами по диагонали; следовательно она обратима: существует матрица $\mathbf{U}^{-1}$. Поэтому представление (15) действительно имеет место, при этом $\mathsf{b} = \mathsf{g}\mathbf{U}^{-1}$.

Замечание 1. Поскольку матрица $\mathbf{U}$ верхнетреугольная, ряд (15) всегда сходится, ибо для каждая компонента g_m элемента g есть конечная сумма по $k = 1, \ldots, m$. ■

Лемма 1. $\mathbf{U}^2 = \mathbf{E}$ (бесконечная единичная матрица), следовательно $\mathbf{U}^{-1} = \mathbf{U}$.

Доказательство. Т.к. матрица $\mathbf{U}$ верхнетреугольная, элемент v_{kn} матрицы $\mathbf{V} := \mathbf{U^2}$ есть конечная сумма

$$v_{kn} := \sum_{j=0}^{\infty} u_{kj} u_{jn} = \sum_{j=k}^{n} Z_k^j Z_j^n \quad .$$

При $k > n$ сумма справа пуста и $v_{kn} = 0$, при $k < n$ также $v_{kn} = 0$, согласно (11); диагональные элементы v_{nn} равны, очевидно, единице. Таким образом,

$$v_{kn} = \delta_{kn}(\text{символ кронекера}) \quad ; \tag{16}$$

это означает, что базис $\mathbf{u}$ ортонормирован, и $\mathbf{V} = \mathbf{E}$. ■

2.2 Двойственность, оператор сопряжения. Матрица $\mathbf{U}$ задает линейное преобразование на пространстве $\mathbf{L}$. Из Леммы 1 следует, что каждый элемент и его образ *взаимно сопряжены*: имеет место *соотношение двойственности*

$$\mathsf{g} = \mathbf{U}\mathsf{b} \quad \Longleftrightarrow \quad \mathsf{b} = \mathbf{U}\mathsf{g} \quad . \tag{17}$$

В этом смысле можно назвать $\mathbf{U}$ *оператором сопряжения.*

3. Биномиальные функции

3.1. Биномиальный базис в пространстве многочленов. Множество всех многочленов представляет собой линейное пространство $\mathcal{P}$. В обычной записи многочлена в виде

$$p(x) = a_0 + a_1 x + \ldots + a_n x^n;$$

вектор коэффициентов $(a_0, \ldots, a_n) \in R^{n+1}$ – это, по существу, координаты многочлена p в базисе степенны́х одночленов (x^k) , $k = 0, 1, \ldots$. Но, аналогично Разделу 2, мы введем в $\mathcal{P}$ другой – *биномиальный* базис; им служит система многочленов (12). Отметим, что степень многочлена π_k равна k.

Система (12) действительно образует базис. В самом деле, всякий многочлен степени n можно представить в виде

$$p(x) = \sum_{j=0}^{n} b_k \pi_k(x) \quad . \tag{18}$$

Для $n = 0$ это очевидно, а при $n \geq 1$ следует по индукции из равенства

$$p(x) = (-1)^n n! a_n \pi_n(x) + q(x) \quad ,$$

где $q(x)$ – многочлен степени не выше $n-1$, который (по индуктивному предположению) может быть разложен по биномиальным полиномам $\pi_0 \ldots \pi_{n-1}$.

Замечание 2. Если рассматривать множество $\mathbf{N}$ как целочисленную *решетку* на вещественной полуоси R_+, то можно сказать, исходя из (13), что полином (функция) π_k есть продолжение на всю прямую R альтернирующих биномиальных коэффициентов Z_k^m , $m \geq k$ (замена верхнего индекса m на непрерывную переменную x). ■

Ниже в данном разделе мы будем иметь дело с функциями вещественного аргумента $x \in R$.

3.2. Биномиальные функции. Введем в рассмотрение класс $\mathcal{F}$ функции f вещественного аргумента $x \in R$, представимых биномиальным рядом

$$f(x) = \sum_{k=0}^{\infty} b_k \pi_k(x) \quad , \tag{19}$$

где $\mathsf{b} = (b_k) \in \mathbf{L}$ – некоторая порождающая последовательность (координаты функции f в биномиальном базисе). Такие функции названы *биномиальными.*

Отметим, что если $x \in \mathbf{N}$, то сумма справа конечна, как и в (15), и поэтому значения $f(m)$, $m \in \mathbf{N}$ определены. Тем самым, с каждой функцией $f \in \mathcal{F}$ связаны две последовательности: первая – это порождающий ее (согласно (19)) элемент, который назовем *генератором* (обозначение $\mathsf{b} = \Gamma f$); вторая – последовательность $\mathsf{g} \in \mathsf{B}$ значений функции на решетке $\mathbf{N}$ (ее *след*, обозначение – $\mathsf{g} = Sf$. Сопоставление (19) с (15) показывает, что $Sf = \mathbf{U}\Gamma f$.

Множество тех значений $x \notin \mathbf{N}$, при которых ряд ряд (19) сходится, назовем *собственной областью* функции f; обозначение $Q = Q_f$. Если это множество пусто, функцию f будем называть *вырожденной.*

Замечание 3. Не следует думать, что отображение $\mathsf{b} \in \mathbf{L} \to f \in \mathcal{F}$ линейно; это не так, поскольку каждая функция f имеет свою собственную область Q_f. Соответственно, $\mathcal{F}$ не является линейным пространством. ■

Сопряженной к функции f называется функция φ, генератором которой является след данной функции, т.е.

$$\varphi(x) = (\mathsf{T}f)(x) := \sum_{k=0}^{\infty} f(k)\pi_k(x) \tag{20}$$

(T – *оператор сопряжения*).

Соотношение двойственности (17) приобретает в классе $\mathcal{F}$ вид

$$\varphi = \mathsf{T}f \quad \Longleftrightarrow \quad f = \mathsf{T}\varphi \quad , \tag{21}$$

при этом

$$\Gamma\varphi = Sf = \mathbf{U}\Gamma f \quad , \quad S\varphi = \Gamma f = \mathbf{U}Sf \quad . \tag{22}$$

3.3. Процедура очистки. Пусть $\mathsf{g} \in \mathsf{B}$ – некоторый фиксированный элемент. Он может служить следом на решетке $\mathbf{N}$ множества разных функций, отличающихся друг от друга колеблющимися волнами между последовательными точками решетки $\mathbf{N}$. В совокупности эти функции образуют *пучок* $\Pi = \Pi(\mathsf{g})$ данного элемента. В этом пучке только одна функция является биноминальной – именно, та, которая дается формулой (19) при $\mathsf{b} = \mathbf{U}\mathsf{g}$. Таким образом, каждый пучок обладает биномиальной функцией и она единственна.

Пусть теперь F – произвольная функция пучка П. Обозначим через $\tilde{F}$) биномиальную функцию соответствующего пучка. Преобразование $F \to \tilde{F}$ можно назвать *процедурой очистки*: оно "очищает" исходную функцию от всяких "лишних" колебаний и "проецирует" ее в класс $\mathcal{F}$. Например, если $F(x) = \sin(n\pi x)$ (n целое), то $\tilde{F}(x) \equiv 0$.

В случае когда $F \in \mathcal{F}$ процедура очистки есть повторное сопряжение, и тогда $\tilde{F} = F$.

Окончательно, F является биномиальной iff [2] $\tilde{F} = F$.

3.4. Область сходимости. По своим свойствам ряд (19) существенно отличается от обычных степенны́х рядов. Это видно хотя бы из того, что степенной ряд имеет радиус сходимости r (так что при $|x| > r$ ряд расходится), в то время как ряд (19) сходится при всех $x = m \in \mathbf{N}$.

[2] Короткое английское iff означает "тогда и только тогда, когда".

Теорема 3 (о сходимости). В точках $x \notin \mathbf{N}$ ряд (19) сходится iff сходится *контрольный* ряд

$$\sum^{\infty} \frac{b_k}{k^{\,1+x}} \tag{23}$$

(нижний предел суммирования несуществен).

Доказательство. Учитывая, что для любого фиксированного $x \in R$ знак величины $\pi_k(x)$ остается постоянным при больших значениях k, теорема непосредственно вытекает из следующей ключевой леммы.

Лемма 2. Если $x \notin \mathbf{N}$ (отрицательные, в том числе целые, значения x допустимы), то

$$|\pi_k(x)| \sim k^{-(x+1)} \qquad k \to \infty \tag{24}$$

(запись $a \sim b$ означает, что отношения a/b и b/a остаются ограниченными).

Доказательство. Имеем, обозначив $y := 1 + x$

$$\pi_k(x) = \prod_{j=1}^{k} \left(1 - \frac{y}{j}\right)$$

Выберем и зафиксируем $m \in \mathbf{N}$ такое, что $m > y$. При $x \notin \mathbf{N}$ будет $\pi_m(x) \neq 0$ и при $k > m$

$$\pi_k(x) = \pi_m(x) \cdot \sigma \qquad \sigma = \sigma_k := \prod_{j=m+1}^{k} \left(1 - \frac{y}{j}\right) \quad .$$

Все сомножители, входящие в произведение σ, положительны, что позволяет для его оценки перейти к логарифмам; находим

$$L := \ln \sigma = \sum_{j=m+1}^{k} \ln\left(1 - \frac{y}{j}\right) = -\sum_{j=m+1}^{k} \left(\frac{y}{j} + \frac{1}{2} \cdot \frac{y^2}{j^2} + \ldots\right) =$$

$$= -y \cdot \sum_{j=m+1}^{k} \frac{1}{j} + O(1) = -yS + O(1) \quad , \quad S = S(k) := \sum_{j=m+1}^{k} \frac{1}{j} \quad ,$$

где $O(1)$ – ограниченная величина при $k \to \infty$.

Осталось оценить сумму S; просуммировав неравенства

$$\frac{1}{j+1} < \int_{j}^{j+1} \frac{ds}{s} = \ln(j+1) - \ln j < \frac{1}{j}$$

для $j = m+1, \ldots, k$, находим

$$S - \frac{1}{m+1} + \frac{1}{k+1} < \ln(k+1) - \ln(m+1) < S \quad .$$

Отсюда получаем, возвращаясь к исходным величинам,

$$S \simeq \ln k \quad , \quad L \simeq -y \ln k \quad , \quad |\pi_k(x)| \simeq \sigma = \exp(L) \sim k^{-y} = k^{-(1+x)}$$

(запись $a \simeq b$ означает, что $a/b \to 1$); это и есть оценка (24). Лемма доказана. ■ (Лемма)

Из Леммы 2 следует, что собственной областью функции f является область Q, в которой сходится ряд (23). Область Q открыта, и непрерывность f в этой области доказывается стандартными методами (ε, δ)–анализа. ■ (Теорема)

4. Некоторые примеры биномиальных функций

О свойствах биномиальных функций пока можно сказать немного; рассмотрим примеры.

4.1. Показательные функции. Пусть $\mathtt{g} = (a^k) \in \mathbf{L}$ – геометрическая прогрессия (*геометрическая последовательность*) с основанием a, порождающая пучок $\Pi = \Pi_{\mathtt{g}}$ ($a \in R$ может быть и отрицательным; при $a > 0$ $\mathtt{g}$ есть след показательной функции $\Phi_a(x) = a^x$, $x \in R$) [3].

Замечание 4. При $a = 0$ считаем

$$g_0 = 0^0 = 1 \quad , \quad g_k = 0 \qquad k \geq 1 \quad . \tag{25}$$

При таком соглашении все последующие утверждения, справедливые при малых значениях a, остаются таковыми же и при $a = 0$; в частности, определена функция

$$\Phi_0(x) := 0^x = \begin{cases} 1 & x = 0 \\ 0 & x > 0 \end{cases} \qquad x \in R_+ \quad , \tag{26}$$

являющаяся пределом при $a \to +0$ фрагмента показательной функции Φ_a на полуоси R_+. ■

[3] В нашем понимании показательная функция Φ_a , $a > 0$ определена **на всей оси** R. Функцию a^x $(a > 0)$, определенную только на некотором подмножестве $\mathcal{O} \subset R$, мы называем *фрагментом* показательной функции на множестве $\mathcal{O}$.

Используя элемент g в качестве генератора, определим биномиальную функцию $f \in \mathcal{F}$, так что $\Gamma f = \mathsf{g}$; тогда сопряженная к f функция $\varphi = \mathsf{T}f$ будет, согласно процедуре очистки (п. 3.3), биномиальной функцией пучка П. След функции f есть последовательность

$$f(m) = \sum_{k=0}^{m} a^k \pi_k(m) = \sum_{k=0}^{m} C_k^m (-a)^k = (1-a)^m \qquad m \in \mathbf{N} \quad . \tag{27}$$

Последовательность (27) также геометрическая, и т.к. она является генератором функции φ, отсюда вытекает

Утверждение 1. При любом $a \in R$ геометрические последовательности с основаниями a и $(1-a)$ взаимно сопряжены. Соответственно, сопряжены генерируемые этими последовательностями биномиальные функции. ■

Далее необходимо рассмотреть отдельные случаи.

С1. $|a| > 1$. В этом случае контрольный ряд (23) расходится при $x \notin \mathbf{N}$, поэтому f вырождена и дается формулой (27).

С2. $|a| < 1$. Ряд (23) сходится при всех $x \in R$. В этом случае вид функции f дается следующей леммой.

Лемма 3. При $a \in (-1, 1)$

$$\sum_{k=0}^{\infty} a^k \pi_k(x) = (1-a)^x \qquad x \in R \quad . \tag{28}$$

Доказательство. . Левая часть (28) определена, т.к. сходимость ряда следует из Теоремы 3. Для каждого фиксированного $x \in R$ рассмотрим правую часть **как функцию от $\boldsymbol{a}$ (!)** ; обозначим ее $H(a)$. Дифференцируя k раз, приходим к выражению

$$H^{(k)}(a) = (-1)^k (1-a)^{x-k} \cdot \prod_{i=0}^{k-1} (x-i) \quad .$$

Полагая здесь $a = 0$ и сопоставляя с (12), получаем $H^{(k)}(0) = k! \cdot \pi_k(x)$; поэтому левая часть (28) может быть записана виде

$$\sum_{k=0}^{\infty} \frac{1}{k!} H^{(k)}(0) a^k \quad . \tag{29}$$

При $a \in (0, 1)$ (29) представляет собой разложение в ряд Тейлора функции H в точке $a = 0$; радиус сходимости ряда равен единице. Таким образом, при любом $x \in R$ равенство (28) выполняется при всех $a \in (-1, 1)$. $\blacksquare$

Поскольку (по построению) $f \in \mathcal{F}$, и в данном случае $f = \Phi_{1-a}$, из Леммы 3 вытекает, что показательная функция Φ_{1-a} с основанием $(1-a) \in (0,2)$ биномиальна и порождается геометрическим генератором с основанием $a \in (-1, 1)$.

Далее, т.к. след (27), служащий генератором сопряженной функции φ, также является геометрической последовательностью с основанием $a' := 1 - a$, то к функции φ применимы те же рассуждения, что и к f с заменой a на a'. Поэтому при $|a'| > 1$ (т.е. при $a \notin [0, 2]$) φ вырождена; в случае $a' \in (-1, 1)$, в силу Леммы 3, $\varphi(x) = (1 - a')^x = a^x = \Phi_a(x)$, $a \in (0, 2)$.

С3. $a = 1$. Гармонический ряд (23) сходится при $x > 0$, причем в силу Леммы 3, по непрерывности $f(x) = \lim\limits_{a \to +0} a^x = 0$; поэтому $f = \Phi_0 = (26)$. Сопряженная функция, равная по определению π_0, есть, одновременно, показательная функция Φ_1.

С4. $a = -1$. По теореме Лейбница о знакопеременных рядах ряд (23) сходится при $x > -1$ (необходимое условие); следовательно, биномиальная функция f является (опять в силу Леммы 3, по непрерывности) фрагментом показательной функции Φ_2 на множестве $\mathcal{O} = \{x \geq -1\}$.

Собирая все сказанное, получаем в итоге

Утверждение 2. Пусть геометрический элемент $\mathsf{g} = (a^k) \in \mathsf{B}$ задает пучок Π_{g}, $f \in \mathcal{F}$ – генерируемая элементом g функция, и $\varphi := \mathsf{T}f$ – биномиальная функция пучка.

1) При $a \notin [0, 2]$) функция φ вырождена, и последовательность ее значений на решетке $\mathbf{N}$ есть g; функция f также вырождена, и ее значения на $\mathbf{N}$ даются формулой (27), в которой $(1 - a) \notin [-1, 1]$.

2) При $a \in (0, 2)$ φ есть показательная функция $\Phi_a(x) = a^x$, определенная на всей оси $x \in R$. При этом, если $a \in (0, 1)$, то f – также показательная функция Φ_{1-a}; если $a = 1$, то $f = \Phi_0$; если $a \in (1, 2)$, то f вырождена со значениями (27) (геометрический след с отрицательным основанием $(1 - a) \in (-1, 0)$).

3) При $a = 0$ $f = \Phi_0$, φ – показательная функция Φ_1, являющаяся базовым полиномом π_0.

4) При $a = 2$ f есть фрагмент показательной функции Φ_2 в области $\mathcal{O} = \{x \geq -1\}$; φ – вырожденнная функция, генерируемая геометрической последовательностью $((-1)^k)$. $\blacksquare$

Дополнение к Лемме 3. Если, следуя Замечанию 2, считать $\pi_k(x)$ альтернирущим биномиальным коэффициентом с непрерывным верхним индексом x, то сумму в левой части (28) можно записать в форме

$$\sum_{k=0}^{\infty} C_k^x (-a)^k \quad ,$$

и тогда она представляет собой **бином Ньютона степени $\boldsymbol{x}$ (!)** (с бесконечным числом членов), для выражения $(1-a)^x$.

Поэтому равенство (28) можно интерпретировать двояким образом: и как разложение в ряд Тейлора функции H, и как "бином Ньютона" с вещественным показателем степени. $\blacksquare$

4.2. Многочлены. Согласно п. 3.1, всякий многочлен p может быть представлен в биномиальном базисе (следовательно $p \in \mathcal{F}$) , при этом порождающий его элемент b *финитен* число его ненулевых компонент конечно, оно равно степени многочлена. Что можно сказать о сопряженной к p функции $f = \mathsf{T}p$?

Т.к. генератор f есть след p, то он не финитен и, следовательно, f не является многочленом. С другой стороны, $Sf = \Gamma p =$ конечная последовательность коэффициентов в (18). В частности, $\mathsf{T}\pi_k$ есть функция, след которой отличен от нуля только единицей на k -м месте (отметим, что это прямо следует из (16)); найдем эту функцию.

Полином $\pi_0(x) \equiv 1$ является показательной функцией с основанием $a = 1$, поэтому его сопряженной является функция Φ_0 (см. (26)) с областью определения R_+. Из соотношения

$$\pi_k(x) = -\frac{x}{k}\pi_{k-1}(x-1) \qquad k \geq 1 \quad ,$$

следует

$$\mathsf{T}\pi_m(x) := \sum_{k=0}^{\infty} \pi_m(k)\pi_k(x) = \sum_{k=m}^{\infty} \pi_m(k)\pi_k(x) = \tag{30}$$

$$\sum_{k=m}^{\infty} \left[-\frac{k}{m}\pi_{m-1}(k-1)\right] \cdot \left[-\frac{x}{k}\pi_{k-1}(x-1)\right] = \qquad (i := k-1)$$

$$= \frac{x}{m} \cdot \sum_{i=m-1}^{\infty} \pi_{m-1}(i)\pi_i(x-1) = \frac{x}{m} \cdot \mathsf{T}\pi_{m-1}(x-1) \qquad m \geq 1 \quad ,$$

т.е.

$$\mathsf{T}\pi_m(x) = \frac{x}{m} \cdot \mathsf{T}\pi_{m-1}(x-1) \qquad m \geq 1 \ , \ x \in R \quad .$$

Поэтому (по индукции) функция $\mathsf{T}\pi_m$ определена в области $\{0, \ldots, m\} \cup \{x > m\}$ и отлична от тождественного нуля в этой области только единицей в точке $x = m$.

Отсюда вытекает

Утверждение 3. Функция f, сопряженная к многочлену p , представимому в виде (18), дается формулой

$$f(m) = b_m \qquad m = 0, \ldots, n \quad , \quad f(x) = 0 \qquad x > n \quad . \quad \blacksquare$$

Литература

1. Холл М. Комбинаторика. – М.: Мир, 1970.

2. Бут Э.Д. Численные методы. – М.: ГИФМЛ, 1959.

3. Кречмар В.А. Задачник по алгебре. – М.: Наука, 1968.

Лист аннотаций

Андрюшкевич О.А., Денисова И.М. Особенности развития аутсорсинга в России /Анализ и моделирование экономических процессов. Сборник статей под ред. В.З.Беленького, вып. 8. – М.: ЦЭМИ РАН, 2011, с. 7–28.

Анализируются особенности процесса формирования в России современного рынка аутсорсинговых услуг в области информационных технологий и бизнес-процессов. Основное внимание уделяется факторам, препятствующим развитию отечественного рынка аутсорсинга.

Трофимова Н.А., Разумовская В.А. Модифицированная гравитационная модель трудовой миграции /Анализ и моделирование экономических процессов. Сборник статей под ред. В.З.Беленького, вып. 8. – М.: ЦЭМИ РАН, 2011, с. 29–42.

Обосновывается возможность применения гравитационных моделей для моделирования процессов трудовой миграции из стран СНГ в Россию. На модифицированной модели проведены прогнозные расчеты миграционных потоков в Центральный федеральный округ.

Гогулин М.А., Четвериков В.М. Модель ценовой конкуренции на рынке однородного товара /Анализ и моделирование экономических процессов. Сборник статей под ред. В.З.Беленького, вып. 8. – М.: ЦЭМИ РАН, 2011, с. 43–60.

Рассматривается рынок определенного товара. При нескоординированном снижении цен отдельными продавцами-конкурентами объем потребительского спроса растет; изучается вопрос об изменении при этом доходов продавцов.

Белкина Т.А. Теоремы достаточности для вероятности неразорения в динамических моделях страхования с учетом инвестиций /Анализ и моделирование экономических процессов. Сборник статей под ред. В.З.Беленького, вып. 8. – М.: ЦЭМИ РАН, 2011, с. 61–74.

Для моделей указанного типа формулируются утверждения, позволяющие избежать доказательства верхних и нижних оценок вероятности неразорения для обоснования ее асимптотических представлений.

Паламарчук Е.С. Управление динамикой равновесной цены в экономике с мультипликативной неопределенностью /Анализ и моделирование экономических процессов. Сборник статей под ред. В.З.Беленького, вып. 8. – М.: ЦЭМИ РАН, 2011, с. 75–88.

В рассматриваемой модели функционал качества включает дисконтирующую функцию, учитывающую, в мультипликативной форме, потери, возникающие при отклонении текущей цены от эталонной траектории планового периода. Исследуются вероятностные свойства оптимального в среднем управления. Отдельно рассматривается случай бесконечного планового горизонта.

Клеппер Л.Я. Вероятности осложнений в органах и тканях при терапии с неоднородным облучением /Анализ и моделирование экономических процессов. Сборник статей под ред. В.З.Беленького, вып. 8. – М.: ЦЭМИ РАН, 2011, с. 89–102.

Дается систематическое описание и анализ основных наработанных к настоящему времени математических моделей дозовых режимов лучевой терапии. Основное внимание уделено формулам для расчета вероятностей осложнений при однородном и неоднородном способах облучения.

Смоляк С.А. Об одном функциональном уравнении /Анализ и моделирование экономических процессов. Сборник статей под ред. В.З.Беленького, вып. 8. – М.: ЦЭМИ РАН, 2011, с. 103–112.

Дается решение функционального уравнения, обобщающего на случай n переменных аналогичное двумерное уравнение, связанное с задачей оптимизации инвестирования инновационных проектов в условиях неопределенности.

Беленький В.З. Биномиальные функции /Анализ и моделирование экономических процессов. Сборник статей под ред. В.З.Беленького, вып. 8. – М.: ЦЭМИ РАН, 2011, с. 113–126.

Математическая заметка о свойствах функций, представимых бесконечными рядами, построенными не в степенно́м базисе, а в базисе биномиальных (разностных) полиномов. Строится бином Ньютона с произвольным (возможно нецелым) показателем степени.

List of abstracts

O.A. Andryushkevich, I.M. Denisova. Development features of Russian outsourcing.

The formation features of the modern outsourcing service market in Russia are analyzed. The service markets for information technologies and business processes are considered. The emphasis is on the obstacles factors for the development of the Russian outsourcing market.

N.A. Trofimova, V.A. Razumovskaya. A modified gravity model of labor migration.

The possibility of applying the gravity models to simulate the labor migration processes from the CIS-countries to Russia is substantiated. A modified model is used to fortecast the migration flows to the Central Federal District.

M.A. Gogulin, V.M. Chetverikov. Price competition model on a market of homogeneous wares.

A market of some homogeneous wares is considered. The customer demand capacity increases the uncoordinated price reduction by some competitive vendors; the resulting income variation for vendors is analyzed.

T.A. Belkina. Sufficient theorems for the nonruin probability in the dynamical insurance models with consideration of investments.

Several statements are formulated for insurance models of mentioned type to avoid a proof of upper and lower estimates for the ruin probability, which are used to substantiate its asymptotic representations.

E.S. Palamarchuk. Control of the equilibrium price dynamics in an economic with multiplicative uncertainty.

In the model under study, the criteria functional contains a discounting factor with the loss function in the multiplicative form. The probabilistic properties of the control optimal in the mean are analyzed. The case of an infinite planning horizon is discussed.

L.Ya. Klepper. Complication probability in organs and tissues under the therapy with inhomogeneous irradiation.

The basic modern mathematical models of dosage regimes used in radiation therapy are systematically described and analyzed. The emphasis is on the formulas for calculating the complication probabilities under the homogeneous and inhomogeneous ways of irradiation.

S.A. Smolyak. On some functional equation.

A two-dimensional equation associated with the optimization problem for the investment of innovation projects under uncertainty is generalized to the n-dimensional case. The solution of resulting functional equation is obtained.

V.Z. Belenky. Binomial functions.

The properties of the functions representable by infinite series in the basis of binomial (difference) polynomials (instead the powers series) are studied. The Newton's binom with arbitrary real degree is constructed.

ОБ АВТОРАХ

Андрюшкевич Ольга Анатольевна
кандидат экон. наук, старший научный сотрудник ЦЭМИ

Беленький Виталий Зиновьевич
доктор физ.-мат. наук, профессор, зав. лабораторией ЦЭМИ

Белкина Татьяна Андреевна
кандидат физ.-мат. наук, доцент, зав. лабораторией ЦЭМИ

Гогулин Михаил Александрович
студент МГИЭМ, кафедра математической экономики

Денисова Ирина Михайловна
кандидат экон. наук, старший научный сотрудник ЦЭМИ

Клеппер Лев Яковлевич
доктор экон. наук, главный научный сотрудник ЦЭМИ

Паламарчук Екатерина Сергеевна
младший научный сотрудник ЦЭМИ

Разумовская Виктория Александровна
выпускница ГАУГН 2011 г.

Смоляк Сергей Абрамович
доктор экон. наук, главный научный сотрудник ЦЭМИ

Трофимова Наталия Аристарховна
кандидат экон. наук, доцент, старший научный сотрудник ЦЭМИ

Четвериков Виктор Михайлович – доктор физ.-мат. наук, профессор, зав. кафедрой математической экономики МГИЭМ

Объем 8.2 п.л. Тираж 120 экз.