Гэри Смит
Джей Кордес

ЛОВУШКИ ПРИЗРАЧНЫХ ЗАКОНОМЕРНОСТЕЙ В АНАЛИЗЕ ДАННЫХ

ДМК ПРЕСС
издательство

Гэри Смит и Джей Кордес

Ловушки призрачных закономерностей в анализе данных

THE PHANTOM PATTERN PROBLEM

The Mirage of Big Data

Gary Smith and Jay Cordes

OXFORD
UNIVERSITY PRESS

ЛОВУШКИ ПРИЗРАЧНЫХ ЗАКОНОМЕРНОСТЕЙ В АНАЛИЗЕ ДАННЫХ

Гэри Смит и Джей Кордес

Москва, 2025

УДК 001.103+001.12
ББК 16.35
С50

С50 Гэри Смит и Джей Кордес
Ловушки призрачных закономерностей в анализе данных / пер. с англ. А. В. Логунова. – М.: ДМК Пресс, 2025. – 220 с.: ил.

ISBN 978-5-93700-433-8

В современном мире на нас обрушивается лавина данных и стремительно растет число закономерностей, которые мы в них обнаруживаем. От природы человеку свойственно верить, что каждая закономерность наполнена смыслом. На самом деле вероятность того, что найденная закономерность окажется полезной, стремительно приближается к нулю. Эта увлекательная книга благодаря бесчисленным примерам поможет читателю не поддаваться дурману данных, не прибегать к бесполезным инвестиционным стратегиям и не бояться делать прививки.

Издание адресовано всем, кто использует данные в работе – ученым, аналитикам, инженерам, консультантам, а также любознательным читателям, которые хотят научиться правильно оценивать связь между событиями.

УДК 001.103+001.12
ББК 16.35

Authorized Russian translation of the title The Phantom Pattern Problem, original ISBN 978-0-19-886416-5, printed and distributed under permission of Oxford University Press, 2020.

Все права защищены. Любая часть этой книги не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами без письменного разрешения владельцев авторских прав.

Материал, изложенный в данной книге, многократно проверен. Но, поскольку вероятность технических ошибок все равно существует, издательство не может гарантировать абсолютную точность и правильность приводимых сведений. В связи с этим издательство не несет ответственности за возможные ошибки, связанные с использованием книги.

ISBN 978-0-19886-416-5 (англ.)
ISBN 978-5-93700-433-8 (рус.)

Copyright © Gary Smith and Jay Cordes, 2020
© Оформление, перевод на русский язык, издание, ДМК Пресс, 2025

Эта книга посвящается нашим семьям

Оглавление

От издательства.....9

Введение. Конечно, вы шутите.....10

Глава 1. Выживаемость потных обработчиков закономерностей.....13

Выживаемость самых потных.....13
Выживаемость самых умных.....14
Распознавание закономерностей.....16
Детские закономерности.....18
Яблоко от яблони недалеко упало.....20
Закономерности повсюду, пока они не исчезнут.....21
Распознавание лиц.....22
Обманчивые закономерности.....23
Самый удачливый ребенок на свете.....24
Как не быть введенным в заблуждение призрачными закономерностями..... 25

Глава 2. Предсказывание того, что предсказуемо.....27

Что такое причинная следственность?.....27
Желтые камешки.....29
Предсказание.....29
Дважды отмерь, один раз отрежь.....31
Лимонники лечат цингу лаймами.....32
Золотой стандарт.....35
A/B-тесты.....36
Рассуждение «после этого, следовательно, из-за этого».....37
От хорошего к так себе.....38
Сплетни.....41
Если любишь свою работу.....43
Продуманные тикеры.....45
Как не быть введенным в заблуждение призрачными закономерностями..... 50

Глава 3. Оболванены и обмануты.....51

Малярия.....51
Суеверия.....53
Формирование закономерностей в воображении.....54
Памятные закономерности на букву «М».....54
Не дроби карты.....55
Если верить, то это обязательно произойдет.....57
Понедельник, 6-е число.....57
Число 47 повсюду среди нас.....61

Нумерология 63
Космические совпадения 65
Это все о нас 66
Закон Боде 68
Закон Мура 71
Как не быть введенным в заблуждение призрачными закономерностями 73

Глава 4. Одурачены снова и снова 75

Подбрасывания и подделки 75
Хмельные шаги 77
Серии на подъеме 80
Вы придираетесь, что ли? 85
Притормози и перетасуй 86
Поселенцы 88
Скопления онкологических заболеваний 89
Чем меньше, тем лучше (и хуже) 91
Стандартизированные тесты 92
Насильственные преступления 93
Мы – первые (или, может быть, вторые) 94
Как не быть введенным в заблуждение призрачными закономерностями 97

Глава 5. Парадокс больших данных 99

Вступление 99
Глубокая переработка данных 100
Вневыборочные данные 102
Вытеснение истинных переменных 103
Множественная регрессия 104
Невежи 106
Матери-иммигрантки и их дочери 106
Пора с этим завязывать 108
Предсказание обменного курса 112
Глубокая переработка твитов Трампа 114
Как не быть введенным в заблуждение призрачными закономерностями 116

Глава 6. Бесплодные поиски 117

Использование Google Trends для отслеживания рынка акций 118
Всматриваясь в звезды 121
Использование Google Trends для отслеживания цен на биткойн 123
Предсказывание безработицы 126
Китайский жилищный пузырь 128
Как не быть введенным в заблуждение призрачными закономерностями 131

Глава 7. Кризис воспроизводимости 132

Неряшливая наука 133
Ты – это то, что ест твоя мать 137
Зомбические исследования 138
Кризис воспроизводимости: тематическое исследование 139
Подсчитайте их 142
Отчет Чарни 143
Скептицизм против отрицательства 145
Как не быть введенным в заблуждение призрачными закономерностями 146

Глава 8. Кто-то вмешался? 147

Новая кола 148
Передача Биллу Клинтону правды 151
Секреты рынка акций 152
Наилучший месяц для покупки акций 153
Оптимизация портфеля 155
Система месяца 159
Бросание кубиков на доллары 160
В чем подвох? 163
Поход по магазинам за доллары 166
Как не быть введенным в заблуждение призрачными закономерностями 171

Глава 9. Видение вещей такими, какие они есть 172

Принцип Питера 178
Google 178
Контролируемые эксперименты 181
Естественные эксперименты 184
Обойденные вниманием женщины доктора Спока 186
Сначала теория, потом данные 188
Чрезмерные ожидания 192
Как не быть введенным в заблуждение призрачными закономерностями 194

Эпилог. Все дело в Байесе 195

Теорема Байеса 196
Медицинские диагнозы 197
Правило Байеса в зале суда 199
Закономерности и апостериорные вероятности 202

Библиография 204

Предметный указатель 215

От издательства

Отзывы и пожелания

Мы всегда рады отзывам наших читателей. Расскажите нам, что вы думаете об этой книге – что понравилось или, может быть, не понравилось. Отзывы важны для нас, чтобы выпускать книги, которые будут для вас максимально полезны.

Вы можете написать отзыв на нашем сайте **www.dmkpress.com**, зайдя на страницу книги и оставив комментарий в разделе «Отзывы и рецензии». Также можно послать письмо главному редактору по адресу **dmkpress@gmail.com**; при этом укажите название книги в теме письма.

Если вы являетесь экспертом в какой-либо области и заинтересованы в написании новой книги, заполните форму на нашем веб-сайте по адресу **http://dmkpress.com/authors/publish_book/** или напишите в издательство по адресу **dmkpress@gmail.com**.

Список опечаток

Хотя мы приняли все возможные меры для того, чтобы обеспечить высокое качество наших текстов, ошибки все равно случаются. Если вы найдете ошибку в одной из наших книг, мы будем очень благодарны, если вы сообщите о ней главному редактору по адресу **dmkpress@gmail.com**. Сделав это, вы избавите других читателей от недопонимания и поможете нам улучшить последующие издания этой книги.

Нарушение авторских прав

Пиратство в интернете по-прежнему остается насущной проблемой. Издательство «ДМК Пресс» очень серьезно относится к вопросам защиты авторских прав и лицензирования. Если вы столкнетесь в интернете с незаконной публикацией какой-либо из наших книг, пожалуйста, пришлите нам ссылку на интернетресурс, чтобы мы могли применить санкции.

Ссылку на подозрительные материалы можно прислать по адресу электронной почты **dmkpress@gmail.com**.

Мы высоко ценим любую помощь по защите наших авторов, благодаря которой мы можем предоставлять вам качественные материалы.

Конечно, вы шутите

В октябре 2001 года председатель совета директоров компании Apple Стив Джобс представил свой революционный портативный музыкальный плеер со встроенным жестким диском iPod. «Если уж говорить о музыке, то всегда иметь при себе всю библиотеку компакт-дисков – это качественный скачок. В вашем кармане может поместиться вся ваша музыкальная библиотека». Несмотря на цену в 399 долл., продажи были феноменальными, достигнув тридцати девяти миллионов в 2006 году, прежде чем их затмило появление смартфона iPhone в 2007 году.

На рис. I.1 показано, что резкий рост продаж плеера iPod в 2005 и 2006 годах совпал с увеличением числа убийств в Соединенных Штатах. Люди убивали друг друга, чтобы завладеть плеером iPod? Неужели слушатели плеера сходили с ума от непрекращающейся музыки, а затем убивали друзей и незнакомцев?

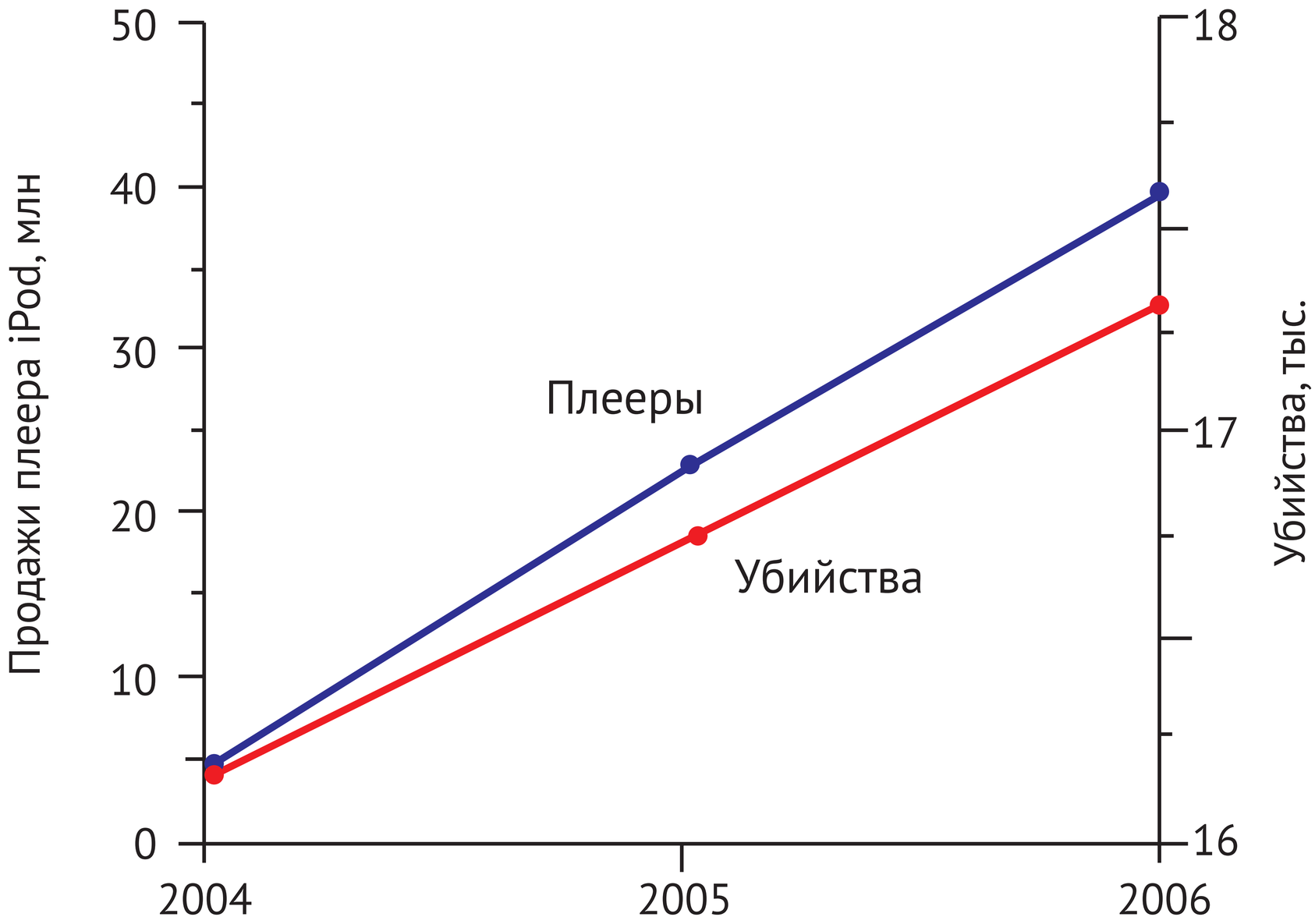

Рис. I.1. Продажи музыкального плеера iPod и количество убийств

Когда мы показали рис. I.1 нашей знакомой, она сразу же отреагировала возгласом «Конечно, вы шутите». Да, мы шутим, но у нас есть основание

шутить. В 2007 году авторитетный вашингтонский аналитический центр Urban Institute опубликовал отчет об увеличении числа убийств в 2006 и 2007 годах:

> Увеличение числа насильственных преступлений и взрывной рост продаж плееров и других портативных мультимедийных устройств – это не просто случайно возникшее совпадение. Мы предполагаем, что за последние два года Америка, возможно, пережила волну *iПреступности*.

В отличие от нас они не шутили.

Всех нас неоднократно предупреждали, что корреляция не есть причинная следственность, но слишком часто эти предупреждения мы игнорируем. От наших далеких предков мы унаследовали непреодолимое желание искать закономерности[1] и поддаваться их очарованию. Мы смеемся над некоторыми явно нелепыми корреляциями; например, количество юристов в Неваде статистически коррелирует с количеством людей, которые умерли, споткнувшись о собственные ноги. Однако другие корреляции, такие как продажи музыкальных плееров iPod и убийства, имеют соблазнительную привлекательность. Если уважаемых исследователей из Urban Institute можно соблазнить надуманными корреляциями, то это может сделать любой из нас.

Тысячи людей убивали друг друга не для того, чтобы украсть их плееры iPod, и тысячи слушателей плееров iPod не были доведены до безумия жаждой убийства. Просто случилось так, что в 2005 и 2006 годах произошло увеличение числа убийств и продаж плееров iPod, как и многое другое. Случайная взаимосвязь между убийствами и продажами плееров iPod продлилась недолго. В 2007 году количество убийств сократилось и с тех пор продолжает снижаться, хотя продажи плееров iPod продолжали расти еще несколько лет, пока они не стали мизерными по сравнению с продажами смартфонов iPhone.

Корреляция между убийствами и продажами плееров iPod особенно смехотворна по той причине, что основана на данных всего за три года. С убийствами будет сильно коррелировать все, что стабильно росло (или снижалось) в течение этого двухлетнего периода, – например, продажи мороженого в США.

На рис. I.2 показано, что корреляция между продажами мороженого и убийствами была такая же, как и корреляция между продажами плееров iPod и убийствами. Люди убивали друг друга, чтобы завладеть музыкальными плеерами или же рожками с мороженым? Почти наверняка ни то ни другое, и все же два опытных исследователя из Urban Institute допустили, что корреляция была причинно-следственной связью.

Как так случилось, что в современную эпоху больших данных и мощных компьютеров эксперты могут быть такими глупыми? По иронии судьбы большие данные и мощные компьютеры являются частью проблемы.

Нас всех воспитывали так, чтобы мы были одурачены, – чтобы мы были привлечены блестящими закономерностями и сверкающими взаимосвязями. И большие данные, и мощные компьютеры подпитывают эту зависимость, потому что с ними так легко находить подобные финтифлюшки, – и они так-

---

[1] В оригинале авторы используют термин *pattern*, который означает закономерность, регулярность, повторяющийся узор, образец, фигуру, схему или шаблон. – *Прим. перев.*

же обеспечивают, чтобы бóльшая часть найденного была так же бесполезна, как и надуманное заявление о том, что плееры iPod увеличивают количество убийств.

Мы должны противостоять соблазну, чтобы не быть одураченными призрачными закономерностями.

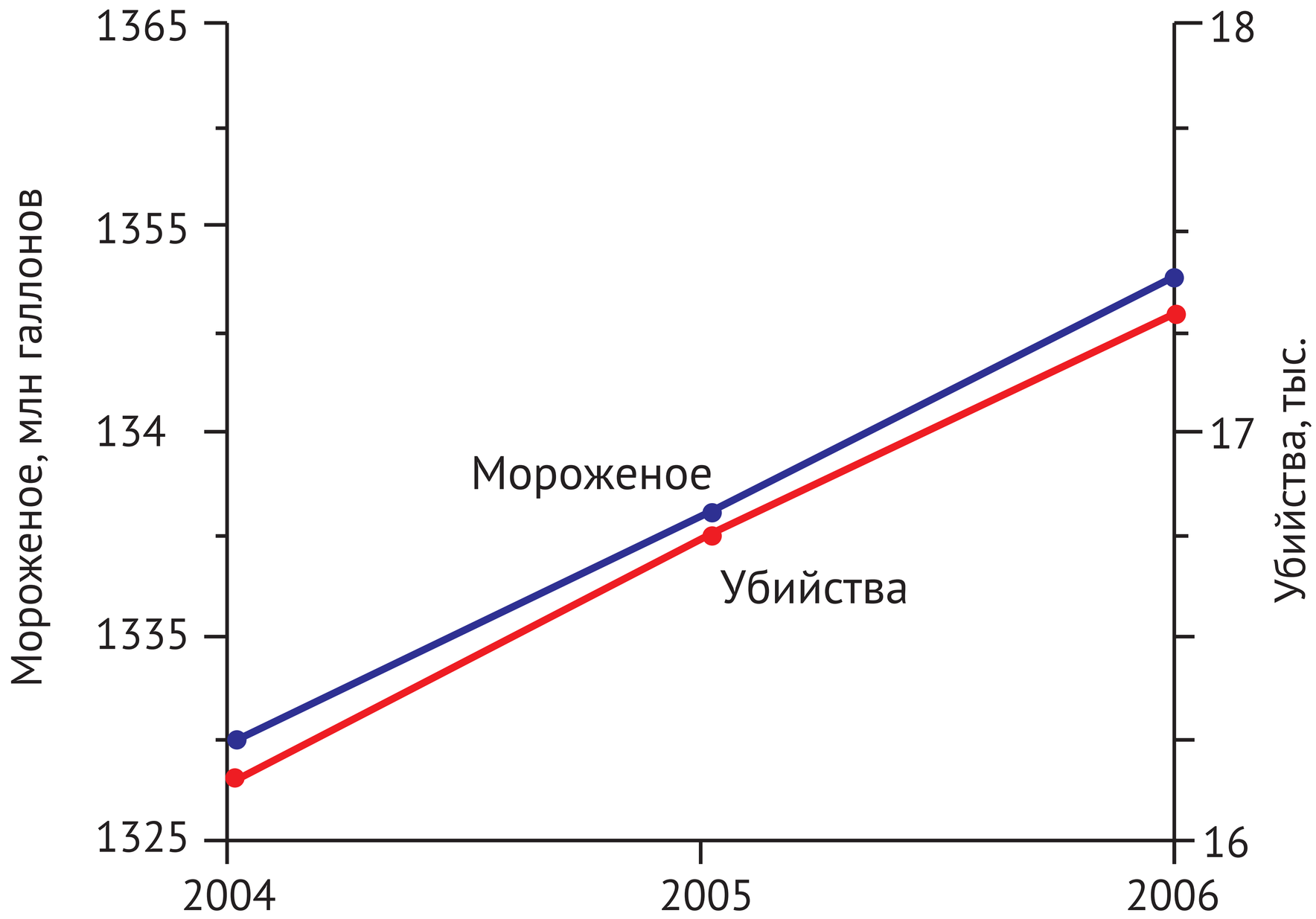

Рис. I.2. Продажи мороженого и убийства

Глава 1

Выживаемость потных обработчиков закономерностей

> **По сравнению с другими животными люди не отличаются особой силой. У нас нет ни острых зубов, ни когтей, ни клювов. У нас нет ни кувалдообразных рогов, ни клыков, ни хвостов. У нас нет брони. Мы не великие пловцы и не спринтеры. Каким образом нашим далеким предкам удалось не только выжить, но и стать хозяевами планеты?**

Нашему восхождению способствовали два фактора: эффективное охлаждение организма и распознавание закономерностей.

Выживаемость самых потных

Наше первое большое преимущество заключается в том, что мы рождены бегать. Наши доисторические предки ходили пешком и бегали повсюду, охотясь за съедобными существами и убегая от зверей, которые хотели их съесть. Люди не быстры, но обладают исключительной выносливостью. Древние люди, которые могли загонять свою добычу, изматывая ее, ели много мяса, что придавало им бóльшую силу и повышало шансы на спаривание и передачу своих генов детям. На протяжении бесчисленных поколений, занятых погоней, пиршествами и спариванием, люди, которые не могли угнаться за ними, были стерты из генофонда.

Охотный промысел на основе бега на выносливость, называется *охотой на изнурение* (persistence hunting) – преследованием добычи до тех пор, пока она больше не сможет бежать. Отчасти наш успех объясняется тем, что два миллиона лет назад наши тела эволюционировали таким образом, чтобы способствовать бегу на выносливость, задолго до того, как наши более поздние предки изобрели стрелы, копья и другое метательное оружие. Среди прочих приспособлений, выявленных профессором антропологии Гарварда Дэниелом Либерманом (Daniel Lieberman):

> Мы развили длинные, упругие сухожилия в ногах и ступнях, которые функционируют как большие резинки, накапливая энергию и высвобождая ее с каждым шагом во время бега, уменьшая количество энергии, необходимое для того, чтобы сделать еще один шаг. Существует также несколько приспособлений, помогающих сохранять устойчивость нашего тела во время бега, например то, как мы уравновешиваем каждый шаг взмахами рук, большие ягодичные мышцы, которые удерживают верхнюю часть тела

в вертикальном положении, и эластичные связки на шее, помогающие удерживать голову ровно.

Даже талия человека, более тонкая и гибкая, чем у наших родственников-приматов, позволяет нам поворачивать верхнюю часть тела во время бега, чтобы уравновешивать слегка смещенные от центра усилия, прилагаемые каждой ногой.

Ключом к успешному бегу на выносливость являются не просто сильные ноги и хорошее равновесие (они есть у многих животных), но и защита организма от перегрева, и люди в этом преуспевают. Даже при сильных колебаниях внешних условий человеческое тело очень хорошо поддерживает внутреннюю температуру в узком безопасном диапазоне от 36,5 °C до 37,5 °C (от 97,7 °F до 99,5 °F). Если температура человеческого тела падает до опасно низкого уровня, скажем 35 °C (95 °F), то возникает *гипотермия*, и, с другой стороны, если температура тела повышается до опасно высокого уровня, скажем 38 °C (100,4 °F), то возникает *гипертермия*.

Тепловое истощение, сопровождающееся головокружением, слабостью, нарушением координации и нагрузкой на сердечно-сосудистую систему, которые могут приводить к тепловому удару и смерти, чаще всего вызывается интенсивными физическими нагрузками в жарких условиях, например бегом или занятиями активными видами спорта в необычно жаркие дни. К счастью, у людей есть отличная встроенная система охлаждения. У нас относительно нет волосяного покрова, мы можем регулировать частоту дыхания во время бега, и мы потеем больше на квадратный дюйм, чем любой другой биологический вид. Напротив, волосатые животные, которые охлаждают себя, тяжело дыша, быстро перегреваются – вот почему многие отдыхают в полдень и почему люди способны пересиливать их во время дневной погони.

В некоторых частях Африки, Мексики и Сибири до сих пор практикуется охота на изнурение. В 1980-х и 1990-х годах один антрополог был свидетелем того, как местные охотники в Ботсване загоняли крупных антилоп. Охота может длиться пять часов и проходить более двадцати миль при температуре выше 100 °F. Антилопа убегала от охотников-людей, отдыхала, когда начинала перегреваться, а затем снова бежала, когда люди ее догоняли. Этот цикл повторялся до тех пор, пока не наступала гипертермия и антилопа не падала в обморок, потому что была слишком перегрета, чтобы продолжать бег.

В 2013 году BBC сообщила, что кенийский пастух загнал двух гепардов, которые убивали его коз. Он не стрелял в них из ружья, он просто загнал их до их падения на землю. Он объяснил, что однажды утром «я пил чай, когда увидел, как они убивают еще одну козу». Он подождал до полудня, а затем отправился в путь с несколькими молодыми людьми из своей деревни. После четырехмильной погони они связали измученных гепардов веревками и передали их Службе охраны дикой природы Кении. Наши предки гордились бы ими, хотя и недоумевали бы, почему они не съели гепардов.

Выживаемость самых умных

Оказывается, выносливость, размер мозга и умственные способности взаимосвязаны. Польза от физических упражнений не передается по наследству

напрямую, но люди, клетки мозга и когнитивное мышление которых лучше всего реагировали на физические нагрузки, имели эволюционное преимущество перед другими.

Из поколения в поколение происходил естественный отбор – выживали самые умные. Самые умные процветали, спаривались и передавали свои интеллектуальные способности своим детям, из которых самые умные с большей вероятностью выживали, спаривались и заводили детей. Окончательные последствия этого полезного цикла заключаются в том, что по сравнению с размерами нашего тела человеческий мозг теперь в три раза больше, чем мозг других млекопитающих, а человеческий интеллект намного превосходит интеллект других животных.

Физические упражнения, похоже, делают нас умнее. Лично мы всегда считали, что должны тренировать свой мозг так же, как тренируем свое тело, читая, играя в карточные игры и решая головоломки. Последние исследования показывают, что на самом деле силу мозга можно увеличивать в большей степени за счет тренировки тела, чем за счет тренировки мозга!

В одном интересном эксперименте мышей поместили на несколько месяцев в четыре разные среды обитания. Некоторые мыши жили в простой клетке без игрушек и ели обычную пищу. Вторая группа жила в разноцветных клетках с зеркалами и качелями и ела изысканную пищу (по мышиным меркам). У третьей группы не было ни игрушек, ни вкусной еды, но было беговое колесо для упражнений. Четвертая группа жила насыщенной жизнью, у нее были красочные игрушки, вкусная еда и беговое колесо.

Всем мышам перед экспериментом и после него ставили интеллектуальные задачи. Единственное, что делало их умнее, – это беговое колесо. Причудливые игрушки и еда не оказывали никакого влияния ни на мышей, у которых были беговые колеса, ни на тех, у кого их не было. Физические упражнения неизменно делали мышей умнее независимо от того, были ли у них интересные игрушки и изысканная еда или не было игрушек и была безвкусная пища.

Как это работает? Во-первых, физические упражнения снабжают мозг кровью и кислородом, которые питают клетки мозга. Физические упражнения также увеличивают выработку белка под названием BDNF (brain-derived neurotrophic factor, т. е. нейротрофический фактор головного мозга), который помогает клеткам мозга расти и соединяться друг с другом.

В дополнение к бегающим в колесах мышам к такому же выводу привели многочисленные эксперименты с людьми. В одном исследовании группе молодых людей был предложен тест на запоминание. Затем некоторые из них сидели неподвижно в течение тридцати минут, в то время как другие крутили педали на велотренажерах в изнурительном темпе. При повторном тестировании мужчины, которые сидели неподвижно, показали те же результаты, что и раньше. Однако у мужчин, которые занимались физическими упражнениями, уровень белка BDNF был выше, и они справились с тестом на запоминание лучше. Многие аналогичные исследования показали, что физические упражнения помогают клеткам мозга расти и выживать, а также улучшают память и обучение.

Похоже, что это работает для всех возрастных уровней.

Исследование, проведенное с участием подростков, показало, что они лучше справлялись с тестами на решение задач после тридцатиминутных упражнений

в день. Исследование, проведенное с участием людей среднего возраста, показало, что после четырех месяцев силовых и аэробных тренировок у них были более высокие результаты по тестам на запоминание и логическое мышление.

Исследование, проведенное среди людей старше пятидесяти с проблемами памяти, показало, что они лучше справлялись с умственными тестами после регулярных физических упражнений, выполнявшихся ими в течение шести месяцев. Исследование пожилых людей показало, что занятия спортом три раза в неделю на треть снижают вероятность развития болезни Альцгеймера. Исследование пациентов с болезнью Паркинсона показало заметное улучшение состояния после занятий на велотренажере.

Врачи давно считают, что физические упражнения замедляют воздействие старения на память («Куда я положил ключи?») и обучение («Как работает этот новый телефон?»). Теперь кажется, что физические упражнения действительно могут обращать процесс вспять, не только замедляя отмирание клеток мозга, но и выращивая новые. Например, одно исследование показало, что у людей, которые занимались физическими упражнениями в течение трех месяцев, выросли новые нейроны, в особенности в той области мозга, которая отвечает за память и обучение.

Распознавание закономерностей

По сравнению с другими животными эволюция, по-видимому, благоприятствовала нашим далеким предкам, которые обладали исключительной выносливостью и незаурядным мозгом. Исключительная выносливость полезна для охоты. Для чего нужны незаурядные мозги?

Во-первых, наш сверхразмерный мозг дополняет нашу замечательную выносливость. Охота на изнурение усиливается благодаря интеллекту, навыкам общения и коллективной работе, которые, возможно, способствовали развитию человеческого интеллекта, коммуникативных и социальных навыков. Лучшие охотники-люди, как правило, были умнее среднего человека, что позволяло им выслеживать добычу, планировать стратегии групповой охоты на более крупных и быстрых животных и претворять эти планы в жизнь. Те, кто был умнее среднего, стали доминировать в генофонде, тогда как те, кто был менее умен, как правило, были менее успешными охотниками, с большей вероятностью умирали молодыми и были менее способны привлекать плодовитых партнерш.

Дело не только в размере нашего мозга. У многих китов, слонов и дельфинов мозг крупнее человеческого, но многие ли из них могут писать прекрасные стихи и вдохновляющие симфонии? Многие ли из них могут проектировать автомобили и строить машины, которые собирают машины?

Крупная часть того, что подразумевается под интеллектом, – это распознавание закономерностей, и оно принесло огромную эволюционную пользу людям. Вот очень неполный список очевидно полезных закономерностей, которые имели значение для выживания и размножения тех, кто их распознавал:

- паническое бегство зебр сигнализирует о хищниках;
- следуя за слонами, можно проследить путь до водопоя;

- день и ночь сменяют друг друга;
- существуют сезоны вегетации;
- темные тучи предвещают дождь;
- приливы приходят и уходят;
- некоторые продукты являются съедобными, другие – ядовитыми;
- некоторые животные являются добычей, другие – хищниками;
- можно выявлять плодовитых партнеров.

Благодаря распознаванию закономерностей люди получили эволюционное преимущество перед другими животными. Те, кто были более способны распознавать признаки опасности и плодовитости, с большей вероятностью передавали свои способности к распознаванию закономерностей будущим поколениям. И действительно существует мнение, что когнитивное превосходство человека над всеми другими животными обусловлено в основном нашим эволюционным развитием в области превосходящей обработки закономерностей.

Многие доисторические закономерности, которые были замечены нашими предками, были, безусловно, бесполезны. Например, во время солнечного затмения люди могли петь или танцевать, а когда солнце появлялось снова, делать вывод, что богу солнца понравились пение и танцы. Такие неверные толкования понятны, но не имеют больших последствий. Сильное эволюционное преимущество распознавания полезных закономерностей легко компенсируется относительно небольшими издержками, связанными с введением в заблуждение бесполезными закономерностями.

Люди обладают некоторыми ценными навыками распознавания закономерностей вместе с другими животными, такими как:

- распознавание – определение своего вида и угрожающих действий;
- передача сигналов – общение с помощью жестов;
- составление карт – использование ориентиров для запоминания местоположения пищи, укрытий и хищников.

Навыки распознавания закономерностей, в которых люди намного превосходят других животных, включают:

- коммуникацию – письменные и устные языки, которые передают подробную информацию;
- изобретательство – создание сложных инструментов и других средств достижения конкретных целей;
- искусство – создание эстетически привлекательных текстов, рисунков, музыки и скульптур;
- воображение будущего – определение сложных возможных последствий конкретных действий;
- волшебство – добровольную приостановку неверия, принятие невозможных мыслей.

Во многих отношениях интеллект часто предусматривает способность принимать правильные решения, основанные на обнаружении закономерностей.

Некоторые считают, что существуют разные виды интеллекта, например умение разбираться в математике, в словах и в людях, но все указанные умения улучшаются благодаря распознаванию закономерностей, например распознаванию широко применимых математических принципов, распознаванию эффективных способов связывания слов или распознаванию настроений и эмоций других людей. Единственное, что объединяет описанные выше предположительно разные типы интеллекта, – это обработка закономерностей.

Некоторые проводят различие между интеллектом и творческой изобретательностью, но опять же важна обработка закономерностей. Термин *интеллект* можно определить как использование накопленной информации и логических рассуждений для определения правильного ответа на вопрос или наиболее перспективного плана действий. Очевидно, что умение распознавать закономерности необходимо для определения наилучшего ответа на вопрос или возможных последствий конкретных действий. Термин *творческая изобретательность*, или *креативность*, напротив, можно определить как умение находить неожиданные, нестандартные ответы, действия и последствия. Решения должны быть разумными. Шутки должны быть творческими. Однако творческая изобретательность также выигрывает от распознавания закономерностей. Знание закономерности помогает генерировать оригинальные идеи, которые не вписываются в закономерность. Шутки нередко бывают смешными именно потому, что их изюминка – это не то, чего мы привыкли ожидать.

Вот победитель конкурса на самую смешную шутку в мире:

> Двое охотников из Нью-Джерси шли по лесу, когда один из них упал на землю. Казалось, он не дышит, а его глаза закатились. Другой парень достает свой мобильный телефон и звонит в службу спасения. Хватая ртом воздух, он обращается к оператору: «Мой друг мертв! Что я могу сделать?» Оператор успокаивающим голосом говорит: «Просто успокойтесь. Я могу помочь. Сначала давайте убедимся, что он мертв». Наступает тишина, затем раздается выстрел. В трубке снова звучит голос парня. Он спрашивает: «Так, и что теперь?»

Детские закономерности

Наши способности к распознаванию закономерностей проявляются в раннем возрасте, когда карапузы и ползуны используют образы и звуки, чтобы выявлять объекты и развивать языковые навыки. Малыши быстро понимают, чем отличаются собаки, стулья и вода, и учатся предвидеть последствия разных действий и событий.

Как ползуны учатся ходить? Они делают тысячи шагов и тысячи раз падают, пока не уяснят закономерность – сочетание действий, которое позволяет им держаться в вертикальном положении, передвигаясь на двух ногах.

Как дети учатся писать слова по буквам и строить предложения? Распознавая закономерности. Они запоминают, что звуки *собака* и *кошка* пишутся как С О Б А К А и К О Ш К А. Когда в доме больше одной собаки или кошки, надо заменять А в конце на И. Если нужно говорить как о собаках, так и о кошках, то ставить И между этими двумя словами. Овладение языком, включая правописание, грамматику и все остальное, можно описать как распознавание законо-

мерностей, и тот факт, что люди столь хорошо умеют распознавать закономерности, является одной из причин, по которой наши языковые навыки развиты намного больше, чем у других животных.

С математикой примерно то же самое. Что такое счет, как не запоминание закономерностей: 1, 2, 3 и 11, 12, 13? Что такое арифметика, как не запоминание правил? Что такое компьютерное программирование, как не применение общих принципов (закономерностей, если хотите) к новым задачам?

Одна знакомая недавно рассказала нам, что ее четырехлетний малыш, худенький и истощенный, всегда был слишком занят, чтобы поесть. Когда его усаживали на стул за обеденным столом, он стискивал зубы и изо всех сил старался освободиться – бегал по дому и играл со всем, что ему попадалось под руку. В конце концов она придумала, как его кормить. Когда она показывала ему интересное видео, он открывал рот и получал пищу. Это было похоже на эксперименты Павлова с собаками: видео продолжалось, и рот открывался. Однако она признала, что именно она была обучена этой закономерностью – включать видео, чтобы он поел.

Одна из дочек Гэри в очень раннем возрасте узнала, что, говоря «Эта-а» и указывая на что-то обеими руками, она заставляла Гэри и его жену подносить все, что находится в непосредственной близости от того места, куда были нацелены пальцы ребенка. У малышки еще не развились языковые навыки. Она просто знала, что звук «Эта-а», сопровождаемый указательным пальцем, приносит результат.

Возможно, этот ребенок видел, как взрослый указывал пальцем и издавал звуки, которые включали в себя «это», после чего что-то передавалось этому человеку. Если это происходило более одного раза, то последовательность событий закрепила эту закономерность в сознании ребенка. Указанное чувство, несомненно, усилилось, когда малышка указала пальцем и произнесла «Эта-а», и ее родители поспешили дать ей то, что она хотела. Малышка усвоила правило «укажи – и получишь», но родители также были обучены следовать закономерности «принеси – и успокоится».

Гэри и его жена, безусловно, почувствовали облегчение, когда малышка освоила более сложные закономерности, взломав речевой код – выучив определенные звуки (т. е. слова), ассоциирующиеся с желаемыми вещами, и научившись связывать слова вместе, образуя предложения.

Иногда быстрое развитие навыков распознавания закономерностей у ребенка может истолковываться взрослыми неверно. Например, совершенно верно, что ребенок, который плачет и привлекает к себе внимание, учится тому, что плач привлекает внимание, и вполне возможно, что дети, которые быстро усваивают эту закономерность, умнее среднего человека. Некоторые исследователи детского развития однажды подсчитали количество плача тридцати восьми младенцев в возрасте от четырех до семи дней, а затем измерили коэффициент интеллекта (IQ) каждого ребенка в возрасте трех лет. Корреляция, показанная на рис. 1.1, составила 0,45, что является статистически значимым показателем. Если бы плач и коэффициент интеллекта были независимы, то шанс такой высокой корреляции был бы меньше 1/200.

Нет, это не значит, что можно повысить коэффициент интеллекта ребенка, заставляя его плакать! Более правдоподобным объяснением является то, что

подвижные, любознательные малыши хотят больше внимания и быстро усваивают закономерность: плач привлекает внимание.

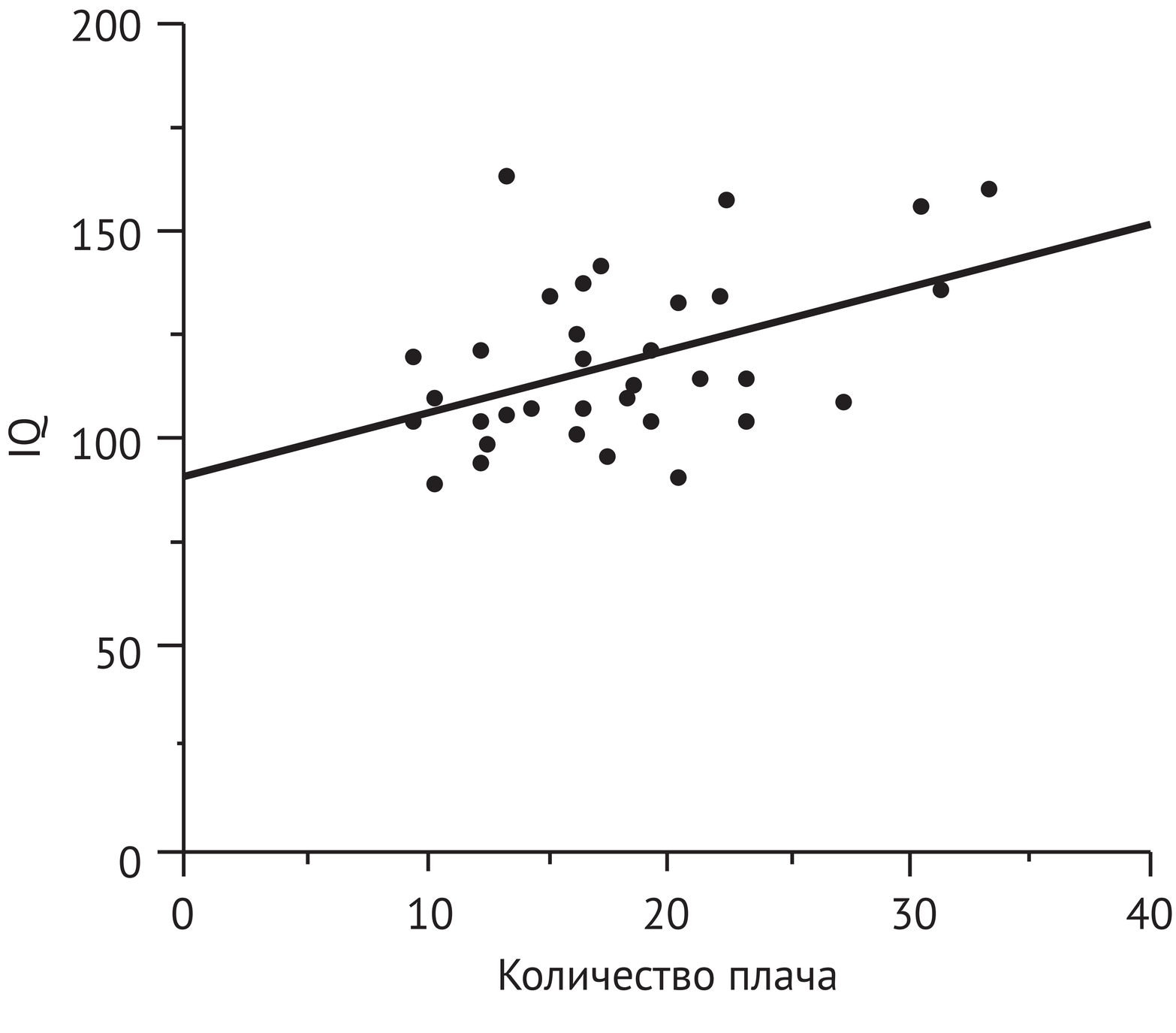

Рис. 1.1. Делает ли плач умнее?

Яблоко от яблони недалеко упало

Человеческий интеллект опирается на распознавание закономерностей – на вспоминание и применение закономерностей, хранящихся в нашем великолепном мозге. Наличие экстраординарного интеллекта часто предусматривает выявление закономерностей, которые другие не замечают, или понимание того, как незнакомая ситуация соотносится со знакомой.

Одним из величайших моментов «ага!» в истории науки стало наблюдение сэра Исаака Ньютона за тем, как яблоко падало с дерева в саду его матери. Ньютон родился на ферме недалеко от Грэнтема, Англия, и учился в Кембриджском университете, когда в 1665 году из-за вспышки бубонной чумы университет был закрыт и Ньютону пришлось вернуться на семейную ферму. (Вдохновил ли гениев карантин из-за COVID-19?)

Позже Ньютон рассказал, что он сидел в яблоневом саду, когда увидел, как на землю упало спелое яблоко (нет, оно не ударило его по голове). Он задался вопросом, почему, если земля круглая и у нее нет ни верха, ни низа, яблоки всегда падают на землю независимо от того, где находится яблоня на планете? И не только это, яблоки всегда падают перпендикулярно земле. Он объяснил эту закономерность тем, что яблоко притягивалось к земле гравитационной силой.

В Кембридже Ньютон пытался понять, почему Луна вращается вокруг земли, а не улетает в космос. Теперь у него был ответ. Закономерность яблока

и закономерность Луны взаимосвязаны! Гравитационные силы действуют на планеты так же, как и на яблоки, и распространяются на огромные расстояния.

Используя данные о наблюдаемых движениях небесных тел в Солнечной системе, Ньютон в конце концов смог сформулировать удивительно простой закон всемирного тяготения: каждое небесное тело притягивается к любому другому с силой, пропорциональной произведению их масс и обратно пропорциональной квадрату расстояния между ними.

Закон всемирного тяготения Ньютона предсказывает траектории запускаемых в воздух футбольных мячей, стрел, ракет и других объектов. Он также объясняет движение приливных и отливных волн и траектории движения астрономических тел.

Когда в 1781 году была открыта планета Уран, ее орбита несколько отличалась от орбиты, предсказанной законом всемирного тяготения Ньютона, если только на ее орбиту не влияла еще одна, пока что не открытая планета. Разумеется, в конечном итоге была открыта планета Нептун, и именно там, где предсказывал закон Ньютона.

Замеченная Ньютоном закономерность привела его к формулировке важного правила, которое помогает нам понимать мир и делать полезные предсказания. Конечно, закон всемирного тяготения Ньютона в конечном счете был вытеснен общей теорией относительности Эйнштейна, но это уже другая история (и другая закономерность).

Закономерности повсюду, пока они не исчезнут

Закономерность усиливается, когда она повторяется. Чем чаще малыш видит людей, которые сидят на стульях, тем сильнее его убеждение в том, что стулья созданы для того, чтобы на них сидели. Чем чаще стул называют стулом, тем сильнее убеждение в то, что это слово и предмет взаимосвязаны.

Закономерностью может быть расстановка предметов (например, двух палочек или ножа, вилки и ложки для еды) или последовательность событий (например, подброшенный вверх мяч, который падает обратно на землю). Закономерность может побудить нас ожидать, что она продолжится (например, официанты в ресторане раздают меню после того, как рассадят посетителей), и замечать отклонения или выбросы, которые не укладываются в эту закономерность. Мы замечаем ситуации, когда есть нож и ложка, но нет вилки, или когда официант не раздает меню.

Игроки в покер замечают подсказки своих оппонентов. Спортивные команды замечают тенденции. Подсказки и тенденции также могут использоваться для обмана. Опытный игрок в покер может заставить оппонента думать, что перед тем, как начать блефовать, он на мгновение закрывает глаза, а затем с закрытыми глазами нажать на педаль, сбросив чудовищную комбинацию карт.

Футбольная команда может провести весь сезон, не сфальсифицировав ни одного гола, а затем сфальсифицировать его в финальной игре чемпионата. Футболисты и тренеры пытаются выявлять тенденции игроков и команды, ча-

сами изучая видеозаписи предыдущих игр. Затем во время игры футбольные защитники могут распознавать игровую комбинацию по тому, как игроки выстраиваются и как двигаются некоторые ключевые игроки до и после начала комбинации. Полузащитники распознают действия защитников точно так же. Именно по этой причине команды просматривают видеозаписи своих собственных игр – для того чтобы выявлять тенденции, которые, возможно, были выявлены их соперниками, а затем удивлять оппонентов, нарушая эти тенденции в критически важных ситуациях.

Один из самых известных матчей в истории Суперкубка по американскому футболу состоялся в матче XLIX в 2015 году. «Нью-Инглэнд Пэтриотс» в игре с «Сиэтл Сихокс» вели со счетом 28:24, но «Сиэтл» завладели мячом на расстоянии одного ярда от линии «Нью-Инглэнд» за двадцать секунд до конца игры. Отличный защитник «Сиэтла» Маршоун Линч по прозвищу Звериный режим проносил мяч двадцать четыре раза на 102 ярда (по 4,25 ярда за попытку), и комментаторы и болельщики предположили, что Линчу отдадут мяч, и он пробьет себе дорогу в концевую зону, чтобы выиграть игру.

Однако тренеры «Сихокс» решили изменить сложившуюся ситуацию и отдать быструю передачу другому игроку после того, как тот резко пробился в концевую зону. Если бы эта неожиданная игровая комбинация сработала, то тренеров можно было бы похвалить за нестандартное мышление. Однако запасной защитник «Пэтриотс» Малком Батлер, новичок, не заявленный на драфт, который вышел на поле во втором тайме, увидел, как «Сиэтл» выстроились в линию (в форме «стека из двух принимающих»), и вспомнил игровую комбинацию, которую «Сиэтл» часто проводили таким образом. На тренировках он несколько раз пытался (безуспешно) отстоять эту комбинацию, но на этот раз он перехватил передачу, принеся команде «Пэтриотс» четвертую победу в Суперкубке. Позже Батлер сказал, что «во время подготовки я запомнил комбинацию, в которой они играли… Я просто опередил его на пути и выиграл в противоборстве». Это был первый пас, который Батлер когда-либо перехватывал, и это произошло потому, что «Сиэтл» попытались нарушить одну закономерность, а Батлер распознал другую.

«Сиэтлу» не удалось успешно замаскировать свою завершающую игровую комбинацию Суперкубка, но эта игра в кошки-мышки может пойти в любую сторону. В Суперкубке НФЛ 2019 года полузащитник «Новой Англии» Том Брэди, игравший в своем девятнадцатом сезоне (и, возможно, величайший полузащитник НФЛ всех времен), неоднократно оказывался одураченным, когда команда соперника «Лос-Анджелес Рэмс» выстраивалась в линию, как будто их игроки готовились к противоборству один на один, а затем переключались в зонную защиту после того, как мяч попадал к Брэди.

Вне всяких сомнений, принимать решения на основе закономерностей весьма полезно. Также бывает полезно принимать решения и на основе того, как соперники истолковывают закономерности и реагируют на них.

Распознавание лиц

Люди обладают невероятной способностью распознавать лица, потому что мы инстинктивно выявляем закономерности, которые отличают одно лицо от дру-

гого. Мы ожидаем, что рот, нос, два глаза и другие черты лица будут находиться в определенных местах и иметь определенные размеры и форму. Мы быстро замечаем отклонения от общей закономерности – большие уши, подбородок с ямочкой, густые брови – примерно так же, как карикатуристы подчеркивают необычные черты лица.

Эти различия между конкретным лицом и общей закономерностью называются отличительными чертами, потому что не сходства, а как раз различия позволяют нашему мозгу мгновенно распознавать людей, даже если лицо частично скрыто очками или тенями.

Людей могут сбивать с толку лица, которые слишком сильно отличаются от того, что мы ожидаем. Нам трудно распознавать лица, если брови полностью удалены, и мы не умеем распознавать перевернутые лица, потому что те черты лица, которые мы помним, почти неизменно предусматривают надлежаще ориентированных людей с бровями. Мы сбиты с толку, потому что знаем и понимаем черты лица. Мы знаем, что такое брови, и ожидаем их увидеть. Мы знаем, что такое глаза, носы и рты, и ожидаем увидеть их на лице именно в таком порядке, сверху вниз.

В противовес этому компьютерные алгоритмы распознавания лиц очень хрупки, поскольку в них используются математические представления пикселей, из которых состоят компьютерные изображения, и не имеют реального понятия о том, что представляет набор пикселей. Компьютерные алгоритмы не знают, что такое глаза, носы и рты; они просто замечают изменения в пикселях при сканировании цифрового изображения.

Такое непонимание может приводить к забавным ошибкам, когда алгоритм пытается сопоставить пиксельные закономерности на разных изображениях. Один из самых современных алгоритмов ошибочно принял мужчину-компьютерщика в необычных очках за женщину-кинозвезду. Изображение кошки с несколькими измененными пикселями было ошибочно принято за миску гуакамоле. Черепаха была ошибочно принята за винтовку.

Обманчивые закономерности

Иногда наши инстинктивные связанные с поиском закономерностей желания могут соблазнять нас видеть воображаемые или бессмысленные закономерности. В 2004 году онлайновое казино заплатило 28 000 долл. за приготовленный на гриле сырный сэндвич десятилетней давности (с откушенным от него кусочком), на котором, как утверждалось, был облик Девы Марии. Сообщалось, что на картофельных чипсах и рентгеновских снимках зубов были обнаружены облики Иисуса. Мать Тереза была замечена на булочке-чиннамон с корицей (пресловутой nun bun, т. е «булочке монахини»).

При наличии достаточного количества бутербродов, картофельных чипсов, рентгеновских снимков зубов и булочек-чиннамон обязательно получатся какие-то необычные формы, которые можно было бы посчитать как что-то реальное (рис. 1.2).

Рис. 1.2. Семья Джея обнаружила президентское печенье

Самый удачливый ребенок на свете

Хотя большинство магазинов 7-Eleven сейчас открыты круглосуточно, свое название они получили в 1946 году, когда начали работать с 7 утра до 11 вечера, продавая напитки, закуски, подгузники и другие предметы первой необходимости. Их фирменный напиток представляет собой кашеобразную ледяную смесь из воды, углекислого газа и ароматизированного сиропа, которую называют «слурпи» (ср. с русским «сёрбать») из-за звука, который напиток издает, когда пьешь его через соломинку.

С 2002 года магазины 7-Eleven раздают бесплатные сладости 11 июля, потому что месяц и число (в американском формате дат) – 7/11 – звучат как название бренда. В июле 2019 года произошла странная вещь, утренний выпуск Национального общественного радио озаглавил эту историю как «Странные новости». CNN сообщил об этой странной новости следующим образом:

> День 7/11, как правило, означает бесплатные угощения для всех, но в этом году для одной семьи из Миссури празднование стало более особенным, чем обычно. Рэйчел Лэнгфорд из Сент-Луиса родила девочку 11 июля – да, именно 7/11. И это еще не все. Малышка Жем Браун родилась в 19:11, а ее вес составлял 7 фунтов и 11 унций.

Представитель NBC сказал, что Жем была «самым удачливым ребенком на свете» и «с таким же успехом ее можно было бы окрестить Леди Удача».

Среди лавины комментариев к этой истории было три главные ветки.

- Это было крайне маловероятно.
  - Ничего себе, вот это да, сумасшедшее совпадение!
  - Кто-нибудь знает духовный смысл всего этого?

- Она должна получить что-нибудь бесплатно от 7-Eleven.
  - Пожизненный запас продуктов и напитков от 7/11.
  - Она должна получать бесплатные слурпы всю жизнь!!
- Надо использовать цифры 7 и 11 при покупке лотерейных билетов.
  - Лучше бы они разыгрывали этот номер каждый день.
  - Потрясающе… Давайте я сбегаю в 7-Eleven и поиграю в 711.

Правда, нашими любимыми были вот эти:

- Пожалуйста, не называйте ее слурпи!
- И почему это должно быть новостью?

И правда, почему это *должно быть* новостью? В среднем каждую минуту в Соединенных Штатах рождается восемь младенцев. В том, что ребенок родился 11 июля или в течение какой-то определенной минуты, нет ничего примечательного. Выберите минуту, любую минуту, и, скорее всего, за эту минуту родится восемь младенцев. Каждый день 7:11 случается два раза, в 7:11 утра и в 7:11 вечера, так что каждый день, включая 11 июля, ожидаемо будет иметься шестнадцать новорожденных в 7:11.

Заявленный вес при рождении – 7 фунтов 11 унций – делает Жем немного более необычной, но это не такой уж необычный вес при рождении. Скептик может также задаться вопросом, не был ли заявленный вес при рождении немного увеличен или уменьшен, чтобы привлечь внимание общественности по всей стране и, возможно, получить несколько бесплатных подарков от 7-Eleven, – скептицизм, подкрепленный новостными сообщениями о том, что родители ребенка планировали связаться с 7-Eleven.

Конечно же, компания 7-Eleven вручила семье подарочную корзину и пожертвовала 7 111 долл. в фонд колледжа. Теперь, может быть, компания сможет отслеживать всех остальных новорожденных 7/11 и делать то же самое.

Как не быть введенным в заблуждение призрачными закономерностями

Можно вывести человека из каменного века, но невозможно вывести каменный век из человека. Умение распознавать закономерности сослужило хорошую службу нашим предкам и, несомненно, во многом стало причиной того, что мы эволюционировали из слабаков в хозяев. Сегодня мы жестко запрограммированы замечать закономерности, и этот врожденный поиск закономерностей часто помогает нам понимать окружающий мир и принимать более обоснованные решения.

Однако закономерности не всегда непогрешимы. Иногда это иллюзия (например, лик Девы Марии на приготовленном на гриле сырном бутерброде). Иногда это бессмысленное совпадение, важность которого мы преувеличиваем (например, роды в 7:11 11 июля). Иногда они приносят вред (например, покупка лотерейных билетов, потому что мы думаем, что удачная случайность предсказывает выигрышный номер).

Наши предки обнаруживали закономерности в физическом мире, используя свои пять базовых чувств: осязание, зрение, слух, обоняние и вкус. Сейчас нас захлестывает лавина данных, и эти данные гораздо более абстрактны, сложны и трудно истолковываемы, чем информация, обрабатывавшаяся нашими далекими предками. Сегодня бóльшая часть получаемой и обрабатываемой нами информации представляет собой цифровое представление реальных явлений, таких как данные о доходах, расходах, уровне преступности и ценах на акции. И все чаще поиск закономерностей осуществляется компьютерными алгоритмами, которые ничего не знают о реальном мире. Количество возможных закономерностей, которые можно выявлять, по сравнению с количеством действительно полезных, растет экспоненциально, а это означает, что шансы на то, что обнаруженная закономерность окажется полезной, быстро приближаются к нулю. Мы можем быть легко одурачены призрачными закономерностями.

Глава 2

Предсказывание того, что предсказуемо

В 1930-х годах французский антрополог Клод Леви-Стросс (Claude Levi-Strauss) был свидетелем массовой резни в бразильских джунглях, которая произошла из-за неправильного толкования закономерности:

[Индеец из намбиквары] с высокой температурой явился в [протестантскую] миссию, и ему прилюдно дали две таблетки аспирина, которые он проглотил; после этого он искупался в реке, у него началось воспаление легких, и он умер. Поскольку намбиквара – отличные отравители, они пришли к выводу, что их соплеменник был убит; они предприняли ответную атаку, в ходе которой были убиты шесть членов миссии, включая двухлетнего ребенка.

Этот случай произошел из-за катастрофической путаницы между корреляцией и причинной следственностью. Сколько бы нам ни говорили, что корреляция не является причинно-следственной связью, наша унаследованная любовь к закономерностям заставляет (слишком часто) думать, что бессмысленная закономерность наполнена смыслом.

Что такое причинная следственность?

Что подразумевается под причинной следственностью? Причинно-следственная связь означает, что одно (причина) влияет на другое (следствие). Удар ногой по футбольному мячу приводит мяч в движение. Сила причины часто влияет на силу следствия. Удар по мячу с большей силой заставляет его двигаться дальше.

Если требуется предсказать дальность полета мяча, то следует учитывать и другие факторы, в том числе вес мяча и параметры ветра. Причинно-следственных связей может быть несколько. Вес мяча и ветер не приводят его в движение, но они влияют на расстояние и направление движения мяча после удара.

Полезные взаимосвязи тоже не обязательно должны быть идеальными на 100 %. Когда ученые приходят к выводу, что «курение табака служит причиной рака легких», это не означает, что буквально у каждого, кто выкуривает сигарету, вскоре разовьется рак легких. Вместо этого они имеют в виду что-то вроде «у мужчины, выкуривающего в среднем более пяти сигарет в

день, шанс развития рака легких в течение жизни в 100 раз выше, чем у некурящего мужчины».

Возможно, было бы полезно использовать слово, отличное от словосочетания *причинно-следственный*, или использовать широкое определение понятия причинной следственности, чтобы охватить различные обоснованные толкования. Независимо от того, как ее обозначать, решающее различие заключается в статистических корреляциях, которые относятся к случайно возникающим совпадениям, и статистических корреляциях, которые наполнены смысла, потому что для корреляции есть глубинное основание.

Если А коррелирует с Б, то существует несколько возможных объяснений.

1. А служит причиной Б. Богатые люди склонны тратить больше, потому что у них больше денег, а не потому, что трата большего количества денег делает их богатыми.
2. Б служит причиной А. Цены на акции, как правило, растут, когда компания объявляет о неожиданно значительном увеличении прибыли, потому что более высокие доходы повышают ценность акций компании, а не потому, что более высокие цены на акции приводят к росту прибыли компании.
3. А служит причиной В, а В служит причиной А. У игроков в крикет хорошая зрительно-моторная координация, потому что хорошая зрительно-моторная координация делает их более качественными игроками, а игра в крикет улучшает зрительно-моторную координацию.
4. Что-то еще служит причиной А и Б. Учащиеся, получившие высокие баллы за один тест по математике, как правило, получают высокие баллы за второй тест по математике не потому, что один балл влияет на другой, а потому, что оба балла отражают их математические способности.
5. Это случайно возникшее совпадение. В 2005 и 2006 годах количество убийств и продажи музыкальных плееров iPod увеличились.

С практической точки зрения, когда мы замечаем закономерность, обычно встает вопрос, существует ли какое-то глубинное основание у этой закономерности, которая, как ожидается, сохранится, чтобы ее можно было использовать для составления надежных предсказаний. Емкая поговорка гласит:

> Для того чтобы что-то предсказать, оно должно быть предсказуемым.

В случае первых четырех перечисленных выше типов закономерностей ответ будет да, существует реальная причинно-следственная структура, которая позволяет делать полезные предсказания. В случае пятого типа закономерностей ответ – нет. А не служит причиной В; В служит причиной А; и нет С, которое служило бы причиной А и В.

Возможно, вместо словосочетания *причинно-следственный* было бы лучше использовать словосочетание *наполненный смыслом*, чтобы пояснять, что, когда мы говорим о наличии причинно-следственной связи, мы не ограничиваемся тем, что А служит причиной В. Наполненная смыслом закономерность

имеет глубинное причинно-следственное объяснение и может использоваться для составления полезных предсказаний. Бессмысленная закономерность – это призрачная закономерность, которая возникает по случайному совпадению и не имеет предсказательной ценности.

Например, древние египтяне заметили, что ежегодному разливу Нила регулярно предшествовало появление Сириуса – самой яркой звезды, видимой с Земли, – которая появлялась на восточном горизонте непосредственно перед восходом солнца. Сириус не был причиной наводнения, но он был полезным предсказателем, потому что у него было глубинное основание: каждый год 19 июля Сириус всходил перед рассветом, а проливные дожди, начинавшиеся в мае на Эфиопском нагорье, служили причиной разлива Нила в конце июля.

Желтые камешки

Представьте, что тысячи мелких камешков различных размеров, форм, цветов и плотности создаются с помощью 3D-принтера, который запрограммирован таким образом, что их характеристики определяются независимо друг от друга. Исследователь, изучающий выборку из 100 таких камешков, может обнаружить, что по случайно возникшему совпадению желтые камешки в этой выборке – по чистой удаче – чаще бывают бугристыми, чем камешки других цветов.

Когда исследователь соберет вторую выборку камешков, созданную с помощью этого 3D-принтера, вряд ли желтый цвет будет хорошим показателем неровности, поскольку 3D-принтер задает цвет и форму самостоятельно. Корреляция между желтым цветом и бугристостью в изначальной выборке была просто случайно возникшим совпадением, и поэтому она исчезла.

Теперь предположим вместо этого, что на дне озера найдено несколько камешков и что *есть* какое-то научное основание, по которому бугристая галька в этом озере имеет тенденцию быть желтой. В этом случае корреляция между бугристостью и желтизной является полезным предсказателем, даже если мы не совсем понимаем причину, по которой бугристые камешки часто бывают желтыми. Возможно, желтые камешки имеют тенденцию быть бугристыми, потому что они привлекают что-то обитающее в озере, которое любит их грызть. А может быть, бугристые камешки более благоприятны для роста желтых водорослей. Или, может быть, желтым и бугристым оказался определенный вид мягкой породы. Принципиальное отличие этого сценария от случайного 3D-принтера заключается в том, что закономерность наполнена смыслом, потому что существует глубинное причинно-следственное основание, по которому бугристые камешки часто бывают желтыми. Вовсе не нужно знать точно, в чем причина, но как раз наличие реального основания наполняет эту взаимосвязь смыслом и делает ее полезной для составления предсказаний. В противном случае эта закономерность будет хрупкой и бесполезной.

Предсказание

Некоторые утверждают, что если предсказание является целью, то причинно-следственная связь не нужна. Достаточно корреляции. Если между влия-

тельными лицами в Facebook и сердечными приступами есть корреляция, то причинно-следственное объяснение не требуется. Достаточно знать, что они взаимосвязаны, потому что одно предсказывает другое. Проблема с этим аргументом состоит в том, что если у закономерности нет основания, то, скорее всего, эта корреляция является временной и кажущейся – как случайные камешки, созданные с помощью 3D-принтера, и как сердечные приступы и влиятельные лица в Facebook.

Один из энтузиастов корреляции приводит такой пример:

> Для того чтобы увидеть разницу между предсказанием и причинно-следственным выводом, представьте себе, что у вас есть набор данных, содержащий данные о ценах и заполняемости гостиниц. <...> Сначала представьте, что гостиничная сеть хочет составить оценку заполняемости гостиниц конкурентов на основе общедоступных цен. Это предсказательная задача. <...> Более высокие опубликованные цены указывают на более высокий уровень заполняемости, поскольку гостиницы, как правило, повышают свои цены по мере их заполнения (с использованием программного обеспечения по управлению финансовой отдачей). И наоборот, представьте, что гостиничная сеть хочет оценить, как изменилась бы заполняемость, если бы гостиница повсеместно повысила цены. <...> Это вопрос причинно-следственного вывода. Очевидно, что даже притом, что цены и заполняемость в типичном наборе данных положительно коррелируют, мы бы не пришли к выводу, что повышение цен увеличит заполняемость.

Для того чтобы предсказания были полезными, они должны надежно опираться на свежие данные, а стабильно надежные предсказания на основе свежих данных нуждаются в наличии причинно-следственной структуры. В приведенном выше примере с гостиницами статистическая корреляция между уровнями цен и уровнями заполняемости не является случайностью; она отражает реальную глубинную взаимосвязь.

Конечно же, можно неверно истолковать взаимосвязь между A и B как A служит причиной B, когда на самом деле все наоборот, и именно B служит причиной A. Когда идет дождь, люди ходят с открытыми зонтиками над головой. Это происходит потому, что дождь служит причиной того, что люди раскрывают зонтики, а не потому, что раскрытие зонтиков служит причиной дождя. Точно так же повышенный спрос на гостиничные номера, как правило, служит причиной роста цен, но повышение цен не увеличивает спрос. Зонты – верный признак дождя, а повышенные цены в гостиницах – верный признак того, что номера заполнены, но нам нужно быть осмотрительными в истолковании причинной следственности.

В противовес сказанному на рис. 2.1 показано, что существует статистическая корреляция между ценами на авокадо в Сан-Франциско в 2015 году и поисками информации о знаке зодиака Дева в Google, причем статистическая взаимосвязь является настолько сильной, что существует лишь один шанс из 10 000, что они оказались бы так тесно связаны по чистой случайности. Для того чтобы цены на авокадо были связаны с поисками информации о Деве, нет никакого глубинного основания, поэтому это было мимолетное совпадение, которое является совершенно бесполезным для составления предсказаний – как в отношении поисков информации о знаке Дева, так и в отношении цен на авокадо.

Недавно Гэри получил электронное письмо от человека, работающего в китайской компании, занимающейся информационными технологиями (слегка отредактированное для большей ясности):

> Моя работа заключается в обработке данных и предсказании будущего. Обычно я использую хороший алгоритм, такой как машинное обучение. Все мои предсказания основаны на исторических данных, но, похоже, исторические данные не очень надежны для предсказания предстоящих событий для компаний. Можете ли вы предложить более подходящий алгоритм, который позволит делать более точные предсказания?

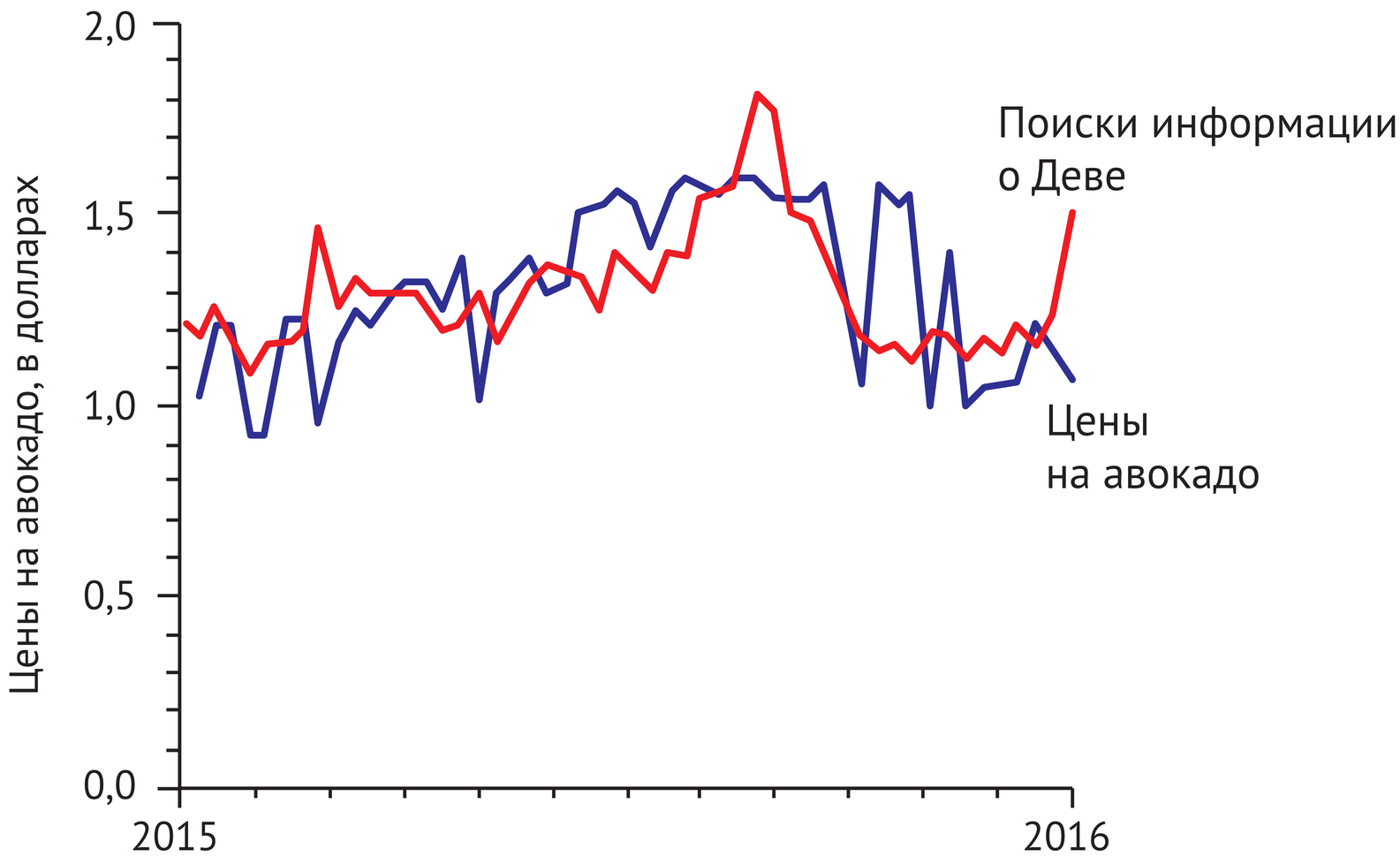

Рис. 2.1. Предсказывание цен на авокадо по поискам информации о знаке Дева

Проблема с этим запросом состоит в убеждении, что причина, по которой обнаруженные закономерности исчезают именно тогда, когда они нужны, заключается в какой-то проблеме с компьютерной программой. Нет. Проблема заключается в том, что у закономерности должно быть реальное глубинное основание, и ни одна компьютерная программа не может нам этого сказать. Компьютеры способны использовать числа, чтобы вычислять другие числа, но они не понимают смысл этих чисел. В примере с авокадо/Девой компьютерные алгоритмы не знают, с каким-либо наполненным смыслом пониманием, что такое авокадо и что такое поиски информации о знаке Дева, поэтому у них нет возможности оценить достоверность взаимосвязи между ценами на авокадо и поисках информации о Деве. Другой алгоритм не смог бы сделать более качественные предсказание цен на авокадо, основываясь на поисках информации о знаке Дева. Для того чтобы делать полезные предсказания, нужно обращать внимание на наполненные смыслом взаимосвязи.

Дважды отмерь, один раз отрежь

Исследовательницу данных, работающую в частном колледже, попросили выяснить, почему значительное число студентов покидали колледж до того, как

выполняли требования, предъявляемые к получению степени. Она покопалась в данных и обнаружила, что пакет финансовой помощи был действительно хорошим предсказателем – студенты, которые обращались за помощью, но не получали ее, покидали колледж чаще всего в течение первых двух лет.

Финансовая помощь предоставляется в виде сниженных расценок. Если колледж обычно взимает 50 000 долл. и предоставляет студенту финансовую помощь в размере 20 000 долл., то это 40%-ная скидка. Колледж в любом случае получает 30 000 долл. (и даже больше, если финансовая помощь является ссудой).

Исследовательница данных произвела приблизительные расчеты пакетов помощи, которые можно было бы предложить студентам, чтобы не давать им бросать учебу и при этом максимизировать доходы колледжа. Когда она показала свой анализ декану приемной комиссии, тот случайно упомянул, что при отчислении студента из колледжа деканат обращается к базе данных колледжа и сбрасывает пакет помощи студенту до нуля, так как колледж не оказывает помощь студентам, которые покинули колледж.

Вот так-так! Существовало причинно-следственное объяснение, но истолкование было обратным. Дело не в том, что отсутствие финансовой помощи служило причиной отчисления студентов. Дело в том, что уход из колледжа служил причиной обнуления финансовой помощи. Иногда важно найти правильное причинно-следственное объяснение. Допущение того, что корреляции будет достаточно, может оказаться дорогостоящей ошибкой.

Лимонники лечат цингу лаймами

Цинга – это жуткое заболевание, причиной которого, как известно, является длительное отсутствие витамина С в рационе. Когда-то цинга была главной причиной смерти моряков, которые отправлялись в длительные плавания, не употребляя в пищу фрукты или овощи, содержащие витамин С.

Васко да Гама главным образом из-за цинги потерял 68 % своей команды во время своего путешествия из Португалии в Индию в 1497–1499 годах; Магеллан потерял 90 % во время своего путешествия из Испании на Филиппины в 1519–1522 годах; а англичанин Джордж Энсон (George Anson) потерял 65 % в течение первых десяти месяцев своего кругосветного плавания в 1740–1744 годах. В целом, по оценкам, в период с 1500 по 1800 год от цинги умерло более двух миллионов моряков.

На протяжении веков считалось, что цинга представляет собой расстройство пищеварения, причиной которого служили самые разные вещи, в том числе чрезмерная работа моряков, употребление соленого мяса и грязной воды. Среди рекомендуемых средств лечения было употребление свежих фруктов и овощей, белого вина, сульфатов, уксуса, морской воды, пива и различных специй.

В 1601 году английский торговец и капер по имени Джеймс Ланкастер (James Lancaster) возглавил флотилию из четырех кораблей, совершившую пятимесячное путешествие из Англии в южную Африку. Моряки на его корабле ежедневно получали по глотку лимонного сока в бутылках и прибыли в добром здравии, в то время как морякам на трех других кораблях не давали лимонного сока, и большинство из них заболели цингой. Ланкастер был убежден в силе

лимонного сока и использовал его в своих будущих путешествиях, но другие отнеслись к этому скептически и не стали этого делать.

В 1747 году английский врач по имени Джеймс Линд (James Lind) провел эксперимент, находясь на борту линейного корабля HMS Salisbury. Он отобрал двенадцать больных цингой, которые были «настолько похожи, насколько это было возможно». Всех их кормили обычной кашей на воде, бараньим бульоном, вареными сухарями и другой неаппетитной матросской пищей. Кроме того, Линд разделил двенадцать пациентов на шесть групп по два человека, чтобы сравнить эффективность шести рекомендованных методов лечения. Двое пациентов каждый день выпивали по кварте крепкого сидра; двое принимали по двадцать пять капель сульфата; двоим давали по две ложки уксуса три раза в день, двое пили морскую воду; двоим давали по два апельсина и лимону через день, а еще двоим давали отвар, в состав которого входили чеснок, мирра, горчица и корень редиса.

Линд пришел к выводу, что:

> наиболее внезапный и заметный положительный эффект был отмечен при употреблении апельсинов и лимонов; один из тех, кто их употреблял, к концу шести дней был готов к работе. <…> Другой выздоровел лучше всех в своем состоянии, и… ему было поручено ухаживать за остальными больными.

К сожалению, его эксперимент не получил широкого освещения, и Линд мало что сделал для его популяризации. Позже он написал, что «моим уделом было давать наставления; власть исполнять их – в руках других».

Медицинское сообщество продолжало считать, что причиной цинги является нарушение работы пищеварительной системы из-за тяжелой работы моряков и неправильного питания, и ее можно вылечить с помощью «газированных напитков», содержащих серную кислоту, алкоголь и специи.

По странному совпадению в 1795 году, через год после смерти Линда, английский военно-морской врач Гилберт Блейн (Gilbert Blane) убедил британский военно-морской флот добавлять лимонный сок в ежедневный рацион моряков, употреблявших ром, и после этого цинга на британских кораблях практически исчезла. (Потребовалось еще 100 лет, чтобы широкое потребление витамина C привело к искоренению цинги на суше[2].)

Эксперимент Линда примечателен тем, что он стал первым примером анализа, основанного на фактических данных, – серьезной попытки провести тщательное клиническое исследование с целью оценки эффективности рекомендованных средств. Также поучительно выявить направления, по которым его эксперимент можно было бы улучшить. Например:

1) должна была быть контрольная группа, которая не подвергалась специальной процедуре. Исследование Линда показало, что те пациенты, которые получали цитрусовые, чувствовали себя лучше, чем те, кому давали морскую воду, но, возможно, это было связано с вредным воздействием морской воды, а не с полезным действием цитрусовых.

---

[2] Лимонники (limeys) – это прозвище британских моряков после открытия средства от цинги. – *Прим. перев.*

Группа студентов колледжа однажды сообщила, что после того, как им была сделана экспериментальная вакцинация, количество простудных заболеваний сократилось на 73 % по сравнению с предыдущим годом. Это было бы удивительно, если бы не тот факт, что контрольная группа сообщила о снижении заболеваемости на 63 %!

2) распределение предполагаемых лекарств должно было быть определено путем случайной жеребьевки. Возможно, Линд считал цитрусовые наиболее перспективным лекарством и подсознательно давал цитрусовые самым здоровым пациентам.

Однажды исследование показало, что дети, которым была сделана экспериментальная вакцинация, имели на 80 % меньше шансов умереть от туберкулеза по сравнению с детьми, которые не получили вакцину. Однако отобранные для вакцинации дети были выбраны потому, что можно было легко получить согласие родителей, и они, возможно, систематически отличались от детей в семьях, которые отказывались давать согласие. Последующее исследование, в котором участвовали только те дети, которые получили согласие родителей, не выявило различий между показателями смертности среди вакцинированных и не вакцинированных детей;

3) было бы лучше, если бы эксперимент был двойным слепым, т. е. чтобы ни Линд, ни пациенты не знали, что они получают. В противном случае Линд мог увидеть то, что хотел увидеть, а пациенты могли сказать Линду то, что, по их мнению, он хотел услышать. В современных условиях некоторым пациентам можно давать витамин С в таблетках, в то время как другим дают плацебо, которое выглядит и имеет вкус витамина С, но на самом деле является инертным веществом.

В ходе одного эксперимента на старших курсах экспериментальной психологии каждому из двенадцати студентов дали по пять крыс для тестирования в лабиринте. Шести студентам сказали, что их крысы «соображают в прохождении лабиринта», потому что они были выведены от крыс, которые очень хорошо справлялись с тестами в лабиринте; остальным шести студентам сказали, что их крысы «не соображают в прохождении лабиринта». На самом деле никакого специального разведения не проводилось. Студентам дали случайно выбранных обычных крыс. Тем не менее, когда студенты тестировали крыс в прохождении лабиринта, крысы, которых назвали «соображающими в прохождении лабиринта», получили более высокие оценки, и их оценки со временем увеличивались еще больше, что указывало на то, что они были умнее и обучались быстрее, чем несообразительные крысы, но на самом деле это отражало систематически смещенные (предвзятые) ожидания исследователей;

4) пациентов должно было быть больше. Два принимающих лекарство пациента являются эпизодическими, а не убедительными свидетельствами. Имея в наличие десятки или сотни случайно подразделенных

пациентов, можно рассчитать шансы того, что наблюдаемые различия обусловлены чисто удачным стечением обстоятельств.

Если двое из двенадцати изученных Линдом пациентов выздоровели, то шансы того, что – по чистой удаче – они получили то же самое лекарство, составляет 9 %, что является многообещающим, но не убедительным свидетельством в пользу эффективности лечения.

Золотой стандарт

Золотым стандартом медицинских тестов является рандомизированное контролируемое испытание (РКИ), удовлетворяющее четырем направлениям, по которым эксперимент Линда можно было бы улучшить.

1. *Контролируемый*: контрольная группа получает плацебо в дополнение к группе, подвергающейся лечебной процедуре, в результате чего можно сравнивать лечебную процедуру с отсутствием лечебной процедуры, не беспокоясь об эффекте плацебо или естественной способности организма к выздоровлению.
2. *Рандомизированный*: испытуемые случайным образом распределяются между процедурной и контрольной группами, в результате чего не нужно беспокоиться о том, отличаются ли люди, подвергающиеся лечебной процедуре, от тех, кто ей не подвергается. При достаточно большой выборке различия между испытуемыми будут усредняться.
3. *Слепой*: тест является дважды слепым, в результате чего ни на испытуемых, ни на исследователей не влияет знание о том, кто подвергается процедуре, а кто нет.
4. *Крупный*: имеется достаточно данных, чтобы сделать наполненные смыслом выводы; в случае относительно редких заболеваний могут потребоваться тысячи пациентов.

Первые два условия, собственно говоря, и дали название РКИ; два других условия придают тестам больше убедительности.

После завершения исследования статистическим вопросом становится вероятность того, что по чистой случайности разница между двумя группами окажется столь же большой, как и фактически наблюдаемая. Большинство исследователей считают «статистически значимой» вероятность менее 0,05. Различия между процедурной и контрольной группами считаются статистически убедительными, если вероятность их возникновения по чистой удаче составляет менее одного шанса из двадцати.

Тесты РКИ могут (и должны) использоваться вне медицинских исследований и не ограничиваться одной «лечебной процедурой». При изучении разных способов преподавания математики можно рассматривать три новых подхода в дополнение к существующему. При исследовании присадок к топливу можно рассматривать четыре возможности в дополнение к топливу без присадок.

Самое замечательное в тестах РКИ то, что они могут демонстрировать причинно-следственные взаимосвязи. Если получающие витамин С пациенты излечиваются от цинги, в то время как получающие плацебо пациенты не излечиваются, то очевидно, что витамин С ответственен за исчезновение цинги.

Приведенный выше пример также иллюстрирует два других аспекта причинной следственности. Во-первых, причинная следственность может быть не определенностью, а тенденцией. Можно с полным основанием утверждать, что витамин С снижает шансы развития цинги и/или увеличивает шансы на выздоровление от цинги, но при этом не утверждать, что он эффективен в 100 % случаев. Утверждение «хроническое употребление алкоголя может стать причиной заболевания печени» не означает, что у всех, кто употребляет алкоголь, развиваются заболевания печени, а только то, что у людей, употребляющих большое количество алкоголя, чаще всего развиваются заболевания печени.

Во-вторых, бывает полезно сказать, что «нечто служит причиной чего-то», не зная точно, как это происходит. Линду не нужно было знать причину, по которой апельсины и лимоны борются с цингой, чтобы их прописывать в качестве лечения. Сегодня мы понимаем, как витамин С творит свое волшебство, однако есть много других случаев, когда можно быть уверенным в том, что А служит причиной В, не зная точно, как именно.

A/B-тесты

В нашем мире, в котором доминирует интернет, возникает масса каверзных вопросов об эффективности различных «лечебных процедур». Будет ли веб-сайт продавать больше гаджетов, если использовать другой цвет фона? Привлечет ли новостная рассылка больше подписчиков, если на главной странице веб-сайта будет размещена информация об одобрении? Привлечет ли онлайновый клуб больше участников, если в нем будут размещены фотографии сотрудников клуба?

На эти и многие другие вопросы можно ответить с помощью РКИ, которые обычно называют *A/B-тестами* в ситуациях, когда они используются в интернет-маркетинге. A/B в названии указывает на то, что A сравнивается с B, хотя также могут быть сравнения с C, D и E.

Давайте рассмотрим вопрос о цвете фона. Возможно, компания использовала белый фон на своей главной странице и рассматривает возможность перехода на небесно-голубой фон. A/B-тест должен проводиться в соответствии с протоколом всех РКИ.

1. *Контролируемый*: страница с небесно-голубым фоном является процедурной, а страница с белым фоном – контрольной.
2. *Рандомизированный*: когда пользователь вводит URL-адрес веб-сайта, используется генератор случайных событий, который определяет, куда пользователь будет отправлен: на небесно-голубую страницу либо на белую.

3. *Слепой*: исследователи напрямую не отслеживают, кто на какую страницу отправляется, и пользователи не знают, что они являются частью эксперимента.
4. *Крупный*: основываясь на данных о прошлых продажах, эксперимент будет проводиться до тех пор, пока не наберется, возможно, достаточное количество продаж, чтобы провести наполненное смыслом сравнение двух фоновых цветов.

После сбора данных применяется статистический тест, чтобы определить, можно ли обоснованно объяснить наблюдаемую разницу в продажах случайностью, связанной с отправкой людей на разные страницы, или же это свидетельство того, что цвет фона действительно играет роль.

Как и другие РКИ, A/B-тесты можно использовать для демонстрации причины и следствия. Предположим, что при неизменности всего остального новый цвет фона сгенерировал гораздо больший объем продаж, чем изначальный цвет фона. Наиболее убедительным объяснением будет то, что изменение цвета фона привело к увеличению продаж. Как и во всех тестах РКИ, причинная следственность может быть не определенностью, а тенденцией. Веб-сайт не перешел от нулевого процента продаж к 100%-му, но шансы на продажу возросли. И опять же мы почти наверняка не знаем точно причину, по которой новый цвет фона понравился покупателям больше, – да нам и не нужно ее знать, – но мы можем сделать вывод о том, что изменение цвета привело к увеличению продаж.

A/B-тестирование в интернете стало настолько распространенным явлением, что в интернет-маркетинге бытует шутка о том, что A/B-тестирование – это аббревиатура от английского выражения «Always Be Testing». (Всегда должно тестироваться.)

Рассуждение «после этого, следовательно, из-за этого»

Когда одно событие предшествует другому, возникает естественная склонность думать, что первое событие является хорошим предвестником второго события. Это не обязательно бывает именно так. Много лет назад в рекламе курортов Club Med было сделано вот такое ироничное замечание:

> Я сидел у воды после напряженного дня, проведенного за теннисом и виндсерфингом, – рассказал отдыхавший здесь биржевой брокер из Филадельфии. – Слушал классическую музыку с группой других людей – это была «Музыка воды» Генделя – и вдруг посмотрел на океан. Я не мог поверить своим глазам.
>
> Прилив, который шел уже несколько часов, на самом деле остановился как вкопанный, развернулся и начал отступать…

Латинская фраза *post hoc ergo propter hoc* («после этого, следовательно, из-за этого») описывает логическую ошибку допущения о том, что раз одно событие следует за другим, то оно должно быть обусловлено первым событием. Воспро-

изведение классической музыки предшествовало наступлению отлива, но не стало его причиной.

Логическая ошибка «после этого, следовательно, из-за этого» особенно вредоносна тем, что в отличие от совпадающих по времени корреляций порядок следования событий, A перед B, убедительно намекает о причинной следственности.

Лауреат Нобелевской премии по экономике Клайв Грейнджер (Clive Granger) наиболее известен благодаря тесту на причинную следственность, который он предложил, когда не представлялось возможным проводить РКИ. Существует ли причинно-следственная взаимосвязь между процентными ставками и ценами на акции? Между уровнем безработицы и исходами президентских выборов? Мы все благодарны экономистам за то, что они не могут манипулировать процентными ставками и уровнем безработицы, чтобы собирать данные для своих исследований.

Вместо этого Грейнджер предложил обратиться к данным наблюдений и посмотреть, помогает ли одна переменная предсказывать будущие значения другой. Если процентные ставки помогают предсказывать будущие цены на акции, но не наоборот, то считается, что изменения процентных ставок служат причиной изменений цен на акции. Такой вывод, очевидно, является примером логической ошибки «после этого, следовательно, из-за этого», поэтому вместо истинной *причинной следственности* указанный тест обычно называют тестом на *причинную следственность Грейнджера*. Таким образом, процентные ставки по Грейнджеру являются причиной цен на акции.

Петухи начинают кукарекать перед рассветом, заявляя о своих территориальных претензиях. Тест Грейнджера показал бы, что кукареканье начинается до того, как небо начинает светлеть, но небо не становится светлым до того, как закукарекают петухи. Следовательно, в результате теста можно сделать вывод, что кукареканье служит причиной того, что небо начинает светлеть, а восходящее солнце не служит причиной кукареканья. Разумеется, правда заключается в том, что ни то ни другое не служит причиной друг друга в любом наполненном смыслом понимании этого слова. Петухи кукарекают в ожидании восхода солнца благодаря циркадному ритму, контролируемому внутренними часами. Петухи, которых круглосуточно держат при слабом освещении, кукарекают (приблизительно) каждые двадцать четыре часа, когда они ожидают восхода солнца.

Без тестов РКИ очень трудно устанавливать истинную причинную следственность.

От хорошего к так себе

Испытания РКИ не всегда возможны. Если мы хотим узнать причину, по которой доходность акций одних компаний выше, чем у других, мы не можем просто взять на себя управление компаниями, изменить методы их управления и посмотреть, что произойдет. Вместо этого приходится довольствоваться данными наблюдений – мы наблюдаем, как разные компании занимаются разными вещами, и пытаемся сделать выводы. Это стоящая цель, но есть веские основания для осторожности. Вспомните последовательность прове-

дения РКИ: выбрать процедурную и контрольную группы, собрать данные и сравнить результаты. Ошибочно будет делать обратное: искать закономерности в данных наблюдений, а затем относиться к обнаруженным закономерностям так, как если бы они были получены в результате испытания РКИ. Такой подход называется *выдвижением гипотез после того, как результаты известны* (англ. *Hypothesizing After the Results are Known*, аббр. HARKing). Резкое звучание английской аббревиатуры отражает его опасности.

Классическим примером является исследование 2001 года, получившее известность как «От хорошего к отличному». Гэри много писал об этом исследовании, но мы пересказываем историю здесь, потому что это очень наглядный пример выдвижения гипотез после того, как результаты известны. После тщательного изучения котировок акций 1435 компаний за 1965–2000 годы Джим Коллинз (Jim Collins) определил одиннадцать компаний, которые превзошли рынок в целом:

Abbott Laboratories	Kimberly-Clark	Pitney Bowes
Circuit City	Kroger	Walgreens
Fannie Mae	Nucor	Wells Fargo
Gillette	Philip Morris	

Портфель из этих одиннадцати акций приносил бы финансовый возврат на вложенный капитал в размере 19,2 % в годовом исчислении в период с 1965 по 2000 год по сравнению с финансовым возвратом в размере 12,2 % по рынку в целом. После определения этих одиннадцати акций Коллинз проанализировал их общие характеристики и сообщил о пяти обнаруженных им общих темах.

1. Уровень лидерства 5: лидеры, скромны в личном плане, но профессионально стремятся сделать компанию великой.
2. Сначала кто, а затем что: важнее принимать на работу нужных людей, чем иметь хороший бизнес-план.
3. Смотреть правде в глаза: при принятии правильных решений учитываются все факты.
4. Концепция ежа: лучше быть мастером в одной области, чем браться за все.
5. Формировать видение своей компании: адаптировать методы работы и стратегии, но не отказываться от ключевых ценностей компании.

Эти пять характеристик правдоподобны, а названия запоминаются, но делать наблюдения о прошлом гораздо проще, чем делать предсказания о будущем.

Свидетельства в пользу этих пяти характеристик были бы убедительными, если бы Коллинз *заранее* выявил компании, которые обладают или не обладают этими характеристиками (процедурная и контрольная группы), а затем проследил за их успехом. Вместо этого он заглянул в результаты и обнаружил пять закономерностей. Он выдвинул гипотезы после того, как результаты были известны, и возгордился этим:

> Важно понимать, что все изложенные в этой книге концепции были разработаны нами путем эмпирического выведения непосредственно из данных. Мы не начинали этот проект с проверки или с доказательства теории. Мы стремились построить теорию с нуля, выводя непосредственно из документальных сведений.

Неудивительно, что одиннадцать акций Коллинза после их отбора показали не такие высокие результаты, как до их отбора. На рис. 2.2 показана результативность портфеля из его одиннадцати акций по сравнению с рынком в целом, начиная с инвестиций в размере 1 долл. США на 1 января 2002 года, вскоре после публикации книги «От хорошего к отличному» (Good to Great). Годовая возвратность (норма прибыли) портфеля от-хорошего-к-отличному составила 4,7 % и рыночного портфеля – 6,9 %.

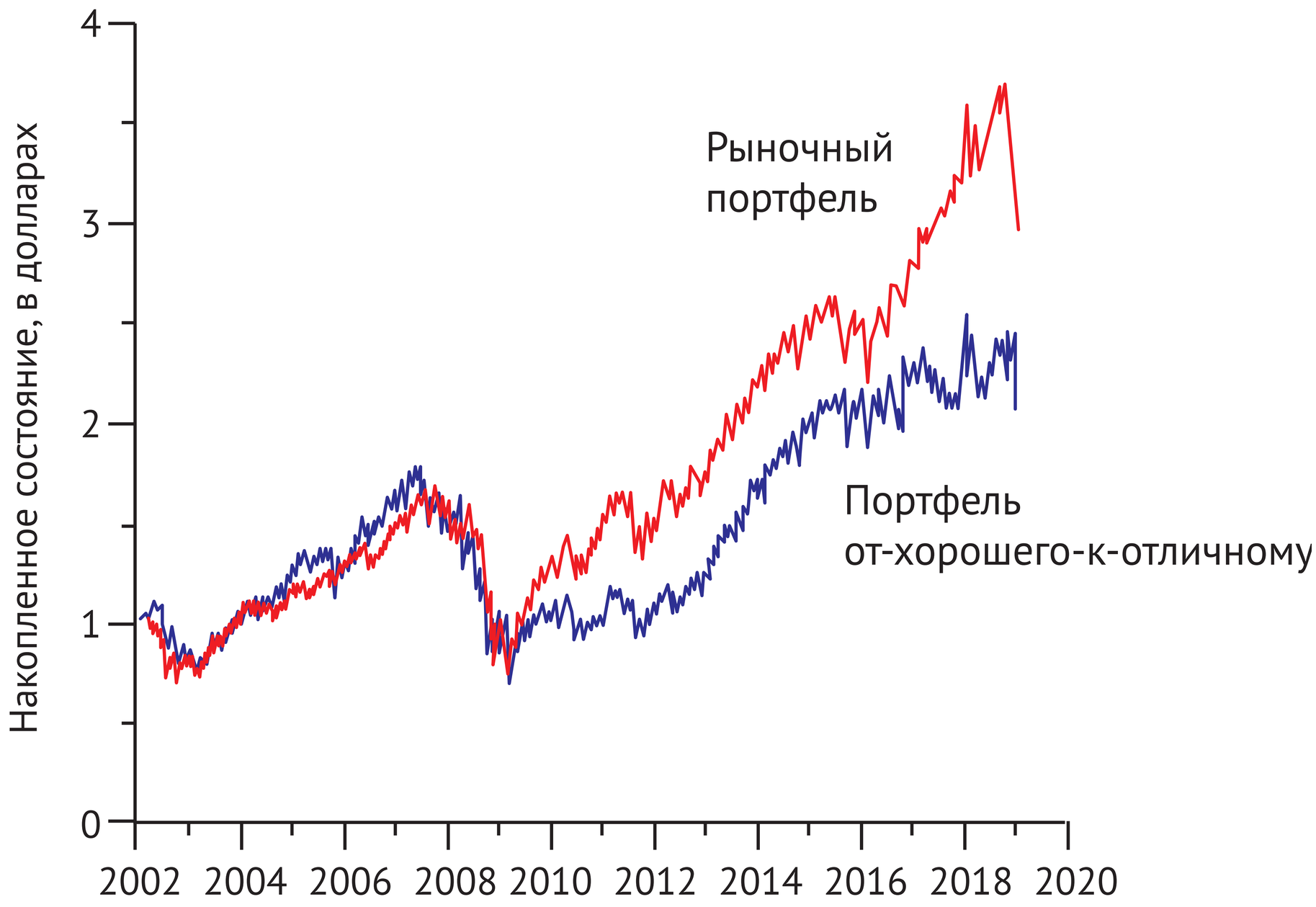

Рис. 2.2. От отличного к так себе

Проблема выдвижения гипотез после того, как результаты известны, характерна для формул/секретов/рецептов в отношении того, как стать богатым, иметь прочный брак, дожить до 100 лет и т. д. и т. п., которые основаны на ретроспективных исследованиях состоятельных людей, прочных браков и долгой жизни.

Например, в 2019 году новостной веб-сайт Business Insider сообщил: «По словам директора по исследованиям в рамках Института рынка изобилия (Affluent Market Institute) Сары Стэнли Фаллоу, если вы обладаете определенным набором характеристик, то у вас больше шансов стать богатым». Госпожа Фаллоу опросила более 600 американских миллионеров и выделила шесть характеристик, которые она называет «факторами богатства», коррелирующими с богатством, например бережливость и уверенность в управлении

финансами. За исключением того, что у нее не пять, а шесть характеристик, получилось в точности как в книге «От хорошего к отличному» и так же бесполезно.

Если мы считаем, что знаем какие-то секреты успеха, то правильным способом проверить нашу теорию было бы выявить людей с этими чертами и посмотреть, как они будут себя вести в течение следующих десяти, двадцати или пятидесяти лет. В противном случае вместо того, чтобы предсказывать будущее, мы просто всматриваемся в прошлое.

Сплетни

Вот пример исследования, которое шло в другом направлении – от теории к данным, а не наоборот.

Легендарный инвестор Филип Фишер (Philip Fisher) расхваливал ценность «сплетней» в своей классической книге 1958 года «Обыкновенные акции и необыкновенные прибыли» (см. библиографию). Сплетни собираются путем опроса управляющих, сотрудников, клиентов и поставщиков компании, а также знающих людей в отрасли, чтобы выявлять способные компании с хорошими перспективами роста. Возможно, Уолл-стрит слишком сосредоточен на цифрах (объемах продаж, доходе, прибылях), которые сильно колеблются от квартала к кварталу, и не уделяет достаточного внимания глубинным качествам, которые делают компанию успешной.

Сплетни трудно измерить, но Гэри придумал интересный способ. С 1983 года журнал Fortune ежегодно публикует список компаний, пользующихся наибольшим почитанием, на основе опросов тысяч руководителей, директоров и аналитиков по безопасности. В первую десятку (по порядку) в 2020 году вошли: Apple, Amazon, Microsoft, Walt Disney, Berkshire Hathaway, Starbucks, Alphabet (Google), JPMorgan Chase, Costco и Salesforce.

В 2006 году Гэри и еще один студент проанализировали динамику портфеля акций десяти крупнейших компаний по рейтингу журнала Fortune за период с 1983 по 2004 год. Формирование портфеля «Наиболее почитаемые» началось с инвестирования равной суммы в каждую из десяти самых почитаемых компаний 1983 года, причем инвестиции делались в день официальной публикации журнала, т. е. через несколько дней после того, как журнал поступал в продажу. Когда был опубликован список самых почитаемых компаний 1984 года, акции 1983 года продавались, а вырученные средства вкладывались в десятку лучших компаний 1984 года и т. д., пока не были сделаны окончательные инвестиции в десятку наиболее почитаемых компаний 2004 года.

Стратегия «Наиболее почитаемые» превосходила индекс S&P 500 на 2,2 процентных пункта в год с соответствующей годовой возвратностью 15,4 % против 13,2 %. Каждый доллар, первоначально вложенный в портфель «Наиболее почитаемые» и рассчитанный по формуле сложных процентов за двадцать два года, вырос бы до 23,29 долл. по сравнению с 15,18 долл. за доллар, первоначально вложенный в S&P 500.

Вряд ли эта разница является своего рода премией за риск, поскольку в качестве наиболее почитаемых в Америке компаний были выбраны крупные и

финансово устойчивые организации, а их акции чаще всего рассматриваются инвесторами как очень безопасные. По обычным статистическим меркам они были более безопасными. Разница в возвратности тоже не связана с выдающимися показателями нескольких компаний. Почти 60 % акций наиболее почитаемых компаний превышают индекс S&P 500.

Возможно, Фишер был прав. Для того чтобы превзойти рынок, нужно сосредотачиваться на сплетнях – нематериальных активах, которые не отражаются в балансовом отчете компании, – а опрос журнала Fortune – это самая настоящая сплетня.

Гэри недавно пересмотрел эту стратегию и обнаружил, что стратегия «Наиболее почитаемые» продолжает опережать индекс S&P 500. Маржа за последующий четырнадцатилетний период 2005–2018 годов фактически была немного выше, чем за начальный период – 10,09 % против 7,27 %, что составляет разницу в 2,82 процентных пункта.

За весь тридцатишестилетний период стратегия «Наиболее почитаемые» приносила 13,30 % годовой прибыли по сравнению с 10,83 % для индекса S&P 500. На рис. 2.3 показано, что каждый доллар, первоначально вложенный в портфель «Наиболее почитаемые» в 1983 году, вырос бы до 89,48 долл. по сравнению с 40,58 долл. для S&P500.

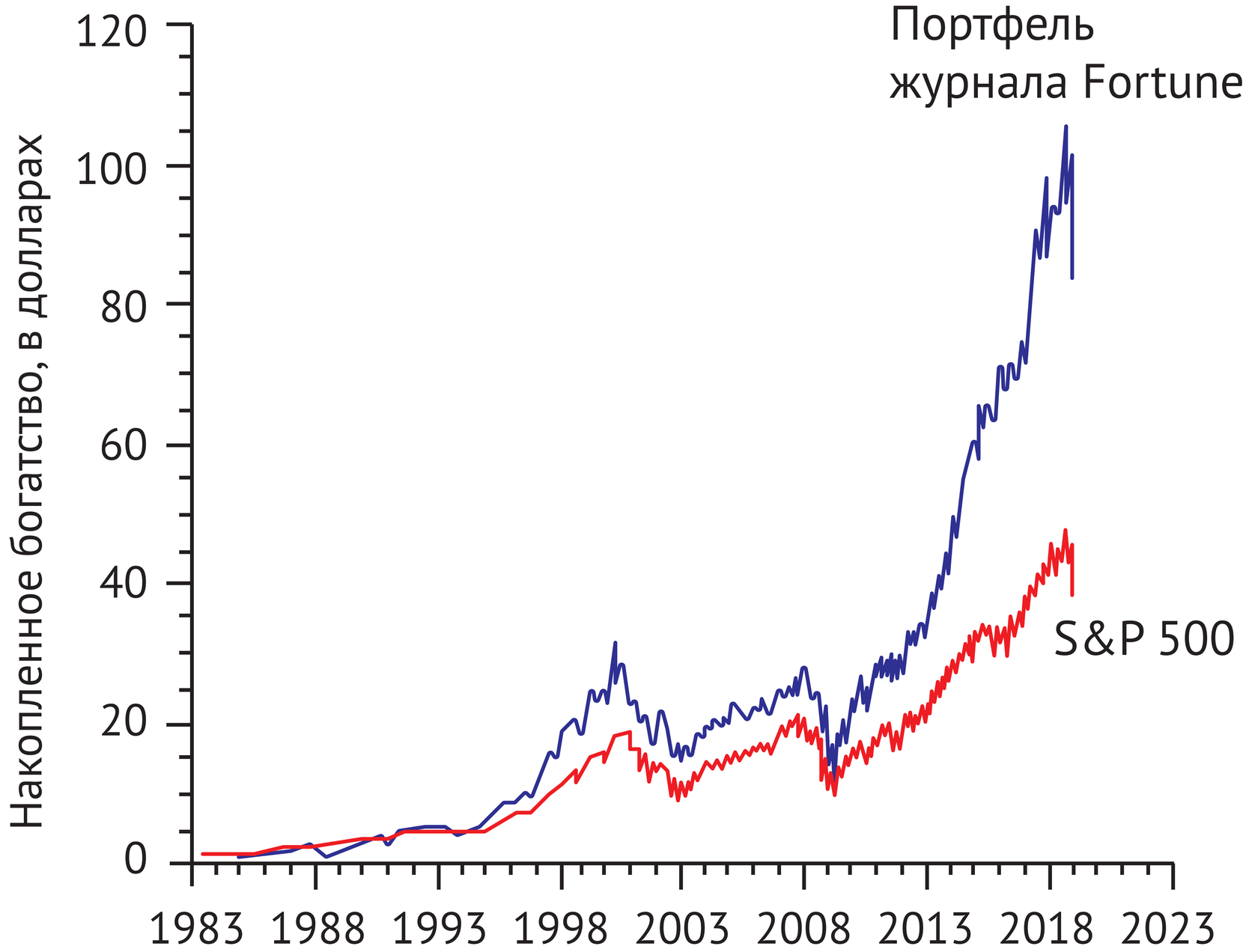

Рис. 2.3. Портфель «Наиболее почитаемые», в сравнении с индексом S&P 500, с 1983 по 2018 год

Из табл. 2.1 видно, что эти высокие показатели не зависели от ежегодных покупок, совершаемых в дату публикации. Совершать покупки через пять, десять, пятнадцать или даже двадцать дней после даты публикации все равно было бы выгодной стратегией.

Таблица 2.1. Годовая возвратность от покупок, совершаемых после даты публикации журнала Fortune, 1983–2018 гг.

Спустя несколько дней после публикации	Портфель «Наиболее почитаемые»	S&P 500
0	13,30 %	10,83 %
5	13,53 %	10,76 %
10	13,64 %	10,83 %
15	13,91 %	11,09 %
20	13,68 %	10,87 %

Описанное выше исследование возвратности акций отличает от исследования «От хорошего к отличному» то, что оно было основано на правдоподобной теории, которая была разработана до рассмотрения данных, использованных для проверки теории, а затем была повторена на дополнительных данных. У исследований гораздо больше шансов оказаться полезными, когда теория опережает данные, а не наоборот.

Если любишь свою работу

Один студент как-то спросил Гэри, где ему стоит работать: в сфере управленческого консалтинга либо инвестиционного банковского дела. Консалтинг считается менее утомительным занятием, но банковское дело потенциально гораздо перспективнее. Гэри ответил пословицей, которую часто приписывают Конфуцию: «Если ты любишь свою работу, ты не проработаешь ни дня в своей жизни». Идеальная работа – это та, на которой вы просыпаетесь по утрам с желанием приступить к работе, работа, с которой вы никогда не захотите уходить на пенсию, которую вы (почти) готовы выполнять бесплатно. Слишком многие люди устраиваются на работу, которую ненавидят, считая дни до выхода на пенсию. Это не тот способ проживать свою жизнь.

Отсюда следует, что люди, которые любят свою работу, скорее всего, будут более продуктивны, а компания, в которой работают люди, любящие свою работу, скорее всего, будет успешней. Возможно, как и в случае с наиболее почитаемыми компаниями, выявление фирм, сотрудники которых любят свою работу, является ценным источником информации.

В целях подтверждения этой теории Гэри и двое его студентов, Сара Ландау и Оуэн Роузбек, изучили ежегодный список пятидесяти «Лучших мест для работы», составляемый компанией Glassdoor. Запущенный в 2008 году веб-сайт компании Glassdoor позволяет нынешним и бывшим сотрудникам оценивать свои компании по пятибалльной шкале от 1 (наименее привлекательная) до 5 (наиболее привлекательная). На данный момент на нем собрано более пятидесяти миллионов отзывов почти о миллионе компаний.

Веб-сайт Glassdoor усредняет рейтинги каждый декабрь (при этом недавние рейтинги имеют больший вес, чем рейтинги за последние годы) и публично объявляет «Лучшие места для работы». Каждый год с 2009 по 2017 годы составлялись отчеты о пятидесяти лучших компаниях; в 2018 году были представ-

лены данные о 100 лучших компаниях, но в целях обеспечения стабильности коллектив Гэри рассматривал только пятьдесят лучших в том году.

Ежегодно около 40 % входящих в число лидеров компаний являются частными фирмами или дочерними компаниями более крупных компаний; например, в рейтинге 2018 года лидером по количеству голосов стала Facebook; следующие три были частные компании (Bain, Boston Consulting и In-N-Out Burger), за которыми пятой следовала Google.

Очевидно, что это не идеальная система. Сотрудники компании Glassdoor проверяют все отзывы (и примерно 20 % из них выбраковываются), однако нет никакого практического способа обеспечить, чтобы отзывы были честными или отражали точку зрения других сотрудников. Люди, которые относятся к компании, в которой они работают, хорошо, чаще тратят время на написание отзывов (и более склонны к преувеличениям), чем люди, которым в основном безразлично.

Тем не менее в этих рейтингах может быть какая-то польза. В целях проверки своей теории Гэри, Сара и Оуэн сравнили показатели портфеля лучших мест для работы за десять лет с индексом S&P 500. В начале каждого года портфель лучших мест инвестировал равную сумму в публично торгуемые акции лучших компаний для работы, которые были объявлены в декабре, и эти акции находились во владении до начала следующего года.

На рис. 2.4 показано, что первоначальная инвестиция в размере 1 долл. в портфель лучших мест на 1 января 2009 года выросла бы до 5,52 долл. на 31 декабря 2018 года, что составляет 19%-ную возвратность в годовом исчислении, в то время как инвестиция в размере 1 долл. в S&P 500 выросла бы до 3,42 долл., что соответствует 13%-ной годовой возвратности.

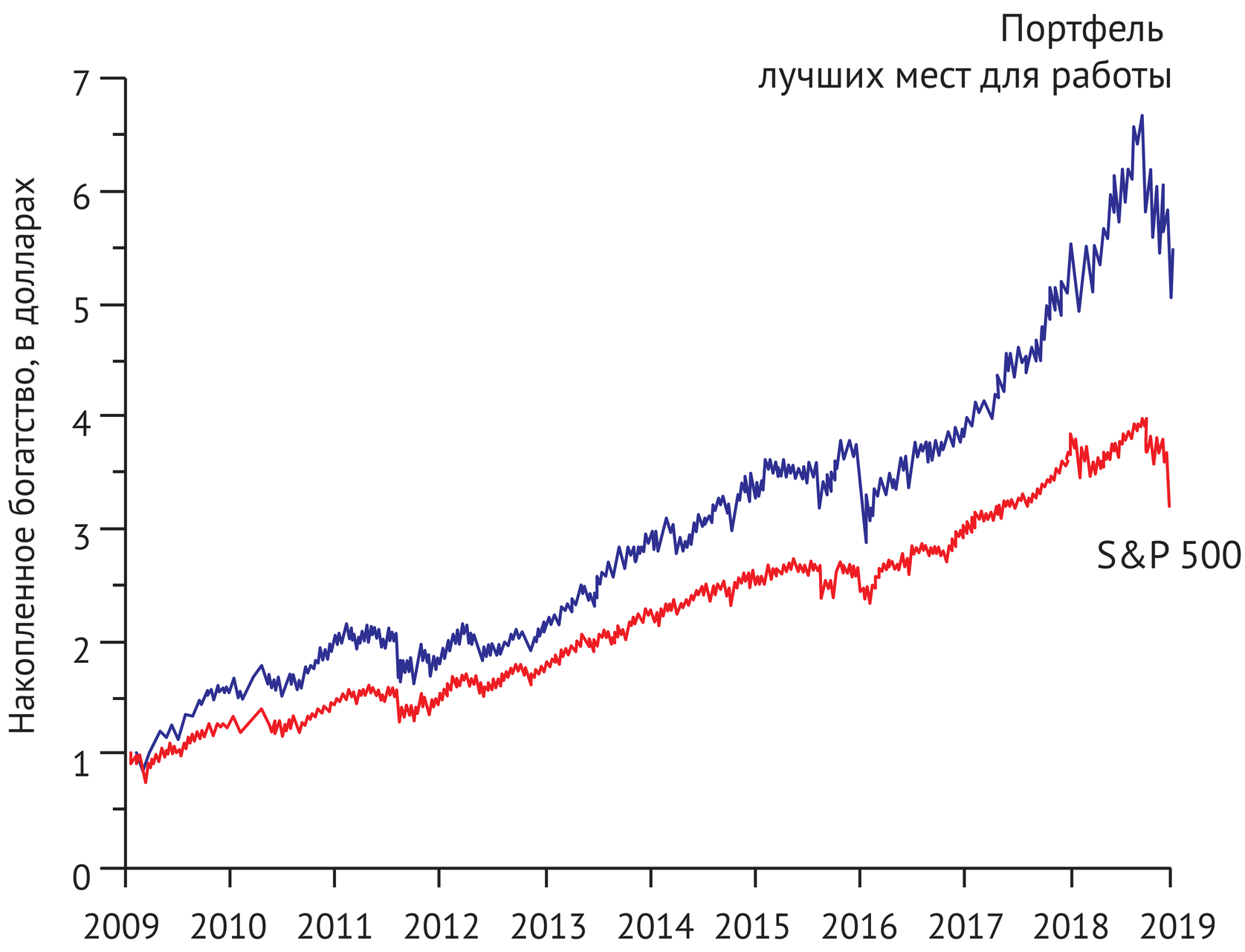

Рис. 2.4. Счастливые работники, счастливые акционеры?

Вполне возможно, что рейтинги от сотрудников являются полезными сплетнями.

Продуманные тикеры

А вот более провокационное исследование, которое тоже было основано на теории.

Корпоративные акции традиционно обозначаются тикерными символами, или биржевыми кодами бегущей строки (названными так потому, что информация о торгах обычно передавалась на биржевом аппарате, передающем биржевые коды компаний и котировки их ценных бумаг в бегущей строке; такие коды называются тикерными символами или просто тикерами). Компании выбирают свои тикерные символы и традиционно используют аббревиатуры названия компании, например AAPL для Apple и GOOG для Google. Иногда тикерные символы становятся настолько привычными, что компании становятся известны по их названиям. International Business Machines (тикер: IBM) теперь повсеместно называется IBM, а Minnesota Mining and Manufacturing (тикер: MMM) в 2002 году сменила свое официальное название на 3M.

За последние несколько десятилетий десятки компаний отказались от традиционного сокращения названия и выбрали символы, которые соответствуют тому, чем занимается компания. Некоторые из них запоминаются своей дерзкой находчивостью; например, выбор компании Southwest Airlines в пользу LUV в качестве тикерного символа был связан с тем фактом, что ее штаб-квартира находится в Лав Филд в Далласе, и компания Southwest Airlines хотела позиционировать себя как авиакомпанию, «построенную на любви».

Те, кто верит, что рынок акций «эффективен», а цена акции точно отражает ее истинную стоимость, отвергают идею о том, что на цены акций могут влиять такие поверхностные вещи, как тикерные символы. Однако инвесторы не всегда так рациональны, как полагают энтузиасты эффективного рынка, и сюда входит влияние тикерных символов. Например, было несколько примечательных случаев, когда инвесторы по ошибке покупали или продавали не те акции, потому что они были сбиты с толку тикерным символом.

Во время работы над диссертацией по экономике в Гарварде Майкл Рашес (Michael Rashes) заметил, что предложение о поглощении компании MCI Communications вызвало скачок цен на акции с тикером MCI. К сожалению, тикерным символом компании MCI Communications был MCIC! Инвесторы по ошибке бросились покупать акции совершенно не связанной компании MassMutual Corporate Investors в совершенно другой отрасли, потому что у нее был тикерный символ MCI. Рашес собрал несколько других примеров и опубликовал статью в журнале «Journal of Finance» под замечательным названием «Инвесторы, находясь в массовом замешательстве, принимают явно неосведомленные решения: (MCI-MCIC)» (см. библиографию).

Инвестиционные решения иногда искажаются ошибками и некорректными суждениями. В конце концов, мы всего лишь люди.

Существуют многочисленные свидетельства того, что суждения человека формируются в зависимости от легкости, с которой информация обрабатывается и запоминается. Например, такие утверждения, как «Осорно находится

в Чили», скорее всего, будут признаваться правдивыми, если они написаны с более удобочитаемой цветностью, как и хорошо зарифмованные афоризмы, например «На языке медок, а на сердце ледок» против «На языке медок, а на сердце вражда»[3].

Кроме того, неоднократно демонстрировалось, что переживания, вызывающие положительное эмоциональное возбуждение, запоминаются чаще и что люди чаще испытывают положительные чувства по поводу вещей, связанных с этим положительным опытом.

Указанные аргументы свидетельствуют о том, что легко произносимые и продуманные тикерные символы легче запоминаются и оцениваются положительно, что может оказывать влияние на цены акций. Мы не можем провести A/B-тест, рандомно назначив выбранным компаниям продуманные и скучные тикерные символы и посмотрев, не будет ли отличаться возвратность акций, но зато такие компании можно выявлять, а потом смотреть на результативность их акций.

В 2006 году Гэри и двое студентов проанализировали 33 000 тикерных символов, используемых компаниями в прошлом и настоящем, в поисках символов, которые можно было бы счесть заслуживающими внимания. Девяносто три процента отобранных совпали. Они объединили списки и удалили тикеры, которые были просто сокращением названия компании (например, BEAR от Bear Automotive Service Equipment), и оставили намеренно продуманные заголовки (GRRR для сафари-парков Lion Country и MOO для United Stockyards).

Они распространили 100 опросов с отобранным списком из 358 тикерных символов, названий компаний, кратким описанием деятельности каждой компании и следующими ниже инструкциями:

> Торговля акциями осуществляется с использованием биржевых символов бегущей строки – тикеров. Некоторые из них – это просто название компании (GM, IBM); некоторые – узнаваемые аббревиатуры названия компании (MSFT для Microsoft, CSCO для Cisco); а некоторые – непроизносимые сокращения (BZH для Beazer Homes, PXG для Phoenix Footwear Group). Некоторые компании выбирают символы, которые непосредственно связаны с их деятельностью; например, компания, производящая футбольное снаряжение, может выбрать GOAL; служба знакомств в интернете может выбрать LOVE.
>
> Из прилагаемого списка тикерных символов выберите, пожалуйста, 25 самых ярких, симпатичных и запоминающихся.

Из списка опрошенных были намеренно исключены опытные специалисты в области инвестиций, поскольку на их выбор могли повлиять знания об инвестиционной деятельности включенных в список компаний.

Для каждого торгового дня с начала 1984 года (когда символы с продуманным тикером начали приобретать популярность) до конца 2005 года они рассчитывали ежедневную возвратность портфеля из восьмидесяти двух акций с продуманным тикером, набравших наибольшее количество голосов в опросе. В качестве контрольной группы они использовали совокупную результативность рынка акций.

---

[3] В оригинале Woes unite foes (Беды объединяют недругов) против Woes unite enemies (Беды объединяют врагов). – *Прим. перев.*

На рис. 2.5 показано, что до 1993 года портфель продуманных тикеров немного отставал от рыночного портфеля, а затем резко вырвался вперед. В целом вычисленная по формуле сложных процентов годовая возвратность портфеля продуманных тикеров составила 23,5 %, а рыночного портфеля – 12,0 %. Из-за влияния сложных процентов за этот двадцатидвухлетний период один вложенный в рынок доллар вырос бы до 12,17 долл., в то время как один вложенный в портфель продуманных тикеров доллар вырос бы до 104,69 долл.

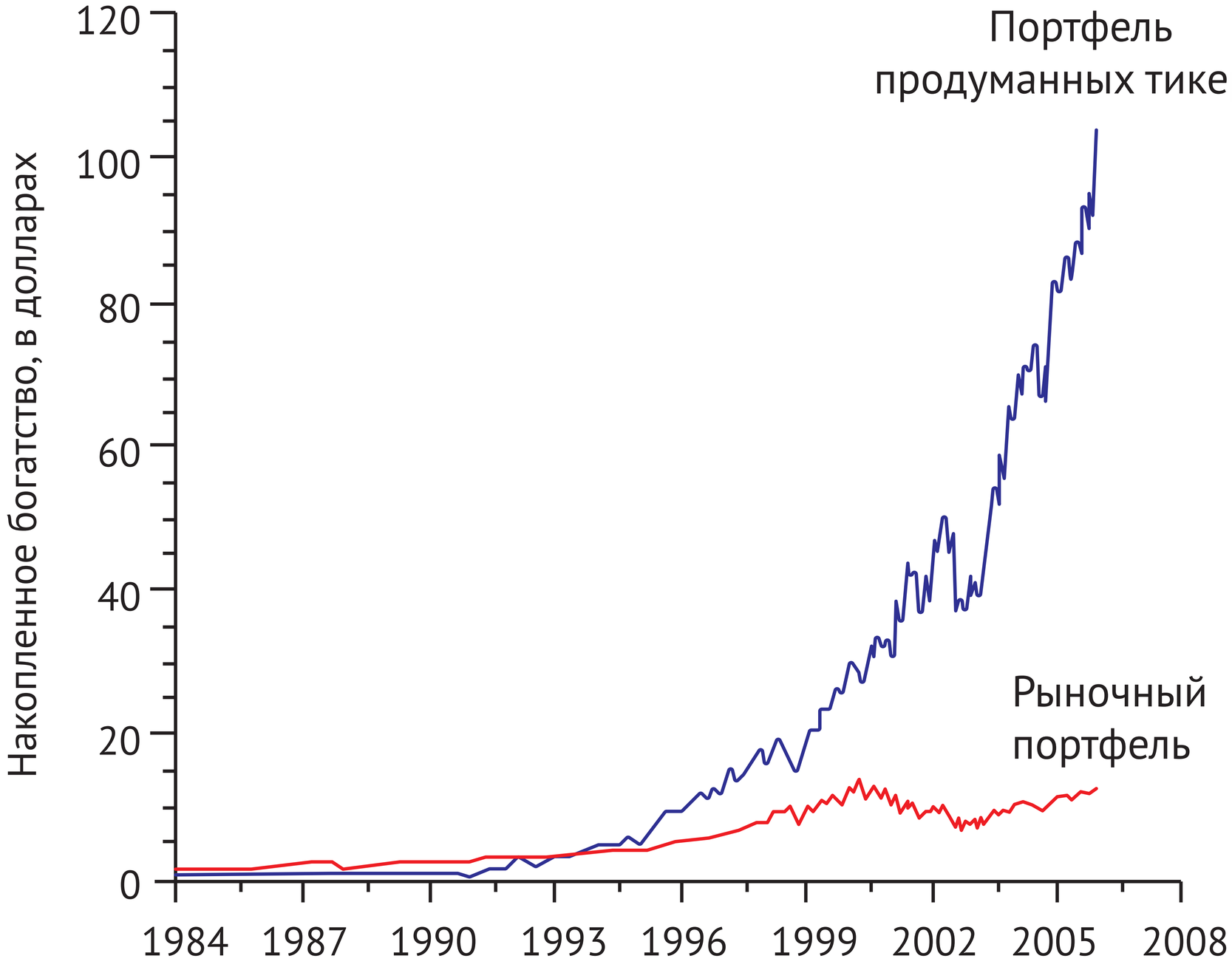

Рис. 2.5. Результативность портфеля продуманных тикеров за период с 1984 по 2005 год

Превосходящая рынок результативность была достигнута не потому, что акции продуманных тикеров были сосредоточены в одной отрасли. Восемьдесят две компании с продуманными тикерами охватывали тридцать одну из восьмидесяти одной отраслевой категории, используемой правительством США, при этом наибольшая концентрация была у восьми компаний, работающих в ресторанах и питейных заведениях, из которых четыре превзошли рынок, а четыре – нет. Успех портфеля продуманных тикеров также не был обусловлен выдающейся результативностью малого числа акций продуманных тикеров: рынок превзошли 65 % акций продуманных тикеров.

Хотя Гэри и его студенты пытались исключить профессионалов отрасли, которые могли быть знакомы с акциями продуманных тикеров и их результативностью, они, возможно, непреднамеренно включили в список людей, которые знали, как развивались некоторые из этих акций в течение изучаемого периода.

В поисках еще более убедительных свидетельств в пользу своей стратегии в 2019 году Гэри и еще двое студентов вновь проанализировали результатив-

ность этого портфеля продуманных тикеров. Люди, опрошенные в 2006 году и выбравшие восемьдесят две акции продуманных тикеров, возможно, знали что-то о том, как эти акции росли до 2006 года, но они никак не могли знать, как эти акции будут расти после завершения опроса.

Как и в предыдущие двадцать два года, с 1984 по 2005 год, портфель продуманных тикеров значительно превосходил рыночный портфель в течение последующих тринадцати лет, с 2006 по 2018 год. На рис. 2.6 показано, что, начав с 1 долл. в первый торговый день 2006 года, портфель продуманных тикеров вырос до 5,03 долл., т. е. показал компаундированную годовую возвратность в 13,2 %, в то время как рыночный портфель в конце 2018 года вырос до 2,47 долл. и показал компаундированную годовую возвратность в 7,2 %.

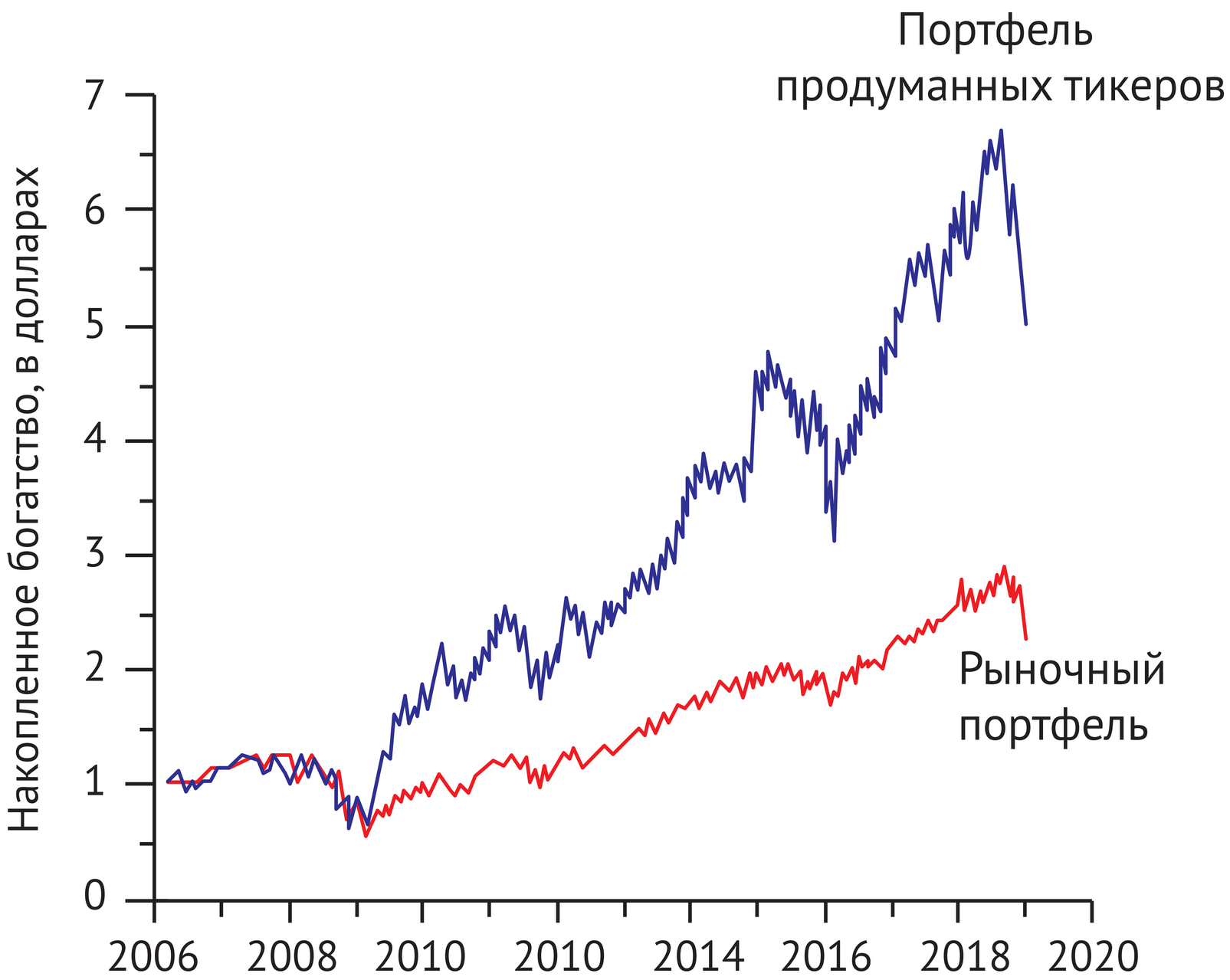

Рис. 2.6. Динамика портфеля продуманных тикеров с 2006 по 2017 год

По странному совпадению один из бывших студентов Гэри, Майкл Соломон, связался с ним после прочтения оригинальной статьи о продуманных тикерах. В 2000 году Майкл работал в частной инвестиционной компании Leonard Green, которая приобрела компанию-оператора сети ветеринарных клиник и диагностических лабораторий VCA Antech. В 2001 году Леонард Грин реорганизовал VCA Antech и провел публичное размещение акций. Майкл предложил использовать тикерный символ WOOF[4], и они согласились.

Инвестор, рассматривающий компании, занимающиеся домашними животными, может столкнуться с VCA Antech и едва ли заметит тикерный символ, если это будет что-то скучное и непроизносимое, например VCAA. Но фактический символ, WOOF, легко запоминается и звучит забавно. Возможно, через

[4] ГАВ на русском. – *Прим. перев.*

несколько дней, недель или месяцев этот инвестор подумает об инвестировании в такую компанию и вспомнит символ WOOF.

Когда в 2001 году акции были размещены на бирже под тикером WOOF, некоторые финансовые эксперты отнеслись к этому с удивлением и скептицизмом. Колонка веб-сайта MarketWatch была озаглавлена «Ветеринарное IPO залаяло на рынке» (Veterinary IPO barking in market). В новостной ленте информационного агентства Dow Jones[5] говорилось, что «Первичное публичное размещение акций компании VCA Antech Inc., символом акций которой является WOOF, было, надо сказать, удачной сделкой». На минуточку, любая огласка – это хорошая реклама, не так ли?

На рис. 2.7 показано, что в течение следующих шестнадцати лет VCA уверенно лидировала на рынке, пока не была приобретена кондитерской компанией Mars. Это приобретение стало причиной резкого роста цен в 2017 году. За более чем шестнадцать лет, в течение которых акции VCA Antech торговались под тикером WOOF, годовая доходность их акций составила 19,4 % по сравнению с 7,2 % для рынка в целом.

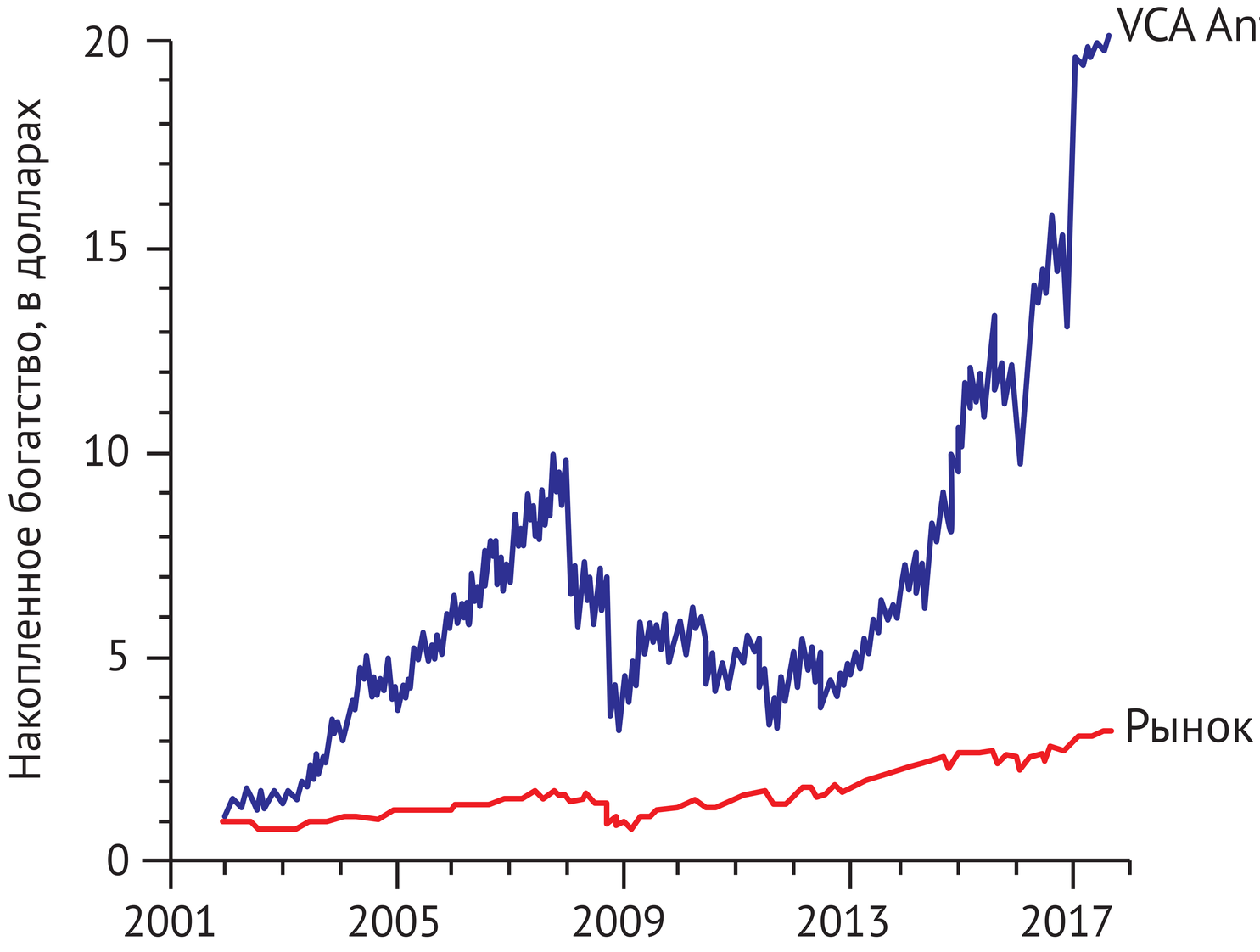

Рис. 2.7. Woof! Woof!

Испытаний РКИ не было, поэтому мы не знаем наверняка, что за превосходную результативность компании VCA Antech и других акций умных тикеров был ответственен тикерный символ, но свидетельства выглядят довольно убедительно. В идее был толк, и, в отличие от «От хорошего к отличному», она была разработана, и акции были отобраны до анализа их результативности – протокол, подкрепленный тем фактом, что спустя дюжину лет после изначального исследования была проведена ревизия результативности, и акции продуманных тикеров по-прежнему превосходили рынок.

---

5 См. **https://www.dowjones.com/business-intelligence/newswires/**. – *Прим. перев.*

Как не быть введенным в заблуждение призрачными закономерностями

Случайно возникающие корреляции бесполезны для составления предсказаний. Для того чтобы что-то предсказывать, это должно быть предсказуемо вследствие глубинной причинно-следственной структуры. Для корреляции должно быть реальное основание: A служит причиной B; B служит причиной A; A служит причиной B, а B служит причиной A; или что-то еще служит причиной A и B. Корреляции без причинной следственности означают предсказания без надежды.

Причинную следственность можно установить с помощью РКИ – испытания, в котором есть как процедурная группа, так и контрольная группа, и испытуемые случайно распределяются по двум группам. В дополнение к этому тест должен быть двойным слепым, чтобы ни испытуемые, ни исследователи не знали, кто входит в процедурную группу. Наконец, должно быть достаточно данных, чтобы иметь возможность делать наполненные смыслом выводы. Настоящим испытаниям РКИ, как правило, можно доверять, так как они основаны на правдоподобных теориях и тщательно тестируются.

Нередко проводить испытания РКИ невозможно. И приходится довольствоваться данными наблюдений. Надлежащее исследование детализирует теорию, которую необходимо протестировать, и только потом обращается к данным. Отыскание закономерности после обращения к данным – дело коварное и, скорее всего, закончится плохо – бесполезной и временной случайно возникшей корреляцией.

Глава 3

Оболванены и обмануты

На протяжении веков жители Новых Гебридов верили, что вши на теле делают человека здоровым. Эта народная мудрость основана на закономерности: у здоровых людей часто бывают вши, а у нездоровых – обычно нет. Однако оказалось, что не отсутствие вшей делает людей нездоровыми, а тот факт, что у нездоровых людей нередко бывает лихорадка, которая отгоняет вшей.

Мы запрограммированы замечать (и даже активно искать) закономерности, которые могут быть использованы для выявления полезных продуктов питания, предупреждения об опасности и принятия правильных решений. К сожалению, обнаруживаемые нами закономерности часто являются обманчивыми иллюзиями. И на протяжении всей истории люди не раз были одурачены призрачными закономерностями.

Малярия

Малярия существует уже тысячи лет. Знаменитый римский врач Квинт Серен Саммоник (Quintus Serenus Sammonicus), живший во втором веке нашей эры, советовал больным малярией носить на шее листок бумаги с треугольной надписью, показанной на рис. 3.1. В верхней строке написано мистическое слово *абракадабра*. В каждой последующей строке удаляется последняя буква предыдущей строки, и буквы сдвигаются вправо, в результате чего каждая буква повторяется, перемещаясь вниз и вправо, и получается полное слово *абракадабра*, идущее от последней строки вверх и вправо. Больным малярией советовали носить этот амулет в течение девяти дней, а затем бросить его через плечо в ручей, текущий на восток.

Неудивительно, что некоторые пациенты находили утешение в четкой закономерности, образованной буквами абракадабры. Вряд ли они с такой же доверчивостью отнеслись бы к случайному набору букв.

После этой бестолковщины некоторые пациенты, несомненно, выздоравливали, и их выздоровление приписывалось (задним числом) предписанию врача. Не смейтесь. Даже сегодня некоторые пациенты принимают неэффективные лекарства, выздоравливают и приписывают свое выздоровление бесполезным лекарствам.

Летом 1969 года брат Гэри рассказал ему, что его друг зарабатывает приличную сумму денег, продавая «энергетические камни» прохожим возле пирса 39 в Сан-Франциско. Это был обычный гравий, который он собирал на ближайшей

парковке и промывал водой из-под крана. Тем летом он бывал там почти каждый день, и никто так и не вернулся, чтобы пожаловаться, что их энергетические камни не работают.

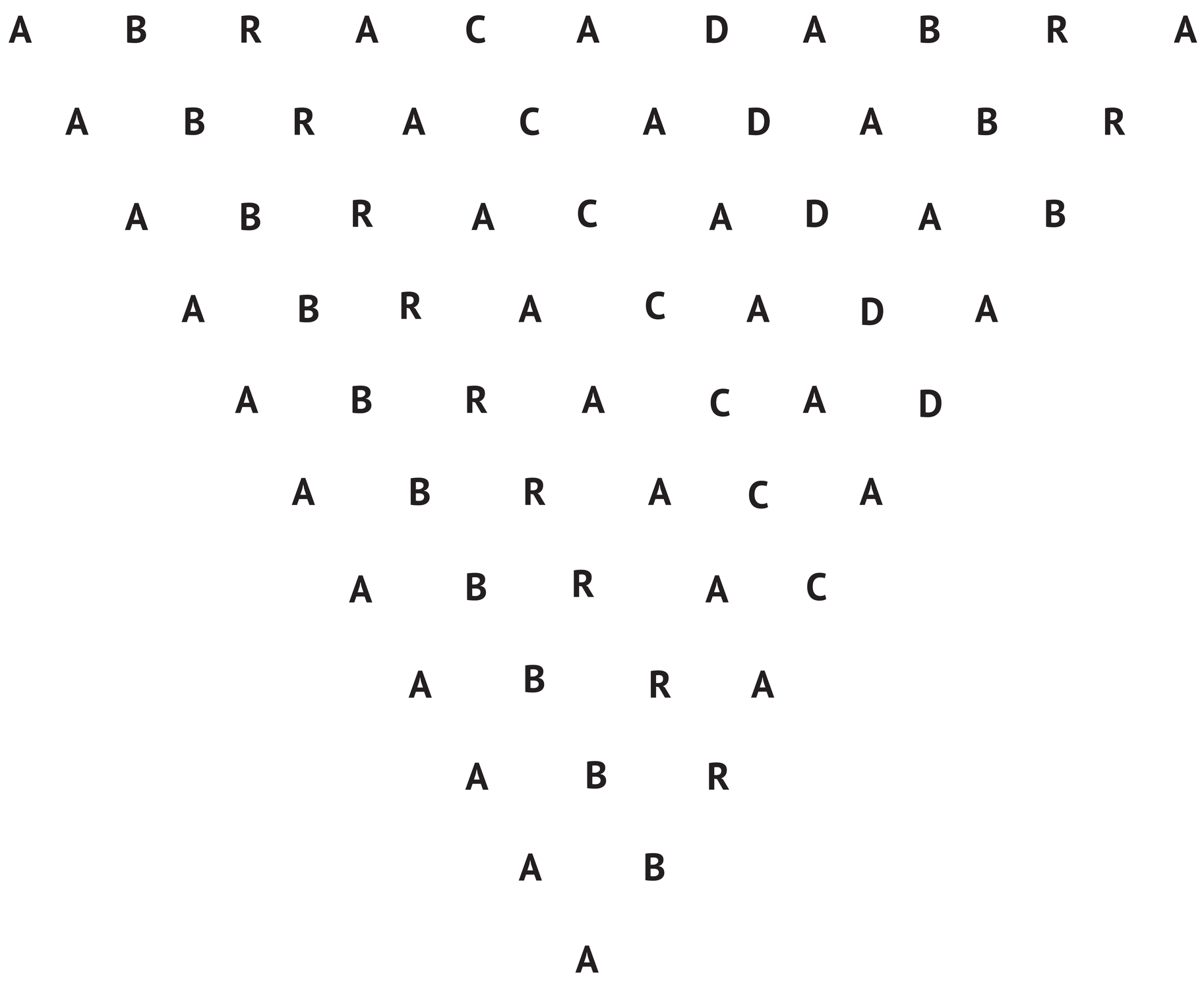

Рис. 3.1. Абракадабра

В 400 году до н. э. Гиппократ заметил, что люди, страдающие от того, что мы теперь называем малярией, часто жили в болотистых местностях. Он описывал малярию как «болотную лихорадку», которая вызывалась *миазмами*, т. е. «загрязненными» гниющими органическими веществами. Миазмы также описывались как «плохой воздух» или «ночной воздух», потому что ночью запахи были особенно неприятными. Считается, что слово «малярия» произошло от итальянского *mal'aria* (плохой воздух), хотя также считалось, что причиной многих других заболеваний, включая холеру, служили миазмы. Теория о миазмах была основана на вводящих в заблуждение закономерностях.

Что касается малярии, то комары любят болотистые местности и выходят ночью, и в конечном итоге было доказано, что указанная болезнь передается не от загрязненного воздуха, а от укусов малярийных комаров *anopheles*.

Британский врач сэр Рональд Росс (Ronald Ross) родился в Индии в 1857 году и был сыном генерала британской индийской армии. Он был вдохновлен сэром Патриком Мэнсоном (Patrick Manson), считавшимся основателем области тропической медицины, на исследование вопроса о том, не передается ли малярийный паразит комарами. В 1897 году, находясь в Секундерабаде, Индия, он извлек личинок нескольких комаров в лаборатории, чтобы убедиться, что

они не были заражены внешним миром. Затем он заплатил больному малярией восемь анн за то, что тот позволил восьми комарам себя укусить. Когда Росс препарировал комаров, он обнаружил малярийных паразитов, которые росли в их желудках, что было убедительным доказательством того, что комары заразились этим паразитом, когда кусали инфицированных людей. (Для проведения рандомизированного контролируемого исследования ему также следовало обследовать контрольную группу комаров, которым не давали кусать больного малярией, чтобы исключить возможность передачи малярийных паразитов от родителей-комаров потомству.)

Он отпраздновал свое открытие, написав стихотворение, в котором были такие строки[6]:

> Со слезами и тяжелым вздохом
> Я нашел твои коварные зерна,
> О всеубивающая Смерть.
> Я знаю, эта мелочь
> Спасет мириады людей.
> О Смерть, где твое жало?
> Твоя победа, о Могила?

В ходе более поздних экспериментов с птицами Росс продемонстрировал, что малярийные комары *anopheles* передают паразита, кусая инфицированных птиц, а затем здоровых, тем самым завершая цикл распространения малярии.

В 1902 году Росс стал первым британским нобелевским лауреатом, получив Нобелевскую премию по физиологии и медицине за свою работу по малярии.

Суеверия

Центральной частью эпохи Просвещения XVIII века было прославление научного метода – настойчивости в том, что убеждения не следует принимать некритически, а следует проверять эмпирически. Необоснованные суеверия плохо поддаются научной проверке. Тем не менее от многих суеверий трудно избавиться, даже если в них нет никакого толка, и у нас нет ничего, кроме догадок об их происхождении. Возможно, если черная кошка перебежит вам дорогу, это будет считаться дурной приметой, потому что черные кошки ассоциируются с ведьмами. Возможно, христиане стучали по дереву на удачу, потому что это напоминало им о кресте, на котором был распят Иисус. Плохая примета при открытии зонта в помещении, вероятно, пришла от древних египтян, которые носили зонты от солнца и думали, что бог солнца Ра обидится, если зонт откроют в помещении, где в нем нет необходимости. Кто знает причину, по которой считается удачей загадать желание на звезду (в особенности на падающую звезду) или найти четырехлистный клевер?

---

6 В оригинале:
With tears and toiling breath,
I find thy cunning seeds,
O million-murdering Death.
I know this little thing
A myriad men will save.
O Death, where is thy sting?
Thy victory, O Grave?

Несомненно, многие люди рассказывают о суевериях для развлечения, но на самом деле в них не верят. Другие, возможно, в них верят, но их убеждения относительно безобидны. В любом случае большинство суеверий трудно проверить эмпирически. Как сформировать процедурную и контрольную группы? Если постучать по дереву и ждать удачи, то как долго нужно будет ждать? Что считается удачей?

Несомненно, избирательное припоминание и склонность искать подтверждения своей точки зрения только усиливают привлекательность суеверий. Мы склонны запоминать попадания, которые укладываются в наши убеждения, и забывать промахи, которые им противоречат. Если в игре выигрывает новичок, то мы поем «новичку везет», потому что это подтверждает нашу веру в удачу новичка. Если новичок проигрывает, то мы приглашаем его сыграть еще раз (и забываем об этом свидетельстве в пользу удачи новичка). Если кто-то проходит под лестницей и через несколько часов, дней или недель случается что-то плохое, то мы можем вспомнить о лестнице. Если ничего плохого не происходит, то мы забываем о лестнице.

Формирование закономерностей в воображении

Наша зависимость от закономерностей настолько сильна, что иногда мы их себе воображаем. У Гэри есть родственник (Джим), который без ума от «Бостон Селтикс», но Джим перестал смотреть их матчи по телевизору после того, как заметил, что в нескольких матчах «Селтикс» играли плохо, пока он смотрел. Он сглазил «Селтикс»! По логике вещей игроки «Селтикс» не знают и не заботятся о том, смотрит ли Джим матчи или нет, поэтому его привычки к просмотру никоим образом не могут повлиять на их игру. И все же он верит.

Мы не знаем (и не хотим знать), какие психологические факторы ответственны за мысли людей о том, что они – ходячие проклятия и все, к чему они прикасаются, превращается в мусор. Однако мы почти уверены, что избирательное припоминание – это большая часть истории. Некоторые люди чаще помнят хорошее, другие – чаще плохое. Конечно же, были моменты, когда Джим наблюдал за хорошей игрой «Селтикс», но вот Джиму запоминались именно те моменты, когда «Селтикс» играли плохо, пока он наблюдал за их игрой.

Гэри однажды предложил провести контролируемый эксперимент. Выбрать игру, а затем рандомно отобрать промежутки времени, которые можно смотреть или не смотреть, и взглянуть, будет ли разница. Джим отказался. Возможно, этот эксперимент был слишком занудным, или, возможно, Джим не хотел, чтобы сила его проклятия подвергалась сомнению.

Памятные закономерности на букву «М»

Накануне президентских выборов в США в 1992 году в понедельник вечером «Миннесота Викингз» играла в футбол с «Чикаго Беарз». В начале телевизионной трансляции дикторы разоблачили индикатор «Викингов»: когда «Миннесота Викингз» выигрывает матч в понедельник перед президентскими выбо-

рами, на выборах всегда побеждает кандидат от республиканцев. Оказывается, этот индикатор был основан ровно на двух наблюдениях: проигрыше «Викингов» в 1976 году до избрания демократа Джимми Картера и победе «Викингов» в 1988 году до избрания республиканца Джорджа Буша. Третий раз был не из приятных: «Миннесота» выиграла в 1992 году, а Джордж Буш проиграл демократу Биллу Клинтону.

Один из студентов Гэри, Гуннар Кнапп, обнаружил несколько других закономерностей. Сенатор от Монтаны Майк Мэнсфилд был родом из Миссулы, штат Монтана, родился в марте, служил в морской пехоте, женился на женщине по имени Морин, учился в горной школе Монтаны и дольше всех был мажоритарным лидером (лидером большинства) в Сенате. Когда он ушел из Сената в отставку, его сменил Джон Мелчер, который до службы в армии учился в университете Миннесоты. Гуннар тоже обнаружил, что в округе Мидлсекс, Массачусетс, был конгрессмен по фамилии Макдональд из города Малден. Можно было бы продолжить, но мы не будем.

Не дроби карты

Джей и его друг Рори однажды были в Лас-Вегасе и сели за стол для игры в блэкджек, чтобы поболтать с другом Рори, который играл за столом в одиночестве. Карты внезапно пошли против друга Рори, и вскоре он встал и, не сказав ни слова, отошел к другому столу. Не было никаких сомнений в том, что он пытался избежать проклятия, которое Джей и Рори на него наложили. Когда Рори позже подошел к нему и сказал, что у того из кармана выпал бумажник, друг велел Рори отойти от него подальше. После того как Рори отдал ему бумажник и ушел, друг дико замахал руками в воздухе, пытаясь отогнать невезение, как будто это был запах, который Рори оставил после себя.

Любой, кто когда-либо играл в казино, знает, что суеверия живы и здоровы. Игроки в блэк-джек зачастую по-настоящему раздражаются, если игрок принимает неверное решение, которое меняет будущие карты, сдаваемые другим игрокам (например, неразумно берет карту из колоды, которая в противном случае досталась бы жалующемуся игроку). Такие неудачные исходы запоминаются, а положительный резонанс неправильного выбора забывается.

Игроки с большим опытом игры в блэк-джек иногда бывают очень взволнованы, когда другой игрок играет так, как он бы никогда не сыграл. Джей узнал об этом из первых уст на своей первой работе после окончания колледжа. Он играл в блэк-джек очень мало, но прочитал классическую книгу Эдварда Торпа «Обыграть дилера». Джей обсуждал блэк-джек с коллегой по имени Робби, и между ними разгорелся спор по поводу относительно неясного вопроса – что лучше: остаться при своих или сделать сплит двух девяток против дилера, у которого на руках шестерка.

В блэк-джеке цель состоит в том, чтобы подобраться ближе к двадцати одному очку (не превысив это число), чем у дилера, который должен взять еще одну карту, если у него в общей сложности шестнадцать или меньше. Получить две девятки (всего восемнадцать), когда у дилера на руках шесть, – это здорово для

вас, потому что фигурные карты (которые считаются за десять) являются наиболее вероятными картами в колоде. Это означает, что дилер, скорее всего, застрянет с самой неблагоприятной суммой очков из возможных (шестнадцать) и, скорее всего, получит больше двадцати одного. Например, если вы остаетесь при своих восемнадцати очках, а у дилера шестнадцать, то вы проиграете только в том случае, если его следующая карта будет тройкой, четверкой или пятеркой (двойка – это ничья). В интуитивном плане игра заключается в том, чтобы «остаться при своих» восемнадцати и дождаться, когда неудачливый дилер проиграет с суммой очков выше двадцати одного.

Однако есть одна сложность: если у вас есть пара, то разрешается раздробить свои карты на две отдельные руки[7] (и выиграть или проиграть в два раза больше). Здесь вы можете раздробить свою пару девяток (сделать сплит) на две руки, в каждой из которых будет по девятке, чтобы в каждую руку можно было раздать дополнительные карты. Если вам повезет и в каждой руке будет по одной фигурной карте, то у вас будет две сильные руки с девятнадцатью картами вместо одной хорошей руки с восемнадцатью. Ведь синица в руке лучше, чем журавль в небе?

Основываясь на собственном игровом опыте, Робби твердо верил, что упускать синицу из рук нельзя. Джей с ним не согласился, основываясь на том, что он прочитал, и он не мог поверить, что Робби больше доверял своим несовершенным воспоминаниям всяких россказней, чем тщательно продуманному анализу.

Робби был в полном недоумении, потому что, как бы красноречиво и страстно он ни доказывал свою правоту, Джей доверял ему не больше чем уважаемому математику, проанализировавшему миллионы компьютерных симуляций.

Потом Джек, будучи актуарием высшего звена компании, случайно оказался рядом с рабочей кабинкой. Он принадлежал к высшему руководству и не был известен своим общением с работниками, стоящими по иерархии ниже, что делало этот разговор еще более неожиданным:

Робби: Эй, Джек, ты же умный парень, иди сюда!

Джек: (подходит, выглядя смущенным и колеблясь)

Робби: Давай так, в блэк-джеке дилер показывает шестерку, а у тебя две девятки. Что ты делаешь?

Джек: Я не знаю, я не играю в карты.

Робби: Что ж, позволь мне тебе сказать: ТЫ НЕ ДРОБИШЬ КАРТЫ!!

Джей до сих пор смеется над нелепым видом Робби, который оказался бодибилдером и кричал на гораздо более мелкого и высокопоставленного управленца, понятия не имевшего, откуда берется его гнев.

Будьте осторожны с людьми, которые полагаются на россказни и избирательное припоминание, в особенности если они, возможно, принимают стероиды.

---

[7] Рука (hand), или комбинация, раздача, – это набор карт игрока или дилера на руках. – *Прим. перев.*

Если верить, то это обязательно произойдет

Иногда суеверие становится самоисполняющимся пророчеством. Существует множество вариантов этой классической истории: инструктор по танцам убеждает начинающую танцовщицу, что ее волшебные туфельки сделают ее великой танцовщицей. Затем она приходит в отчаяние, когда забывает взять свои туфельки на важный концерт. Инструктор проливает свет, говоря, что в ее обуви нет ничего волшебного. Осознав, что она хорошая танцовщица, девушка выступает блестяще.

В реальной жизни Пеле, который считается лучшим футболистом всех времен, однажды провел несколько нехарактерно неудачных матчей и решил, что это из-за того, что игровая майка, которую он подарил болельщику, должно быть, была удачной. Пеле посоветовал другу сделать все возможное, чтобы найти свою удачную футболку, и, конечно же, когда футболка была возвращена Пеле, к нему вернулся и успех. Друг не сказал Пеле, что ему не удалось найти удачную футболку; поэтому он просто подарил Пеле ту же футболку, в которой тот играл во время серии неудачных матчей.

Когда Гэри проводит выпускной экзамен на своих занятиях, студенты иногда приходят на экзамен в деловой одежде, другие студенты приходят в пижамах. Возможно, они думают, что их необычная одежда принесет им удачу. Возможно, они чувствуют себя увереннее, когда одеваются для успеха или носят удобную одежду.

Уверенность, несомненно, важна, и, возможно, иногда бывает полезной даже неуместная уверенность.

Понедельник, 6-е число

Практически в каждой культуре есть сулящие удачу и неудачу суеверия, которые существуют уже на протяжении стольких поколений, что никто точно не знает, как они возникли. Многие из этих суеверий связаны со удачными и неудачными числами, несмотря на то что, казалось бы, произвольные различия между удачными и неудачными числами в разных культурах убедительно свидетельствуют о том, что эти суеверия не имеют рациональной основы. Если бы число было действительно неудачным, то оно было бы неудачным во всем мире.

Число 13 считается неудачным во многих западных культурах. Некоторые говорят, что 13 является неудачным, потому что предавший Иисуса Христа Иуда Искариот был тринадцатым человеком, пришедшим на Тайную вечерю. Другие говорят, что в календарном году бывает тринадцать полнолуний, а полнолуния вызывают в людях странное поведение. Третьи говорят, что тринадцать – плохое число по сравнению с двенадцатью, которое является особым числом, потому что в году двенадцать месяцев, двенадцать знаков зодиака, двенадцать апостолов, и число двенадцать является последним перед подростковым возрастом (тринадцать, четырнадцать и т. д.). Какой бы ни была причина, страх перед числом 13 бывает настолько непреодолимым, что у него есть название: *трискайдекафобия*.

Многие высотные здания в западных странах имеют двенадцатый и четырнадцатый этажи, но не тринадцатый, как показано на рис. 3.2. Иногда Джея так и подмывает подняться на четырнадцатый этаж и сказать: «Знаете, если этот этаж называется четырнадцатым, это еще не значит, что так оно и есть. Под вами двенадцать этажей, а это значит, что на самом деле вы на тринадцатом». Пока что он был слишком вежлив, чтобы это сделать.

Рис. 3.2. Заметили, что чего-то не хватает?
Автор: Sgerbic – самофотография, общественное достояние

Пятница, 13-е число, которая случается от одного до трех раз в году, конечно, считается особенно неудачной, возможно, потому что Иисус был распят в пятницу (хотя и не в пятницу тринадцатого). Однако в Италии 13 считается удачным числом, а вот 17 – неудачным, и именно пятница, 17-е число, считается неудачным днем. В Греции и многих испаноязычных странах неудачным считается не пятница, а вторник, 13-е число.

Говорят, что греки число 3 тоже считают неудачным, потому что «неудача приходит по три раза», тогда как шведы считают 3 удачным числом, потому что «все хорошее должно приходить по три раза». Как мы уже говорили, если бы числа 3 действительно были везучими или невезучими, то это было бы единообразно так – не повезло в Швеции и не повезет в Греции.

В японском, китайском и кантонском диалектах произношение слов *четыре* и *смерть* очень похоже («си»), что делает число 4 неудачным, в то время как произношение слова *восемь* похоже на слова *богатство* или *процветание*, что

делает число 8 удачным. (Кстати, на самом деле из-за заимствованного из китайского неудачного произношения в японском языке есть альтернативный способ произношения слова *четыре* («ён»).)

В Азии нередко отсутствует четвертый этаж. В больницах Китая и Японии не хватает даже палат с номером 4. Многие люди прилагают немало усилий, чтобы получать домашние адреса, номера телефонов и автомобильные номерные знаки. Некоторые родственники Гэри отказались даже рассматривать возможность покупки привлекательного в других отношениях дома, потому что номер дома был неудачным.

Джей работал аналитиком данных на 44-м этаже здания в центре Лос-Анджелеса. Оказывается, 44 по-китайски произносится как «сиши си» (sìshísì), что звучит очень похоже на «си ши си», что буквально означает «умри, да, умри» (sǐ shì sǐ). Не помогает и распространенность грамматической конструкции «да» между повторяющимися словами, которая приравнивается к утверждению, что это действительно так, поэтому произнесение «сорок четыре» звучит как «должен умереть». Излишне говорить, что в китайском языке число 44 не считается удачным.

Одна из американо-китайских коллег Джея не сказала своей матери, что она работает на сорок четвертом этаже. Потенциальный китайский сотрудник отклонил предложение о работе из-за адреса!

Еще один коллега рассказал забавную историю, которую он узнал во время посещения Диснейленда в Китае. Руководители Диснейленда поначалу были удивлены, что никто не покупал их именные зеленые шляпы Робин Гуда и Питера Пэна.

Затем они узнали, что в китайской культуре зеленая шляпа означает, что ваш супруг вам изменяет. Когда Джей выразил удивление по поводу того, что любители розыгрышей не покупают шляпы с приколом в подарок, его китайский управляющий проницательно заметил, что «возможно, чувство юмора китайцев отличается от вашего».

Число 666 многие считают сатанинским числом, и люди, испытывающие к нему крайнее отвращение, помечаются как страдающие от гексакосиоихексеконтагексафобии[8]. Джей когда-то работал в компании, которой пришлось столкнуться с этим жутким числом. Компания проводила эксперименты на веб-страницах и назначала каждому эксперименту идентификатор теста. Когда некоторые сотрудники заметили, что количество экспериментов перевалило за 600 и приближается к 666, разгорелся спор о том, стоит ли пропустить это число и перейти сразу с 665 на 667. Рори, управляющий по тестированию, отмахнулся от опасений, несмотря на растущее давление со стороны коллег, которые не считали это поводом для смеха. Почему бы ему не отказаться от 666-го, поскольку это было бы недорого и помогло многим людям почувствовать себя лучше? Возможно, он хотел подчеркнуть свою точку зрения и помочь людям справиться с их гексакосиоихексеконтагексафобией. Эксперимент 666 начался и закончился, и (насколько нам известно) никаких демонов вызвано не было.

И все же на углу улицы рядом со зданием автобус сбил пешехода насмерть. Однако это произошло не на следующий день после эксперимента 666; это

[8] Англ. *hexakosioihexekontahexaphobia.*

было утром после того, как поздно вечером было разослано электронное письмо, в котором всем предлагалось разработать свой «запасной план на случай столкновения с автобусом». Делайте с этим что хотите.

Существует скудно мало свидетельств того, что некоторые числа являются особенно удачными или неудачными, хотя это может быть самоисполняющимся пророчеством, если наш страх перед чем-то становится причиной того, чего мы боимся. Например, пациенты, перенесшие операцию в «неудачный» день, могут чувствовать себя плохо из-за эмоционального потрясения, тогда как человек, который в «удачный» день работает усерднее, может добиться большего. Также, несомненно, существует сочетание избирательного припоминания и склонности искать подтверждения своей точки зрения. Мы чаще замечаем и припоминаем ситуации, когда что-то плохое случается в «неудачный» день, чем когда это случается в другие дни.

Некоторые предприимчивые люди продают доверчивым простакам путеводители по мечтам, обещая превращать мечты человека в выигрышные лотерейные номера. В одном путеводителе говорится, что он поможет «научиться раскрывать силу своей мечты, превращая образы в удачные числа и пробуя их в лотерее или игре пауэрбол. Вы будете поражены, насколько ваши мечты могут сделать вас богатым». Например, читателям рекомендуется ставить на число 34, если они мечтают о стейке, и на число 10, если они мечтают о яйцах. Кто знает, что они должны сделать, если им нужно угадать шесть чисел, а они мечтают только о стейке и яйцах.

Этот дерзкий самиздатовский путеводитель на сорока четырех страницах в мягкой обложке продается за 29,99 долл.

Поскольку выигрышные номера выбираются случайным образом и не зависят от чьих-либо желаний, мы не удивлены, что автор решил продавать путеводитель по мечтам вместо того, чтобы покупать лотерейные билеты. Нас также огорчают люди со скромным достатком, которые покупают путеводители по мечтам и лотерейные билеты.

В 1987 году, в год, когда было три пятницы тринадцатого числа, главный экономист банка в Филадельфии сообщил, что за последние сорок лет было еще шесть лет, когда было три пятницы тринадцатого числа, и в течение трех этих лет начинался экономический спад. Мы не думаем, что он шутил. Мы действительно считаем, что у него было изрядно много свободного времени, и он был одурачен призрачной закономерностью. Каким-то образом 1987 год обошелся без рецессии.

Иногда люди просто замечают закономерности. В других случаях они активно их ищут. В статье, опубликованной в престижном журнале British Medical Journal, сравнивалось количество поступлений в медицинские учреждения в юго-западном районе Темзы в Англии в шестую пятницу тринадцатого числа, которые произошли за четырехлетний период, с количеством поступлений в предыдущую пятницу шестого числа. Сначала они сравнили количество поступлений в отделения неотложной помощи в связи с несчастными случаями и отравлениями шестого и тринадцатого числа и не нашли ничего статистически убедительного. Поэтому они просмотрели все случаи поступлений в больницы в связи с несчастными случаями и отравлениями и снова ничего не нашли. Затем они разделили случаи поступлений на пять подкатегорий, пока-

занных в табл. 3.1: случайные падения, транспортировка, отравления, травмы из-за животных и растений и неустановленные случаи, будь то случайные или преднамеренные.

Таблица 3.1. Количество поступлений жителей юго-западной части Темзы в разбивке по видам несчастных случаев

Причина	Пятница, 6-е	Пятница, 13-е
Падение	370	343
Транспортировка	45	65
Отравление	37	33
Животные	1	3
Причины не установлены	1	4
Итого	**454**	**440**

В целом на 6-е число было зарегистрировано больше поступлений, но была одна категория – транспортировка, – в которой число поступлений было выше на 13-е число. Поэтому они завершили свое исследование мрачным предупреждением: «Пятница 13-е для некоторых является неудачной. Риск госпитализации в результате транспортного происшествия может возрасти на целых 52 %. Рекомендуется оставаться дома».

Это наглядный пример поведения «ищите закономерность, и вы ее найдете». Несмотря на то что 6-го числа случаев поступлений в медучреждения было больше, чем 13-го, исследователи продолжили искать какую-то категорию, любую, пока не нашли то, что хотели найти.

Число 47 повсюду среди нас

Гэри преподает в колледже Помона, а Джей окончил данный колледж, поэтому авторы, естественно, испытывают сильную привязанность к волшебному числу колледжа 47. В 1964 году легендарный профессор статистики по имени Дональд Бентли (Donald Bentley) продемонстрировал своим студентам причудливое геометрическое доказательство утверждения о том, что все числа равны 47, очевидно, в поддержку студенческого проекта, целью которого было составление списка случаев, когда наблюдалось число 47.

Как и в случае со всеми удачными/неудачными числами, подавляющая часть истории с числом 47 представляет собой пример избирательного припоминания. Каждый день нас бомбардируют числами, и мы замечаем те, которые считаем удачными или неудачными, которые соответствуют нашему дню рождения или отражают какое-то другое совпадение. Выпускники Помоны находятся в состоянии повышенной готовности по поводу числа 47 и пропускают числа 46 и 48 незамеченными. К колледжу Помоны можно подъехать с запада, свернув на съезд 47 с автострады Сан-Бернардино. Подъезжая с востока, можно свернуть на съезд 48, но кого волнует 48? Как вариант, можно проехать по автостраде Футхилл и сойти на съезде 50 или 52 в зависимости от того, откуда ехать: с запада или востока, но опять же какая разница?

В верхнем ряду органа зала Лайман-холл в Помоне 47 труб; не спрашивайте о других рядах. Выпускник университета Помоны Ричард Чемберлен был 47-м человеком в очереди на спасение в фильме «Ад в поднебесье» (The Towering Inferno); не обращайте внимания на другие его фильмы. Здание студенческого общежития Мадд-Блейсделл в Помоне было построено в 1947 году, и на посвященной ему мемориальной доске написано 47 букв; не обращайте внимания на 83 других здания в Помоне.

Глядя за пределы Помоны:

- роман Толстого «Крейцерова соната» назван в честь 47-го опуса Бетховена;
- Новый Завет приписывает Иисусу 47 чудес;
- теорема Пифагора является 47-й констатацией «Элементов» Евклида;
- Цезарь провозгласил свой девиз «пришел, увидел, победил» в 47 году до н. э.;
- тропики Рака и Козерога расположены на расстоянии 47° друг от друга.

Довольно впечатляюще, если только вы не задумаетесь об этом (а мы надеемся, что вы задумаетесь). Сколько миллионов, миллиардов или триллионов раз числа от 1 до 100 появлялись на протяжении истории и в нашей повседневной жизни? Наверняка существует очень большое количество 47 (а также 46, 48 и других чисел). Поищите любое число, и вы его найдете.

Существует целое общество «47», в котором люди сообщают о своих случаях наблюдения числа 47. Например:

> Хоккейный номер моего друга Тима – 47. Позже он нам сказал, что начал замечать довольно частую встречаемость числа 47. Сначала это была просто шутка, но потом я начал это замечать. Такие вещи, как набрать 47 очков в игре в дартс, найти телефонные номера, в которых есть 47…

Поищите число 47, и вы его найдете. (После редактирования этого раздела Джей заметил, что батарея его телефона разрядилась на 47 %! В следующий раз, когда он редактировал этот раздел, он проверил еще раз, и результат составил 43 %. Так близко!)

В дополнение к этому, как и в случае со всеми удачными/неудачными числами, часть истории с числом 47 является самоисполняющимся пророчеством. Студенты из Помоны использовали его довольно либерально. Выпускник университета Помоны Джо Меноски был сценаристом и сопродюсером многих серий «Звездного пути» и щедро разбрасывал в них 47 (и обратное ему число 74): «Энтерпрайз» был построен в секторе 47, экипаж останавливается на подпространственной ретрансляционной станции 47, всего было уничтожено 47 клингонских кораблей, на планете осталось 47 выживших, а один человек уменьшился до 47 см.

Режиссер и продюсер «Звездного пути» Джей Джей Абрамс подхватил эстафету и продолжил традицию в других своих постановках. Фильм «Миссия невыполнима: Протокол Фантом» заканчивается на 47-м пирсе. Тепловой генератор в «Звездных войнах: Пробуждение силы» находится в 47-м участке.

В следующий раз, когда вы заметите число 47, подумайте о том, является ли это избирательным припоминанием либо еще одним примером распространения данного числа студентом из Помоны. Только не чувствуйте себя обязанным сообщать об этом в общество «47».

Нумерология

Западная нумерология берет свое начало от греческого философа, математика и мистика Пифагора, которому приписывают много математических открытий, включая его теорему, которую мы изучали в школе, но, похоже, он не был инициатором того, что сейчас называется нумерологией.

Современная нумерология переводит полное имя человека в математическое число, используя трансляционный код, приведенный в табл. 3.2, который назначает числа буквам латинского алфавита. Так, буквам «a», «j» и «s» назначается число 1.

Таблица 3.2. Нумерологический код

1	2	3	4	5	6	7	8	9
a	b	c	d	e	f	g	h	i
j	k	l	m	n	o	p	q	r
s	t	u	v	w	x	y	z	

Для того чтобы определить число имени человека, используется нумерологический код, который определяет число для каждой буквы полного имени человека, а затем сумму цифр. Используя Гэри в качестве примера («Gary Nance Smith»), получаем:

Gary: 7 + 1 + 9 + 7 = 24,

Nance: 5 +1 + 5 + 3 + 5 = 19,

Smith: 1 + 4 + 9 + 2 + 8 = 24,

Gary Nance Smith: 24 + 19 + 24 = 67.

Затем отдельные цифры числа полного имени складываются до тех пор, пока не будет получено однозначное корневое число:

6 + 7 = 1,

1 + 3 = 4.

Число имени Гэри равно 4. Это его *число судьбы*, которое, как говорят, раскрывает таланты и способности, с которыми он родился. Как ни странно, если бы Гэри сменил свое имя на законных основаниях, то его коренное число тоже изменилось бы, и у него, вероятно, были бы другие таланты и способности. Джея при рождении звали Джеймс, поэтому у него, вероятно, раздвоение личности.

У Гэри также есть число при рождении, основанное на дате его рождения, – 11 ноября 1945 года:

$$(1 + 1) + (1 + 1) + (1 + 9 + 4 + 5) = 23,$$

$$2 + 3 = 5.$$

Число даты рождения Гэри равно 5, которое, как говорят, соответствует его жизненному пути. Как ни странно, у всех, кто родился в один и тот же день, одинаковый жизненный путь.

Затем числа имени и даты рождения преобразовываются в человеческие черты: (1) лидер, (2) посредник, (3) коммуникатор, (4) учитель, (5) искатель свободы, (6) воспитатель, (7) искатель, (8) амбициозный и (9) гуманист.

Гэри, очевидно, был рожден, чтобы стать учителем, и его жизненный путь – путь искателя свободы, что его вполне устраивает, хотя он был бы рад любой другой возможности.

Различные нумерологи используют несколько разные слова для обозначения этих черт. Неудивительно, что назначенные слова являются загадочными и двусмысленными, поэтому большинство людей находят их правдоподобными и успокаивающими независимо от их имени или даты рождения.

На своих занятиях по статистике Гэри время от времени проводит интересный эксперимент. В первый день занятий он просит студентов заполнить анкету, в которой указаны их пол, дата рождения и несколько вопросов, которые он будет использовать на последующих занятиях, например «Сколько часов вы спали за последние сутки?». Иногда он приходит на второй урок с набором астрологических показаний, основанных на дате рождения каждого ученика. Он просит каждого ученика внимательно изучить показания и поставить им оценку (от A до F) в зависимости от их точности.

В подавляющем большинстве оценки выходят A и B, хотя Гэри иногда раздает случайно определенные результаты, а в других случаях выдает всем одинаковые. Это еще один пример поведения «ищите и найдете».

Нумерологию и астрологию можно рассматривать как дешевое развлечение, однако если люди принимают неверные решения исходя из своих нумерологических или астрологических показаний, то это может оборачиваться реальными издержками.

Друг рассказал Джею о супружеской паре, которая обычно не принимала никаких важных решений до тех пор, пока не справлялась с большой книгой по астрологии, заполненной сложными графиками, показывающими почасовые энергетические уровни. Однажды, когда они отправились в автосалон, чтобы купить машину, им дали понять, что продажа должна будет завершена до 17:00, но именно в этот промежуток времени энергоуровни изменялись в худшую сторону. (Неужели они превратятся в тыквы?) Продавец заверил, что все будет готово до 17:00.

Когда часы пробили 17 часов вечера, они были в нескольких минутах от подписания окончательных документов. К удивлению продавца, пара извинилась и сказала ему, что им придется вернуться на следующий день, чтобы завершить сделку. Они уехали из автосалона без машины и вернулись на следующий день, когда энергетические уровни снова стали положительными. Если они оценили свое время, то это суеверие обошлось им в немалую сумму денег.

Наше врожденное стремление к порядку в нашей жизни предрасполагает нас благосклонно относиться к анализам и советам, основанным на астроло-

гических показаниях, числах судьбы и других закономерностях или псевдозакономерностях, которые помогают нам находить утешение перед лицом столь большой неопределенности в отношении мира и самих себя.

В научном мире нумерология пользуется столь низким уважением, что некоторые циничные ученые отвергают открытые их коллегами надуманные закономерности как «нумерологию».

Космические совпадения

Многие люди потратили бесчисленное количество часов, обнаруживая необычные совпадения в практически бесконечном потоке чисел, которые служат мерой различных аспектов Вселенной. Например:

- радиус Луны = 1080 миль = 3(360) = 3(1/2)(1)(2)(3)(4)(5)(6);
- радиус Земли = 3960 миль = 11(360) = 11(1/2)(1)(2)(3)(4)(5)(6);
- радиус Луны + радиус Земли = 5040 миль = (1)(2)(3)(4)(5)(6)(7);
- диаметр Земли + диаметр Луны = 2(3960) + 2(1080) = 10 080, что является количеством минут в неделе.

Все это поразительно; однако экваториальный радиус Луны составляет 1080 миль, а полярный – 1079 миль, в то время как экваториальный радиус Земли составляет 3963 мили, а полярный – 3950 миль. Искатели закономерностей используют приближенные числа 1080 и 3960, поскольку они кратны 360, что удобно для расчета (1/2)(1)(2)(3)(4)(5)(6).

Еще одна особенность заключается в том, что сумма цифр диаметров Солнца, Земли и Луны равна 9:

- диаметр Солнца составляет 864 000 миль = 8 + 6 + 4 + 0 + 0 + 0 = 18 и 1 + 8 = 9;
- диаметр Земли составляет 7920 миль = 7 + 9 + 2 + 0 = 18 и 1 + 8 = 9;
- диаметр Луны составляет 2160 миль = 2 + 1 + 6 + 0 = 9.

Число 9 является особенным, потому что оно встречается во многих духовных и мистических контекстах, например в девяти человеческих чертах в нумерологии, девяти чертах личности в эннеаграмме, девяти библейских дарах Бога, числе Брахмы (Творца в индуизме) и, конечно же, в нашей книге «Девять подводных камней науки о данных». Число 9 также обладает замечательным математическим свойством: сумма цифр любого числа, умноженная на девять, равна девяти.

Однако, для того чтобы сумма цифр диаметров Солнца, Земли и Луны была равна 9, приходится выкручиваться и изворачиваться, так как фактические средние диаметры составляют соответственно 864 938, 7918 и 2159 миль.

Искатели закономерностей могут не только рыться в практически неограниченном количестве результатов измерений, но и использовать мили, когда подходят они, и километры, когда они работают лучше. Для совпадений, связанных со временем, они могут использовать самые разные единицы измерения, включая годы, месяцы, недели, дни, часы, минуты и секунды.

Воодушевленные всевозможными закономерностями, созданными благодаря такой гибкости, некоторые искатели закономерностей интерпретируют математические курьезы, на изучение которых они потратили долгие часы, как свидетельство в пользу того, что Вселенную создал Бог, поскольку Бог наверняка использовал бы тщательно организованный генеральный план, а не размещал объекты произвольного размера в произвольных местах. Так, один коллекционер закономерностей заявил, что:

> наша работа состоит в том, чтобы попытаться выяснить, какую систему использовал Создатель, и, к удивлению, оказалось, что он использовал простую решетку размером 9×11 и соотношения 7 и 11, а также 14.

Одного искателя закономерностей особенно впечатлил тот факт, что Солнце и Луна имеют очень разный размер, но из-за их соответствующего расстояния от Земли они кажутся одинакового размера:

> Создатель решил расположить планеты на правильном расстоянии друг от друга, чтобы в определенных случаях мы могли наблюдать удивительную гармонию работы мастера.

Это поразительное совпадение, но оно не вечно. В краткосрочной перспективе из-за гравитационных сил расстояние от Луны до Земли постоянно меняется на протяжении ее обращения по орбите. В долгосрочной перспективе Луна когда-то была намного ближе к Земле, а теперь постепенно от нее удаляется (примерно на 3,8 см в год). Как однажды сказал Томас Хаксли, мы имеем «уничтожение красивой гипотезы уродливым фактом».

Это все о нас

Легко понять, почему люди когда-то верили, что Солнце вращается вокруг Земли. Каждый день они своими глазами видят, что Солнце восходит на востоке, движется по небу и заходит на западе. Это самая надежная закономерность, на которую только можно надеяться.

Вера в то, что мы являемся центром Вселенной, а Солнце, Луна и звезды вращаются вокруг нас, также соответствует нашему чувству собственной важности. Эта вера была настолько сильна, что закреплена в Библии: «И утвердил Господь землю на основании ее; не поколеблется вовек».

Аристарх впервые предложил гелиоцентрическую систему за 1700 лет до того, как Николай Коперник написал свою книгу «О вращении небесных сфер». Дело не в том, что Солнце вращается вокруг Земли, а в том, что Земля вращается вокруг своей оси в восточном направлении, создавая иллюзию того, что Солнце движется вокруг Земли с востока на запад. Гелиоцентрическая модель Аристарха должна была завоевать сторонников своей простотой и тем фактом, что она объясняла странные зигзагообразные движения планет, если смотреть с движущейся Земли. Конкурирующая геоцентрическая модель Птолемея предусматривала «деференты», «эпициклы», «эксцентрики» и «экванты», которые

нуждались в том, чтобы планеты двигались по сложному набору окружностей, вложенных в окружности. Король Кастилии и Леона Альфонсо X однажды пожаловался: «Если бы Господь Всемогущий посоветовался со мной перед тем, как приступать к Творению, я бы посоветовал что-нибудь попроще».

Ирония чрезмерно сложной геоцентрической модели заключается в том, что она действительно работала! Она предсказывала орбиты планет лучше, чем гелиоцентрическая модель, потому что в то время предполагалось, что орбиты должны быть круглыми (на самом деле они слегка эллипсоидальные). Указанные вложенные круги позволяли модели максимально аппроксимировать эллипсоидальные движения планет и обеспечивать надежные и точные предсказания.

Только спустя поколение после выхода книги Коперника Иоганн Кеплер (Johannes Kepler) внимательно изучил данные, собранные Тихо Браге (Tycho Brahe), и определил, что (1) орбиты имеют эллиптическую форму, (2) планеты движутся с разной скоростью и (3) Солнце находится не совсем в центре этих орбит. С учетом этих новых фактов гелиоцентрическая модель теперь давала более точные предсказания, чем геоцентрическая модель. Но она по-прежнему не была принята.

Сегодня Галилей известен как отец астрономии, а его знаменитый мощный телескоп внес последние штрихи. Открытие спутников, вращающихся вокруг Юпитера, опровергло утверждение о том, что все вращается вокруг Земли. Кроме того, Галилей наблюдал фазы Венеры (колебания света на поверхности планеты), которые были предсказаны гелиоцентрической моделью Коперника.

Это должно было бы положить конец спорам, если бы католическая церковь не была непреклонно привержена библейскому утверждению о том, что Земля неподвижна и стоит на своем основании. В 1616 году вера в гелиоцентрическую модель была объявлена еретической, и началась знаменитая битва между Галилеем и католической церковью.

Галилей был знаком с папой Урбаном VIII со времен его учебы в Пизанском университете и после нескольких бесед почувствовал, что получил благословение папы на написание книги, в которой излагались различные конкурирующие точки зрения и доказывалось превосходство гелиоцентрической модели. Выпущенная им книга, вероятно, оказалась совсем не такой, как ожидал папа римский, и, вместо того чтобы уладить спор, подлила масла в огонь.

Диалог Галилея о двух главных мировых системах, казалось, высмеивал позицию церкви через персонажа Симплицио, который был представлен как человек настолько умный, насколько позволяло его имя. Некоторые из диалогов Симплицио имели поразительное сходство с заявлениями папы римского об астрономии и, возможно, придавали особую актуальность требованиям инквизиции, чтобы Галилей прибыл в Рим и предстал перед судом.

Галилей отправился в Рим и более двух недель встречался со своими обвинителями. Его заставили отречься от своих взглядов и приговорили не к пыткам, а к бессрочному домашнему аресту. Менее чем через десять лет Галилей умер, так и не дождавшись широкого признания своей гелиоцентрической модели. Иллюзорное представление о том, что Солнце встает на востоке и заходит на западе, которое люди наблюдали каждый день, было слишком укоренившимся, чтобы большинство людей могли взглянуть на это с другой стороны.

Закон Боде

Расстояния внутри нашей Солнечной системы измеряются в астрономических единицах (АЕ) – среднем расстоянии между Землей и Солнцем. В XVIII веке два немецких астронома, Иоганн Титиус (Johann Titius) и Иоганн Боде (Johann Bode), заметили закономерность, показанную в табл. 3.3, в расстояниях четырех ближайших к Солнцу планет.

Таблица 3.3. Расстояния четырех ближайших к Солнцу планет

Порядок *n*	Планета	Расстояние (АЕ)	Закономерность
1	Меркурий	0,39	0,4
2	Венера	0,72	0,4 + 0,3
3	Земля	1,00	0,4 + 0,3(2)
4	Марс	1,52	0,4 + 0,3(2^2)

Указанная закономерность может быть выражена математически в виде так называемого закона Титиуса–Боде, или просто закона Боде:

$$\textit{расстояние} = 0{,}4 + 0{,}3(2^{n-2}),\ \text{если } n = 2\ (\text{Венера}),\ 3\ (\text{Земля})\ \text{или}\ 4\ (\text{Марс}).$$

Уравнение закона Боде, $\textit{расстояние} = 0{,}4 + 0{,}3(2^{n-2})$, не ахти какой закон, так как на самом деле оно дает только расстояния Венеры и Марса относительно Земли и не применимо к Меркурию, если только *n* не установлено произвольно равным минус бесконечности.

Однако табл. 3.4 и рис. 3.3 показывают, что этот закон также неплохо работал бы для двух других известных в то время планет – Юпитера и Сатурна, – если бы между Марсом и Юпитером находилась планета.

Таблица 3.4. Шесть известных планет на момент разработки закона Боде

Порядок *n*	Планета	Расстояние (АЕ)	$D = 0{,}4 + 0{,}3(2^{n-1})$
1	Меркурий	0,39	0,55
2	Венера	0,72	0,70
3	Земля	1,00	1,00
4	Марс	1,52	1,60
5	?		2,80
6	Юпитер	5,20	5,20
7	Сатурн	9,55	10,00

Это была закономерность, которая не имела никакого глубинного основания, но она подтвердила веру многих в то, что Бог расположил планеты в соответствии с преднамеренным порядком. Отметив разрыв между Марсом и Юпитером, Боде задался вопросом: «Можно ли поверить, что Создатель Вселенной оставил это пространство пустым? Конечно, нет». И Титиус, и Боде рекомендовали астрономам использовать обнаруженную ими закономерность для поиска новых планет.

Рис. 3.3. Закон Боде служит хорошим предзнаменованием

В 1781 году достаточно близко к тому месту, где, согласно закону Боде, должна была находиться следующая планета после Сатурна, был открыт Уран, а в 1801 году между Марсом и Юпитером недалеко от того места, где, согласно закону Боде, должна была находиться пятая планета от Солнца, была обнаружена Церера. Несомненно, это было поразительным свидетельством Божьего замысла о Вселенной.

Если вы не слышали о планете Церера, то это потому, что она больше не считается планетой. В 1930 году закон Боде был еще больше подорван открытием Нептуна и Плутона, расположенных далеко от того места, где, согласно закону Боде, они должны находиться (табл. 3.5 и рис. 3.4).

Таблица 3.5. Закон Боде работает для семи из десяти (?) планет

Порядок *n*	Планета	Расстояние (АЕ)	$D = 0,4 + 0,3(2^{n-1})$
1	Меркурий	0,39	0,55
2	Венера	0,72	0,70
3	Земля	1,00	1,00
4	Марс	1,52	1,60
5	Церера (?)	2,77	2,80
6	Юпитер	5,20	5,20
7	Сатурн	9,55	10,00
8	Уран	19,22	19,60
9	Нептун	30,11	38,80
10	Плутон	39,54	77,20

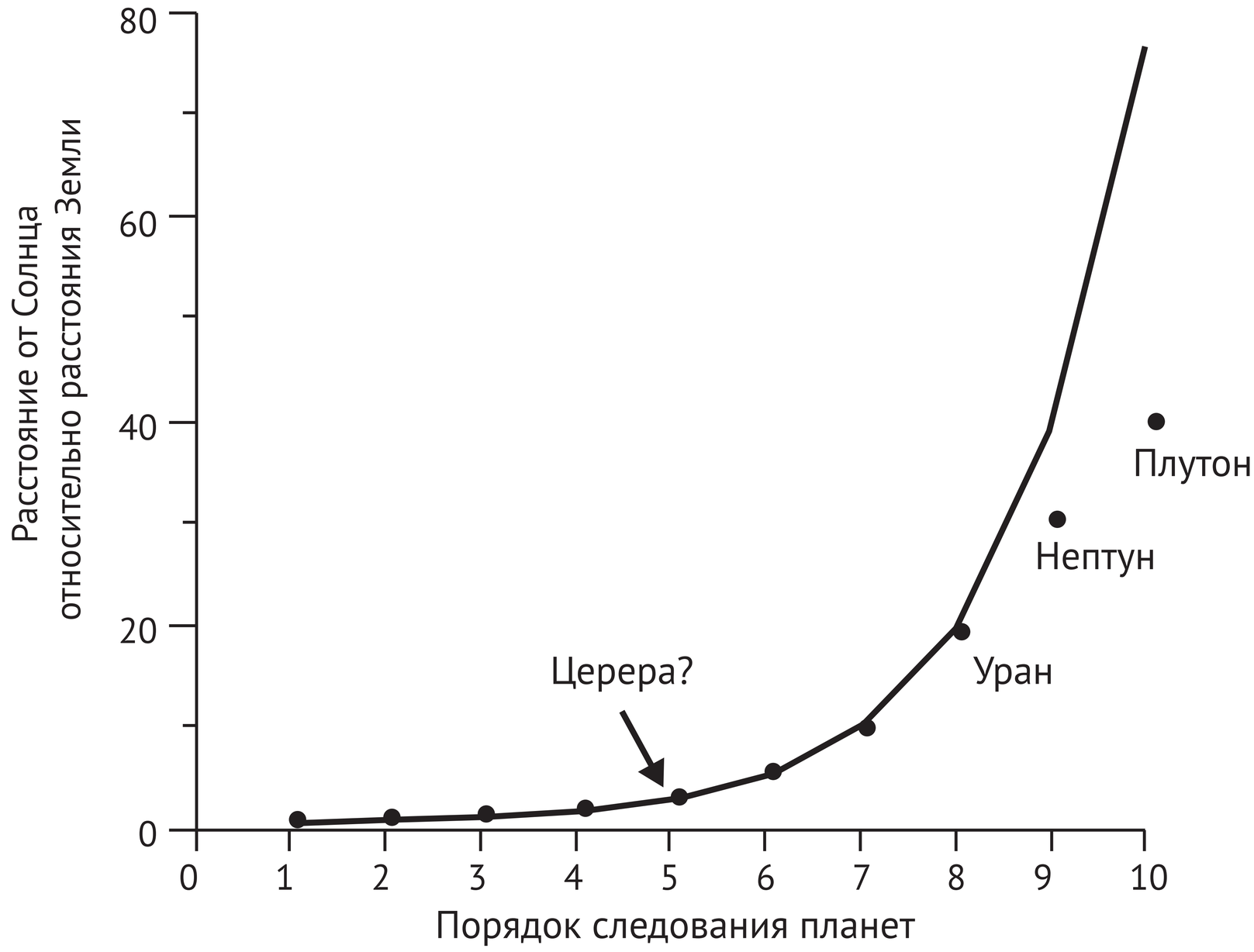

Рис. 3.4. Закон Боде работает для Цереры, не являющейся планетой, но не для Нептуна и Плутона

Несколько человек пытались реанимировать закон Боде, видоизменив его, в том числе С. Б. Ульман (S. B. Ullman), который предложил приведенное ниже сложное расширение:

расстояние = 0,4,	если n = 1 (Меркурий);
$0{,}4 + 0{,}3(2^{n-2})$,	если n = 2 (Венера), 3 (Земля), 4 (Марс), 5 (Церера), 6 (Юпитер), 7 (Сатурн) или 8 (Уран);
$0{,}4 + 0{,}3(2^{n-2}) - ((n - 8)3)^2$,	если n = 9 (Нептун) или 10 (Плутон).

Подобные благонамеренные усилия являются примером того, что сейчас называется *переподгонкой*, – неустанным добавлением сложности, направленным исключительно на то, чтобы модель лучше вписывалась в данные. Критикуя переподгонку, великий математик Джон фон Нейман (John von Neumann) однажды сказал: «С четырьмя параметрами я смогу вписать слона, а с пятью – заставить его махать хоботом». На рис. 3.5 показана четырехпараметрическая модель, которая действительно похожа на слона.

Окончательным результатом переподгонки нередко является, как в случае с модифицированным законом Боде, модель, которая хорошо вписывается в данные, но не имеет глубинной рифмы или основы – поэтому она плохо работает при составлении предсказаний за пределами ограниченной области, в которой с ней манипулировали и вносили в нее правки, чтобы она вписывалась в данные.

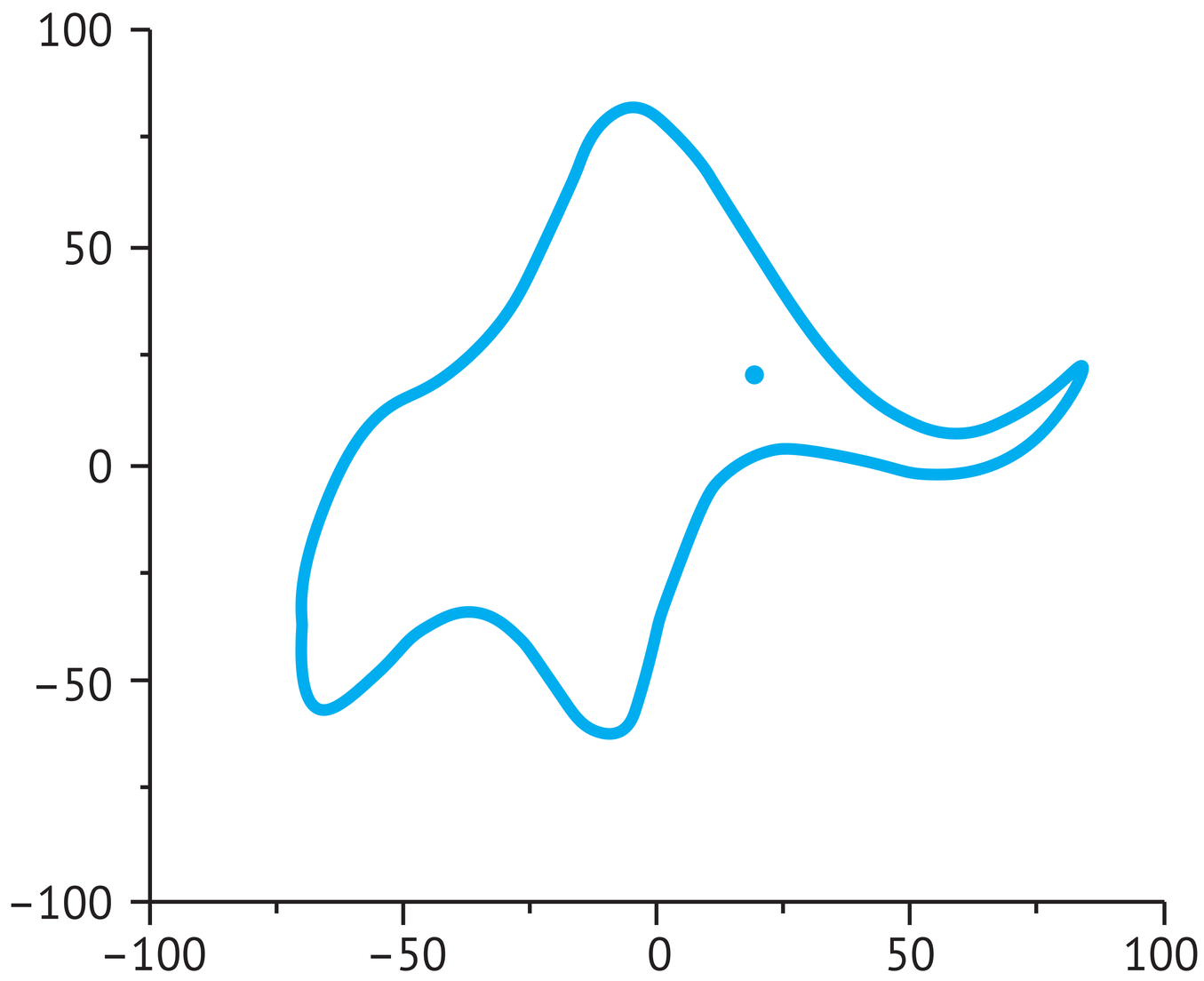

Рис. 3.5. Слон Джона фон Неймана

В 2008 году Международный астрономический союз (МАС) понизил статус Плутона до карликовой планеты и признал еще четыре карликовые планеты: Цереру, Хаумеа, Макемаке и Эриду. Теперь закон Боде оказался неразрешимой загадкой. Если опустить карликовые планеты, такие как Церера и Плутон, то закон Боде не работает для Юпитера, Сатурна, Урана и Нептуна. Если включить пять карликовых планет, то закон Боде работает для Цереры, Юпитера, Сатурна и Урана, но не работает для Нептуна, Плутона и трех карликовых планет.

Что же остается делать? Лучшим ответом, без сомнения, было бы признать, что на протяжении миллиардов лет наша солнечная система стабилизировалась благодаря хорошо отстоящим друг от друга планетам, которые не были разрушены гравитационным притяжением между планетами, но было бы ошибкой думать, что это расстояние соответствует математическому «универсальному закону». Людей утешает мысль о том, что все управляется числовыми закономерностями, которые еще только ждут своего открытия. Однако безосновательные закономерности ненадежны. Закон Боде – интересная закономерность, применение которой неизвестно.

Закон Мура

В 1965 году соучредитель компаний Fairchild Semiconductor и Intel Гордон Мур (Gordon Moore) написал статью с провокационным названием «Втискивание большего количества компонентов в чипы интегральных схем» (Cramming more components onto integrated circuit chips). Используя данные за четыре года о количестве транзисторов на квадратный дюйм в интегральных схемах, он построил график, подобный рис. 3.6. Единицы измерения, отложенные на вертикальной оси, – это натуральные логарифмы количества транзисторов на квадратный дюйм, вследствие чего наклон вписанной линии показывает ско-

рость роста. Здесь темпы роста составили примерно 1, или 100 %, что означает ежегодное удвоение.

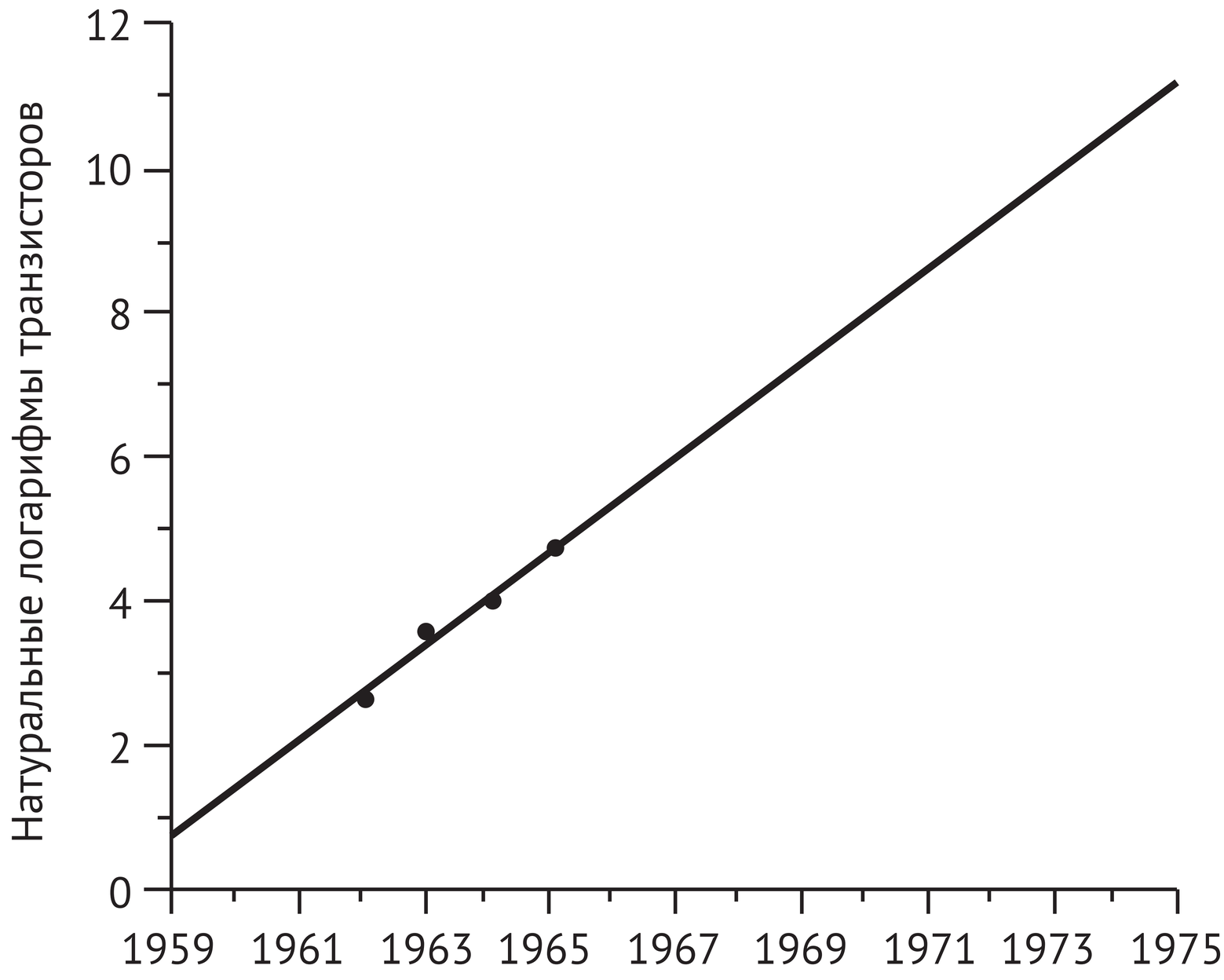

Рис. 3.6. Ежегодное удвоение

Указанный быстрый рост числа транзисторов стал известен как закон Мура, хотя Мур и не утверждал, что открытая им закономерность является физическим законом, подобным закону сохранения материи или законам термодинамики. Это было всего лишь эмпирическое наблюдение. Но «ежегодное удвоение» является незабываемо простым правилом с невероятными последствиями.

Мур отметил, что ежегодное удвоение количества транзисторов на квадратный дюйм в период с 1965 по 1975 год увеличило бы количество транзисторов на квадратный дюйм в ошеломляющие 1000 раз, до 65 000, и заявил, что втиснуть такое количество транзисторов в схему на самом деле было возможно:

> Безусловно, в краткосрочной перспективе можно ожидать, что этот показатель сохранится, если не увеличится. В долгосрочной перспективе темпы роста несколько более неопределенны, хотя нет оснований полагать, что они не будут оставаться практически неизменными в течение как минимум 10 лет. Это означает, что к 1975 году количество компонентов в интегральной схеме при минимальных затратах составит 65 000.
>
> Считаю, что такая большая схема может быть собрана на одной пластине.

Фактическая цифра в 1975 году оказалась примерно в десять раз меньше предсказания Мура, что заставило его пересмотреть свои расчеты темпов роста с ежегодного удвоения до удвоения каждые два года.

На рис. 3.7 показано, что количество транзисторов на дюйм продолжает поразительно увеличиваться в течение более чем пятидесяти лет, в среднем удваиваясь каждые два года. Провокационное, но неверное предсказание Мура о том, что в 1975 году в схему можно было бы втиснуть 65 000 транзисторов, в 2018 году стало карликовым по сравнению с 9 млрд транзисторов на квадратный дюйм.

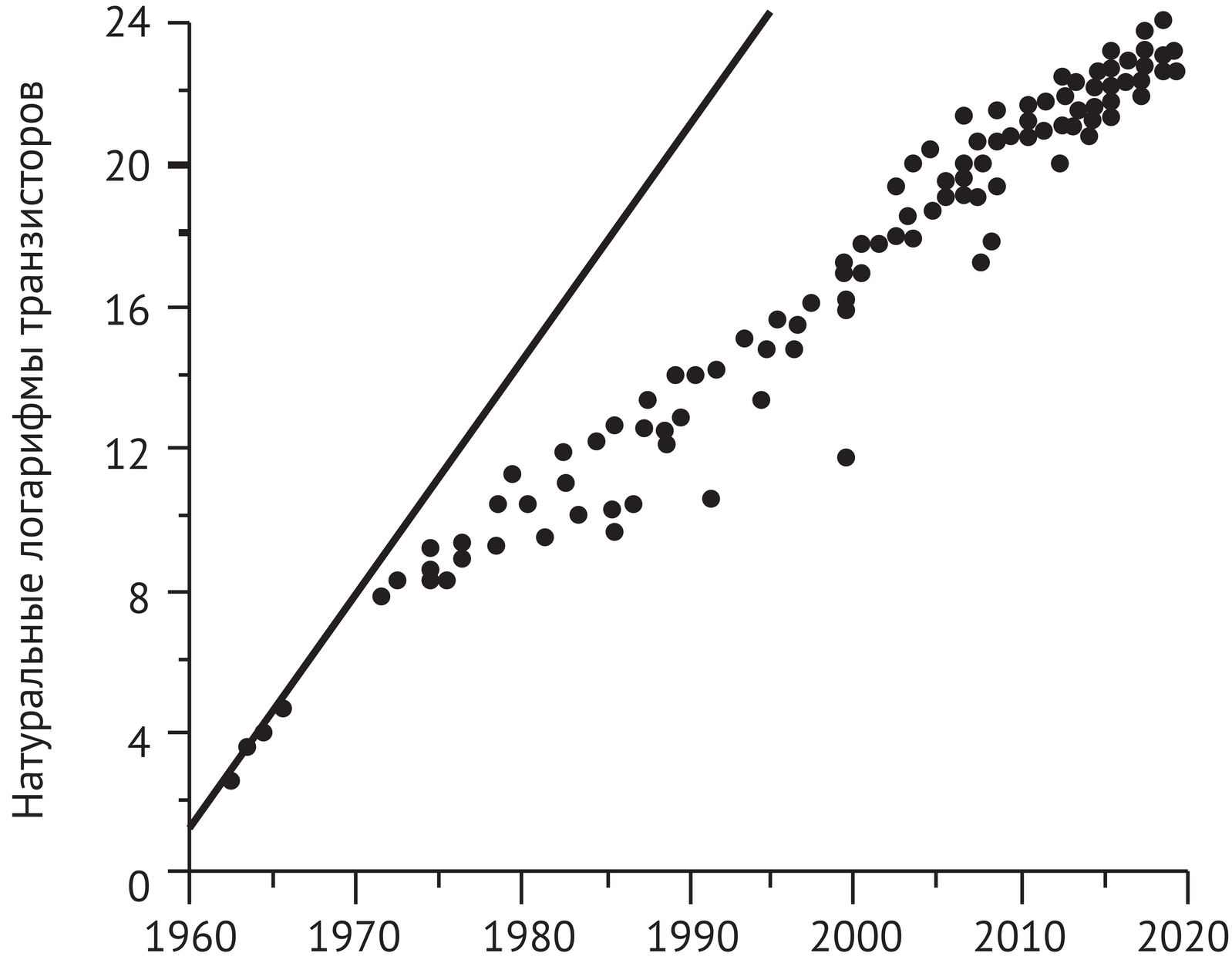

Рис. 3.7. Удвоение каждые два года

В интервью 2015 года Мур сказал: «Думаю, что закон Мура умрет в ближайшие десять лет или около того, но это неудивительно». В 2016 году ведущий инженер Intel заявил, что действие закона Мура может закончиться через четыре или пять лет. Многие считают, что будущие чипы будут намного умнее, а не намного меньше.

Закон Мура – это простое, но слишком точное описание общей закономерности. Удивительно, как из года в год компьютерные компоненты становятся намного мощнее и дешевле, но было бы ошибкой думать, что этот прогресс можно описать простым математическим законом.

Безосновательные закономерности ненадежны.

Как не быть введенным в заблуждение призрачными закономерностями

Мы запрограммированы на то, чтобы замечать закономерности, активно их искать и поддаваться их влиянию. Иногда они оказываются полезными, в других случаях они нас одурманивают и обманывают. Наша привязанность

к закономерностям настолько сильна, что она пережила эпоху Просвещения и победу научного метода – без сомнения, поощряемая и подстрекаемая избирательным припоминанием и склонностью искать подтверждения своей точки зрения. Мы запоминаем, когда закономерность сохраняется, и подтверждаем нашу веру в нее. Мы забываем или отбалтываемся от тех моментов, когда этого не происходит.

Мы все еще находимся во власти глупых суеверий и очарованы числовыми совпадениями. Мы по-прежнему считаем, что одни числа являются удачными, а другие – неудачными, хотя в разных культурах числа считаются удачными и неудачными по-разному. Мы по-прежнему считаем некоторые числа особенными и замечаем их повсюду вокруг себя. Мы по-прежнему превращаем числовые закономерности в законы и экстраполируем случайные везения на уверенные предсказания.

Очарование закономерностей трудно игнорировать. Перед искушением трудно устоять. Первый шаг – распознать соблазн.

Глава 4

Одурачены снова и снова

Мы унаследовали от наших далеких предков ошибочный взгляд на случайность. Мы думаем, что случайные события, такие как подбрасывание монеты, должны чередоваться, а не выпадать сериями по нескольку орлов или решек подряд. Так, за подбрасыванием монеты, в результате которого выпадает орел, вскоре должна последовать решка. В целом за жизненными взлетами должны следовать падения, за высотами – низины, за хорошим – плохое. Если бы все было так, то эти события не были бы случайными, поскольку появление орла увеличивало бы шансы появления решки, что вовсе не является случайностью. Когда что-то происходит по-настоящему случайно, могут происходить серии, и они происходят.

Подбрасывания и подделки

Гэри и Джерри получили по монетке, чтобы подбросить их десять раз и записать результаты, которые приведены в табл. 4.1. Один участник, следуя правилам, подбросил монетку десять раз. Другой не стал возиться и записал десять воображаемых подбрасываний. Кто подбрасывал и кто подделывал?

Таблица 4.1 Десять подбрасываний и подделок

Гэри	Джерри
О	О
О	Р
О	Р
Р	О
Р	Р
О	О
О	О
О	Р
О	О
Р	Р

Еще один способ визуализации этих данных показан на рис. 4.1. Какая последовательность подбрасываний кажется вам более случайной?

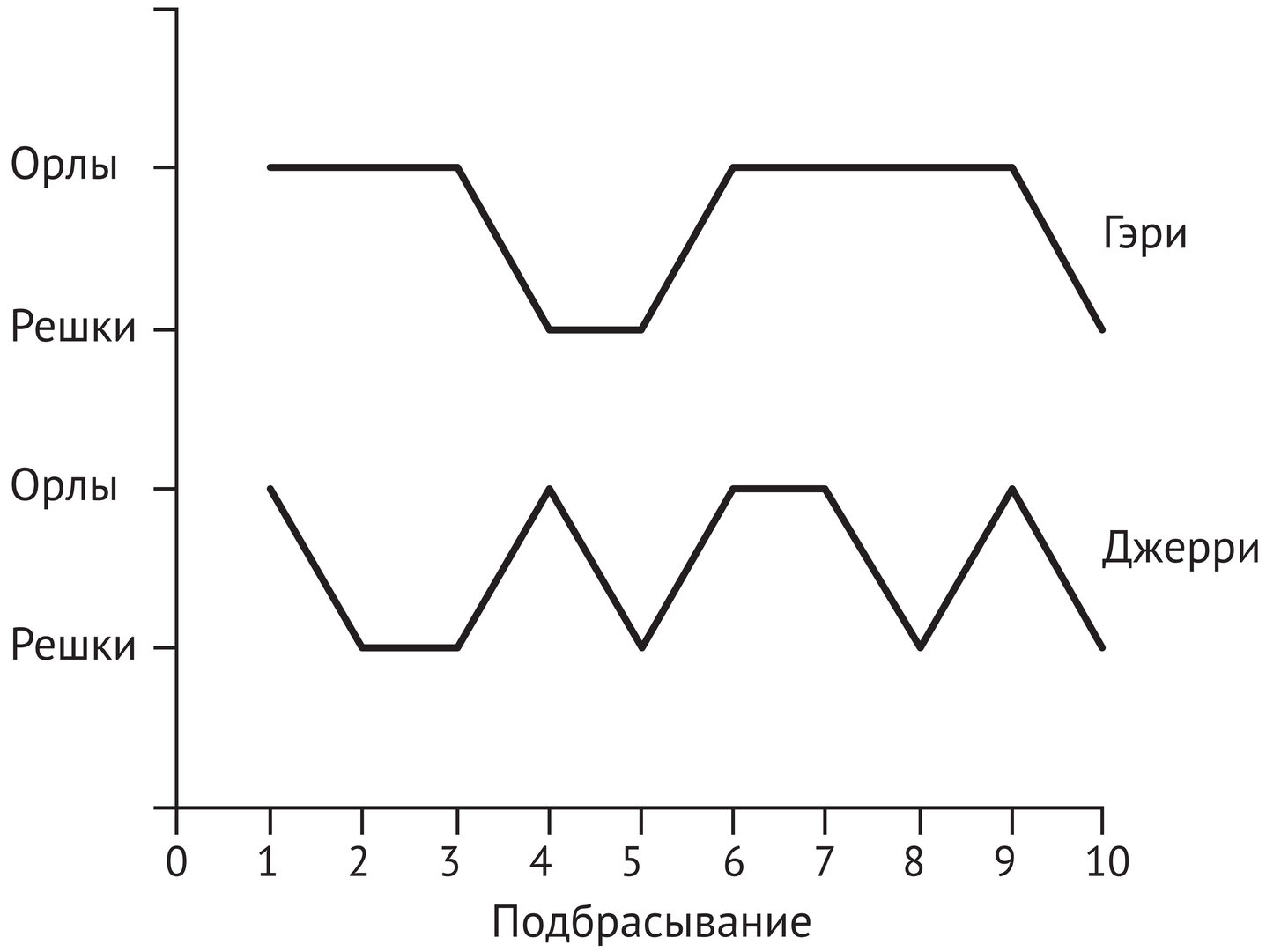

Рис. 4.1. Сможете ли вы распознать подделку?

Большинство смотрящих на эти результаты людей думают, что Джерри подбрасывал, а Гэри смошенничал. Джерри сообщил о 50 % орлов, тогда как Гэри – о 70 %. Гэри сообщил, что у него была серия из четырех выпадений орла подряд и еще одна серия из трех выпадений орла подряд, а у Джерри самыми длинными сериями были два выпадения орла подряд и две решки подряд.

Несмотря на то что все выглядит иначе, у Гэри подбрасывания были настоящими. Мы знаем об этом, потому что на недавнем уроке статистики он подбросил монету десять раз, и результаты приведены в табл. 4.1. Мы знаем, что подбрасывания Джерри – поддельные, потому что мы попросили его представить десять подбрасываний монеты.

Настоящие подбрасывания Гэри не являются чем-то необычным. Воображаемые подбрасывания Джерри совершенно ненастоящие. Существует всего 25%-ный шанс, что десять подбрасываний дадут пять орлов и пять решек. В трех случаях из четырех возникнет тот или иной дисбаланс.

Что касается серий, то настоящие результаты Гэри возникают гораздо чаще, чем поддельные результаты Джерри. При десяти подбрасываниях монеты существует лишь 17%-ный шанс, что самая длинная серия выпадений орла или решки подряд будет равна двум, тогда как существует 46%-ный шанс серии из четырех или больше. Настоящая самая длинная серия Гэри из четырех гораздо вероятнее, чем заявленная самая длинная серия Джерри из двух.

Чем больше мы смотрим на данные, тем больше шансов обнаружить серии и другие поразительные закономерности. Если подбросить монету десять раз, то вряд ли выпадет серия из десяти орлов или десяти решек. Но если подбросить монету 1000 раз, то шанс, что где-то в ходе этих подбрасываний выпадет серия из десяти или более орлов или решек, составляет 62 %. Закономерности, которые кажутся вряд ли возможными, на самом деле очень возможны, в особенности при наличии большого количества данных.

Заблуждение о том, что случайные данные не имеют серий, может послужить причиной разного рода неприятностей. Когда мы видим серии, возникает соблазн сделать поспешный вывод о том, что происходит что-то настоящее. Если игрок выигрывает четыре партии подряд, мы считаем, что он, должно быть, в ударе и, скорее всего, продолжит выигрывать. Если специалист по подбору акций делает четыре правильных предсказания подряд, то мы считаем, что он, должно быть, – гуру, которому стоит заплатить за совет. На самом деле им обоим, вероятно, повезло. Чем больше игр мы смотрим и чем больше за кем-то из игроков следим, тем больше шансов, что кому-то повезет.

Spotify, iTunes и другие цифровые музыкальные плееры предлагают режим случайного воспроизведения, в котором встроенный алгоритм рандомно отбирает песни из пользовательского списка воспроизведения. Когда они только появились, алгоритмы были по-настоящему случайными, поскольку каждая песня в списке воспроизведения имела равные шансы быть выбранной для воспроизведения следующей. При таком количестве людей, воспроизводящих так много музыки, неизбежно возникали серии, в которых один и тот же исполнитель или жанр воспроизводился несколько раз подряд, что приводило к жалобам на то, что с алгоритмом что-то было не так.

С алгоритмами как раз все было в порядке, а вот с восприятием пользователями того, как выглядит случайный отбор, – нет. В итоге компании решили модифицировать свои алгоритмы, чтобы вместо того, чтобы быть действительно случайными, порядок воспроизведения песен стал больше соответствовать ожиданиям пользователей.

Хмельные шаги

В табл. 4.1 и на рис. 4.1 показаны подбрасывания монеты, сделанные Гэри, – орел или решка. На рис. 4.2 показана кумулятивная разница между количеством орлов и решек. Исходы вовсе не выглядят случайными, хотя это были десять совершенно обычных подбрасываний монеты. Похоже, что разница между количеством орлов и решек имеет восходящий тренд, что предвещает растущий дисбаланс между орлами и решками. Но не следует забывать, что мы имеем дело с подбрасываниями монеты, и вот что происходит при таких случайных событиях, как подбрасывания монеты. Исходы могут двигаться случайно в том или ином направлении в течение длительных периодов времени.

На рис. 4.2 показано так называемое *случайное блуждание*. В любой момент времени разница между количеством орлов и решек может пойти вверх (если монета выпадет орлом) или вниз (если монета выпадет решкой). Блуждание является случайным, потому что направление, в котором оно движется, не зависит от предыдущих движений, точно так же как и следующий шаг пьяного человека, который может быть сделан в любом направлении и не зависит от предыдущих шагов.

Парадоксальность случайных блужданий заключается в том, что, хотя каждый шаг является случайным и не зависит от предыдущих, блуждание по чистой удаче может отклоняться в ту или иную сторону и образовывать поразительные закономерности. Например, многие технические аналитики изучают графики цен на драгоценные металлы, акции и другие инвестиционные ин-

струменты, надеясь предсказать, в какую сторону пойдут цены в следующий раз. Несмотря на то что цены подчиняются случайному блужданию, – и, следовательно, непредсказуемы, – закономерности неизбежны.

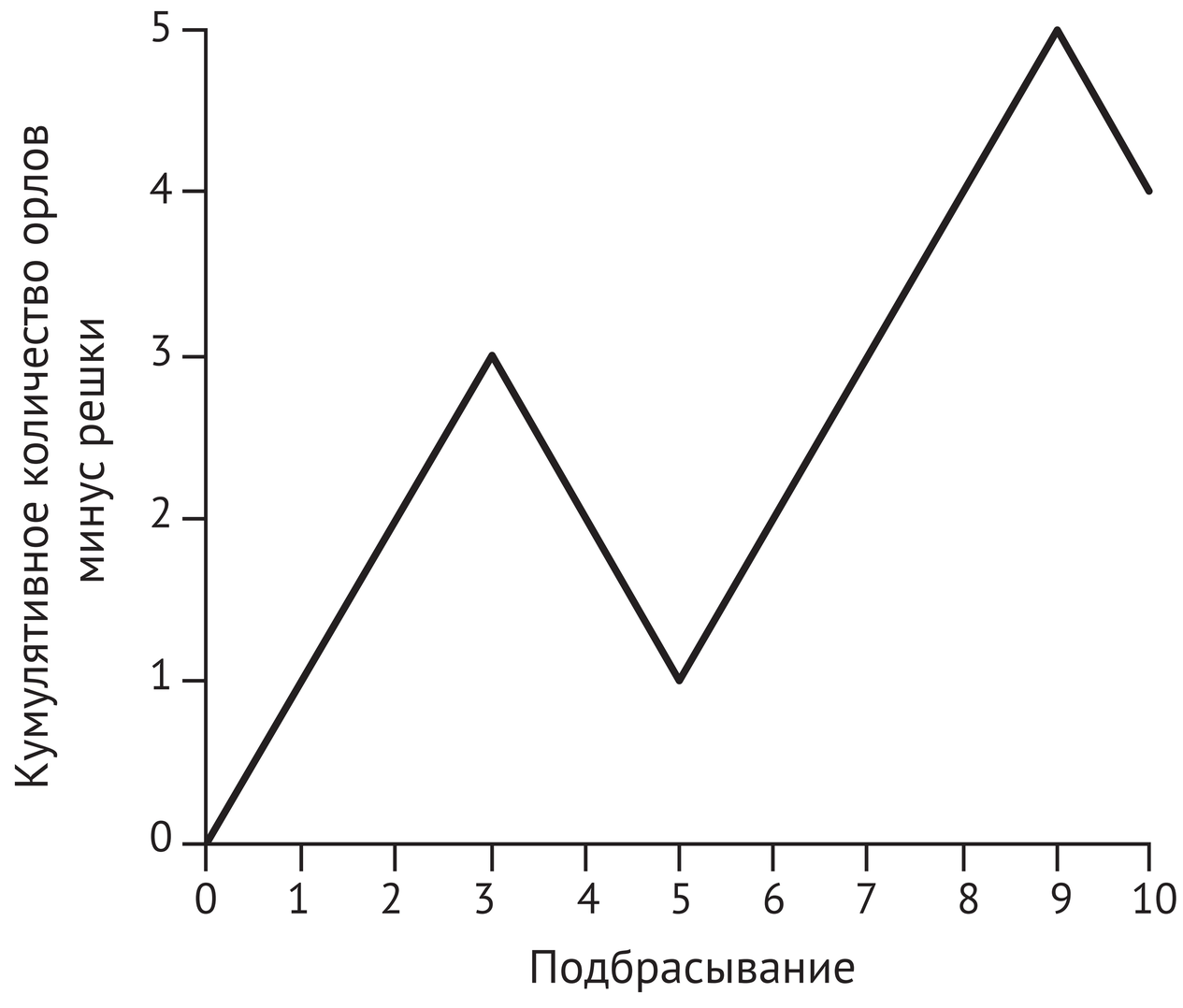

Рис. 4.2. Кумулятивная разница между количеством орлов и решек у Гэри

В ценовом графике открытие–максимум–минимум–закрытие для ежедневных цен на золото и другие инвестиционные инструменты используется последовательность дневных вертикальных линий, каждая из которых охватывает минимальную и максимальную цены в тот день, хеш-метка слева от вертикальной линии показывает цену на момент открытия биржевых торгов в тот день, и хеш-метка с правой стороны показывает цену на момент закрытия торгов.

На рис. 4.3 и 4.4 показаны два графика открытие–максимум–минимум–закрытие для ежедневных цен на золото за 100-дневный период. Один из этих графиков представляет собой настоящий график цен на золото; другой – поддельный график, построенный с учетом цены в 1000 долл., и затем каждый воображаемый день, подбрасывая электронную монету 25 раз, при этом цена идет вверх, когда монета выпадает орлом, и идет вниз, когда монета падает решкой. Указанный эксперимент был повторен 100 раз, что для этих воображаемых цен на золото соответствует 100 торговым дням.

Какая диаграмма, рис. 4.3 или рис. 4.4, настоящая, а какая поддельная? Мы не наносили никаких цифр на вертикальную ось, потому что не хотели давать никаких подсказок. Вопрос состоит в том, можно ли вообще, рассматривая эти два графика открытие–максимум–минимум–закрытие, с уверенностью отличить реальные цены на золото от поддельных.

Ответ – нет, и в этом суть. Когда технические аналитики изучают подобные графики, они нередко находят закономерности (например, восходящие каналы и уровни поддержки), которые, по их мнению, наполнены смыслом. Чего они недооценивают в полной мере – так это того, что даже случайные блуждания, при которых будущие движения цен полностью независимы от измене-

ний цен в прошлом, могут генерировать закономерности. Если мы не можем различить, возникла ли закономерность на основе настоящих цен или случайных подбрасываний монеты, то при всем желании она не может быть полезна для предсказаний цен.

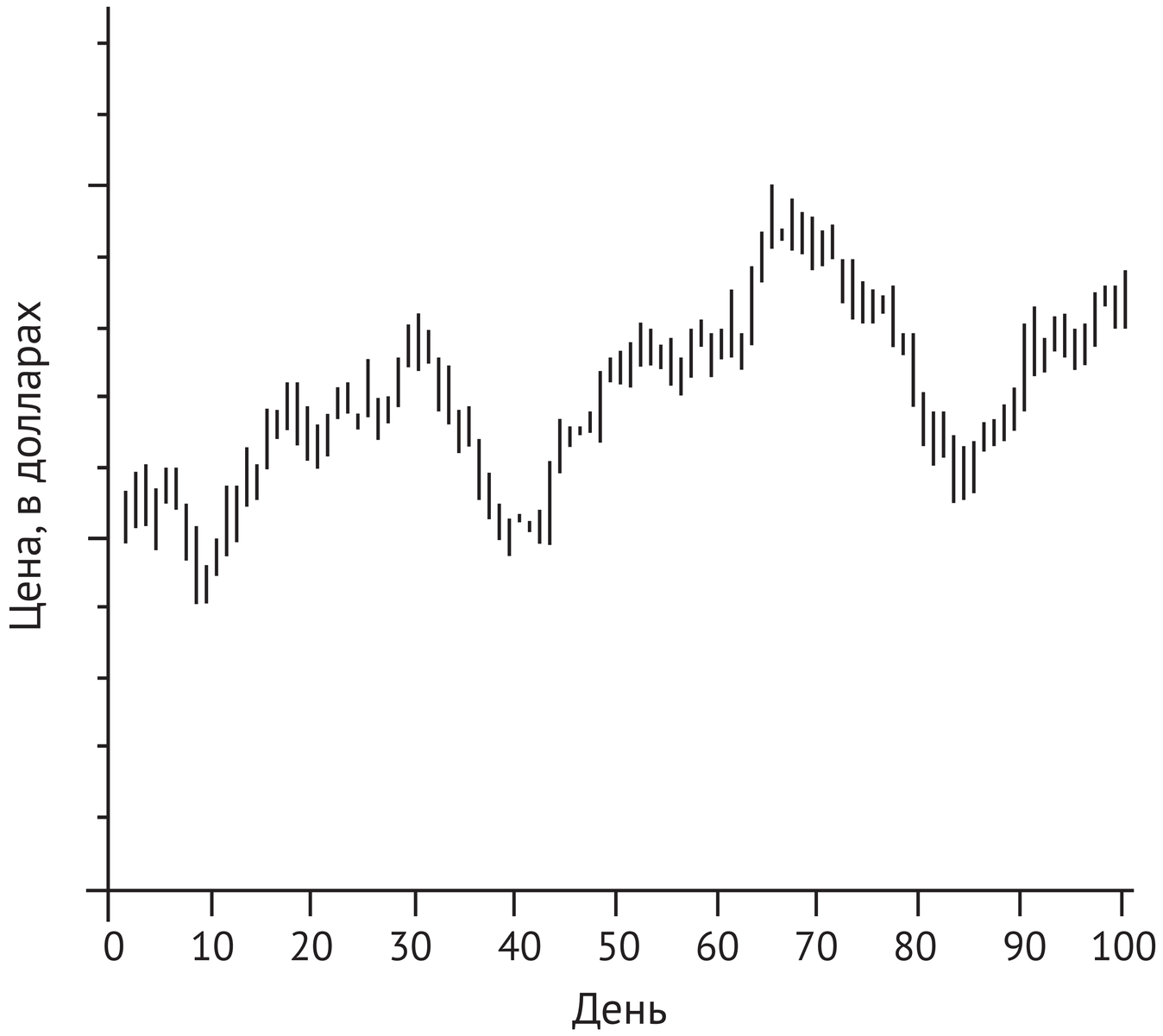

Рис. 4.3. Цены на золото (или подбрасывания монеты), торгующиеся в восходящем канале

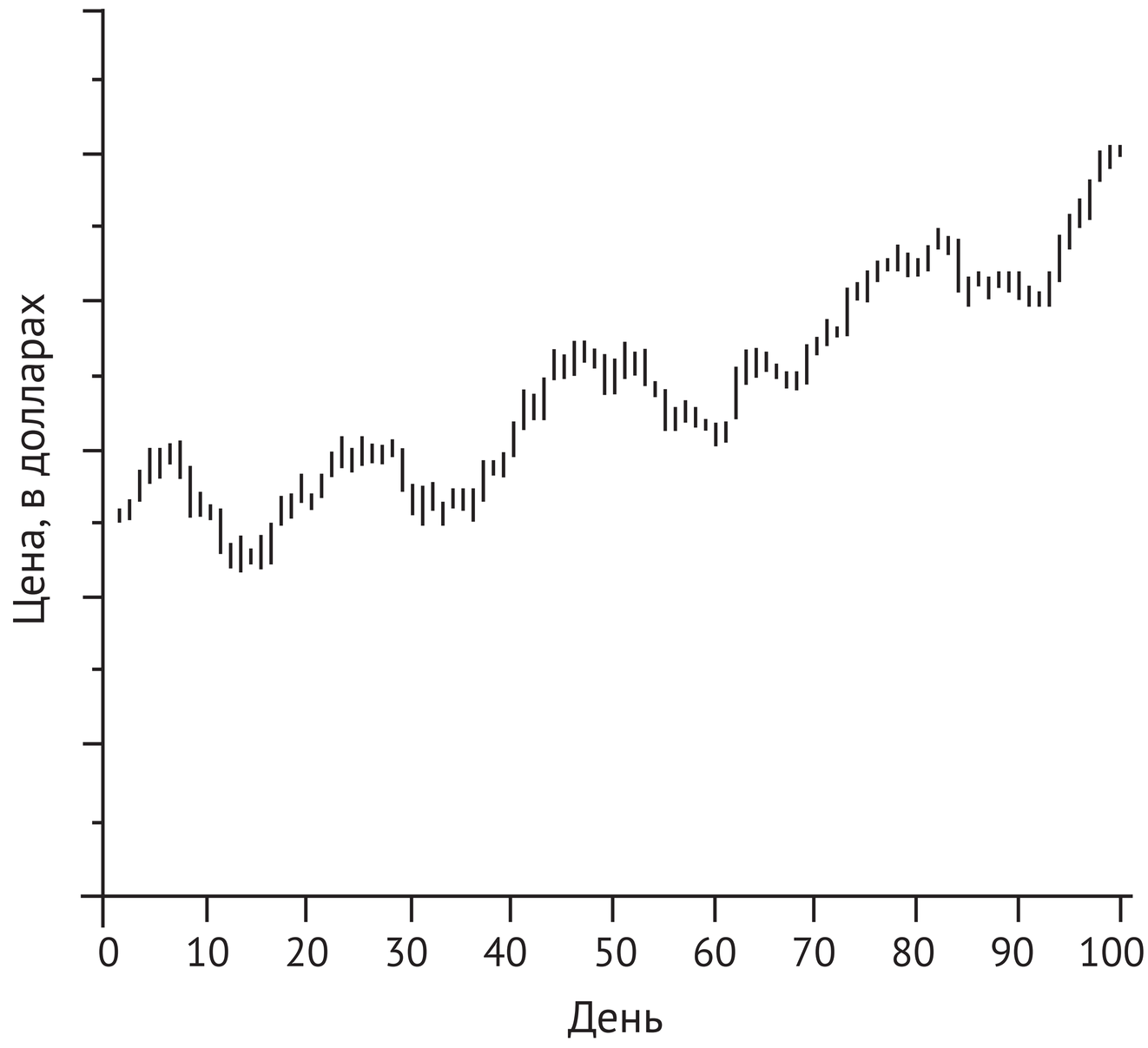

Рис. 4.4. Цены на золото (или подбрасывания монеты), торгующиеся в восходящем канале

(КСТАТИ, рис. 4.3 – настоящий; рис. 4.4 – поддельный.)

Серии на подъеме

Многие спортсмены и болельщики считают, что у игроков иногда бывают промежутки, когда они *в ударе*, из-за чего шансы на успех временно поднимаются, или *на спаде*, из-за чего шансы на успех временно падают. Например, они видят, как баскетболист, который обычно попадает в 50 % бросков, находится в ударе и попадает несколько раз подряд. Игрок Национальной баскетбольной ассоциации (НБА) Первис Шорт, который однажды набрал 59 очков в матче НБА, утверждал, что «ты живешь в своем собственном мире. Это трудно описать. Но корзина кажется такой широкой. Что бы ты ни делал, ты знаешь, что мяч попадет в кольцо».

Насколько реальны промежутки, когда игроки в ударе и на спаде, или это иллюзия? Мы знаем, что существует 46%-ный шанс, что у подброшенной десять раз монеты выпадет как минимум четыре орла подряд или четыре решки подряд, и мы знаем, что такие комбинации бессмысленны, потому что шансы на выпадение орла или решки при следующем подбрасывании остаются равными пятьдесят на пятьдесят. Со спортивными результативностями точно так же? Это лишь временные серии, представляющие собой не более чем бессмысленное совпадение?

На эти вопросы трудно ответить, потому что, в отличие от подбрасывания монеты, условия во время большинства спортивных соревнований постоянно меняются. В одной игре НБА игрок может попытаться сделать бросок на один фут из-под кольца и в следующей игре бросок на двадцать четыре фута в прыжке, или же его могут плотно блокировать в одной игре и слабо – в следующей.

Гэри провел исследования игр в боулинг и в подковки, условия которых остаются стабильными, и пришел к выводу, что, хотя игроки могут и не иметь ударных промежутков, у них все же бывают промежутки *на подъеме*, так как шансы на выбивание всех 10 кеглей и точного метания подковы в цель после предыдущих успехов несколько возрастают. Возможно, это связано с тем, что, в отличие от монет, люди обладают памятью и нередко работают лучше, когда уверены в себе.

Вполне возможно, что уверенные в себе спортсмены выступают лучше, чем сомневающиеся, и эту идею подтверждают экспериментальные свидетельства. В одном исследовании измерялась сила рук двадцати четырех студентов колледжа, которых затем разделили на пары на соревнованиях по армрестлингу, причем оба соперника получали неверную информацию о том, у кого из них руки сильнее. Более слабый человек выигрывал десять из двенадцати поединков. В аналогичном эксперименте, включавшем соревнование на мышечную выносливость ног, студенты-участники показывали более высокие результаты, когда им (неверно) говорили, что их соперник недавно перенес операцию на колене, чем когда им (неверно) говорили, что их соперник был легкоатлетом из университетской команды.

Гэри обратил внимание на боулинг и подковки, потому что броски шара и метания подковки выполняются в схожих условиях, с относительно небольшим промежутком времени между попытками. На ежегодном соревновании НБА по трехочковым броскам тоже существуют относительно стабильные ус-

ловия. Матч всех звезд НБА – это выставочный матч в середине сезона, в котором участвуют две дюжины лучших игроков лиги. Акцент делается на зрелищных атакующих действиях, которые развлекают болельщиков, а защита в основном старается не мешать. В 2019 году итоговый счет в Матче всех звезд составил 178:164, по сравнению со средним показателем в 111 очков за игру в регулярном чемпионате.

В дополнение к Матчу всех звезд во время Уик-энда всех звезд проводится несколько других мероприятий, в том числе соревнование по заколачиванию мяча из-под кольца и по трехочковым броскам. В соревновании по трехочковым броскам приглашенные игроки по очереди совершают двадцать пять дальних бросков, по пять бросков с каждой из пяти позиций. Три лучших снайпера квалификационного раунда переходят в чемпионский раунд, где они снова совершают двадцать пять бросков. (До 2000 года соревнование состояло из трех раундов.)

В соревновании 2019 года участвовало десять человек, и Стеф Карри, Джо Харрис и Бадди Хилд вышли в чемпионский раунд. Харрис победил, попав девятнадцать раз из двадцати пяти, в том числе ошеломляющие двенадцать подряд. В социальных сетях появилось модное выражение «в ударе». Харрис определенно был в ударе. Или нет? Если игрок попадает несколько раз подряд, то это не обязательно свидетельствует о том, что у него наступила ударная серия. Возможно, он просто очень хороший снайпер.

Харрис великолепно забивает трехочковые. В регулярном чемпионате он лидировал в лиге, реализовав 47,4 % своих трехочковых бросков. Соревнование всех звезд по трехочковым броскам проходят легче, чем регулярный чемпионат, потому что в нем нет защитников. В квалификационном и чемпионском раундах Харрис реализовал тридцать шесть из пятидесяти (72 %) своих попыток. Однако вероятность того, что кто-то, даже такой хороший, как Харрис, сможет попасть двенадцать раз из двенадцати, очень мала. В случае с Харрисом, если допустить, что вероятность успеха каждого броска составляет 72 %, то вероятность того, что он попадет двенадцать раз из двенадцати, составляет 0,019, или менее 2 %.

Стандартный порог статистической значимости в научных исследованиях составляет 5 %. Если происходит что-то, что имеет вероятность произойти случайно менее 5 %, то мы вправе заключить, что это нечто большее, чем просто случайность. Серия Харриса, очевидно, было статистически значимым свидетельством в пользу того, что он был в ударе.

Не так быстро. В чемпионском раунде Харрис сделал двадцать пять бросков, и его серия из двенадцати попаданий подряд началась с третьего броска. Мы должны рассчитать вероятность того, что в какой-то момент из этих двадцати пяти он попадет двенадцать раз подряд, а не вероятность того, что, если он сделает только двенадцать бросков, он попадет все двенадцать.

Можно рассчитать вероятность того, что в какой-то момент раунда из двадцати пяти бросков Харрис попадет двенадцать раз подряд, учитывая тот факт, что он попал девятнадцать раз и промазал шесть. Если бы его двадцать пять бросков были независимыми, без ударных и упадочных промежутков, то каждая возможная комбинации из девятнадцати попаданий и шести промахов была бы равновероятной. На рис. 4.5 показаны вероятности серий различной

длины. Шансы серии хотя бы из двенадцати составляют 6,8 %, что по правилу пяти процентов не совсем статистически значимо.

Даже этот расчет, равный 6,8 %, вводит в заблуждение. Было десять игроков и тринадцать сетов по двадцать пять бросков (десять квалификационных сетов и три чемпионских сета). Было бы неправильно поступать избирательно, концентрируясь на чемпионском сете Харриса после того, как стали известны результаты, и делать вид, что мы не смотрели на остальные двенадцать сетов.

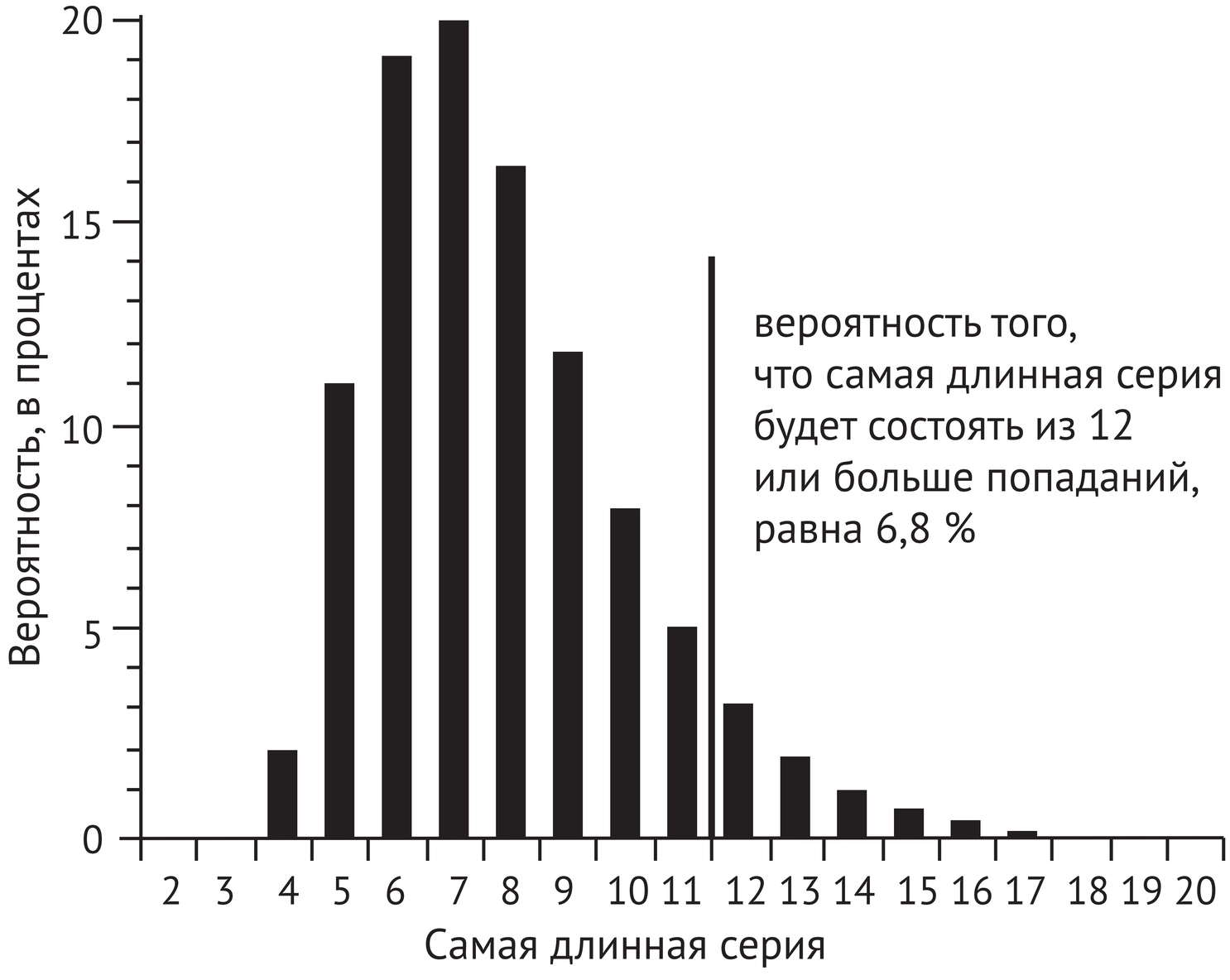

Рис. 4.5. Вероятность самой длинной серии для Джо Харриса в чемпионском раунде

В табл. 4.2 приведены сопоставимые расчеты для каждого игрока в квалификационном раунде и для трех игроков в чемпионском раунде. Серия Харриса наиболее примечательна. Самые длинные серии большинства игроков ничем не примечательны.

Учитывая, что в общей сложности было сыграно тринадцать сетов, правильно спросить, какова вероятность того, что по чистой случайности хотя бы у одного игрока наступит серия с шансом менее 5 %. Эта вероятность составляет отрезвляющие 49 %. Если бы каждый бросок был независимым, без ударных и упадочных серий, то существует 49%-ный шанс того, что хотя бы у одного игрока случится серия, которая сама по себе имеет шанс менее 5 %.

Именно поэтому ошибочно искать закономерность после того, как данные были собраны. В данном случае неверно рассматривать все тринадцать сетов и выделять чемпионский раунд Харриса в качестве убедительного свидетельства в пользу того, что у баскетболистов бывают ударные серии.

Когда мы рассматриваем все данные о соревновании 2019 года целиком, не исключено, что хотя бы у одного игрока будет статистически впечатляющая серия. Тот факт, что лучшим найденным нами результатом была одна серия, которая не была статистически значимой, во всяком случае, свидетельствует против утверждения о том, что у баскетболистов бывают ударные серии.

Таблица 4.2. Соревнование по трехочковым броскам Всех звезд НБА, 2019

Игрок	Команда	Всего бросков	Самая длинная серия	Вероятность
Квалификационный раунд				
Джо Харрис	Бруклин Нетс	17	8	0,202
Кемба Уокер	Шарлотт Хорнетс	11	4	0,814
Крис Миддлтон	Милуоки Бакс	8	7	0,359
Сет Карри	Портленд Трэйл Блэйзерс	13	4	0,766
Дэмиан Лиллард	Портленд Трэйл Блэйзерс	13	4	0,766
Бадди Хилд	Сакраменто Кингз	18	7	0,511
Дэнни Грин	Торонто Рэпторс	16	5	0,714
Дирк Новицки	Даллас Маверикс	12	5	0,412
Стеф Карри	Голден Стэйт Уорриорз	19	10	0,198
Девин Букер	Финикс Санз	15	7	0,149
Чемпионский раунд				
Джо Харрис	Бруклин Нетс	19	12	0,067
Бадди Хилд	Сакраменто Кингз	13	5	0,412
Стеф Карри	Голден Стэйт Уорриорз	18	9	0,190

В проведенных до настоящего момента расчетах рассматривался вопрос о том, находится ли игрок в ударе *во время* раунда, в случае с Харрисом, увязывая двенадцать из девятнадцати попаданий, которые он сделал в чемпионском раунде. Еще игрок может показывать, что он в ударе, добиваясь большего успеха в одном раунде, чем в других; например, в течение трех лет соревнований он попадал в 50 % бросков в пяти раундах, но в одном яростном раунде набрал 80 %. Для Джо Харриса 2019 год стал его первым соревнованием, и он попал семнадцать раз из двадцати пяти бросков в первом раунде с серией из восьми попаданий и девятнадцать раз из двадцати пяти бросков в чемпионском раунде с серией из двенадцати попаданий. Во всяком случае он был последователен, но мы не можем сделать большого вывода из двух раундов данных.

А что можно сказать о других игроках в истории этого соревнования? Соревнования по трехочковым броскам в НБА проводятся с 1986 года, и за это время 124 игрока в общей сложности сделали 389 25-бросковых сетов. Пожалуй, самый высокий результат был достигнут в 1991 году, когда в полуфинальном раунде Крейг Ходжес попал двадцать один раз из двадцати пяти бросков, включая поразительные девятнадцать подряд, – оба достижения остаются абсолютными рекордами. Ходжес также очень хорошо отбросался в первом раунде самого первого соревнования по трехочковым броскам в 1986 году, когда он попал двадцать раз из двадцати пяти.

Ходжес выступал в НБА в течение десяти лет и трижды лидировал в лиге по точности трехочковых бросков. За карьеру точность его трехочковых бросков составила 40 %. Ходжес восемь раз участвовал в соревнованиях по трехочко-

вым броскам, трижды одержав победу и дважды заняв второе место. Он был отличным трехочковым снайпером, и у нас есть данные о девятнадцати раундах соревнований с его участием.

Мы знаем подробности его полуфинального раунда 1991 года, но у нас нет полной информации о нескольких других его раундах. Некоторые из используемых в соревновании двадцати пяти мячей являются красно-бело-синими «денежными мячами», которые оцениваются в два очка по сравнению с одним очком за стандартные мячи. Во многих ранних соревнованиях зарегистрированные результаты показывают количество набранных очков каждым игроком, но не количество попаданий или последовательность, в которой они были сделаны.

Один фанат трехочковых связался с Ходжесом напрямую, пытаясь узнать информацию о его выступлении в 1986 году, и получил следующий ответ:

> Мир всем… это Крейг Ходжес… не уверен, где вы сможете найти это видео… также не уверен, сколько шотов я сделал в самом первом раунде первого в своей истории соревнования. <…> Не уверен, поможет ли это. <…> Мир всем, Ходжес.

Нам удалось найти газетную заметку, в которой сообщалось, что в 1986 году Ходжес сделал двадцать метких бросков из двадцати пяти в первом раунде, но мы не смогли найти информацию о нескольких других соревнованиях.

В целом Ходжес набрал 321/570 = 0,563 возможных очка в своих девятнадцати раундах. Если допустить, что он попадал в 56,3 % всех своих 475 бросков, то получится 267,5 бросков, которые мы округлили до 268. Тогда наш первый набор вопросов в отношении игрока, который попал 268 раз из 475 в девятнадцати раундах, какова вероятность того, что у него (a) будет хотя бы один раунд, в котором он попадет двадцать один раз из двадцати пяти бросков, и (b) будет иметь хотя бы один раунд, в котором он попадет девятнадцать раз подряд? Наш второй набор вопросов заключается в следующем: учитывая, что в соревновании приняли участие 124 игрока, какова вероятность того, что хотя бы один из них сделает что-то столь же необычное, как Ходжес?

Ответы приведены в табл. 4.3. Во-первых, учитывая совокупную результативность Ходжеса, маловероятно, что он выступит так же хорошо, как в полуфинале 1991 года, в особенности повторит свою серию из девятнадцати попаданий. Во-вторых, учитывая, что в соревновании приняли участие 124 игрока, вряд ли кто-то из них совершит что-то столь же невозможное, как серия Ходжеса из девятнадцати попаданий.

Таблица 4.3 Ходжес в ударе в полуфинальном раунде 1991 года

Результат	Вероятность того, что Ходжес выступил бы хорошо	Вероятность того, что кто-то сделал бы что-то настолько же необычное
21 из 25	0,0026716	0,2823
19 раз подряд	0,0000477	0,0059

Один из важных выводов состоит в том, что задача оценки наличия у баскетболистов ударных и упадочных серий не из простых! Поверхностные призна-

ки, такие как то, что игрок попадает несколько раз подряд, вполне могут быть результатом случайности – точно так же как монета может выпасть орлом несколько раз подряд. Даже когда мы выявляем что-то действительно выдающееся, например Джо Харриса, сделавшего девятнадцать попаданий из двадцати пяти бросков, в том числе двенадцать подряд, в 2019 году, это может просто объясняться тем фактом, что мы выдернули эту результативность избирательно – например, подбросили десять монет 1000 раз и отметили, что все десять монет однажды выпали орлом.

Тем не менее существуют статистические способы объяснить избирательность, и, когда мы это сделаем, серия Крейга Ходжеса из девятнадцати метких бросков подряд в соревновании 1991 года станет слишком невероятной, чтобы ее можно было объяснить везением или избирательностью. Он был в ударе.

Ударные промежутки случаются не каждый день, но они случаются.

Вы придираетесь, что ли?

В 2018 году футболист команды «Каролина Пантерс» Эрик Рид (Eric Reid) пожаловался на то, что Национальная футбольная лига (НФЛ) использовала свою программу якобы случайного тестирования на наркотики в качестве мишени в отношении него. За одиннадцать недель он прошел тестирование семь раз, и простым объяснением было то, что НФЛ к нему придиралась из-за того, что он был первым игроком, который присоединился к Колину Каперник (Colin Kaepernick), опустившись на колени во время исполнения национального гимна, и его также несколько раз штрафовали за чрезмерно агрессивные подкаты.

Рид сказал, что «это, должно быть, статистически невозможно. Я не математик, но это не может быть случайностью». Один из комментаторов согласился: «Вероятность того, что какой-либо игрок будет проходить *случайное* тестирование [столько раз], невероятно мала». Тренер Рида поддержал своего игрока: «Если бы мое имя упоминалось так часто, то я бы купил лотерейный билет».

Вскоре к нему присоединились другие игроки, жалуясь, что их проверяли после грубого подката или глупой шутки. Спортивная новостная платформа Yahoo Sports сообщила, что вероятность такого частого выбора Рида для тестирования составляла всего 0,17 %.

Эта ситуация похожа на подбрасывание монеты и трехочковые броски. До того, как начнутся подбрасывания или броски, вероятность того, что возникнет та или иная закономерность, невелика, но после того, как подбрасывания или броски произойдут, вероятность того, что возникнет какая-то закономерность, высока. В начале сезона вероятность того, что Эрик Рид пройдет многократное тестирование, невелика, но после одиннадцати недель тестирования вероятность того, что кто-то пройдет многократное тестирование, велика.

Для выбора имен игроков, которые будут тестироваться, независимая лаборатория использует компьютерный алгоритм, и НФЛ и профсоюз игроков НФЛ провели совместное расследование по поводу жалобы Рида и пришли к выводу, что тесты Рида действительно были случайными.

Давайте проверим вероятности. Первый тест Рида на наркотики был обязательным после того, как он подписал контракт с «Пантерами». После этого первоначального теста из каждой команды НФЛ каждую неделю случайно отбираются десять игроков, и Рид был выбран шесть раз за одиннадцать недель. Если бы мы выбрали Рида до начала сезона, то вероятность того, что его имя всплывет шесть или более раз, составила бы 0,001766. (Это вероятность, о которой сообщает Yahoo Sports.)

Однако мы этого не сделали. Мы выбрали имя Рида после того, как было сыграно одиннадцать игр. В составе «Каролины» семьдесят два игрока, которые могут пройти тестирование, и любой из них мог быть выбран несколько раз. Вероятность того, что кто-то будет выбран как минимум шесть раз, выше, чем вероятность того, что Рид будет выбран шесть раз. В частности, вероятность того, что хотя бы один игрок «Каролины» будет выбран шесть или более раз, составляет 0,119. В дополнение к этому существует тридцать две команды НФЛ, и довольно велики шансы на то, что какой-то игрок из какой-нибудь команды будет выбран шесть раз. Вероятность того, что хотя бы один игрок НФЛ пройдет тестирование не менее шести раз, составляет 0,983 – это почти наверняка.

Как и во многих случайных событиях, после свершившегося факта легко обнаруживать закономерности, которые было бы трудно предсказать заранее. Скорее всего, впоследствии мы обнаружим что-то, что раньше казалось вряд ли возможным.

Притормози и перетасуй

Бридж – это игра, в которую играют стандартной колодой из пятидесяти двух карт, которые перетасовываются и раздаются четырем игрокам из двух противоположных команд по два игрока каждая. В фазе разыгрывания раздачи (т. е. разыгрывания руки) каждая «взятка» состоит из того, что назначенный игрок выступает лидером, кладя на стол карту, а остальные игроки следуют за ним, если смогут. Например, если игрок начинает с пятерки пик, то другие игроки тоже должны играть картой той же масти, если она у них есть.

В бридже существует так много возможных комбинаций карт, что на практике каждая из них отличается от другой, делая игру бесконечно трудной и увлекательной и предусматривая сложные и изощренные стратегии. И действительно, бридж является одной из немногих игр, где компьютерные алгоритмы еще не победили лучших игроков-людей.

Раньше карты в бридже тасовались и сдавались самими игроками. В конце 1970-х и начале 1980-х годов серьезные соревнования начали переходить к автоматической раздаче карт. Сначала игроки жаловались на то, что автоматы были неисправны, потому что они сдавали слишком много экстремальных комбинаций с неравными распределениями. Чаще, чем они помнили, по меньшей мере одному игроку сдавали карты с отсутствием какой-то масти («карты с пропусками») или шесть, семь или более карт одной масти.

К этим жалобам отнеслись серьезно, потому что участвовавшие в этих соревновательных матчах люди обладали многолетним опытом, подтверждаю-

щим их утверждения о том, что автоматически тасуемые раздачи карт имели более экстремальные распределения, чем те, к которым они привыкли. Несколько математиков рассчитали теоретические вероятности и сравнили их с фактическим распределением автоматически тасуемых раздач. Оказалось, что распределение автоматически тасуемых раздач было правильным. Например, в 18 % случаев хотя бы у одного игрока должны отсутствовать какие-то масти; в 50 % случаев хотя бы у одного игрока должно быть шесть или более карт одной масти; и, что примечательно, в 15 % случаев хотя бы у одного игрока должно быть семь или более карт одной масти. Автоматически тасуемые раздачи совпадали с этими частотами.

Проблема была не в тасующих машинах, а в людях, которые тасовали карты.

Как и в случае с подбрасыванием монеты, игроки в бридж недооценивали уровень частоты, с которой случайно отбираемые карты демонстрируют, казалось бы, необычные закономерности, и это неверие укреплялось годами несовершенной тасовки.

Когда разыгрывается бриджевая раздача[9], во многих взятках используются четыре карты одной масти, например четыре пики. В дополнение к этому одна и та же масть часто выпадает два или три раза подряд, в результате чего от восьми до двенадцати карт одной масти группируются вместе. Когда карты собираются в конце розыгрыша, многие карты одной масти, скорее всего, будут скапливаться в группы. Если колоду перетасовать всего два или три раза, чтобы перейти к следующей бриджевой руке («Поторопись и сдавай!»), то некоторые из этих скопившихся мастей останутся практически нетронутыми и будут розданы каждому игроку поровну. Крайним случаем может быть ситуация, когда взятка с четырьмя пиками не разбивается двумя или тремя перетасовками, гарантируя, что при раздаче карт каждый игрок получит одну из этих четырех пик, – это делает невозможным отсутствие пик и затрудняет любому игроку иметь семь или более пик. Проблема заключалась в том, что люди недостаточно тасовали карты!

Статистик и бывший профессиональный фокусник Перси Диаконис (Persi Diaconis) доказал – как в теории, так и на практике, – что если колоду карт разделить на две равные половины и идеально перетасовать карты, чередуя по одной карте из каждой половины, то после восьми идеальных перетасовок колода возвращается к изначальному порядку. Как раз наши несовершенные тасовки приводят к тому, что колода отклоняется от изначального порядка, и требуется несколько неполноценных тасовок, чтобы тщательно перемешать карты. Диаконис и еще один статистик, Дэйв Байер (Dave Bayer), показали, что для рандомизации колоды карт недостаточно двух, трех, четырех или даже пяти несовершенных перетасовок карт человеком. По их эмпирическому правилу обычно требуется семь перетасовок. Шести перетасовок недостаточно, а больше семи не окажет существенного влияния на случайность колоды.

Если требуется получить случайные исходы при игре в бридж, покер и другие карточные игры, то следует притормаживать и тасовать карты семь раз.

---

[9] То есть набор из 13 карт, который получает каждый игрок в игре бридж. – *Прим. перев.*

Поселенцы

«Поселенцы Катана» (The Settlers of Catan) – это потрясающая настольная игра, созданная немецким дизайнером игр Клаусом Тойбером (Klaus Teuber). Она переведена на десятки языков и продана тиражом в десятки миллионов экземпляров. Базовое игровое поле для четырех игроков состоит из девятнадцати шестиугольников (гексов), представляющих ресурсы: три кирпичных гекса, четыре древесных, четыре шерстяных, четыре зерновых, три рудных и один пустынный. Игроки накапливают ресурсы, бросая кости, вытягивая карты, торгуя и определяя местоположение своих поселений и городов. Отчасти ее привлекательность заключается в том, что гексы можно раскладывать практически неограниченным количеством способов, а стратегии игроков зависят от того, как расположены шестиугольники. Правила просты, но стратегии сложны и неуловимы.

Официальные правила игры «Поселенцы Катана» рекомендуют перетасовывать ресурсные гексы, случайно раскладывать их на доске рубашкой вверх, а затем переворачивать. На рис. 4.6 и 4.7 показаны два варианта расположения гексов. Какая доска «Поселенцев» кажется вам более случайной?

Это двухмерная версия десяти подбрасываний монет Гэри и Джерри, в которой вместо десяти монет используется девятнадцать гексов. У нас есть глубоко укоренившаяся тенденция считать, что скопления вряд ли возникают случайно – будь то четыре орла подряд или три древесных гекса подряд. Ошибочно полагать, что орел и решка или гексы «Поселенцев» должны чередоваться.

Рис. 4.6. Несбалансированная доска «Поселенцев»

Рис.4.7. По настоящему случайная доска «Поселенцев»?

Игроки нередко бывают обескуражены, обнаруживая, что случайное расположение приводит к появлению триплета, как в случае с тремя древесными ресурсами на рис. 4.6. Поскольку доска не выглядит случайной, игроки переставляют гексы до тех пор, пока не найдут расположение, подобное рис. 4.7, которое выглядит случайным.

Как и в случае с монетными сериями, случайно расположенные гексы нередко имеют поразительные случайные закономерности. Предсказать конкретную закономерность заранее сложно. Ожидается, что какая-то закономерность обнаружится после свершившегося факта. Существует 29%-ная вероятность того, что на случайно построенной доске «Поселенцев» будет хотя бы один триплет, как показано на рис. 4.6. Существует лишь 4%-ный шанс, что на доске, подобной рис. 4.7, ни у одного из соседних гексов не будет одинакового ресурса.

Игра становится интереснее, если игроки принимают тот факт, что скопления следует ожидать, не ограничивая возможности обманчивыми представлениями о том, как выглядит случайность.

Скопления онкологических заболеваний

Упомянутый выше принцип скопления (или кластеризации) применим к вещам гораздо более серьезным, чем настольные игры. Предположим, что девятнадцать гексов в «Поселенцах Катана» – это девятнадцать малых городов, в каждом из которых проживает 100 человек, и что у каждого человека есть 10%-ный шанс заболеть инвазивным раком в возрасте до шестидесяти лет независимо от того, где он живет.

Мы использовали компьютерный генератор случайных чисел, чтобы определить, заболеет ли каждый воображаемый человек раком в возрасте до шестидесяти лет или нет. На рис. 4.8 показаны результаты компьютерной симуляции. Результаты не были необычными. С учетом 100 человек в каждом городе и 10%-го шанса заболеть раком мы ожидаем, что в среднем у десяти человек в городе разовьется онкология. В нашей симуляции средний показатель по городу составляет ничем не примечательные 10,63, что, как и ожидалось, близко к десяти, но не совсем десять. Также неудивительно, что в одном городе онкология развилась у шестнадцати человек, а в другом – только у пяти. В девятнадцати городах вероятность того, что хотя бы в одном из них будет шестнадцать или более жертв рака, составляет 55 %, а вероятность того, что хотя бы в одном городе их будет пять или меньше, – 66 %. Как и в случае с подбрасыванием монеты и «Поселенцами Катана», кажущиеся необычными результаты не являются чем-то необычным.

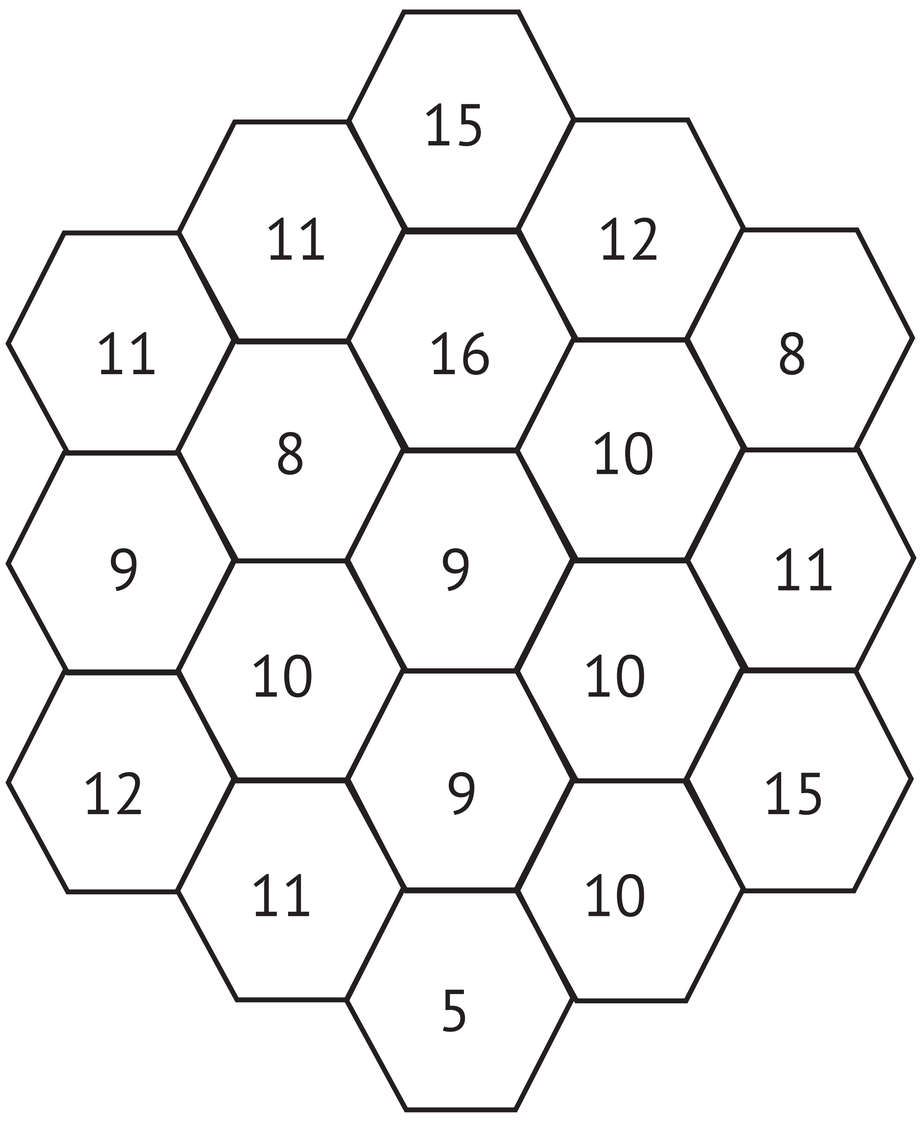

Рис. 4.8. Онкологическая заболеваемость в 19 малых городах

Если бы это были реальные города и мы не понимали, насколько сильно варьируются случайные данные, мы могли бы подумать, что города с шестнадцатью и пятью случаями рака выделяются, и мы могли бы попытаться найти объяснение. Мы могли бы также обратить внимание на тот факт, что в двух соседних городах в центре к северу на карте было пятнадцать и шестнадцать случаев, в то время как в центре к югу – только пять. Немного понаблюдав, мы могли бы обнаружить, что в северной части этого региона есть вышка сотовой связи, и прийти к выводу, что проживание рядом с этой вышкой вызывает рак, а проживание вдали от вышки снижает шансы развития рака.

Если бы городов было больше, то, скорее всего, результаты были бы еще более экстремальными – и все из-за капризной природы удачи. Например, из 1000 городов существует 92%-ный шанс того, что хотя бы в одном городе будет

двадцать или более жертв рака, и 80%-ный шанс того, что хотя бы в одном городе будет двое или меньше. Если бы мы увидели, что в одном городе зарегистрировано двадцать случаев, а в другом – только два, было бы заманчиво поискать объяснение – чем эти города отличаются друг от друга, – и мы наверняка нашли бы различия – возможно, в школах, парках, деревьях, водонапорных башнях или линиях электропередач, – которые кажутся важными, но на самом деле таковыми не являются.

В 1970-х годах появилось широко разрекламированное сообщение о том, что воздействие электромагнитных полей (ЭМП) от линий электропередач служат причиной онкологических заболеваний. Оно было основано на открытии эпидемиолога, сводящегося к тому, что некоторые дома, в которых жили люди, умершие от рака в возрасте до девятнадцати лет, находились вблизи линий электропередач. Реальность такова, что ученым известно об ЭМП очень многое, но нет убедительной теории о том, как ЭМП от линий электропередач могут служить причиной рака. Электромагнитная энергия от линий электропередач намного слабее, чем от лунного света, а магнитное поле слабее магнитного поля Земли. Несмотря на опровергающие это утверждение последующие исследования и эксперименты, многие люди по-прежнему считают, что линии электропередач служат причиной рака.

Когда зубная паста выжата из тюбика, ее трудно вставить обратно.

Чем меньше, тем лучше (и хуже)

Соотносящийся с этим статистический принцип заключается в том, что после разложения крупного набора данных на малые группы данных среди этих малых групп мы, скорее всего, обнаружим закономерности, которые будут казаться необычными. Вот почему наблюдаемые различия между малыми группами, как показано на рис. 4.8, нередко не имеют смысла.

Если сделать миллион подбрасываний монеты и подразделить их, скажем, на 100 000 групп по десять подбрасываний, то некоторые группы с десятью подбрасываниями, скорее всего, будут полностью орлами, а другие группы – решками. Было бы ошибкой думать, что эти скопления орлов и решек вызваны чем-то иным, кроме случайности, – думать, например, что для одних бросков использовалась монета с орлом на обеих сторонах, а для других – монета с решкой на обеих сторонах.

Возможность появления таких расхождений выше, если сравниваются малые группы, скажем группы из десяти человек, чем если бы сравнивались большие группы, скажем группы из 1000 человек, потому что в исходах десяти подбрасываний больше изменчивости, чем в исходах 1000 подбрасываний.

Этот принцип может быть применен ко многим ситуациям, исход в которых зависит от удачи. Например, даже если в малых городах нет ничего изначально хорошего или плохого, случайные исходы, скорее всего, будут экстремальными – хорошими либо плохими – именно в малых городах, чем в крупных. Это относится к успеваемости, статистике преступности, заболеваемости раком и многому другому. Выявление лучших и худших городов на самом деле может сводиться просто к выявлению самых малых городов.

Стандартизированные тесты

С 1998 по 2013 год согласно калифорнийской программе стандартизированного тестирования и отчетности (Standardized Testing and Reporting, аббр. STAR) требовалось, чтобы все учащиеся государственных школ со 2-го по 11-й классы ежегодно проходили общештатное тестирование с использованием стандартизированных тестов. Всем школам был присвоен балл индекса академической успеваемости (Academic Performance Index, аббр. API) с целью ранжирования школ по всему штату, а также с целью сравнения со 100 школами с аналогичными демографическими характеристиками. Баллы API были опубликованы в средствах массовой информации, размещены в интернете и представлены родителям в карточке школьной отчетности.

Баллы API варьировались от 200 до 1000, при этом для каждой школы был установлен целевой показатель в 800 баллов. Для любой школы с уровнем API ниже 800 был установлен одногодичный целевой показатель роста API, равный 5 % от разницы между ее API и 800. Таким образом, школа с API равным 600 имела целевой показатель роста API, равный 610. Целевым показателем школы с API выше 800 было поддержание своего API.

Балл API школы определялся по проценту учащихся в каждом из пяти квинтилей, формируемых общенациональными баллами за тесты. По-настоящему среднестатистическая школа, в которой 20 % учащихся входят в каждый квинтиль, имела бы показатель API 655, что значительно ниже целевого показателя штата в 800 баллов. Школа в Лейк-Уэбегоне, у которой все баллы выше среднего и равномерно распределены между 50-м и 99-м процентилями, получила бы показатель API, равный 890. (Ведущий радиопрограммы «Компаньон по дому в прериях» Гаррисон Кейллор описал вымышленный городок Лейк-Вубегон как место, где «все дети выше среднего уровня». Исследователи в области образования назвали эту невозможность «эффектом Лейк-Уэбегона», чтобы констатировать изъян в заявлениях о том, что все школы должны показывать результаты выше среднего.)

Мы собрали данные индекса API у 394 объединенных школьных округов, которые охватывают классы от детского сада до одиннадцатого класса. В среднем в объединенном школьном округе обучалось 7646 учащихся. Затем мы рассмотрели пять школьных округов с самыми высокими баллами API. Все эти наиболее успешные школьные округа были ниже среднего по численности и насчитывали в среднем 2592 ученика, что составляет лишь треть от среднего школьного округа. Глядя на эти высокорезультативные округа, можно подумать, что учащиеся в малых школьных округах успевают лучше, чем учащиеся в крупных школьных округах.

С другой стороны, если посмотреть на пять школьных округов с самыми низкими баллами API, то можно увидеть, что все они тоже были ниже среднего по численности, в среднем насчитывая всего 138 учащихся. Даже если отбросить школьные округа с численностью учащихся менее 100 человек, средняя численность пяти округов с самыми низкими баллами API составила всего 440 человек.

Так что же в итоге? Малые школьные округа лучше или хуже крупных школьных округов? Глядя на округа с самой высокой и самой низкой результативно-

стью, различить невозможно. Все это лишь подтверждает наше наблюдение о том, что между малыми школьными округами существует большая вариативность, поэтому у них, как правило, самые экстремальные результаты – самые высокие и самые низкие баллы.

В качестве общей меры можно было бы посмотреть на корреляцию между баллами API и размерами округов. Оказывается, что эта корреляция равна 0,01, т. е., по сути, равна нулю. Сосредоточение внимания на школьных округах с лучшей или худшей результативностью, безусловно, вводит в заблуждение.

Насильственные преступления

Считаете ли вы, что города США Сент-Луис, Детройт и Балтимор погрязли в преступности? Согласно официальной статистике ФБР в пятерку городов США с наибольшим количеством насильственных преступлений на душу населения входят города, о которых вы, вероятно, никогда не слышали: Индастри, Вернон, Тависток, Лейксайд и Тетерборо.

Причина, по которой вы мало слышали об этих пораженных преступностью городах, является той же причиной, по которой в них самый высокий уровень преступности – они очень маленькие, и в 2017 году их население составляло соответственно 204, 113, 5, 8 и 69 человек. Когда в городе проживает пять человек, одно преступление дает уровень преступности 20 000 преступлений на 100 000 жителей, что в десять раз превышает уровень преступности в Сент-Луисе, Детройте или Балтиморе.

В табл. 4.4 приведены уровни преступности в восьми городах – пять городов с самыми высокими уровнями преступности и три города, известных своими высокими уровнями преступности. Следует ли из этого сделать вывод, что малые города более опасны, чем крупные? Если сделать такой вывод, то снова будет упущен из виду статистический факт, что экстремальные значения обычно обнаруживаются в малых наборах данных, поскольку в малых наборах данных наблюдается более высокая вариабельность.

Таблица 4.4. Небезопасные города, 2017 год

Город	Количество насильственных преступлений	Численность населения	Преступлений на 100 000 жителей
Индастри, Калифорния	73	204	35 784
Вернон, Калифорния	35	113	30 973
Тависток, Нью-Джерси	1	5	20 000
Лейксайд, Колорадо	1	8	12 500
Тетерборо, Нью-Джерси	5	69	7246
Сент-Луис, Миссури	6461	310 284	2083
Детройт, Мичиган	13 796	670 792	2057
Балтимор, Мэриленд	12 430	613 217	2027

Давайте посмотрим на пять городов с самым низкими уровнями преступности. Неудивительно, что все они – малые. Более чем в 1000 городах вообще не было насильственных преступлений. В шести из этих свободных от преступности городов проживало менее 100 человек, а их средний размер был 1820 человек, что составляет одну двенадцатую от среднего размера городов США.

Являются ли малые города более опасными или менее опасными, чем крупные? Вероятно, ни то ни другое. Корреляция между размером города и уровнем насильственных преступлений составляет незначительные 0,003. Это всего лишь еще один пример того, что в малых наборах данных обычно больше вариации, чем в крупных.

Мы – первые (или, может быть, вторые)

В большинстве спортивных соревнований проводится финальный матч, игра или серия, которые определяют чемпиона. В футболе, крикете и регби каждые четыре года проводятся финальные турниры чемпионата мира. В Соединенных Штатах в футболе, баскетболе и бейсболе ежегодно проводятся Суперкубок, финалы НБА и Мировая серия. После этого игроки и болельщики празднуют победу команды-победителя, а проигравшей команде напоминают, что второе место – это всего лишь ступенька для проигравшего в первенстве.

Действительно ли по одной игре можно определить, какая команда является чемпионом, а какая – проигравшей в первенстве? Вспомните, как много вариации находится в малых выборках. Предположим, что в течение одного года Германия и Франция станут двумя лучшими футбольными командами в мире, и если они сыграют друг с другом 100 раз, то Германия выиграет шестьдесят семь раз и проиграет тридцать три раза. (Болельщики сборной Франции и других команд, пожалуйста, не присылайте нам свои жалобы, это чисто гипотетически.)

Суть в том, что если более сильная команда не выигрывает в 100 % случаев, то у более слабой команды есть шансы на победу. Здесь команда Германии явно является лучшей, и она продемонстрировала бы свое превосходство, если бы сыграла с командой Франции 100 раз. Однако, если Франция и Германия сыграют один матч в финале чемпионата мира, то у Франции будет треть шансов на победу. Если Франция все-таки выиграет, то болельщики будут настаивать на том, что Франция – это лучшая команда, даже несмотря на то, что одна игра является слишком малой выборкой, чтобы определиться с тем, какая команда лучше.

Стоит также помнить, что для попадания на чемпионат каждая команда должна пройти через предварительные матчи. Лучшая команда может даже не дойти до финала! Бразилия выиграла пять чемпионатов мира по футболу, но в 1982 году ее сборная, которая, возможно, была лучшей командой в истории, вылетела из четвертьфинала чемпионата мира, проиграв Италии со счетом 2:3. Польша в 1974 году, Франция в 1986-м и Англия в 1990-м – это всего лишь три команды в длинном списке великих сборных, которые не выходили в финал, не говоря уже о победе в чемпионате.

Как насчет таких видов спорта, как бейсбол, где две команды-финалисты играют до семи игр и первая команда, выигравшая четыре игры, становится

чемпионом? Серия из семи игр, безусловно, дает больше информации, но все еще далека от окончательной.

В Мировой серии по бейсболу 2019 года «Хьюстон Астрос» сыграли вничью с «Вашингтон Нэшнлз». «Нэшнлз» выиграли девяносто три из 162 игр в регулярном чемпионате (57 %), тогда как «Астрос» выиграли 107 игр (66 %). В состав «Астроса» вошел самый ценный игрок Американской лиги (MVP) Алекс Бригман, новичок года Джордан Альварес и двое из трех финалистов премии Сая Янга за лучшую подачу, Геррит Коул и Джастин Верландер (Коул стал победителем).

«Астрос» обыграли могущественную команду «Нью-Йорк Янкиз» (победителя 103 матчей в регулярном чемпионате) со счетом четыре игры к одной, чтобы попасть в Мировую серию, и были явными фаворитами турнира над «Нэшнлз». Коэффициенты ставок давали «Астрос» 68 % шансов на победу в Мировой серии, что является самым высоким показателем с 2007 года.

За день до начала Мировой серии Гэри написал:

> Кто выиграет Мировую серию? Не знаю, но я точно знаю, что бейсбол – это игра, квинтэссенцией которой является удача. Подачи на малой высоте по прямой траектории попадают прямо в игроков на поле. Неудачные подачи затухают во внутреннем поле. Подачи в воздух по дуговой траектории с трудом ловятся и по ним с трудом промахиваются. Подачи называются страйками, а страйки – подачами. Даже лучшие отбивающие делают в два раза больше аутов, чем отбиваний. Даже лучшие команды проигрывают более трети своих игр.
>
> В этом сезоне у «Хьюстон Астрос» был самый высокий процент побед (66 %) в бейсболе, однако они проиграли две игры из шести «Балтимору», который выиграл только треть своих игр – не потому, что команда «Балтимор» была лучшей, а потому, что в этих двух играх ей повезло больше.
>
> «Астрос» – одна из 10 лучших команд этого сезона (наряду с «Янкиз», «Тампа-Бэй», «Миннесотой», «Кливлендом», «Оклендом», «Атлантой», «Вашингтоном», «Сент-Луисом» и «Доджерс»), но кто выиграет серию из 7 матчей между двумя этими командами? Ваша догадка будет ничем не лучше моей – может быть, даже лучше, – но это все равно только догадка.

Гэри напомнил читателям о команде «Окленд Эйс» 1990 года, в которой наиболее ценным игроком был Рики Хендерсон, победителем приза Сая Янга – Боб Уэлч, а вторым призером – Дэйв Стюарт. Их подающий Деннис Экерсли имел показатель полезности питчера (ERA) 0,61, сделав 73 страйк-аута и пять проходов (одну намеренную) в 73 иннингах. У них также были Марк Макгвайр, Хосе Кансеко, Карни Лэнсфорд и полдюжины других звезд.

Выступая в самом сложном дивизионе на тот момент, в Западной американской лиге, они выиграли 103 матча в регулярном чемпионате, третий год подряд лидируя в лиге. Накануне Мировой серии известный спортивный журналист Томас Босуэлл писал в The Washington Post следующее: «Давайте скажем коротко и ясно. Бейсбольный сезон закончился. Никто не сможет победить команду “Окленд Эйс”».

«Эйс» проиграли в Мировой серии в четырех матчах подряд команде «Цинциннати Редз», которые выиграли всего девяносто одну игру в регулярном чемпионате. Обозреватель «Чикаго» Майк Ройко заметил, что это произошло

потому, что в составе «Эйс» было три бывших игрока «Чикаго Кабз». Нет, это произошло потому, что в серии из семи матчей между двумя хорошими командами может случиться все что угодно.

Как насчет «Нью-Йорк Метс» в 1969 году? Они вступили в лигу в 1962 году и проиграли рекордные 120 матчей. За первые шесть лет своего существования они в среднем одерживали 54 победы и 108 раз проигрывали, что составляет 33,3 % побед. В 1968 году они одержали 73 победы при 89 поражениях, а затем по удачной случайности одержали 100 побед в 1969 году. В Мировой серии они встречались с «Балтимор Ориолз», которые выиграли 109 матчей регулярного чемпионата с участием всех звезд, включая Фрэнка Робинсона, Брукса Робинсона и Буга Пауэлла. У них были питчеры Майк Куэльяр (23:11 при показателе полезности 2,38), Джим Палмер (16:4 при показателе полезности 2,34) и Дэйв Макнелли (20:7 при показателе полезности 3,22). У вышедших на замену Эдди Уотта, Пита Ричерта и Дика Холла были показателе полезности 1,65, 2,20 и 1,92.

Подпирая Куэльяра, «Ориолз», как и ожидалось, выиграли первую игру Мировой серии со счетом 4:1. Затем «Ориолз» проиграли следующие четыре игры. В течение следующих четырнадцати лет «Метс» оставались явно посредственной командой, выигравшей 46 % своих матчей, но в течение одного удивительного сезона они стали «Чудо-Метс».

Каковы шансы на то, что по всем объективным меркам лучшая команда проиграет Мировую серию? На удивление высоки. В такой игре, как бейсбол, где так важна удача, короткая серия из семи игр говорит очень мало.

Вероятность победы команды меняется от игры к игре в зависимости от стартового питчера и других факторов, но мы можем получить довольно хорошую оценку совокупной вероятности того, что команда выиграет серию из семи игр, просто допустив постоянную вероятность победы для каждой команды. Предположим, что у «Астрос» есть 60%-ный шанс выиграть любую игру против «Нэшнлз». Эта вероятность довольно щедра, поскольку «Астрос» выиграли 66 % своих матчей в регулярном чемпионате, играя против команд среднего уровня. Но мы будем использовать 60%-ную вероятность победы в игре против «Нэшнлз», чтобы дать шанс команде «Астрос», даже если есть основания для сомнений.

Даже при такой щедрой оценке команды «Астрос» получается, что существует 30%-ный шанс того, что «Нэшнлз» станут чемпионами мира. Если у «Астрос» есть более чем 55%-ный шанс выиграть каждую игру, то у «Нэшнлз» есть 40%-ный шанс открыть шампанское.

Мы также можем перевернуть этот вопрос и спросить, насколько сильно наша оценка команды «Астрос» будет зависеть от того, насколько хорошо она выступит в Мировой серии. Интернет-компании используют формулу под названием правила Байеса, чтобы оценивать и пересматривать вероятность того, что вам понравится тот или иной товар или услуга, на основе информации, которую они собирают о вас. Здесь можно сделать то же самое и пересмотреть нашу оценку вероятности победы команды «Астрос», равную 0,60, основываясь на том, насколько хорошо они выступают в Мировой серии.

Ответа на этот вопрос нет. Рекорд «Астрос» в регулярном чемпионате не гарантирует, что они выиграют серию из семи матчей против «Нэшнлз», а по-

беда или поражение в Мировой серии мало что говорит о том, насколько они хороши. Даже если «Астрос» проиграют четыре игры подряд, наша лучшая оценка вероятности того, что они обыграют «Нэшнлз» в другой игре, лишь незначительно снизится, с 0,600 до 0,594. Если «Астрос» выиграют четыре игры подряд, то их вероятность немного возрастет до 0,604. Все, что находится между, имеет еще меньший эффект. Независимо от того, как пройдет Мировая серия, это едва ли повлияет на нашу оценку команды «Астрос».

Такова природа игры. В такой игре, как бейсбол, где так много шансов, серия из семи игр полностью зависит от удачи. Лучшая команда нередко проигрывает, и победа в Мировой серии очень мало говорит о том, какая команда на самом деле лучше.

Гэри закончил свою статью:

> Наслаждайтесь серией; знаю, что так и будет. Но обратите внимание на удачу и вспомните «Метс» 1969 года и «Эйс» 1990-х.

Чем все закончилось? Первые две игры были сыграны в Хьюстоне, и «Нэшнлз» выиграли обе, что дало повод некоторым экспертам по очевидностям заявить, что они с самого начала знали, что команда «Вашингтон» – лучше. Истина в том, что исход двух игр мало о чем говорит. Затем серия перенеслась в Вашингтон, и «Астрос» выиграли следующие три матча. Теперь эксперты по очевидностям заявили, что «Астрос» снова в строю. Истина в том, что исход трех игр мало о чем говорит.

Серия вернулась в Хьюстон, и «Нэшнлз» выиграли обе игры и Мировую серию. Это была первая серия из семи матчей в истории любого крупного вида спорта, в которой каждая игра была выиграна командой гостей. Серия изобиловала большим количеством удачных моментов – сильно отбитые мячи пролетали достаточно близко к полевому игроку, чтобы быть пойманными, или достаточно далеко, чтобы не быть пойманными; судьи делали правильные и неправильные решения; были мелкие и серьезные травмы.

Если бы «Астрос» и «Нэшнлз» сыграли еще семь матчей подряд, кто бы победил? То немногое, что можно сказать с уверенностью, – это то, что победителем будут не «Янкиз». Еще одна очевидность заключается в том, что, когда две хорошие команды играют один или даже семь раз, этого недостаточно, чтобы выявить, какая команда лучше, – т. е. какая команда вышла бы на первое место, если бы они сыграли друг с другом 100 раз.

Наслаждайтесь матчами чемпионата, но не будьте так строги к проигравшему в первенстве.

Как не быть введенным в заблуждение призрачными закономерностями

Закономерности неизбежны. Серии, скопления и корреляции являются нормой, а не исключением. Кроме того, когда данные разделены на малые группы, мы не должны удивляться, обнаруживая существенную вариацию между группами.

Например, при большом количестве подбрасываний монеты, скорее всего, появятся случайно возникшие скопления орлов и решек. Если подбрасывания разделены на малые группы, то, скорее всего, между группами будут большие случайные различия.

В общенациональных данных о заболеваемости раком, преступности, баллах тестов или о чем-то еще, скорее всего, есть переменчивые скопления. Если данные подразделены на более мелкие географические единицы, такие как города, то между городами, скорее всего, будут иметься ярко выраженные различия, и самые экстремальные результаты, скорее всего, будут найдены в самых малых городах. В спортивных соревнованиях между мало-мальски хорошо подобранными командами исход нескольких игр практически не имеет смысла.

Главная трудность – преодоление унаследованной нами склонности думать, что все закономерности наполнены смыслом; например, считать, что скопления в больших наборах данных или различия между малыми наборами данных представляют нечто реальное, требующее объяснения. Нередко это просто бессмысленное стечение обстоятельств.

Глава 5

Парадокс больших данных

Вступление

Предприятие по ремонту компьютеров наняло авторитетную фирму по аналитике данных с целью создать программу на базе искусственного интеллекта (ИИ), которая сообщала бы техническим специалистам о том, какие детали им следует брать с собой на заявки выполнить техническое обслуживание. Информационно-аналитическая фирма оцифровала огромную базу данных телефонных записей запросов клиентов на обслуживание, а затем создала алгоритм, который позволял отыскивать слова, наиболее точно соотносимые с каждым видом ремонта.

Было проделано немало работы по конвертации звуков в текст, поскольку компьютеры все еще не в совершенстве распознают то, что мы пытаемся сказать, когда бормочем, запинаемся, неправильно произносим слова и говорим с акцентом, но была надежда, что огромное количество собранных ими текстовых данных даст полезные результаты.

И как все обернулось? Емкие слова информационно-аналитической фирмы были такими: «Мы с треском провалились».

Проблема заключалась в том, что многие слова, которые люди используют в телефонных разговорах, содержат очень мало полезной информации. Рассмотрим, например, вот этот фрагмент разговора:

Компания: Здравствуйте. ABC Electronics. Чем могу вам помочь?

Клиент: Здравствуйте! Это Джерри Гарсия. Компьютер моего сына сломался.

Компания: Что с ним не так?

Клиент: Да, я звонил вам некоторое время назад, может быть, два или три месяца назад, не знаю. Нет, знаете, думаю, это было где-то в июле. В любом случае вы нас спасли (saved), но, хм, это другое.

Компания: Вы можете сказать мне, что случилось?

Клиент: Он не может найти (find) кое-что из того, что сохранил (saved), что ему нужно для школы. Он знает, что оно где-то там, но не знает где.

[…]

Компания: Не могли бы вы сообщить нам свой адрес?

Клиент: Конечно, Уэст-Грин (Green), 127. Нас легко найти (find) – это примерно в квартале от методистской церкви на Эппл (Apple).

Компьютерный анализ приведенных выше слов представляет собой масштабную задачу. Слова могут быть неоднозначными и иметь несколько смыс-

лов. Слово «saved» в английском варианте появляется дважды, но относится к проблеме с компьютером только в одном из этих двух случаев. То же самое со словом «найти» (*find*). Слова «зеленый» (*green*) и «яблоко» (*apple*) могут быть связаны с проблемой, но это не так. Компьютерным алгоритмам требуется чертовски много времени, чтобы определить разницу между релевантной и нерелевантной информацией, потому что они буквально не знают, что слова означают.

Более крупная трудность связана с тем, что почти все слова в этом разговоре не имеют никакого отношения к проблеме, которую клиент хочет устранить. Учитывая тысячи слов, используемых в повседневной беседе, компьютерный алгоритм, скорее всего, обнаружит много случайно возникших взаимосвязей. Например, слово «легко» (*easy*) могло необычно часто использоваться во время телефонных звонков по поводу залипания клавиатуры, а слово «знать» (*know*) часто использоваться во время телефонных звонков по поводу неисправности принтера. Поскольку ни одно из этих слов не имеет никакого отношения к проблеме клиента, эти случайные совпадения совершенно бесполезны для выявления причин, по которым клиенты будут звонить в будущем.

Информационно-аналитическая фирма решила эту проблему путем привлечения человеческого опыта – технических специалистов, которые выезжают на вызов для устранения неполадок. Технические специалисты смогли выявлять ключевые слова, которые являются надежными индикаторами компьютерных проблем, требующих устранения, и компоненты, необходимые для устранения этих проблем.

Неуправляемый поиск закономерностей не удался, а вот компьютерные алгоритмы, которым помогал человеческий опыт, преуспели.

Этот пример вряд ли можно назвать аномалией. Было подсчитано, что 85 % всех осуществляемых предприятиями проектов по обработке больших данных проваливаются. Одной из причин этого провала является необоснованное убеждение в том, что все, что нужно предприятию сделать, – это позволить компьютерам выявлять закономерности.

Глубокая переработка данных

Научный метод начинается с фальсифицируемой теории, за которой следует сбор данных для статистической проверки этой теории. *Глубокая переработка данных* (data mining) идет в другом направлении, анализируя данные без привязки к теориям – более того, она рассматривает использование экспертных знаний как нежелательное препятствие, ограничивающее возможности для обнаружения новых знаний.

В статье 2008 года, озаглавленной «Конец теории: лавина данных делает научный метод устаревшим» (*The End of Theory: The data deluge makes the scientific method obsolete*), главный редактор журнала Wired утверждал, что:

> ...петабайты позволяют нам говорить: «Корреляции вполне достаточно». Мы можем перестать искать модели. Мы можем анализировать данные, не выдвигая гипотез о том, что они могут показать. Мы можем использовать данные в самых больших вычислительных кластерах, которые когда-либо видел мир, и поручать статистическим алгоритмам находить закономерности там, где наука не может.

В 2013 году два профессора в области теории и практики вычислительных систем описали глубокую переработку данных как поиск «скрытых закономерностей и секретных корреляций». Под «скрытыми закономерностями и секретными корреляциями» они подразумевают взаимосвязи, которые удивили бы экспертов. Они считают, что, для того чтобы находить вещи, которые экспертам незнакомы, мы должны давать компьютерам обрабатывать данные, не стесняя себя рамками ограниченных знаний экспертов.

Безудержный поиск закономерностей часто сочетает глубокую переработку данных с тем, что нобелевский лауреат Рональд Коуз (Ronald Coase) назвал «*издевательством над данными*» (data torturing): манипулированием, подрезанием или перестановкой данных до тех пор, пока не появится какая-либо закономерность. Некоторые исследователи, руководствуясь благими намерениями, считают, что если глубокая переработка данных может выявлять новые взаимосвязи, то мы не должны ограничивать ее возможности. Надо подразделять данные на подмножества, содержащие закономерности. Отбрасывать данные, которые не позволяют закономерности быть закономерностью. Комбинировать данные, если это создает закономерность. Устранять выбросы, которые нарушают закономерность. Включать выбросы, которые создают закономерность. И продолжать искать до тех пор, пока не будет найдено что-нибудь – что угодно.

В первых строках предисловия к книге, посвященной использованию глубокой переработке данных с целью выявления знаний, профессор в области теории и практики вычислительных систем без видимой иронии написал:

> «Если достаточно долго издеваться над данными, то Природа даст признательные показания», – сказал в 1991 году лауреат Нобелевской премии по экономике Рональд Коуз. Это утверждение по-прежнему справедливо. Однако достичь этой возвышенной цели непросто. Во-первых, во многих приложениях «достаточно долго» на практике может оказаться «слишком долгим» и, следовательно, неприемлемым. Во-вторых, для того чтобы получить «признание» от крупных наборов данных, необходимо использовать самые современные средства «издевательства». В-третьих, природа очень упряма – она не поддается легко либо вообще не желает раскрывать свои секреты.

Коуз задумал свой комментарий не как возвышенную цель, которую необходимо достичь с помощью самых современных инструментов издевательства над данными, а как резкую критику практики обшаривания данных в поисках закономерностей.

Например, после нарезки и разбивки данных компания вполне может обнаружить, что большинство ее лучших женщин-софтостроительниц предпочитают для хот-догов горчицу кетчупу. Однако, когда они дискриминируют соискательниц, которые любят кетчуп, в пользу любительниц горчицы, новые сотрудницы оказываются явно посредственными. Причина этого провала заключается в том, что предпочтение горчице со стороны лучших женщин-софтостроительниц было обнаружено при изучении сотни черт и сотни способов разделения и упорядочивания данных. Неизбежно существовали какие-то случайные корреляции с мастерством софтостроения. Когда корреляция возникает случайно, она исчезает при применении к новым сотрудникам. Она просто бесполезна.

В нашей предыдущей книге «9 подводных камней науки о данных» приведены многочисленные примеры неудачных результатов, которых люди из академических кругов и промышленности добивались, используя модели, добытые в результате глубокой переработки данных, без какой-либо глубинной теории. Один руководитель предприятия неоднократно выражал свое презрение к теории многозначительным комментарием: «Вверх значит вверх». Он считал, что, когда компьютер находит закономерность, логическая причина не требуется. Вверх значит вверх. К сожалению, часто оказывалось, что «вверх значило вниз» или «вверх было ничем», поскольку обнаруженные закономерности, которые должны были увеличивать выручку компании, на самом деле снижали выручку или вообще не имели никакого эффекта.

Заманчиво полагать, что доступность огромных объемов данных увеличивает возможность того, что глубокая переработка данных позволит обнаруживать новые, ранее неизвестные взаимосвязи. Однако реальность такова, что в крупных наборах данных случайные совпадения неизбежны, и чем больше набор данных, тем больше возможность наступления ситуации, когда наша находка является случайным совпадением. Парадокс больших данных заключается в нашей уверенности, что лавина данных поможет нам лучше понимать мир и принимать правильные решения, когда на самом деле чем больше данных мы собираем в поисках закономерностей, тем больше возможность наступления ситуации, когда наша находка будет вводить в заблуждение и окажется бесполезной.

Вневыборочные данные

Опасности глубокой переработки данных нередко проявляются, когда обнаруженная при перетряхивании данных закономерность исчезает при применении к новым данным. Поэтому, казалось бы, эффективным способом классифицирования статистической закономерности как наполненной смыслом или как бессмысленной является разбивка изначальных данных на две части – *внутривыборочные* данные, которые можно использовать для обнаружения моделей, и *вневыборочные* данные, которые откладываются в сторону, чтобы их можно было использовать для тестирования тех моделей, которые были обнаружены с использованием внутривыборочных данных. Если модель, обнаруженная с помощью половины данных, хорошо работает с другой половиной, то это свидетельствует о полезности модели. Указанная процедура разумна, но, к сожалению, не дает никаких гарантий.

Предположим, что мы пытаемся найти способ предсказать результаты футбольных матчей «Ливерпуля» в английской премьер-лиге и делим сезон 2018 года на первую половину (девятнадцать внутривыборочных игр) и вторую половину (девятнадцать вневыборочных игр). Если алгоритм глубокой переработки данных будет смотреть на данные о температуре в сотнях калифорнийских городов за день до матча «Ливерпуля», то он сможет обнаружить, что разница между высокими и низкими температурами в Клермонте, штат Калифорния, является хорошим предсказателем счета «Ливерпуля» в игре.

Если эта статистическая закономерность является чистым совпадением (а это, несомненно, так и есть), то тестирование взаимосвязи на вневыбороч-

ных данных, скорее всего, покажет ее бесполезность для предсказания счета «Ливерпуля» в игре.

Однако если это произойдет, то алгоритм глубокой переработки данных сможет продолжить поиск других закономерностей (в Калифорнии есть много городов и – при необходимости – других штатов) до тех пор, пока не найдет ту, которая позволяет делать успешные предсказания как с помощью внутривыборочных данных, так и с помощью вневыборочных данных, – и он действительно добьется успеха, если рассмотрит достаточно большое количество городов. Мнимые корреляции могут быть обнаружены в первых девятнадцати играх сезона премьер-лиги, и ровно так же они могут быть обнаружены во всех тридцати восьми играх.

Закономерность обычно считается статистически значимой, если шанс ее возникновения по чистой удаче составляет менее 5 %. Это означает, что если мы ошибаемся и сравниваем только группы случайных чисел, то 5 % сравниваемых нами групп будут статистически значимыми!

Пять процентов – это один из двадцати, поэтому мы ожидаем, что одна корреляция из каждых двадцати пройдет внутривыборочный тест, а одна корреляция из 400 пройдет как внутривыборочный, так и вневыборочный тесты. Целеустремленный исследователь, который анализирует 10 000 групп несоотнесенных данных, может рассчитывать найти двадцать пять статистически значимых корреляций как внутривыборочно, так и вневыборочно. В век больших данных существует *гораздо* больше, чем 10 000 наборов данных, которые можно проанализировать, и гораздо больше, чем двадцать пять мнимых корреляций, которые выдержат как внутривыборочный, так и вневыборочный тесты.

Вневыборочные тесты, безусловно, полезны, однако глубокая переработка данных с использованием вневыборочных данных по-прежнему остается глубокой переработкой данных и по-прежнему подвержена тем же ловушкам.

Вытеснение истинных переменных

С тотальной глубокой переработкой данных связана более тонкая проблема, которая усугубляется вневыборочными тестами. Предположим, что алгоритм глубокой переработкой данных используется для отбора предсказательных переменных из набора данных, который включает относительно малое количество «истинных» переменных, имеющих причинно-следственную связь с предсказываемой переменной, а также большое количество «помеховых» переменных, которые не зависят от предсказываемой переменной. Одна из проблем, как мы уже увидели, заключается в том, что некоторые помеховые переменные, скорее всего, по случайному совпадению будут успешными как внутривыборочно, так и вневыборочно, но затем с треском провалятся, когда модель выйдет в эфир с новыми данными.

Более крупная проблема состоит в том, что алгоритм глубокой переработки данных может отбирать помеховые переменные вместо истинных, которые были бы полезны для получения полезных предсказаний. Тестирование и ретестирование модели, добытой в результате глубокой переработки данных, может в конечном итоге выставить наружу помеховые переменные как

бесполезные, но это никогда не вернет истинные переменные, которые были вытеснены помеховыми переменными. Чем больше помеховых переменных было первоначально рассмотрено, тем больше шанс, что некоторые истинные переменные исчезнут без следа.

Множественная регрессия

В целях иллюстрации опасностей глубокой переработки данных мы провели несколько компьютерных симуляций методом Монте-Карло (названных так в честь мекки азартных игр), в которых для получения гипотетических данных использовался компьютерный генератор случайных чисел. Проведя очень большое количество симуляций, мы смогли выявить типичные и нетипичные исходы.

Одна из замечательных особенностей симуляций методом Монте-Карло заключается в том, что, поскольку данные были сгенерированы, мы знаем, какие переменные связаны причинной следственностью, а какие имеют корреляцию только по случайному совпадению – все этого для того, чтобы увидеть, насколько хорошо статистический анализ способен различить разницу.

В табл. 5.1 представлены результаты симуляций, в которых каждая кандидатная объясняющая переменная была помеховой переменной. Каждая переменная, отобранная алгоритмом глубокой переработки данных как полезная, на самом деле была бесполезной, однако глубокая переработка данных неукоснительно выявляла значительное количество переменных, которые сильно коррелировали с целевой переменной. Например, из 100 возможных переменных алгоритм глубокой переработки данных отобрал в среднем 6,63 бесполезных переменных для составления предсказаний.

Таблица 5.1. Симуляции без истинных переменных

Количество кандидатных переменных	Среднее количество отобранных переменных	Внутривыборочная корреляция	Вневыборочная корреляция
5	1,11	0,244	0,000
10	1,25	0,258	0,000
50	3,05	0,385	0,000
100	6,63	0,549	0,000
500	97,79	1,000	0,000

В табл. 5.1 также показано, что с увеличением числа кандидатных объясняющих переменных увеличивается и среднее число отобранных помеховых переменных. Независимо от того, насколько сильно эти переменные коррелируют с целевой переменной, они совершенно бесполезны для будущих предсказаний. Вневыборочные корреляции неизбежно в среднем равны нулю.

Мы провели еще один набор симуляций, в которых пять переменных (истинных переменных) использовались для определения значения целевой переменной. Например, у нас может быть 100 кандидатных переменных, из которых пять определяют целевую переменную, а девяносто пять являются *по-*

меховыми переменными. Если бы глубокая переработка данных сработала, то пять наполненных смысла переменных всегда были бы в наборе переменных, отобранных алгоритмом глубокой переработки данных, а помеховые переменные всегда исключались бы.

Результаты представлены в табл. 5.2. Включение пяти истинных переменных не устранило отбор помеховых переменных; оно просто увеличило количество отобранных переменных. Чем больше число кандидатных переменных, тем больше помеховых переменных включается в набор и тем хуже вневыборочные предсказания. Это является эмпирическим свидетельством в пользу парадокса больших данных:

> Казалось бы, наличие данных для крупного числа переменных поможет нам отыскивать более надежные закономерности; однако чем в большем количестве переменных мы роемся в поисках закономерностей, тем меньше шанс, что наши находки будут полезными.

Таблица 5.2. Симуляции с пятью истинными переменными

Количество кандидатных переменных	Среднее количество отобранных переменных	Внутривыборочная корреляция	Вневыборочная корреляция
5	4,50	0,657	0,606
10	4,74	0,663	0,600
50	6,99	0,714	0,543
100	10,71	0,780	0,478
500	97,84	1,000	0,266

Также обратите внимание, что при рассмотрении меньшего количества помеховых переменных отбирается меньшее количество помеховых переменных. Вместо того чтобы применять алгоритм глубокой переработки данных к сотням, тысячам или даже сотням тысяч неотфильтрованных переменных, было бы лучше использовать человеческий опыт, чтобы исключать как можно больше помеховых переменных. Вот следствие парадокса больших данных:

> Чем больше количество возможных объясняющих переменных, тем важнее опыт человека.

Эти симуляции также документируют то, как избыток помеховых переменных может вытеснять истинные переменные. Например, при 100 кандидатных переменных в 50 % случаев одна или несколько истинных переменных были вытеснены, а в 16 % случаев – две или более переменных. Были даже случаи, когда все пять истинных переменных были вытеснены.

Суть проста. Переменные, обнаруженные с помощью глубокой переработки данных, могут казаться полезными, даже если они не имеют отношения к делу, а истинные переменные могут быть обойдены вниманием и отброшены, даже если они полезны. Оба недостатка подрывают перспективность глубокой переработки данных.

Невежи

Страховая компания создала огромную базу данных с записями каждого телефонного, факсимильного или электронного взаимодействия со своими клиентами. Всякий раз, когда клиент обращался в компанию, сотрудник исправно ставил галочки и заполнял пустые поля, чтобы создать постоянную цифровую запись о контакте. До сих пор эта база данных использовалась только для выставления счетов и других канцелярских целей, но теперь компания надеялась, что информационно-аналитическая фирма сможет всесторонне переработать огромное количество собранных данных и определить надежные предсказатели покупки страхового полиса обратившимся в фирму человеком.

Информационно-аналитическая фирма сделала то, что ей было сказано, и создала алгоритм, который рылся в данных. Поскольку компьютеры не знают смысла слов, алгоритм не учитывал смысл каких-либо полей или пробелов. Его задача, с которой он справился очень хорошо, состояла в том, чтобы просто отыскивать закономерности в полях и пробелах, которые можно было бы использовать для составления надежных предсказаний покупок страховых полисов.

Алгоритм оказался невероятно успешным, или, может, стоит сказать *подозрительно* успешным? Он смог со 100%-ной точностью предсказывать, купит ли обратившийся в фирму человек страховку или нет. Если бы он не был абсолютно успешным, то его, возможно, приняли бы и использовали, не задумываясь, но 100%-ное совершенство вызвало подозрения даже у информационно-аналитической фирмы.

Потребовалось немало детективной работы, чтобы разгадать тайну этого успеха. Все зависело от одного поля в базе данных, которое не было четко обозначено, но имело три возможных варианта ответа: телефон, факс или электронная почта. Если какой-либо из этих ответов был положительным, то клиент приобретал страховку.

Информационно-аналитической компании пришлось провести обратный инжиниринг базы данных, чтобы выяснить последовательность событий, которые приводили к появлению этого поля. Оказалось, что, когда клиенты отменяли страховку, компания использовала это поле для регистрации того, что запрос на отмену отправлялся по телефону, факсу или электронной почте. Для того чтобы аннулировать страховку, клиент должен был приобрести страховку, поэтому между этой переменной и фактом покупки страховки было 100%-ное соответствие. Эта закономерность была на 100 % бесполезна.

Матери-иммигрантки и их дочери

Говорят, что Соединенные Штаты – это страна возможностей, где люди могут делать в своей жизни то, что хотят, где сын иммигрантов с Ямайки может стать государственным секретарем, а ребенок из бедной семьи из Хоупа, штат Арканзас, или сын кенийского студента-стипендиата, могут стать президентом.

Однако многие люди считают, что в США практически отсутствует экономическая мобильность. Майкл Харрингтон (Michael Harrington) однажды на-

писал, что «реальное объяснение того, почему бедные находятся там, где они есть, заключается в том, что они совершили ошибку, родившись не у тех родителей». Судьба определяется при рождении. В статье журнала Business Week 2003 года, озаглавленной «Пробуждение от американской мечты» (Waking Up from the American Dream), и в эссе нобелевского лауреата Пола Кругмана (Paul Krugman) 2004 года, озаглавленном «Смерть Горацио Элджера» (The Death of Horatio Alger), утверждалось, что экономическая мобильность в США – это миф. Бедные семьи оказываются в ловушке в бедных кварталах с распавшимися семьями, плохими школами и культурой низких устремлений. И вырваться из этого порочного круга бедности и страданий очень трудно.

Экономическая мобильность особенно важна для иммигрантов, которые часто несут огромные финансовые и социальные издержки, поскольку надеются улучшить жизнь для себя и своих детей. Являются ли дети иммигрантов в США более обеспеченными, чем их родители-иммигранты?

По версии журналов Business Week и The Wall Street Journal, профессор Джордж Хесус Борхас (George Jesus Borjas) из Гарвардской школы государственного управления имени Джона Ф. Кеннеди был назван ведущим американским иммиграционным экономистом. Борхас утверждает, что в прошлом «в течение одного-двух десятилетий после прибытия иммигрантов их доходы превышали доходы местных жителей с сопоставимым социально-экономическим статусом… дети иммигрантов были даже успешнее своих родителей». Однако Борхас приходит к выводу, что это уже не так: «Недавно прибывшие, вероятно, на протяжении большей части своей трудовой жизни зарабатывают на 20 % меньше, чем коренные жители». Один из его центральных аргументов заключается в том, что этнические кварталы обладают культурными и социально-экономическими особенностями, которые ограничивают межпоколенческую мобильность иммигрантов.

Исследования межпоколенческой мобильности иммигрантов, в том числе исследования Борхаса, как правило, сравнивают средний доход мужчин-иммигрантов в первом поколении в ходе одной переписи населения США со средним доходом мужчин во втором поколении в ходе более поздней переписи. Каждая перепись содержит огромное количество данных, но лучше немного хороших данных, чем много плохих. Сравнение мужчин в первом поколении в ходе одной переписи со средним доходом мужчин во втором поколении в ходе более поздней переписи выявило несколько проблем.

Во-первых, пренебрежение к женщинам вызывает сожаление, поскольку дочери и сыновья могут вести совершенно разный образ жизни. В фильме «Гран Торино» дочь иммигрантов-хмонгов говорит, что «женщины поступают в колледж, а мужчины отправляются в тюрьму». Если это правда, то основанные на опыте сыновей меры по восходящей мобильности могут быть неприменимы к дочерям.

Во-вторых, в этих исследованиях рассматривается индивидуальный доход, но доход домохозяйства, возможно, более актуален, поскольку одним из важных аспектов мобильности является степень, в которой люди вступают в брак с людьми из разных социально-экономических слоев.

В-третьих, в бедных семьях, как правило, больше детей, и это делает средние показатели по поколениям дезориентирующими. Предположим, что

в 2000 году есть два мужчины-иммигранта, один из которых работает инженером-софтостроителем и зарабатывает 380 000 долл., а другой – садовником и зарабатывает 20 000 долл. В 2020 году у первого иммигранта есть один сын, который зарабатывает 760 000 долл.; у второго иммигранта четыре сына, каждый из которых зарабатывает по 40 000 долл. Каждый сын зарабатывает на 100 % больше своего отца, но средний доход второго поколения на 16 % меньше среднего дохода первого поколения.

Вместо того чтобы сравнивать средний показатель по первому поколению со средним показателем по второму поколению, следует сравнивать родителей с их детьми. Разумным индикатором дохода семьи является район, в котором живут люди. Дома в безопасных и чистых районах с привлекательными удобствами (например, хорошие государственные школы) являются более желанными и, следовательно, более дорогими. Поскольку переезд из дома в дом обходится дорого, район проживания домохозяйства бывает хорошим индикатором его экономического статуса.

Гэри получил специальный доступ к записям о рождениях в Калифорнии, которые могут быть использованы для установления связи между матерями-иммигрантками и их дочерьми на сопоставимых этапах их жизни (одинаковый возраст и количество детей) и для выявления их районов проживания. К сожалению, сопоставимых данных по отцам-иммигрантам и их сыновьям нет.

Примерно 85 % взрослых дочерей матерей, родившихся за границей, живут по другим почтовым индексам, чем их матери, и большинство меняющих почтовые индексы дочерей переезжают в более богатые почтовые индексы. По сравнению с дочерьми белых женщин, родившихся в Калифорнии, дочери женщин-иммигранток обладают равной географической мобильностью и большей экономической мобильностью. Разрыв между экономическим статусом дочерей матерей, родившихся за границей, и дочерей матерей, родившихся в Калифорнии, составляет менее половины разрыва между экономическим статусом их матерей.

Нас не должен прельщать объем данных. *Релевантные* данные лучше, чем их *обилие*.

Пора с этим завязывать

На многие интересные вопросы можно ответить, только собрав большое количество данных, наблюдая за поведением людей в интернете. В идеале данные должны быть получены в ходе рандомизированных контролируемых испытаний, таких как A/B-тесты различных дизайнов веб-страниц. Когда это невозможно (а обычно это не так), лучше всего начинать с четкого понимания того, чего вы пытаетесь достичь, а затем найти надежные данные, которые хорошо подходят для этой цели. Перелопачивание огромного количества данных в поисках интересных закономерностей, скорее всего, закончится плачевно.

Вместе с двумя студентами Гэри работал над исследованием, которое имело четкую цель и было сосредоточено на данных, которые могли бы послужить этой цели. Один из студентов проводил бóльшую часть своего времени в колледже, играя в онлайновый покер. Это не было пустой тратой времени, потому что он заработал немного денег и приобрел ценные навыки критического

мышления. Позже он защитил докторскую диссертацию по экономике и теперь работает в американской исследовательской организации Mathematica Policy Research (ныне Mathematica Inc.). Другой студент тоже защитил докторскую диссертацию по экономике и теперь работает в Федеральной резервной системе.

Когда они были студентами в Помоне, они написали совместную дипломную работу (Гэри был сначала научным руководителем, а затем соавтором), используя данные об онлайновом покере, чтобы ответить на интересный исследовательский вопрос. Дэниел Канеман (Daniel Kahneman) и Амос Тверски (Amos Tversky) однажды заявили, что «человек, который не смирился со своими проигрышами, скорее всего, пойдет на риск, который в противном случае был бы для него неприемлем». Игравший в покер студент был лично заинтересован в том, чтобы выяснить, не соблазняются ли игроки в покер, терпящие крупные убытки, делать чрезмерно рискованные ставки в опрометчивой попытке отыграть то, что они потеряли. Он мог бы использовать это понимание для корректировки своей игры против крупных неудачников. Изучение этого вопроса также стало бы проверкой заявления Канемана и Тверски.

В популярной покерной игре «Техасский холдем» (Texas Hold ‘Em) со ставками от 25 до 50 долл. каждая раздача начинается с того, что игрок, сидящий непосредственно слева от дилера, делает малую принудительную ставку (блайнд) в размере 25 долл. в банк, а игрок, сидящий на два места слева от дилера, делает большую принудительную ставку в размере 50 долл. Затем каждому игроку раздаются две «закрытые карты», в которые могут заглядывать только они сами. Игроки, которые еще не внесли деньги в банк, решают, играть им либо сбросить карты. Для того чтобы начать игру, игроки должны либо «согласиться» с большой принудительной ставкой (50 долл.), либо поднять ставку выше 50 долл., вынудив других игроков сделать максимальную ставку или сбросить карты.

После этого начального раунда ставок раздаются три общие карты (флоп), которые видны всем и объединяются с двумя закрытыми картами каждого игрока, образуя комбинацию из пяти карт. Затем следует еще один раунд ставок, за которым следует четвертая общая карта (тёрн), дополнительные ставки, пятая общая карта (ривер) и заключительный раунд ставок. Оставшийся игрок с наилучшей комбинацией из пяти карт, которую можно собрать из двух закрытых карт и пяти общих карт, выигрывает банк.

Гэри и другие студенты не могли тратить время на серьезные игры в покер, подглядывая в карты игроков и регистрируя то, как они играют. По юридическим и финансовым причинам они не могли проводить эксперименты с реальными игроками и на реальные деньги. Однако игравший в покер студент из Помоны был хорошо знаком с онлайновым источниками данных. Full Tilt Poker (Покер во весь опор) – это онлайновая покерная комната, запущенная в июне 2004 года при участии сборной профессионалов покера. Деятельность компании и ее веб-сайта регулировалась Комиссией по азартным играм Канаваке, расположенной на территории мохавков в Канаде. Поскольку веб-сайт находился за пределами юрисдикции США, ему удалось избежать нормативных положений и налогов США (правда, в 2011 году все это изменилось после завершения данного исследования).

Веб-сайт Full Tilt Poker понравился Гэри и двум студентам тем, что игроки рисковали реальными деньгами, а во многих играх были большие принудительные ставки, что отсеивало новичков. На самом деле Full Tilt Poker мог похвастаться самым большим количеством профессиональных игроков в онлайновом покере в мире. Решающим моментом стало то, что компьютерная программа под названием PokerTracker позволила Гэри и студентам регистрировать каждую онлайновую игру – каждую сданную карту, сумму каждой ставки и каждый исход. С января по май 2008 года они регистрировали данные 24 часа в сутки по всем столам с принудительными ставками 25/50 долл., считающимся столами с высокими ставками и привлекающим опытных игроков в покер. В итоге они получили данные о том, как сотни тысяч знающих людей играли от карты к карте, ставя серьезные деньги. Указанные онлайновые данные были буквально самым лучшим способом проверить теорию о том, что игроки меняют свой стиль игры после крупных проигрышей.

Игравший в покер студент знал, что существует общепринятая мера *расшатанности* (looseness): процент раздач, в которых игрок непринудительно делает ставку, чтобы увидеть карты на флопе. Сюда может входить согласие со ставкой либо поднятие ставки, но не входят слепые ставки (блайнды), поскольку они являются принудительными.

Собранные (тайтовые) игроки сбрасывают карты, когда их две закрытые карты оказываются слабыми; расшатанные (лузовые) игроки остаются в игре, надеясь, что удачный флоп усилит их комбинацию. За столами с шестью игроками игроки, как правило, считаются очень собранными, если их проигрыш составляет менее 20 %, и крайне расшатанными, если их проигрыш превышает 50 %. По их данным, средний уровень расшатанности за полными столами с шестью игроками составлял 26 %.

В теоретическом плане у опытных игроков в покер есть свой стиль игры, который, по их мнению, им подходит и который основан на сотнях тысяч сыгранных раздач. Определившись со своей оптимальной стратегией, они должны ее придерживаться независимо от исхода последних нескольких раздач. Если они терпят крупные убытки, то им следует признавать, что им не повезло, и придерживаться своей стратегии. Вопрос исследования заключался в том, становятся ли терпящие крупные убытки игроки менее осторожными и разыгрывают раздачи, которые они обычно сбрасывают, вкладывая деньги в банк, чтобы увидеть флоп.

Гэри и студенты трактовали раздачу, в которой игрок выиграл или проиграл 1000 долл., значительным выигрышем или проигрышем. После крупного выигрыша или проигрыша они наблюдали за поведением игрока в течение следующих двенадцати раздач – двух циклов за столом на шесть игроков. Их данные включали 346 игроков, которые соответствовали различным критериям, при среднем количестве сыгранных раздач, равном 1738. Половина игроков выиграла или проиграла более 200 000 долл., 10 % выиграли или проиграли более 1 млн долл. Как мы уже говорили, это реальные игроки, ставящие серьезные деньги.

Данные касались столов на шесть игроков и столов для противостояний один на один (хедз-ап), за которыми могут играть только два игрока. Иногда за столами для шести игроков были свободные места, и это влияло на стратегию

игроков. Например, шансы, что пара восьмерок в закрытых картах принесет наилучшую комбинацию, уменьшается по мере увеличения числа игроков. Поэтому Гэри и студенты сгруппировали данные в соответствии с количеством игроков за столом. Они не объединяли данные столов для противостояний один на один с данными столов на шесть игроков, за которыми играют только два игрока, потому что у людей, которые предпочитают играть в покер один на один, могут быть другие стили игры, чем у игроков, которые выбирают стол на шесть игроков, но иногда имеют четыре свободных места.

Гэри и студенты обнаружили, что игроки действительно обычно меняли свой стиль игры после выигрыша или проигрыша крупного банка – в частности, они обычно начинали играть менее осмотрительно после крупного проигрыша, очевидно, надеясь на удачные карты, которые быстро сотрут их проигрыш.

Из табл. 5.3 видно, что обычно после крупного проигрыша игроки неизменно ведут себя расшатаннее, чем после крупного выигрыша. Например, при шести игроках за столом 135 игроков после крупного проигрыша играли расшатаннее, чем после крупного выигрыша, тогда как обратное было верно только для 68 игроков. Для того чтобы проверить правильность этого вывода, Гэри и студенты также проанализировали пороговые значения в 250 и 500 долл. для крупного выигрыша или проигрыша и обнаружили, что в любом случае большинство игроков после крупного проигрыша играют расшатаннее. Они также обнаружили, что крупные проигрыши, по-видимому, запоминаются больше, поскольку доля игроков, которые играли расшатаннее после проигрыша, увеличивалась с увеличением размера проигрыша.

Таблица 5.3. Игроки играли расшатаннее после крупного проигрыша больше, чем после крупного выигрыша

Игроков за столом	Количество игроков	Средняя расшатанность	Игроки, которые были расшатаннее после большого выигрыша	Игроки, которые были расшатаннее после большого проигрыша
один на один	228	51	74	154
2	40	46	17	23
3	33	35	11	22
4	75	29	21	54
5	150	26	53	97
6	203	26	68	135

Было ли это изменение стратегии прибыльным? Если опытные игроки с самого начала используют прибыльные стратегии, то менять их будет ошибкой. Это именно то, что они обнаружили. Те игроки, которые после крупного проигрыша играли расшатаннее, были менее прибыльными, чем обычно.

Поскольку у этих исследователей была четко поставленная цель, они смогли сосредоточиться на соответствующих данных и прийти к разумному выводу, который вполне может быть применим далеко за пределами игры в покер. Мы все должны стараться не позволять неудачам подталкивать нас к принятию безрассудных решений. Мы не должны делать рискованных инвестиций из-за

того, что потеряли деньги на рынке акций. Мы не должны заменять подчиненного после одного слабого решения. Мы не должны бросаться в пагубные отношения после разрыва.

Предсказание обменного курса

Гэри однажды подшутил над научным сотрудником (Алексом), который настаивал на том, что данные говорят сами за себя: «Кому нужна причинная следственность, когда есть корреляция?» Алекс был исключительно самоуверен. Он сказал Гэри, что хотел бы, чтобы после его смерти его гроб несли все, кто когда-либо работал с ним над групповыми проектами, и, опустив его в землю, попрощались с ним в последний раз.

Ничто из сказанного Гэри не разубедило Алекса в его непоколебимой вере в то, что для получения надежных предсказаний нужна лишь корреляция. Алекс утверждал, что нам не нужно знать причины, по которым две вещи связаны, и мы не должны тратить время на их выяснение.

Наконец, Гэри предложил пари. Он предоставил Алексу ежедневные данные за десять лет по целевой переменной (Алексу это было неизвестно, это был обменный курс между турецкой лирой и долларом США) и данные по 100 другим переменным, которые можно было бы использовать для предсказания целевой переменной. Для того чтобы гарантировать, что Алекс не будет использовать экспертные знания при построении своей модели, ни одна из переменных не была помечена. Данные были поделены на внутривыборочные данные за пять лет и вневыборочные данные за пять лет, и Алекс мог по своему усмотрению использовать любой алгоритм глубокой переработки данных, чтобы увидеть, сможет ли он найти модель, которая хорошо работала бы как внутривыборочно, так и вневыборочно.

Единственной ставкой была гордость.

Гэри предоставил Алексу реальные данные за десять лет об обменном курсе лиры к доллару, но остальные переменные были просто случайными числами, созданными Гэри. Алекс построил буквально миллионы моделей (75 287 520, если быть точным) на основе данных за первые пять лет, а затем протестировал эти модели на данных за вторые пять лет. Его наилучшая модель для внутривыборочных данных показала корреляцию, равную 0,89 между предсказанными и фактическими значениями целевой переменной, но корреляция для вневыборочных данных составила –0,71! Между предсказанными и фактическими значениями наблюдалась сильная положительная внутривыборочная корреляция и сильная отрицательная вневыборочная корреляция.

Алекс был на время ошеломлен тем, что модель, которая показала такую хорошую результативность внутривыборочно, может показывать такую плохую результативность вневыборочно. Однако у него было много данных, поэтому он продолжил глубокую переработку данных и смог найти несколько моделей, которые показали исключительно хорощщю результативность как внутривыборочно, так и вневыборочно. На рис. 5.1 приведены внутривыборочные и вневыборочные корреляции для всех 3016 моделей, у которых внутривыборочная корреляция превышала 0,85. Из них 248 моделей име-

ли вневыборочные корреляции выше 0,80, а 92 модели имели корреляции выше 0,85. Многие модели имели более высокие корреляции вневыборочно, чем внутривыборочно!

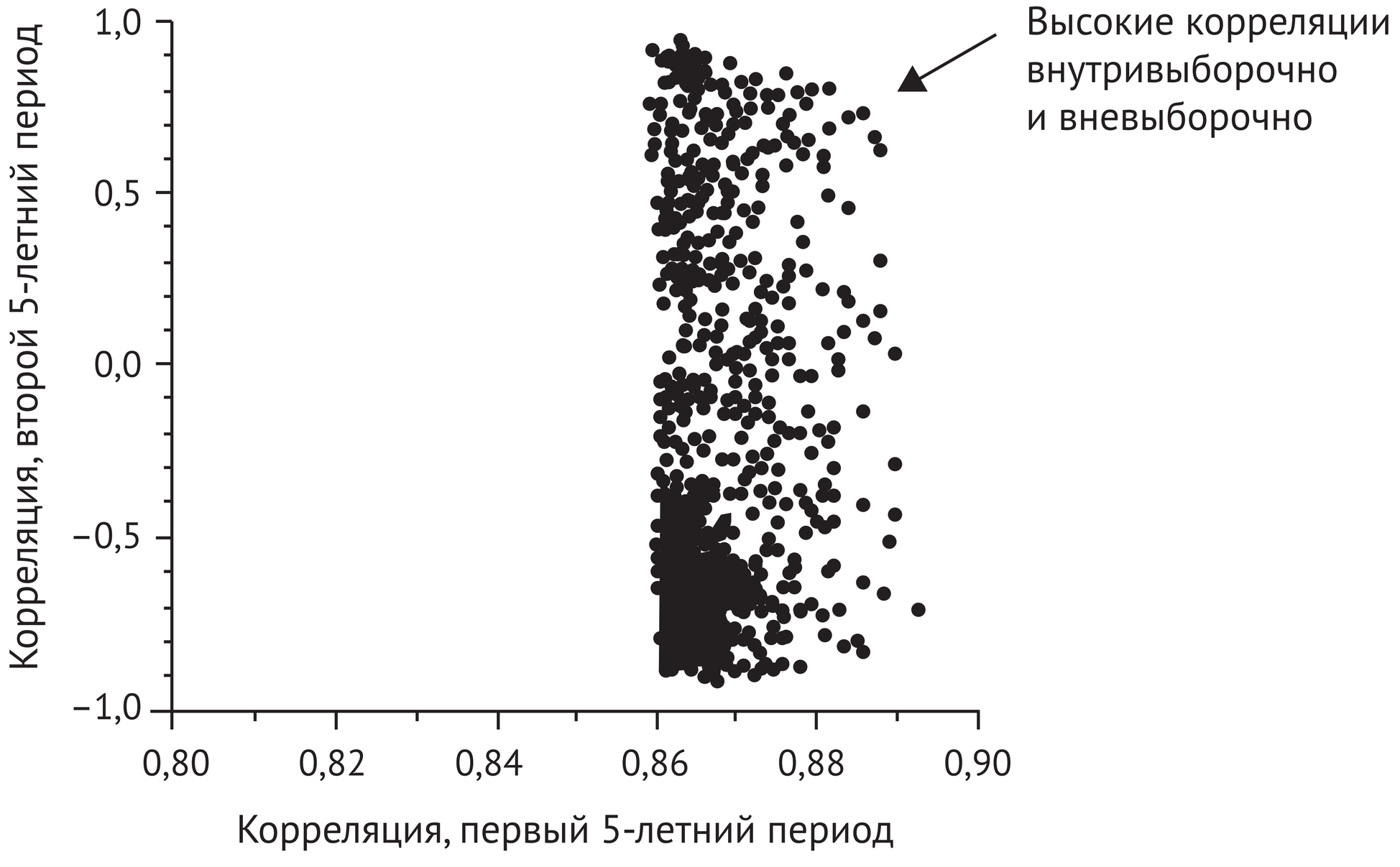

Рис. 5.1. Внутривыборочные и вневыборочные корреляции

Алекс решил использовать модель, у которой была самая высокая вневыборочная корреляция. В этой модели использовались переменные 2, 7, 41, 53, 56 для предсказания целевой переменной и имелась внутривыборочная корреляция 0,87 и вневыборочная корреляция 0,92.

Он с гордостью показал эту модель Гэри, уверенный, что выиграл пари. Однако у Гэри был припрятан секрет. На самом деле у него были данные за пятнадцать лет, с 2003 по 2017 год. Гэри предвидел, что Алекс подвергнет данные глубокой переработке и тестированию, повторяя все до тех пор, пока не отыщет модель, которая будет показывать хорошую результативность как внутривыборочно, так и вневыборочно, поэтому он предоставил Алексу данные за 2003–2007 годы и за 2008–2012 годы и зарезервировал данные за оставшиеся пять лет (2013–1017) для честного испытания модели Алекса. На рис. 5.2 показано, что она с треском провалилась. Корреляция между предсказанным и фактическим значением обменного курса составила ничтожные 0,08. Данные не говорят сами за себя.

Вторым сюрпризом со стороны Гэри стало его открытие, что все 100 возможных объясняющих переменных были не экономическими данными, а случайными числами. Несмотря на то что все показанные на рис. 5.1 внутривыборочные корреляции были выше 0,85 в течение первого пятилетнего периода (2003–2007), средняя корреляция во втором пятилетнем периоде (2008–2012) составила 0,002, т. е. практически была равна нулю. В этом есть толк, поскольку предсказательные переменные были в буквальном смысле случайными

данными, и независимо от того, насколько хорошо любая из них справлялась с внутривыборочными данными, они ожидаемо в среднем будут бесполезны в невыборочно.

Рис. 5.2. Внутривыборочно, вневыборочно и дополнительно вневыборочно

По чистой удаче 248 моделей показали хорошую результативность за оба пятилетних периода. Однако, когда эти 248 моделей, построенных на основе внутривыборочных данных и валидированных на основе вневыборочных данных, были использованы на свежих данных (2013–2017), средняя корреляция снова оказалась практически нулевой.

Гэри извлек два желаемых урока. Во-первых, если будет рассмотрено достаточное количество моделей, то некоторые из них обязательно окажутся полезными независимо от того, насколько (в буквальном смысле) случайными являются предсказательные переменные. Во-вторых, предсказательный успех внутривыборочно и вневыборочно не гарантирует успеха на новых данных.

Глубокая переработка твитов Трампа

В сентябре 2019 года в Twitter было 330 млн активных пользователей и в среднем 500 млн твитов в день, что составляет почти 200 млрд в год. Какой замечательный источник данных, который можно обшаривать и вышелушивать алгоритмами глубокой переработки данных!

Гэри и один из бывших студентов Гэри Джон Зук (John Zuk) решили сосредоточиться на твитах, отправленных Дональдом Трампом. У него в Twitter 66 млн подписчиков, и он занимает самый влиятельный пост в мире, так что, возмож-

но, его твиты имеют реальные последствия. Трамп, безусловно, так думает. Он хвастался своей «великой и непревзойденной мудростью» и описывал себя как «прекрасно выглядящего и умного, настоящего стабильного гения»! Он делится своей мудростью, отправляя в среднем девять твитов в день.

Гэри и Джон не были первыми, кто начал собирать данные о твитах Трампа. Исследование банка Bank of America показало, что рынок акций работает лучше в те дни, когда Трамп пишет меньше. Исследование банка JP Morgan показало, что твиты, содержащие слова «Китай», «миллиард», «продукты», «демократы» и «великий», оказывают статистически значимое влияние на процентные ставки.

Признавая важность репликации, Гэри и Джон разделили три года, прошедшие после победы Трампа на выборах 8 ноября 2016 года, на первые два года (которые использовались для обнаружения знаний) и третий год (который использовался для проверки, что первоначальные результаты подтверждаются вне выборки). За этот период Трамп отправил около 10 000 твитов, содержащих примерно 15 000 уникальных слов.

Ежедневные флуктуации в словоупотреблении Трампа коррелировались с промышленным индексом Доу Джонса через один-пять дней. Многие из этих корреляций не подтвердились вне выборки, но некоторые из них подтвердились. Например, было предсказано, что увеличение использования Трампом слова «спасибо» на одно стандартное отклонение приводит к увеличению среднего индекса Доу Джонса через четыре дня на 377 пунктов. Шанс, что корреляция будет столь же сильной, как та, которая была обнаружена, составлял менее одного на тысячу. Более того, корреляция во вневыборочный период была даже сильнее, чем во внутривыборочный период!

Воодушевленные этим открытием, Гэри и Джон обратились к другим источникам. Трамп давно восхищается президентом России Владимиром Путиным. В июне 2019 года Трамп сказал: «Он отличный парень. <…> Он потрясающий человек». В другой раз он описал их взаимное восхищение: «Путин говорил обо мне хорошие вещи. Он сказал: “Он – лидер, и в этом нет никаких сомнений, он – гений”.». Возможно, твиты Трампа находят отклик в России. Немного глубокой переработки данных раскрыло, что увеличение количества слова «экономика» в твитах Трампа на одно стандартное отклонение предсказуемо приводило к повышению температуры в Москве через пять дней на 2,00 °F. И снова корреляция была еще сильнее вневыборочно, чем внутривыборочно.

Трамп также назвал председателя Северной Кореи Ким Чен Ына «великим лидером» и сказал, что «когда он говорит, его люди встают по стойке смирно. Я хочу, чтобы мои люди поступали так же». На этот раз немного глубокой переработки данных обнаружило, что увеличение числа твитов Трампа со словом «великий» на одно стандартное отклонение предсказуемо приводило к повышению температуры в Пхеньяне через пять дней на 2,79 °F. И снова корреляция была еще более сильной вневыборочно.

Наконец, Гэри и Джон изучили статистическую взаимосвязь между выбором слов Трампом и количеством очков, набранных раннерами бейсбольной команды «Вашингтон Нэшнлз» во время пробежек. Телеканал CBS News сообщил, что, когда Трамп посетил пятую игру Мировой серии 2019 года, на него обрушился «поток свиста и насмешек со стороны, как казалось, большинства зрителей» и скандирования «Заприте его в клетке!». Возможно, свист и насмешки

были вызваны тем, что болельщики «Нэшнлз» знают, что судьба их команды была определена его твитами.

Конечно же, анализ твитов Трампа вынес на поверхность статистически значимую корреляцию между использованием Трампом слова «год» и количеством очков, набранных раннерами национальных команд четыре дня спустя. Увеличение количества публикаций в Твиттере на одно стандартное отклонение в годовом исчислении увеличило количество набранных национальными командами очков на 0,725.

Вот вам и обнаружение знаний. Мы никогда не ожидали, что будут обнаружены такие статистически убедительные, ранее неизвестные взаимосвязи.

Как не быть введенным в заблуждение призрачными закономерностями

Научный метод тестирует теории с помощью данных. Глубокая переработка данных позволяет обходиться без теорий и копаться в данных в поисках закономерностей, что часто сопровождается издевательствами над данными посредством перестановок, манипуляций и пропусков. Заманчиво полагать, что большие данные повышают эффективность глубокой переработки данных. Однако парадокс больших данных гласит, что чем больше данных мы перетряхиваем, тем больше шанс, что найденные нами закономерности будут обманчивыми и бесполезными.

Данные не говорят сами за себя, и вверх – это не всегда вверх.

Глава 6

Бесплодные поиски

Один из сыновей Гэри иногда играет со своими друзьями в игру, когда они идут в ресторан на обед. Все они выключают свои телефоны и кладут их на середину стола. Тот, кто первым проверит свой телефон, должен заплатить за всех. Игра редко длится долго.

Проигравший в этой игре никогда не следит за погодой, новостями или ценами на акции. Неслучайно социальные сети обладают непреодолимой притягательной силой. Люди хотят видеть, что делают другие, и делиться своими достижениями. Проверку своих телефонов сравнивают с пристрастием к азартным играм, поскольку обычно мы испытываем разочарование, но время от времени получаем крупный выигрыш, который заставляет нас возвращаться в надежде на новые острые ощущения.

Для исследователей данных одним из преимуществ интернета является коллекция огромного количества данных из социальных сетей, в которую они могут всматриваться в поисках закономерностей, не выходя из своих офисов. Им больше не нужно проводить эксперименты, наблюдать за поведением или делать опросы.

Мы оба являемся исследователями данных и ценим удобство наличия данных под рукой, но мы также скептически относимся ко многому из того, что в наши дни считается данными. Существует принципиальная разница между данными, собранными в ходе рандомизированных контролируемых испытаний (РКИ), и данными, состоящими из разговоров людей в социальных сетях и веб-страниц, которые они посещают в интернете.

Однако еще более важной проблемой является то, что взрывной рост данных значительно увеличил количество случайных возникающих закономерностей, которые могут обнаруживаться настойчивыми исследователями. Если существует относительно фиксированное количество полезных закономерностей, а количество случайно возникающих закономерностей растет экспоненциально, то с каждым днем соотношение полезных и бесполезных закономерностей обязательно должно приближаться к нулю.

Мы проиллюстрируем наши аргументы, рассмотрев несколько реальных и вымышленных примеров использования поисковых данных Google для выявления бесполезных закономерностей.

Использование Google Trends для отслеживания рынка акций

Давний финансовый обозреватель газеты New York Times Бертон Крейн предложил вот такой соблазнительный совет в отношении того, как зарабатывать большие деньги на рынке акций:

> Поскольку мы знаем, что акции будут как расти, так и падать, мы могли бы также получить от них немного трафика. Человек, который покупает акции по 10 и продает их по 20, зарабатывает 100 %. Но человек, который покупает их по 10, продает по 14 с половиной, выкупает обратно по 12 и продает по 18, выкупает обратно по 15 и продает по 20, зарабатывает 188 %.

Конечно, было бы чрезвычайно выгодно запрыгивать в рынок акций и выпрыгивать из него, быстро покупая до того, как цены пойдут вверх, и продавая до того, как цены пойдут вниз. Увы, исследование за исследованием показывают, что определить время входа и выхода очень трудно по сравнению с простой стратегией «купить и удерживать», когда акции покупаются и никогда не продаются, если только деньги не нужны на дом, пенсию или на что-то еще, на что инвесторы откладывают.

Поскольку акции приносят прибыль, инвесторы, которые то входят в акции, то выходят из них, при этом оказываясь правыми только в половине случаев, в среднем получают более низкий возврат на вложенный капитал, чем если бы они оставались в акциях все это время. Было подсчитано, что инвесторы, которые запрыгивают в рынок и выпрыгивают из него, пытаясь угадать, в какую сторону пойдет рынок дальше, должны быть правы в трех случаях из четырех, чтобы добиться такого же успеха, как если бы они придерживались стратегии «купить и удерживать». Исследование профессиональных инвесторов, проведенное банком Merrill Lynch, показало, что если сравнивать со стратегией «купить и удерживать», то «подавляющее большинство фондов теряют деньги в результате неправильного определения времени входа в рынок и выхода из него».

Эти мрачные результаты не останавливают людей от попыток. В 2013 году три исследователя сообщили, что они нашли новаторский способ определения времени входа и выхода, используя данные поиска в Google, чтобы предсказывать, будет ли индекс Доу Джонса расти или падать. Они задействовали предоставляемый компанией Google онлайновый источник данных Google Trends для еженедельного сбора данных о частоте поиска пользователями 98 различных ключевых слов:

> Мы включили термины, связанные с концепцией рынков акций, при этом некоторые термины были предложены инструментом Google Sets, который выявляет семантически соотносящиеся ключевые слова. Таким образом, используемый набор терминов был выбран не произвольно, поскольку мы намеренно привнесли некоторую финансовую предвзятость.

Использование уничижительного слова «предвзятость» является неудачным, поскольку оно намекает на нечто неправильное в использовании поисковых запросов, связанных с рынком акций. Убежденность в том, что корреляция преобладает над причинной следственностью, предполагает, что путь к открытию новых идей – это поиск закономерностей, не обремененных тем, что, как нам кажется, мы знаем о мире, – обнаружение способов обыграть рынок акций, ища «непредвзятые» слова, которые не имеют никакого отношения к акциям. Роковой изъян слепой стратегии заключается в том, что почти наверняка будут обнаружены случайные закономерности, а одни лишь данные не способны отличить наполненные смыслом закономерности от бессмысленных. Если у нас есть мудрость в отношении чего-то, то, как правило, лучше использовать ее – в данном случае, привнести некоторую финансовую экспертизу.

Для каждого из этих 98 слов исследователи рассчитали еженедельный «индикатор моментума» (или скорости изменения цены), сравнив поисковую активность за конкретную неделю со средней поисковой активностью за предыдущие недели. Этот индикатор является *скользящей средней*, поскольку он меняется еженедельно по мере того, как проходит череда недель. В табл. 6.1 приведен пример, в котором скользящая средняя рассчитывается за предыдущие три недели. (Предоставляемые службой Google Trends значения поиска приведены на шкале от 0 до 100, что отражает относительное словоупотребление в течение рассматриваемого периода.)

Таблица 6.1. Индикатор моментума с использованием трехнедельного среднего значения

Неделя	Значение поиска	Трехнедельная скользящая средняя	Индикатор моментума
1	41		
2	43		
3	45		
4	44	43	1
5	43	44	–1

На четвертой неделе в стилизованном примере в табл. 6.1 скользящая средняя за недели с первой по третью равна 43, а индикатор моментума равен 44 – 43 = 1, означая, что поисковая активность по этому ключевому слову была выше своего среднего значения за последние три недели. На пятой неделе трехнедельная скользящая средняя равна 44, а индикатор моментума – отрицательный, 43 – 44 = –1, поскольку поисковая активность на пятой неделе была ниже своего недавнего среднего значения.

Исследователи проанализировали скользящие средние за период от одной до шести недель для каждого из девяноста восьми ключевых слов и пришли к выводу, что наиболее успешная стратегия торговли акциями основана на ключевом слове «долг», используя трехнедельную скользящую среднюю и следующее ниже правило принятия решений:

- покупать индекс Доу Джонса, если индикатор моментума – отрицательный;
- продавать индекс Доу Джонса, если индикатор моментума – положительный.

Используя данные за семилетний период с 1 января 2004 года по 22 февраля 2011 года, они сообщили, что эта стратегия принесла поразительную годовую возвратность в размере 23,0 % по сравнению с 2,2 % для стратегии «купить и удерживать». Их заключение было таким:

> Наши результаты свидетельствуют о том, что эти предупреждающие знаки в данных об объеме поиска могли бы использоваться при разработке прибыльных торговых стратегий.

Причем никаких причин они не приводят:

> Потребуется дальнейшая работа, чтобы дать исчерпывающее объяснение глубинных психологических механизмов, которые заставляют людей искать такие термины, как «долг», перед тем как продавать акции по более низкой цене.

Фатальная проблема этого исследования заключается в том, что исследователи рассмотрели 98 разных ключевых слов и шесть разных скользящих средних (всего 588 стратегий). Если они рассматривали два правила торговли (покупка при положительном показателе индикатора моментума *или* продажа при отрицательном показателе индикатора моментума), то было обследовано 1176 стратегий. При таком количестве возможностей наверняка были обнаружены какие-то случайные закономерности, что подрывает доверие к тем, о которых сообщалось.

Когда мы смотрим назад, – когда мы предсказываем прошлое, а не будущее, – даже рассмотрение большого количества случайных чисел позволит вывести несколько случайно прибыльных правил торговли.

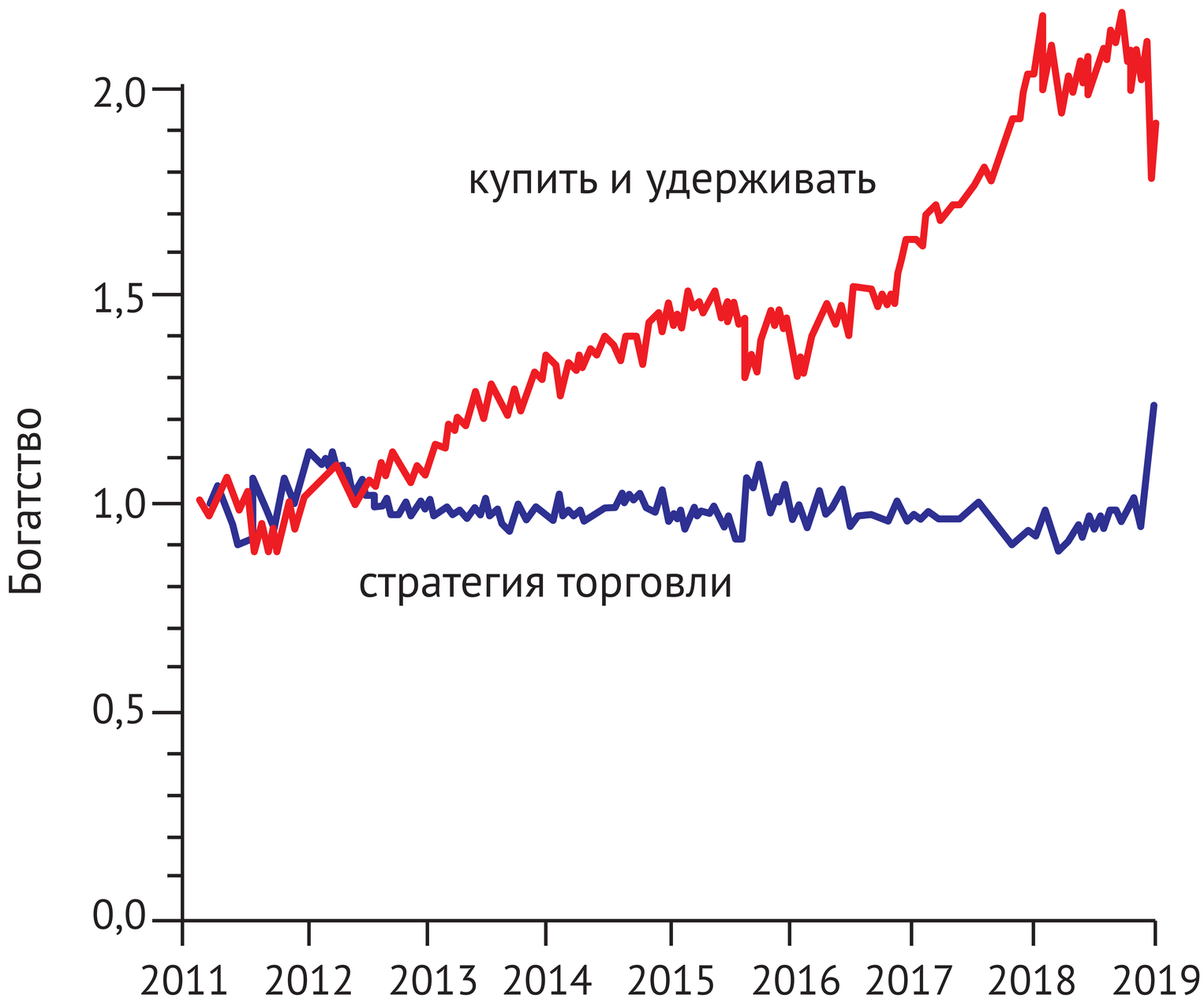

Рис. 6.1. Неуклюжие шатания при входе в рынок и выходе из него

Давайте посмотрим, насколько хорошо их самая успешная модель предсказывала будущее в следующие семь лет, с 22 февраля 2011 года по 31 декабря 2018 года. На рис. 6.1 показаны результаты. Их стратегия на основе слова «долг» обеспечивала годовую возвратность в размере 2,81 % по сравнению с 8,60 % в рамках стратегии «купить и удерживать».

Неудивительно, что их добытая в результате глубокой переработки данных стратегия с треском провалилась на свежих данных. Избирательно выдернутые закономерности обычно исчезают.

Мы не шутим, когда говорим о нашей искренней надежде, что эти исследователи не покупали и не продавали акции, основываясь на своей добытой в результате глубокой переработки данных стратегии. С другой стороны, один циник напомнил о том, что в потере денег есть что-то такое, что действительно заставляет людей обращать пристальное внимание на процесс принятия решений.

Всматриваясь в звезды

Вместо того чтобы полагаться на ключевые слова Google, почему бы не воспользоваться мудростью древних и не предсказать индекс Доу Джонса на основе результатов поиска в Google информации о двенадцати астрологических знаках? А почему бы и нет?

Джей относится к Рыбам. Он говорит, что этот факт объясняет причину, по которой он не верит в астрологию. Он не считает, что Солнце, Луна и планеты влияют на цены акций, не говоря уже о поиске в Google информации об астрологических знаках. Тем не менее мы скачали ежемесячную информацию о результатах поиска в Google за десятилетний период с 2009 по 2018 год двенадцати астрологических знаков: Овна, Тельца, Близнецов, Рака, Льва, Девы, Весов, Скорпиона, Стрельца, Козерога, Водолея и Рыб. Затем разбили данные на пятилетний внутривыборочный период (2009–2013), чтобы создать модель, и пятилетний вневыборочный период (2014–2018), чтобы протестировать модель.

На рис. 6.2 показано, что модель, основанная исключительно на поиске информации об одном астрологическом знаке, Скорпионе, сработала отлично. Корреляция между предсказанными и фактическими значениями индекса Доу Джонса составляет замечательные 0,83.

И угадайте что? Гэри – Скорпион, и он инвестировал и учил инвестированию на протяжении десятилетий!

Если вы подумали – пусть даже на мгновение, – что, возможно, именно поэтому Скорпион коррелирует с индексом Доу Джонса, то сможете понять, как легко люди поддаются соблазну случайно возникающих закономерностей. Гэри не служит причиной роста или падения цен на акции, и Гэри не служит причиной более частого или более редкого использования людьми поисковика Google для поиска информации о знаке Скорпиона.

Одна из двенадцати корреляций между ценами на акции и любыми двенадцатью поисковыми словами должна быть выше, чем другие корреляции, и, какой бы она ни оказалась, мы наверняка сможем придумать объяснение, которое будет творческим, но беспредметным.

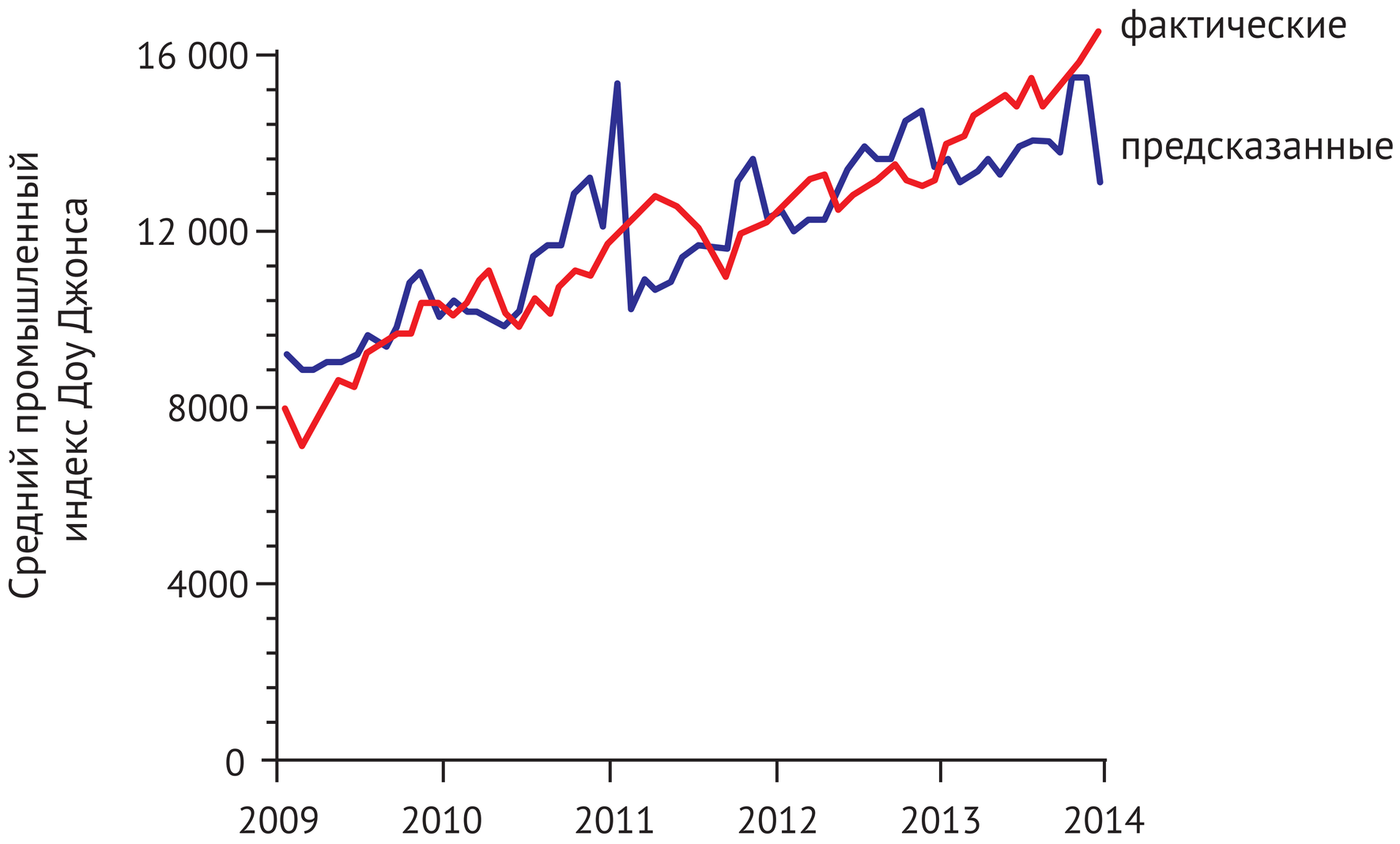

Рис. 6.2. Многообещающий индикатор рынка акций

Прежде чем бросаться покупать или продавать акции на основе результатов поиска информации о знаке Скорпиона, давайте посмотрим на результативность модели вневыборочно. На рис. 6.3 показано, что ответ, если подытожить одним словом, *разочаровывает*. Рынок акций взлетел, в то время как поисковая активность снизилась. В энный раз мы снова были одурачены призрачной закономерностью. Предсказывать прошлое легко, даже с помощью бестолковых моделей, но для надежного предсказания будущего требуются модели, в которых есть толк.

Рис. 6.3. Разочаровывающий предсказатель рынка акций

Использование Google Trends для отслеживания цен на биткойн

Биткойн – самая известная криптовалюта, цифровое средство обмена, которое работает независимо от центральной банковской системы. Как инвестиция биткойн представляет собой чистую спекуляцию. Инвесторы, которые покупают облигации, получают проценты. Инвесторы, которые покупают акции, получают дивиденды. Инвесторы, которые покупают многоквартирные дома, получают арендную плату. Единственный способ для людей, инвестирующих в биткойны, получать прибыль – это продавать свои биткойны дороже, чем они за них заплатили.

В 2018 году профессора экономики Йельского университета Алех Цывински (Aleh Tsyvinski) и Юкун Лю (Yukun Liu) (в то время еще аспирант) попали в заголовки газет благодаря статье, в которой они рекомендовали инвесторам хранить по крайней мере один процент своего портфеля в биткойнах. Они также сообщили, что цены на биткойны можно предсказывать исходя из частоты упоминания слова «биткойн» в поисковых запросах Google. Увеличение количества запросов о биткойнах обычно предшествует росту цен на биткойны. Снижение количества запросов о биткойнах предсказывает снижение цен. Они приписывали эту корреляцию изменениям во «внимании инвесторов».

За изученный ими период времени, с января 2011 года по май 2018 года, корреляция между ежемесячными поисками по термину «биткойн» и рыночной ценой биткойна в первый день следующего месяца составила замечательные 0,78. На рис. 6.4 эта тесная корреляция показана вплоть до весны 2018 года, когда цены на биткойны не падали так сильно, как предсказывалось, из-за снижения числа поисковых запросов по термину «биткойн».

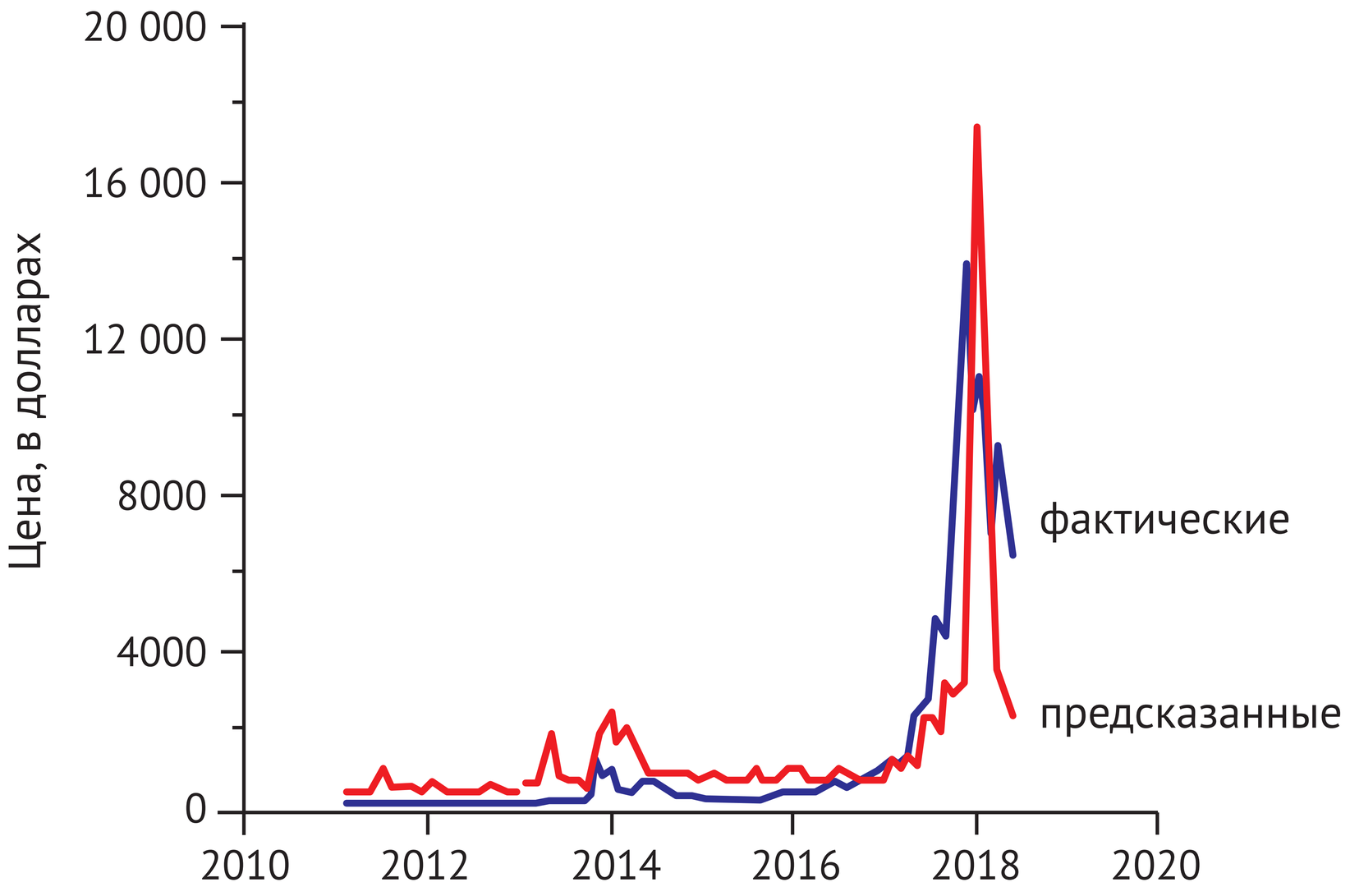

Рис. 6.4. Предсказывание цен на биткойны по результатам поиска по термину «биткойн»

Если бы задокументированная корреляция продолжилась, то это был бы новый способ разбогатеть. В месяцы, когда количество поисковых запросов по биткойну является высоким, покупать в последний день месяца; в месяцы, когда количество поисковых запросов по биткойну является низким, продавать в последний день месяца.

Часть интеллектуальной привлекательности предсказывания цен на биткойн на основе поисковых данных заключается в том, что цифровые данные используются новаторским способом для предсказания цены на цифровую валюту.

Корреляция также имеет поверхностную привлекательность. Цены на биткойн определяются исключительно суммами, которые люди готовы платить за биткойны, и увеличение внимания инвесторов вполне может приводить к повышению готовности платить более высокие цены. С другой стороны, многие люди (если не большинство), которые ищут информацию об этой криптовалюте, могут не быть заинтересованы в покупке. Им просто интересно узнать о новой вещи под названием биткойн. В сфере недвижимости при выставлении дома на продажу людей, которые приходят на день открытых дверей, но не имеют намерения покупать, называют праздно шатающимися зеваками. Наверняка существуют миллионы зевак, приходящих «поглазеть» на биткойны, что делает поиск в Google плохим средством привлечения внимания инвесторов.

Упадут ли цены на биткойн после того, как зеваки перестанут искать больше информации о биткойне? Нет, если только они не были потенциальными инвесторами. Были ли показанные на рис. 6.4 рост и последующее падение активности в поиске по термину «биткойн» результатом того, что инвесторы гонялись за биткойном, а затем убегали от него, или же зеваки узнали о биткойне то, что они хотели узнать, а затем почувствовали, что им не нужно знать еще больше?

На рис. 6.5 показано, что произошло после анализа Цывински/Лю. Поисковая активность стабилизировалась, а цены на биткойн отскочили, что послужило причиной того, что поисковая активность стала плохим предсказателем цен на биткойн.

Изначальная корреляция имела скупое теоретическое основание. Веская причина, по которой между поисковыми запросами по термину «биткойн» и ценами на биткойн должна существовать устойчивая корреляция, отсутствовала. Существовала временная корреляция, возможно, пока зеваки искали информацию о спекулятивном пузыре в ценах на биткойн, но эта корреляция длилась недолго и была бесполезна для предсказывания будущих цен на биткойн.

На рис. 6.6 показана сопоставимая праздная взаимосвязь между ценами на биткойн и поисковыми запросами по термину «Джуманджи», что отсылает к фильму «Джуманджи: Добро пожаловать в джунгли», вышедшему в декабре 2017 года, незадолго до пика пузыря биткойна. За период с января 2011 года по май 2018 года корреляция между ежемесячными поисковыми запросами по термину «Джуманджи» и рыночной ценой биткойна в первый день следующего месяца составила 0,73, что сопоставимо с корреляцией в 0,78 между поисковыми запросами по термину «*биткойн*» и ценами биткойна. Тех, кто верит, что корреляция преобладает над причинной следственностью, не смутит тот факт, что Джуманджи не имеет никакого отношения к биткойну. Корреляции вполне достаточно.

Рис. 6.5. Неверное предсказывание цен на биткойн, исходя из поисковых запросов по термину «биткойн»

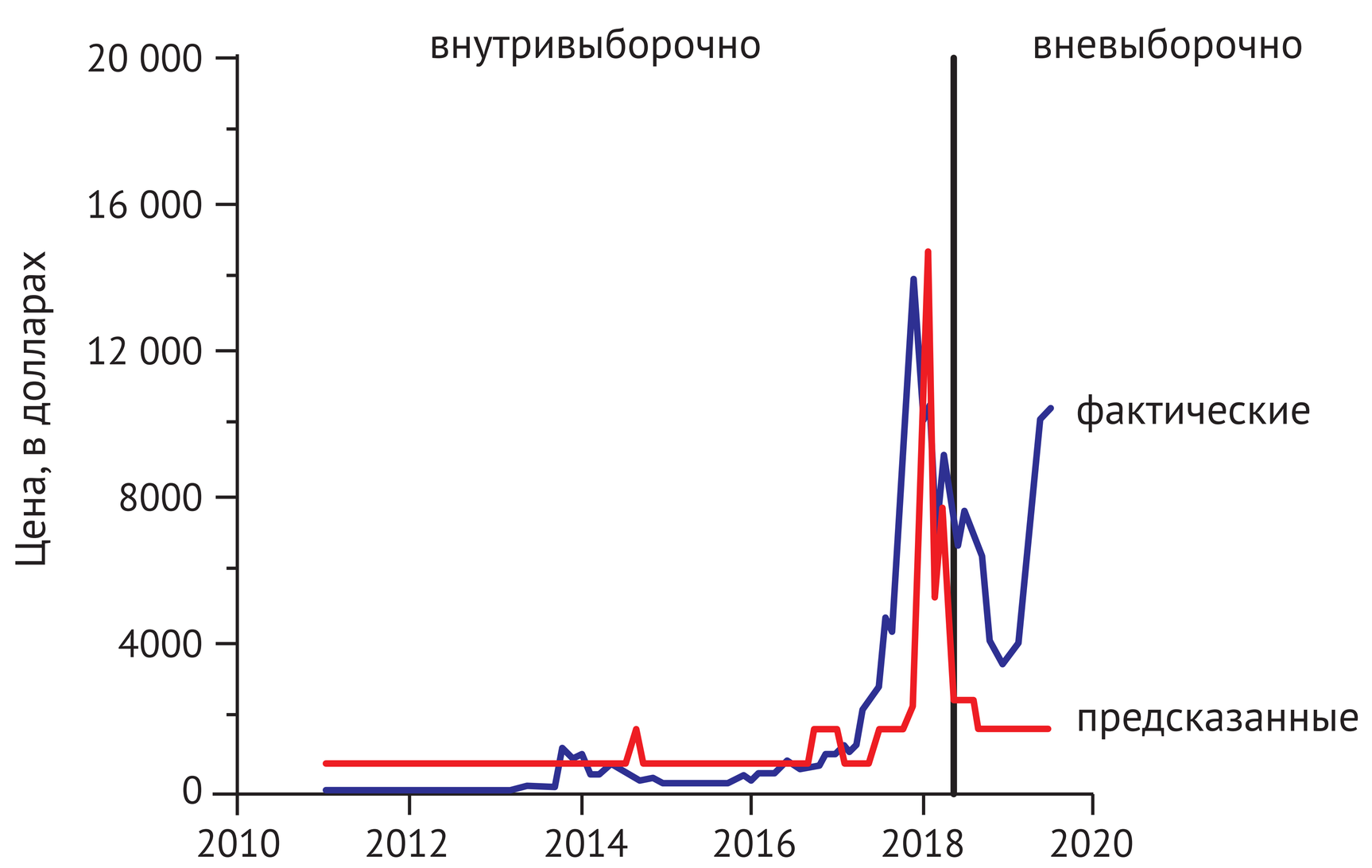

Рис. 6.6. Джуманджи и биткойн

В каждом случае при использовании поисковых запросов по терминам «биткойн» или «Джуманджи» наблюдается тесная внутривыборочная корреляция во время роста и падения цен на биткойн, что позволяет предполагать, что поисковый запрос (по «биткойну» или «Джуманджи») может использоваться для составления прибыльных предсказаний цен на биткойн. Но затем модели полностью сдуваются на отскоке цен на биткойн в 2019 году.

Корреляции находить легко. А вот полезные корреляции – почти неуловимы.

Предсказывание безработицы

Доступность данных из Google, Twitter, Facebook и других платформ социальных сетей предоставляет практически неограниченное количество переменных, которые могут использоваться для глубокой переработки данных, – и фактически неограниченное количество мнимых взаимосвязей, которые могут быть обнаружены.

В целях демонстрации этого утверждения мы отыскали несколько глупых данных, которые можно было бы использовать для предсказывания уровня безработицы в США. Это не заняло много времени. В настольной игре «Поселенцы Катана» есть пять ресурсов: кирпич, древесина, руда, овцы и шерсть. Поскольку данные по безработице публикуются ежемесячно, мы собрали ежемесячные данные из Google Trends по частоте использования этих пяти ресурсов в поисковых запросах Google с января 2004 года – насколько позволяют данные.

Затем мы вычислили модель, в которой для предсказывания уровня безработицы на следующий месяц с января 2004 по декабрь 2011 года использовалось количество ежемесячных поисковых запросов, оставив для вневыборочного тестирования модели период с 2012 по 2019 год.

На рис. 6.7 показано, что эти пять поисковых терминов неплохо справились с предсказыванием уровня безработицы. Корреляция между фактическим уровнем безработицы и предсказанным значением, основываясь на поисковых запросах за предыдущий месяц, составила впечатляющие 0,89. Стоит ли нам искать работу в качестве экономических прогнозистов?

Рис. 6.7. Использование ресурсов для предсказывания безработицы

Напомним, что это не реальные данные по кирпичу, древесине, руде, овцам и шерсти, которые на самом деле могут иметь какое-то отношение к экономике и уровню безработицы. Это частотность, с которой люди искали эти слова в поисковой системе Google. Тем не менее корреляция 0,89 неоспорима. Некоторые переработчики данных наверняка придумают возможные объяснения. Другие, возможно, скажут, что в объяснениях нет необходимости. Вверх значит вверх.

Вместо того чтобы подумать об объяснениях или решать, что объяснения не нужны, давайте посмотрим, как наша модель на основе поиска информации о ресурсах предсказывала безработицу во вневыборочный период с 2012 по 2019 год. Ужасно. На рис. 6.8 показано, что предсказанная безработица колебалась в районе 9 %, в то время как фактическая безработица упала ниже 4 %.

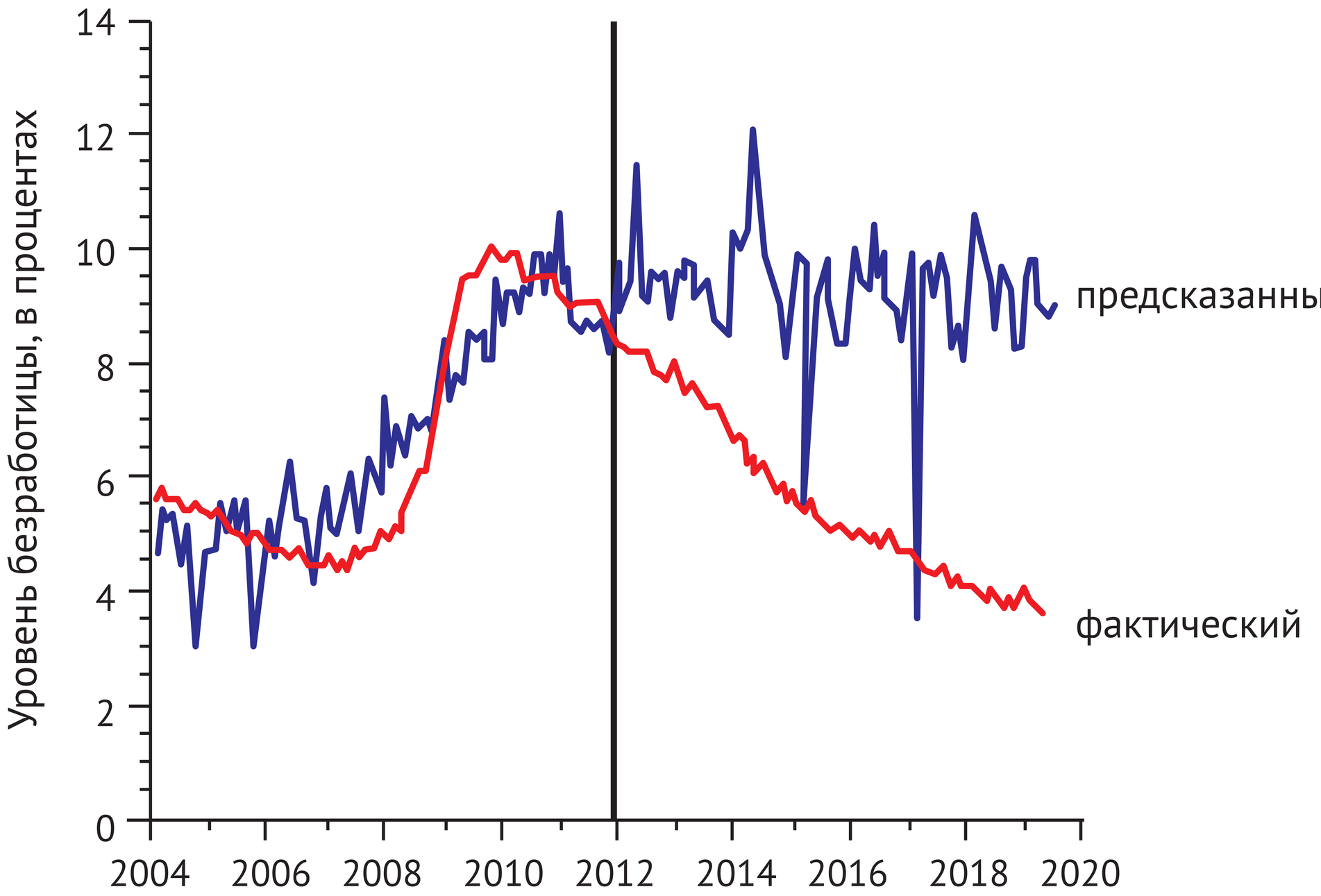

Рис. 6.8. Ресурсы с треском проваливаются вне выборки

Не было абсолютно никаких причин, по которым поиски в Google информации о ресурсах игры «Поселенцы Катана» будут как-то связаны с уровнем безработицы. Оказалось, что они коррелируют по удачной случайности, но случайно возникающие корреляции имеют временный характер и бесполезны. Этот глупый пример приведен как экземпляр более широкой точки зрения. В мире с безграничными данными существует бесконечный запас случайно возникающих корреляций. Их нетрудно находить и трудно не поддаваться соблазну – быть одураченными призрачными закономерностями.

Китайский жилищный пузырь

Мы не утверждаем, что интернет-данные бесполезны. Мы утверждаем, что интернет-данные следует использовать взвешенно. Исследователи должны собирать актуальные, надежные данные с четко поставленной целью, а не просто копаться в данных в поисках закономерностей. Хорошим примером может служить недавнее исследование китайского рынка недвижимости.

В Китае вся земля принадлежит государству, но в 1998 году Государственный совет Китая принял решение приватизировать рынок недвижимости, разрешив местным органам власти продавать права землепользования сроком на несколько десятилетий (обычно на семьдесят лет) застройщикам, которые возводят здания для продажи населению. С 1998 по 2010 год ежегодный объем построенного частного жилья вырос со 140 млн м^2 до более чем 610 млн, а доля домовладельцев в городском жилищном фонде достигла 84,3 %. В последние годы более четверти ВВП Китая было связано со строительством недвижимости.

Одним из обоснований такого спроса на жилье является то, что в Китае уже несколько десятилетий существует значительный дисбаланс по признаку пола, во многом обусловленный укоренившимся культурным предпочтением детей мужского пола и усугубляемый политикой «одного ребенка», действовавшей с 1979 по 2015 год. Владение домом желательно для мужчин, потому что это показатель статуса, который повышает их конкурентоспособность в брачных лотереях. Кроме того, многие китайские родители помогают своим детям покупать жилье. Доля домовладельцев в китайских семьях в возрасте до тридцати пяти лет составляет 55 % по сравнению с 37 % в Соединенных Штатах.

На рис. 6.9 показано, что жилая недвижимость в Китае оказалась более выгодной инвестицией, чем акции: жилье приносило относительно привлекательную возвратность на вложенный капитал при значительно меньшей волатильности. Однако именно такая обращенная в прошлое конъюнктурная перспектива подпитывает спекулятивные пузыри. Люди видят, что цены растут, и спешат покупать, что служит причиной дальнейшего роста цен – до тех пор, пока они не прекратят, что ведет к обвалу цен, потому что нет смысла покупать, если цены не растут.

Некоторые опасаются, что китайский рынок недвижимости является спекулятивным пузырем, и в качестве свидетельства указывают на стремительный рост цен. С 2003 по 2013 год в крупнейших городах Китая – Пекине, Шанхае, Гуанчжоу и Шэньчжэне – цены на жилье с поправкой на инфляцию ежегодно увеличивались на 13,1 %. В городах второго и третьего уровня Китая ежегодный прирост составил соответственно 10,5 и 7,9 %. Китайский миллиардер и магнат недвижимости Ван Цзяньлинь (Wang Jianlin) предупредил, что китайская недвижимость – это «самый большой пузырь в истории».

Гэри и студент Уэсли Лян (Wesley Liang) исследовали этот вопрос. Спекулятивный пузырь возникает, когда цена актива поднимается намного выше величины дохода, генерируемого этим активом, потому что инвесторы покупают актив не ради дохода, а для того, чтобы продать его по более высокой цене, чем они заплатили. В жилом секторе доходом является экономия на арендной плате (стоимости аренды сопоставимого жилья) за вычетом платежей по ипотеке, налогов на имущество, расходов на содержание и других расходов.

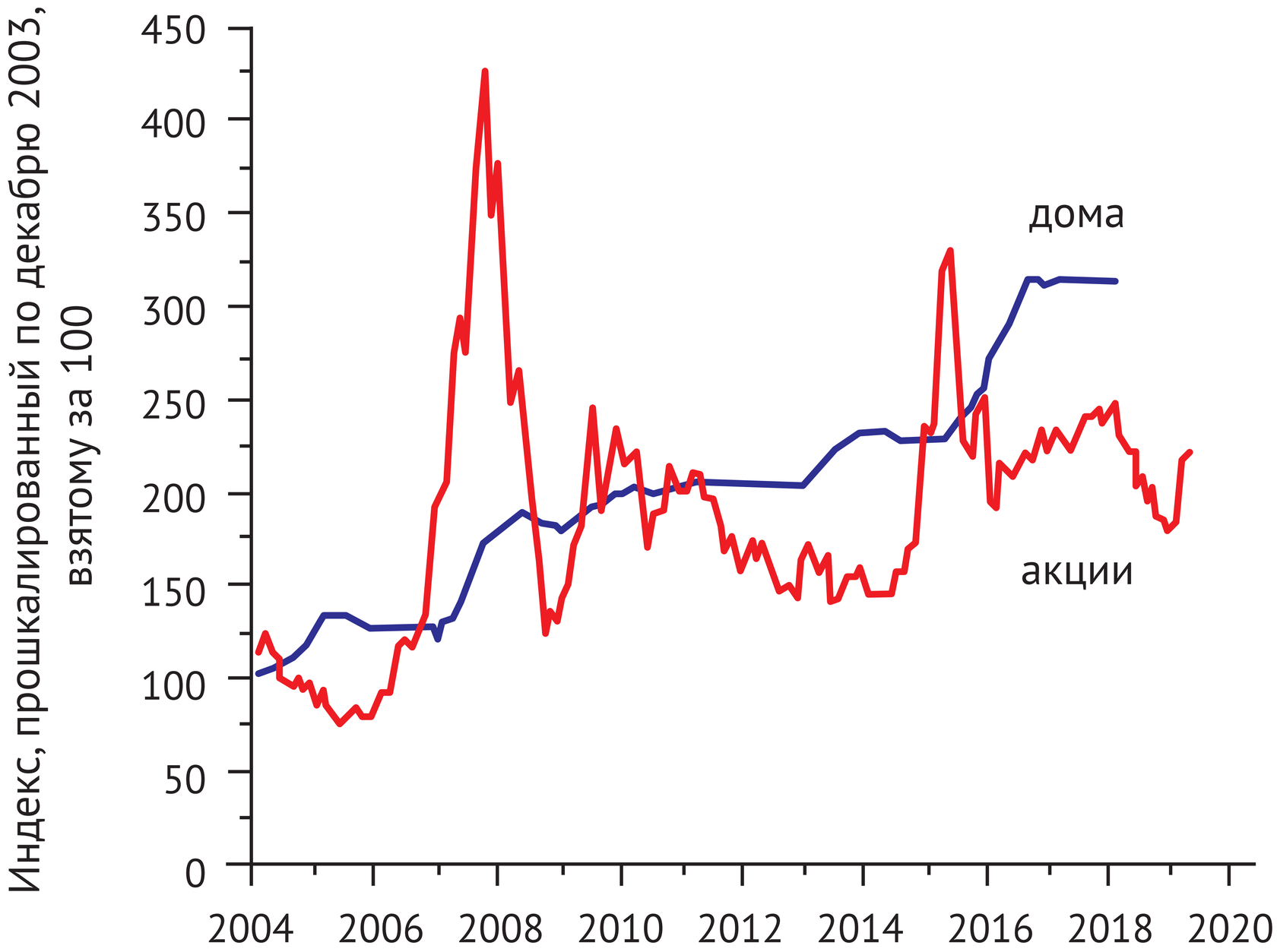

Рис. 6.9. Цены фондовом на рынке Шанхая и цены на подержанные дома в Шанхае

Гэри и Уэсли (которые свободно владеют китайским языком) смогли получить данные об арендной плате и ценах от одной из крупнейших китайских брокерских фирм по недвижимости 链家 (Льенджа). Они провели изучение объектов жилой недвижимости в двух крупнейших городах Китая, Пекине и Шанхае, в течение девяти месяцев с июня 2018 года по февраль 2019 года. Эти данные включают стоимость аренды или продажи, местоположение жилого комплекса, количество комнат, площадь, направление фасада, этаж, общее количество этажей в здании и год постройки.

Веб-сайт компании Льенджа требует, чтобы пользователь вводил китайский номер мобильного телефона на странице входа на веб-сайт. На номер телефона отправляется код подтверждения, который можно использовать на странице входа на веб-сайт. После входа на веб-сайт данные о прошлых транзакциях по продаже и аренде жилья можно найти под названием раздела 查交易, что переводится как «Проверить прошлые транзакции». Затем транзакции по продаже и аренде жилья сопоставляются вручную.

В этом исследовании есть два очень важных момента. Во-первых, Гэри и Уэсли практически не смогли бы собрать эти данные без интернета. Во-вторых, они использовали теоретически валидную модель, чтобы определить, какие данные следует собирать и анализировать; они не лазили по интернету в поисках переменных, которые могли бы коррелировать с ценами на жилье в Китае.

Они смогли идентифицировать более 400 пар одинаковых домов как в Пекине, так и в Шанхае. Из табл. 6.2 видно, что дома, как правило, небольшие и дорогие; однако, чтобы оценить, оправданы ли цены экономией на аренде, Гэри и Уэсли нужно было сравнить, во что обойдется покупка и аренда каждой пары одинаковых домов.

Таблица 6.2. Данные по жилью в Китае

Город	Среднее количество комнат	Медианные квадратные метры	Медианная ежемесячная аренда	Медианная цена	Медианная цена за квадратный метр
Пекин	1,75	61,50	954 долл.	707 750 долл.	11,517 долл.
Шанхай	1,64	55,59	611 долл.	407 750 долл.	6,925 долл.

Первоначальным допущением Гэри и Уэсли было то, что покупатели владеют своими домами вечно, передавая их из поколения в поколение, не потому, что дома никогда не перепродаются, а потому, что они хотели рассчитать уровни возвратности (нормы прибыли), основываясь исключительно на доходах от домов, не делая никаких допущений о будущих ценах на жилье. Это тот же самый подход, который используется стоимостными инвесторами для оценивания уровней возвратности на рынке акций. Гэри и Уэсли также рассчитали уровни возвратности на конечных горизонтах, используя различные допущения о том, как быстро могут расти цены на жилье в течение следующих одного года до тридцати лет.

В табл. 6.3 приведены результаты для бесконечного горизонта и для конечного горизонта с трехпроцентным ежегодным повышением цен на жилье, арендной платы и расходов. Гэри и Уэсли использовали ряд других допущений, чтобы оценить надежность результатов.

Таблица 6.3. Годовые уровни возвратности от инвестиций, в процентах

Период владения	Пекин	Шанхай
Один год	–15,10	–13,20
Десять лет	–2,17	–0,92
Тридцать лет	1,11	1,70
Навсегда	2,50	3,16

Возвратности зачастую оказываются отрицательными и явно непривлекательными по сравнению с китайскими десятилетними государственными облигациями, которые приносят отдачу в размере более трех процентов и имеют значительно меньший риск. Люди, которые покупали дома в Пекине или Шанхае в качестве инвестиций в 2018 и 2019 годах, очевидно, ожидали, что цены на жилье продолжат стремительно расти.

В будущем домовладение в Китае может стать еще менее привлекательным с финансовой точки зрения. По истечении срока действия соглашений о землепользовании владельцы могут подать заявку на продление, но нет никаких гарантий относительно стоимости продления. Возможно, китайское правительство будет взимать сборы за продление исходя из текущих рыночных цен. Однако крупные единовременные выплаты, скорее всего, будут восприняты многими домовладельцами как несправедливые и недоступные по цене. Менее разрушительной альтернативой было бы разрешить ежегодную выплату взносов за продление, подобно налогам на недвижимость, или ввести прямые налоги на недвижимость вместо сборов за продление. Сборы за продление

срока землепользования и/или налоги на недвижимость существенно снизят и без того низкий доход от владения жильем.

Эти результаты свидетельствуют о том, что на рынках жилья Пекина и Шанхая образовался пузырь, и покупателям жилья не следует ожидать продолжения ежегодного роста цен на жилье, выражающегося двузначными цифрами. Если двузначный рост цен продолжится, то пузырь на рынке недвижимости Китая станет еще больше и опаснее.

С другой стороны, возможные последствия жилищного кризиса в Китае выглядят настолько пугающе, что китайское правительство вряд ли будет стоять в стороне и позволит этому случиться. Рынок недвижимости слишком велик, чтобы потерпеть крах. Если из пузыря начнет выходить воздух, то китайское правительство, скорее всего, вмешается с помощью законов, нормативных актов или прямых закупок, чтобы предотвратить крах. Пузырь на рынке недвижимости в Китае, скорее всего, закончится не взрывом, а распусканием нюней.

Как не быть введенным в заблуждение призрачными закономерностями

Интернет предоставляет лавину данных, которые исследователи могут использовать для понимания и предсказания поведения людей. Однако если A/B-тесты не используются, то эти данные получены не из РКИ, которые позволяют тестировать на причинно-следственную связь и исключать спутывающие влияния. В дополнение к этому люди, которые пользуются интернетом в целом и социальными сетями в частности, безусловно, не представляют весомо всю популяцию в целом, поэтому при составлении выводов об общей популяции данные об их деятельности следует использовать осмотрительно.

То, что мы читаем или видим в интернете, не обязательно соответствует действительности. То, что мы делаем в интернете, не обязательно является информативным. Безудержное всматривание в поисковые запросы, обновления, твиты, хештеги, изображения, видео или подписи к ним наверняка выявят практически неограниченное количество призрачных закономерностей, которые возникают совершенно случайно и являются абсолютно бесполезными.

Глава 7

Кризис воспроизводимости

В 2011 году Дэрил Бем (Daryl Bem) опубликовал ставшую знаменитой статью под названием «Чувствуя будущее» (Feeling the Future). Бем сообщил о результатах девяти экспериментов, проведенных им за десять лет, восемь из которых показали статистически значимые свидетельства экстрасенсорного восприятия (evidence of extrasensory perception, аббр. ESP). В ходе одного эксперимента студентам Корнелльского университета показали на мониторе компьютера два занавеса – один с левой стороны, другой с правой – и попросили угадать, за каким из них находится изображение. Новаторский подход Бема к этой игре в угадывание заключался в том, что компьютерный генератор случайных событий выбирал занавес с изображением после того, как ученик угадывал ответ. Когда изображение было эротическим, студенты смогли правильно угадывать в статистически значимых 53 % случаев. Студенты чувствовали будущее. В ходе еще одного эксперимента Бэм обнаружил, что люди лучше справлялись с тестом на воспоминание, если изучали слова, на которых их тестировали после прохождения теста. И снова студенты чувствовали будущее.

Бем был известным социальным психологом, и его статья была опубликована в престижном журнале. Вскоре об этих замечательных исследованиях стало известно по всему миру, включая статью на первой полосе в газете New York Times и интервью для ток-шоу The Colbert Report. Бем пригласил других повторить его результаты и даже предоставил пошаговые инструкции о том, как повторять его эксперименты.

Бем был настолько известен, журнал – настолько уважаем, а результаты – настолько невероятны, что несколько исследователей приняли вызов и попытались повторить эксперименты Бема. Они не нашли никаких свидетельств в пользу того, что люди могут чувствовать будущее. Однако статья Бема не пропала даром. Это было одно из двух событий, которые сделали 2011 год для социальной психологии переломным.

По другую сторону Атлантики сорокапятилетний датский социальный психолог Дидерик Стапель (Diederik Stapel) стал деканом Тилбургского университета и сделал себе имя публикацией десятков статей с такими броскими названиями, как:

«Когда милые хулиганы делают вас счастливыми, а честные продавцы – грустными»;

«Тайная жизнь эмоций»;

«Нет боли – нет выгоды: условия, при которых сравнение в лучшую сторону приводит к повышению эффективности».

В одной из его самых известных статей сообщалось, что употребление мяса делает людей эгоистичными и асоциальными; в другой заявлялось, что беспорядочная окружающая среда делает людей расистами. В неопубликованной работе было обнаружено, что люди, которые заполняют анкету, скорее всего, едят эм-энд-эмс из чашки в ближайшей кофейне, если на чашке написано слово «kapitalisme», а не бессмысленное нагромождение одних и тех же букв.

В отличие от Бема многие статьи Стапеля были правдоподобными. В отличие от Бема многие данные Стапеля были фальшивыми.

В 2011 году Тилбургский университет отстранил Стапеля от работы за фальсификацию данных; 58 его опубликованных статей в настоящее время отозваны, включая упомянутые выше.

Позже Стапель объяснил, что он «хотел слишком многого слишком быстро», но им также двигало «стремление не к истине, а к эстетике и красоте». Он предпочитал упорядоченность и четкие ответы. Ему так сильно нужны были закономерности, что он их придумывал, когда не мог найти.

В начале своего фальсификационного пути он собирал данные, а затем их очищал, меняя несколько цифр здесь и там, пока не получал ответы, которые хотел получить. Позже он не стал утруждать себя сбором данных. Он просто приводил разумные цифры, которые подтверждали его теории. В ходе исследования эм-энд-эмс он сидел рядом с чашкой, полной эм-энд-эмс, и съел, как ему казалось, разумное количество; затем он сделал вид, что провел эксперимент с несколькими добровольцами, и просто придумал цифры, которые гармонировали с его собственным потреблением эм-энд-эмс.

В инцидентах с Бемом и Стапелем в 2011 году есть два очень интересных момента. Во-первых, они не изолированы, а во-вторых, они не так уж сильно отличаются.

Неряшливая наука

Три университета, в которых работал Стапель (Тилбургский университет, Гронингенский университет и Амстердамский университет), созвали следственные комиссии для выяснения масштабов недобросовестных поступков Стапеля. Их итоговый отчет назывался «Порочная наука: мошеннические исследовательские практики социального психолога Дидерика Стапеля», причем первые два слова – «Порочная наука» – являлись ключевой частью названия. Проблемы, которые они обнаружили, были серьезнее, чем выдумки Стапеля. Они лежали в основе того, как нередко проводятся исследования.

Комитеты были удивлены и разочарованы, обнаружив широко распространенную «неряшливую науку»:

«Побочным результатом» расследований комитетов является вывод о том, что в самой дисциплине существует гораздо больше аспектов, чем предполагалось изначально, которые следует считать нежелательными или даже неправильными с точки зрения академических стандартов и научной добросовестности.

Еще одним явным признаком является то, что во время интервью несколько соавторов, которые сами проводили анализ и не все принадлежали к «школе» Стапеля, защищали серьезные и менее серьезные нарушения надлежащего научного метода словами: «Этому я научился на практике; в моем исследовательском окружении все так поступают, и то же самое делают все, с кем мы общаемся на международных конференциях».

Отчет включает в себя длинный раздел, посвященный *систематическому смещению при верификации* (verification bias), подтасовыванию данных с целью получения результатов, подтверждающих желаемый исследователем вывод:

> Одно из самых фундаментальных правил научного исследования заключается в том, что изыскание должно быть спланировано таким образом, чтобы факты, которые могли бы опровергнуть гипотезы исследования, имели по крайней мере равные шансы на появление, как и факты, подтверждающие гипотезы исследования. Нарушения этого фундаментального правила… по сути, делают гипотезы невосприимчивыми к фактам.

Они приводят несколько примеров.

1. Эксперимент повторяется (возможно, с незначительными изменениями) до тех пор, пока не будут получены статистически значимые результаты; это единственный описанный эксперимент.
2. Проводится несколько разных тестов, и сообщаются только статистически значимые результаты.
3. Результаты нескольких экспериментов объединяются, если это объединение дает статистически значимые результаты.
4. Некоторые данные отбрасываются, чтобы произвести статистически значимые результаты, либо вообще не упоминаются в исследовании, либо приводятся под надуманным предлогом, например «студенты просто отвечали тем, что приходило им в голову».
5. Выбросы в данных оставляются, если они необходимы для статистической значимости, без упоминания того факта, что результаты зависят от малого числа аномальных данных.

Когда комитеты получили доступ к полным наборам данных и обнаружили подобную практику в опубликованных статьях, исследователи поспешили найти оправдания. Что еще более обескураживает, исследователи часто считали свои действия совершенно нормальными.

Комитеты также обнаружили, что ведущие журналы не только игнорировали подобные нарушения, но и поощряли их:

> Рецензенты также просили сообщать не обо всех выполненных анализах, например просто не упоминать о каких-либо условиях, для которых не было обнаружено никаких

эффектов, хотя первоначально ожидались эффекты. Иногда рецензенты настаивали на ретроспективных пилотных исследованиях, о которых затем сообщалось, что они были проведены заранее. За счет этого эксперименты и выбор предметов оправдываются преимуществами прошлого периода.

Нередко рецензенты отдавали предпочтение интересному, элегантному, лаконичному и убедительному рассказу, возможно, в ущерб необходимой научной тщательности.

Из выводов комитетов, проводивших расследование по делу Стапеля, вытекало, что, кроме абсурдного предположения, в статье Бема не было ничего необычного. Именно так поступало большинство исследователей в области социальной психологии. Друг Бема Ли Росс сказал, что «уровень доказательства здесь был обычным. Я имею в виду как положительные, так и отрицательные стороны. Я имею в виду, что это был именно тот тип традиционного психологического анализа, который [часто встречается] с теми же недостатками и проблемами, что и у большинства исследований».

Стапель и Бем не так уж сильно отличаются друг от друга. Они оба работают в континууме, который варьируется от незначительной доработки данных до тотальной подтасовки и полной фабрикации. Сильно ли отличается проведение экспериментов до тех пор, пока не будет получено то, что требуется, от быстрого способа изготовления данных, чтобы получить желаемые результаты?

В заголовке статьи в журнале Slate проблема была сформулирована кратко: «Дэрил Бем доказал реальность экстрасенсорного восприятия, а это значит, что наука разрушена». Доверие к научным исследованиям было подорвано тем, что получило название *кризиса воспроизводимости* (или кризиса репликации), поскольку попытки воспроизвести опубликованные исследования часто заканчиваются неудачей; в случае Стапеля – из-за того, что он сфальсифицировал данные; в случае Бема – из-за того, что он манипулировал данными, как ему заблагорассудится.

Например, в исследовании эротических картинок Бем провел свой эксперимент с пятью разными видами картинок и выбрал единственную, которая была статистически значимой. Статистик Эндрю Гельман (Andrew Gelman) называет это «садом с расходящимися дорожками». Если вы будете бродить по саду, выбирая наугад, куда идти, на каждом перепутье, то ваш конечный пункт покажется вам почти волшебным. Каковы шансы, что вы придете именно в это место? И все же вы должны были где-то оказаться. Если бы конечный пункт был указан до того, как вы начали прогулку, то было бы удивительно, что вы найдете к нему дорогу. Однако выявление пункта назначения после того, как вы закончили прогулку, определенно не является удивительным.

В качестве иллюстрации последствий систематического смещения при верификации три исследователя попросили двадцать студентов Пенсильванского университета послушать песню группы The Beatles под названием «When I'm 64». Затем исследователи применили немного неряшливой науки, чтобы продемонстрировать, что студенты стали на полтора года моложе после того, как прослушали эту песню. Они не упоминали Бема, но сходство было очевидным.

Во время одного из интервью Бем заявил без угрызений совести:

> Я начинал эксперимент, и если он ни к чему не приводил, то бросал его и начинал заново с изменениями. Я не очень внимательно следил за тем, какие из них я отбросил, а какие нет. Вероятно, вначале я был очень неряшлив. Думаю, что, возможно, часть критики является вполне обоснованной. Я никогда не был нечестным, но, с другой стороны, критики были правы.

В сентябре 2012 года в электронном письме исследователям в области социальной психологии нобелевский лауреат Дэниел Канеман предупредил, что «я вижу надвигающуюся катастрофу». Несколько лет спустя Гэри был на конференции, где известный социальный психолог сообщил, что его область является эталоном невоспроизводимых исследований, и добавил: «По умолчанию допускаю, что все, что публикуется в моей области, неверно».

Начался кризис воспроизводимости.

По иронии судьбы некоторые теперь задаются вопросом, было ли озорство Бема преднамеренным. Потратил ли он десять лет на подготовку тщательно продуманного розыгрыша, который стал бы звоночком для других специалистов в этой области? Публиковал ли он нелепые результаты и поощрял ли других пытаться повторить их, чтобы навести порядок в социальной психологии? По словам самого Бема в интервью журналу Slate в 2017 году, ответ был отрицательным, но, возможно, это часть розыгрыша:

> Я сторонник строгости, но предпочитаю, чтобы это делали другие. Я понимаю ее важность – некоторым это доставляет удовольствие, – но у меня не хватает на это терпения. <…> Если вы посмотрите на все мои прошлые эксперименты, то они всегда были риторическими приемами. Я собирал данные, чтобы показать, как будет изложена моя точка зрения. Я использовал данные для убеждения и никогда по-настоящему не беспокоился о том, повторится ли это или нет.

Бем и Стапель были социальными психологами, но кризис репликации выходит далеко за рамки социальной психологии. В медицинских исследованиях многие «доказавшие свою эффективность» методы лечения на практике оказываются менее эффективными, чем в ходе экспериментальных испытаний. Эта закономерность настолько распространена, что у нее даже есть название – «эффект затухания».

Например, резерпин был популярным средством для лечения гипертонии, пока в 1974 году исследователи не сообщили, что он значительно повышает риск развития рака молочной железы. Его применение резко сократилось, и сегодня он используется редко. Однако в результате нескольких попыток повторить первоначальные результаты был сделан вывод о том, что связь между резерпином и раком молочной железы была ложной. Один из первых исследователей позже описал первоначальное исследование как «катастрофу с резерпином и раком молочной железы».

Оглядываясь назад, он признал:

> Мы провели в буквальном смысле тысячи сравнений, включающих сотни исходов и сотни (если не тысячи) воздействий. С точки зрения теории вероятности должны были проявиться «статистически значимые» ассоциации, и то, что мы назвали возможной причинно-следственной связью, на самом деле было случайной находкой.

Они шли по саду с расходящимися дорожками.

Опрос, проведенный в 2018 году среди 390 профессиональных статистиков, которые предоставляют статистические консультации медицинским исследователям, показал, что более половины из них оказывалось в ситуациях, когда их просили делать то, что они считали серьезным нарушением надлежащей статистической практики, включая проведение нескольких тестов после обследования данных и ложное представление этих тестов после свершившегося факта как теорий, которые были задуманы до обращения к данным. Опрос, проведенный в 2011 году среди более чем 2000 психологов-исследователей, обнаружил, что большинство из них признались в использовании сомнительных исследовательских практик. Исследование, проведенное в 2015 году одним из лучших научных журналов Nature, показало, что более 70 % опрошенных людей сообщили, что они безуспешно пытались воспроизвести эксперимент другого ученого и более половины из них безуспешно пытались воспроизвести некоторые из своих собственных исследований!

В рамках запущенного Брайаном Носеком (Brian Nosek) проекта по воспроизводимости были проанализированы 100 исследований, опубликованных в трех ведущих журналах по психологии в 2008 году. В 97 из этих исследований были получены значительные результаты, что само по себе наводит на мысль о проблеме. Конечно же, 97 % всех хорошо проведенных исследований не дали значимых результатов. Двести семьдесят исследователей вызвались повторить эти 97 исследований. Только 35 из 97 изначальных выводов были подтверждены, и даже в этом случае эффекты неизменно оказывались меньше, чем сообщалось изначально.

В рамках Проекта по репликации экспериментальной экономики (Experimental Economics Replication Project) была предпринята попытка воспроизвести восемнадцать исследований в области экспериментальной экономики, опубликованных в двух ведущих экономических журналах в 2011–2014 годах. Только одиннадцать (61 %) из последующих исследований показали значимые эффекты в том же направлении, о котором сообщалось изначально.

Одной из причин кризиса воспроизводимости является растущее доверие к алгоритмам глубокой переработки данных в части построения моделей, не основанных на опыте человека. Исходы, полученные с помощью моделей, добытых в результате глубокой переработки данных, по своей природе невоспроизводимы, поскольку они почти наверняка будут содержать в данных закономерности, которые исчезают при тестировании на новых данных.

Ты – это то, что ест твоя мать

В 2008 году группа британских исследователей проанализировала 133 продукта питания, которые потреблялись 740 британскими женщинами до беременности, и пришла к выводу, что у женщин, которые съедали хотя бы одну порцию хлопьев на завтрак в день, было гораздо больше шансов родить мальчиков, по сравнению с женщинами, которые съедали одну или меньше порций в неделю. После того как это замечательное открытие было опубликовано в престижном журнале Proceedings of the Royal Society под броским заголовком «Ты – это то, что ест твоя мать», о нем стало известно по всему миру, и за первую неделю

после публикации оно набрало более 50 000 просмотров в Google. Несомненно, наблюдалось увеличение потребления хлопьев женщинами, которые хотели иметь детей мужского пола, и снижение потребления хлопьев женщинами, которые хотели иметь детей женского пола.

Как быстро заметил еще один исследователь, Стэнли Янг (Stanley Young), необычность этого вывода заключается в том, что пол ребенка определяется тем, какую хромосому несет мужской сперматозоид, X либо Y: «Женский пол не имеет никакого отношения к полу ребенка». Тогда каким образом исследовательская группа обнаружила статистически значимую взаимосвязь между хлопьями и полом? Они проанализировали потребление 132 продуктов питания в три периода времени (до зачатия, через 16 недель после зачатия и от 16 до 28 недель после зачатия), в результате чего было получено в общей сложности 396 переменных. Кроме того, Янг и двое его соавторов написали, что «по-видимому, также существует скрытое множественное тестирование, поскольку было проведено много дополнительных тестов, о которых сообщалось в других статьях». Неудивительно, что сотни тестов выявили статистически значимую взаимосвязь для одного продукта питания в определенный момент времени (хлопьев до зачатия).

Зомбические исследования

Джей получил следующее ниже электронное письмо от друга (который по совпадению был студентом Гэри):

> Риски и эффективность вакцин… являются особенно актуальной темой здесь, в Орегоне, поскольку законодательное собрание штата работает над законопроектом, который исключает возможность немедицинского освобождения школьников от вакцинации. Если ребенок не привит и не имеет медицинской справки, то ему не будет разрешено посещать государственную школу.
>
> Мне сказали, что существуют исследования, подтверждающие оба вывода: вакцины действительно вызывают аутизм и другие аутоиммунные заболевания, и вакцины не вызывают эти состояния. Я сам не проводил никаких исследований. Понимаю, что стержневое исследование, подтверждающее вредность вакцин, было отозвано. В чем же правда?
>
> У меня есть друзья, которые подумывают уехать из штата, если этот законопроект вступит в силу.
>
> Хотел бы разобраться с научной точки зрения, прежде чем рассматривать проблему личной свободы, связанную с таким законом.

Исследование 1998 года, в котором утверждалось, что оно выявило ассоциацию между вакцинами MMR[10] и аутизмом, действительно было отозвано из-за целого ряда статистических недостатков спустя двенадцать долгих лет после его публикации. (Подробности приведены в книге Гэри «Стандартные отклонения: ущербные допущения, растерзанные данные и другие способы лгать, используя статистику» – Standard Deviations: Flawed Assumptions, Tortured Data,

[10] MMR – это комбинированная вакцина против кори, эпидемического паротита (свинки) и краснухи (немецкой кори), обеспечивающая иммунитет от этих трех заболеваний при одной инъекции. – *Прим. перев.*

and Other Ways to Lie with Statistics.) Хотя всегда существует небольшой риск аллергической реакции, нет никаких свидетельств в пользу связи между вакцинами MMR и аутизмом, а опасностей, которые они предотвращают, намного больше, чем любые риски, которые они создают. К сожалению, как и большинство отозванных исследований, это исследование продолжало жить после отзыва. В 2010 году статья была отозвана, а главный редактор журнала заявил, что статья была «абсолютно лживой». Исследователю также было запрещено заниматься медицинской практикой в Англии. Тем не менее некоторые люди все еще верят в ложные заявления газеты. Они превратились в зомбическое исследование.

Многие отозванные статьи, в том числе некоторые из работ Стапеля, до сих пор цитируются исследователями, которые пребывают в блаженном неведении об их исчезновении. То же самое относится и к статьям, которые были дискредитированы, хотя и не отозваны. Многие веб-сайты продолжают сообщать о том, что, как доказано научно, употребление хлопьев до зачатия может увеличивать шансы женщины на рождение мальчика, а отказ от хлопьев может увеличивать шансы на рождение девочки.

Еще одним примером является ныне дискредитированное заявление о том, что ураганы с женскими именами являются более смертоносными, чем ураганы с мужскими именами (полное раскрытие: Гэри был одним из разоблачителей). Данные об ураганах не были сфабрикованы, но в статье были допущены многие из перечисленных выше ошибок, связанных со систематическим смещением при верификации. Например, в документе были указаны ураганы, имевшие место до 1979 года, когда все ураганы имели женские имена, а ураганы были более смертоносными из-за более слабых зданий и меньшего количества заблаговременных предупреждений. В течение периода после 1979 года, когда женские и мужские имена стали чередоваться, не было никакой разницы в смертоносности. Тем не менее с каждым годом по мере приближения сезона ураганов социальные сети буквально кишат страшными предупреждениями об ураганах с женскими именами.

Кризис воспроизводимости: тематическое исследование

В главе 6 мы обсуждали проведенное в 2018 году исследование Олега Цывински и Юкуна Лю, в котором утверждалось, что цены на биткойн можно предсказывать по тому, как часто слово «биткойн» упоминается в поисковых запросах Google. Они также проанализировали корреляции между ценами на биткойн и сотнями других экономических переменных и пришли к выводу, что биткойн «представляет собой класс активов, который можно оценивать с помощью простых финансовых инструментов».

Их исследование, по сути, было безудержной глубокой переработкой данных, поскольку нет никакого логического основания связывать цены на биткойн с чем-либо, кроме настроений инвесторов. В отличие от облигаций, приносящих проценты, акций, приносящих дивиденды, квартир, приносящих арендную плату, предприятий, приносящих прибыль, и других реальных инве-

стиций, биткойн не приносит вообще ничего, поэтому нет убедительного способа оценить биткойн в ключе, в каком инвесторы могут оценивать облигации, акции, квартиры, предприятия и другие реальные инвестиции.

Цены на биткойн могут сильно колебаться, потому что люди покупают биткойны, ожидая, что цена вырастет, и продают, ожидая, что цена упадет, что приводит к еще большему росту и падению цен на биткойн. Такие самореализующиеся ожидания не обязательно должны быть связаны с чем-то измеримым. Возможно, ими движет нечто большее, чем то, что великий британский экономист Джон Мейнард Кейнс (John Maynard Keynes) назвал «животным духом».

Попытки связать цены на биткойн с реальными экономическими переменными обречены на разочарование, однако Лю и Цывински пытались сделать именно это, и их разочарования поучительны. К их чести, в отличие от многих исследований, они не скрывали своих неудавшихся результатов; они были прозрачны в том, сколько переменных они рассматривали: 810! При рассмотрении такого количества переменных настоятельно рекомендуется использовать гораздо более строгий стандарт, чем обычный 5%-ный порог статистической значимости (в качестве крайнего примера физики элементарных частиц считают данные убедительными только в том случае, если вероятность того, что это произойдет случайно, составляет менее одного раза на 3,5 млн). Лю и Цывински скорректировали барьер, но в неверном направлении и посчитали статистически значимой любую ассоциацию, шанс возникновения которой по удачному стечению обстоятельств составляет менее 10 %!

Лю и Цывински использовали данные за период с 1 января 2011 года по 31 мая 2018 года. Для наших репликационных тестов мы использовали вневыборочный период с 1 июня 2018 года по 31 июля 2019 года. К их переменным относятся обменный курс канадского доллара к доллару США, цена на сырую нефть, возвратность акций в сфере здравоохранения и возвратность акций в пивной отрасли. Время от времени предлагаемые ими обоснования редко бывают убедительными.

Например, Лю и Цывински признают, что, в отличие от акций, биткойны не генерируют наличные деньги и не выплачивают дивиденды, поэтому они использовали то, что они называют «похожей метрикой», – количество пользователей биткойн-кошельков:

> Очевидно, что прямой меры дивидендов для криптовалют не существует. Однако по своей сути отношение цены к дивидендам является мерой разрыва между рыночной стоимостью и фундаментальной стоимостью актива. Рыночная стоимость криптовалюты – это просто наблюдаемая цена. Мы определяем фундаментальную стоимость, используя количество пользователей биткойн-кошелька.

Количество пользователей биткойн-кошельков не является заменой выплачиваемых акционерам денежных дивидендов. Этот надуманный заместитель напоминает бесполезные метрики (например, посещаемость веб-сайтов), которые движимые желанием инвесторы придумали во времена пузыря доткомов, чтобы оправдывать постоянно растущие цены на акции.

Одна из вероятных взаимосвязей, которую они реально рассмотрели, – это взаимосвязь между текущей и прошлой возвратностью биткойна. На рис. 7.1

показаны цены на биткойн за исследуемый период. На рис. 7.2 показан объем торгов за тот же период. В результате, по сути, образуется спекулятивный пузырь с постоянно растущими ценами, который вынуждает спекулянтов тратить деньги впустую, оказывая дальнейшее повышательное давление на цены. Рисунки 7.1 и 7.2 убедительно подтверждают идею о том, что спекулянты бросились покупать биткойн, когда цена на него выросла, а затем сбежали.

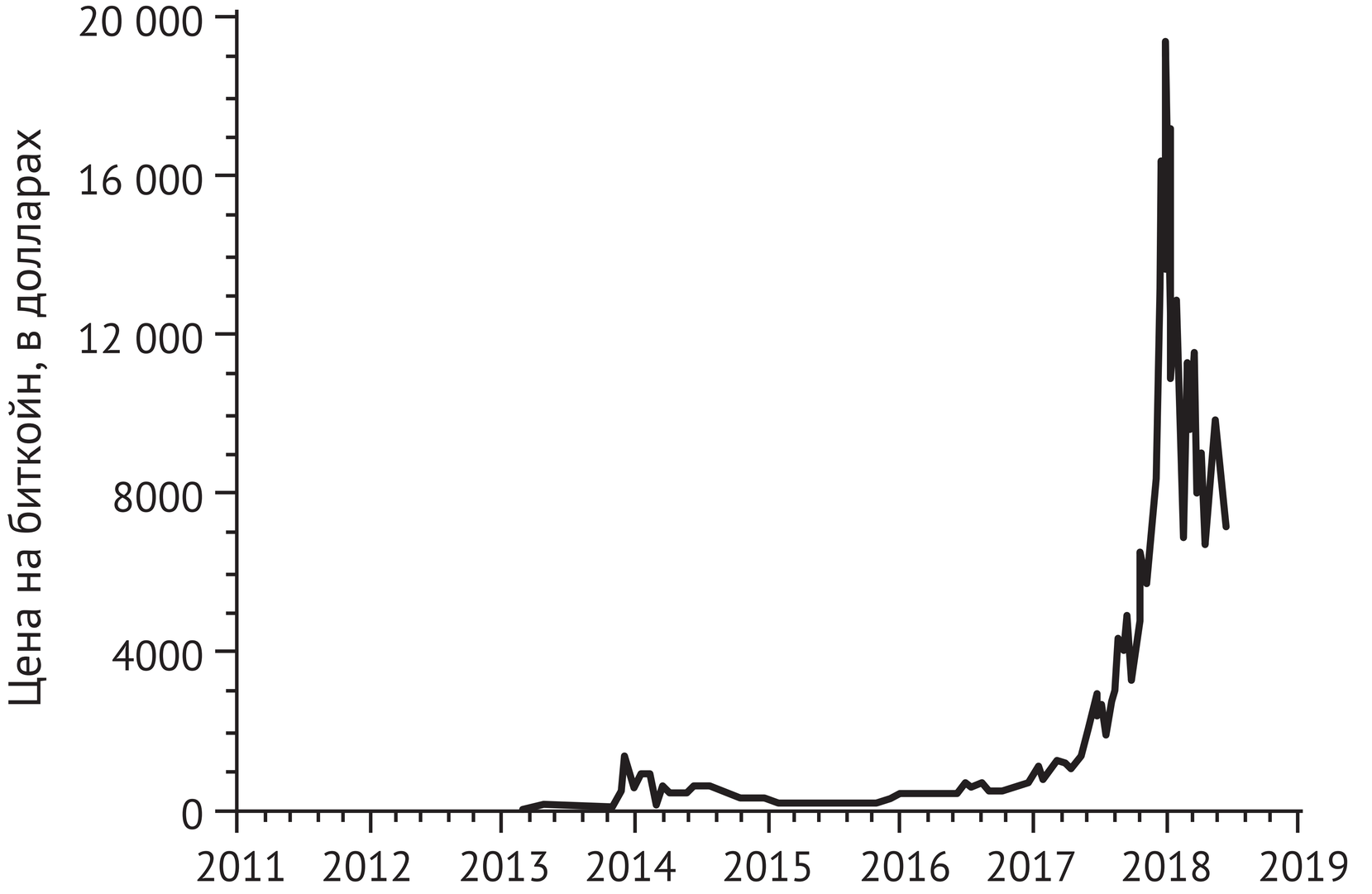

Рис. 7.1. Пузырь биткойна

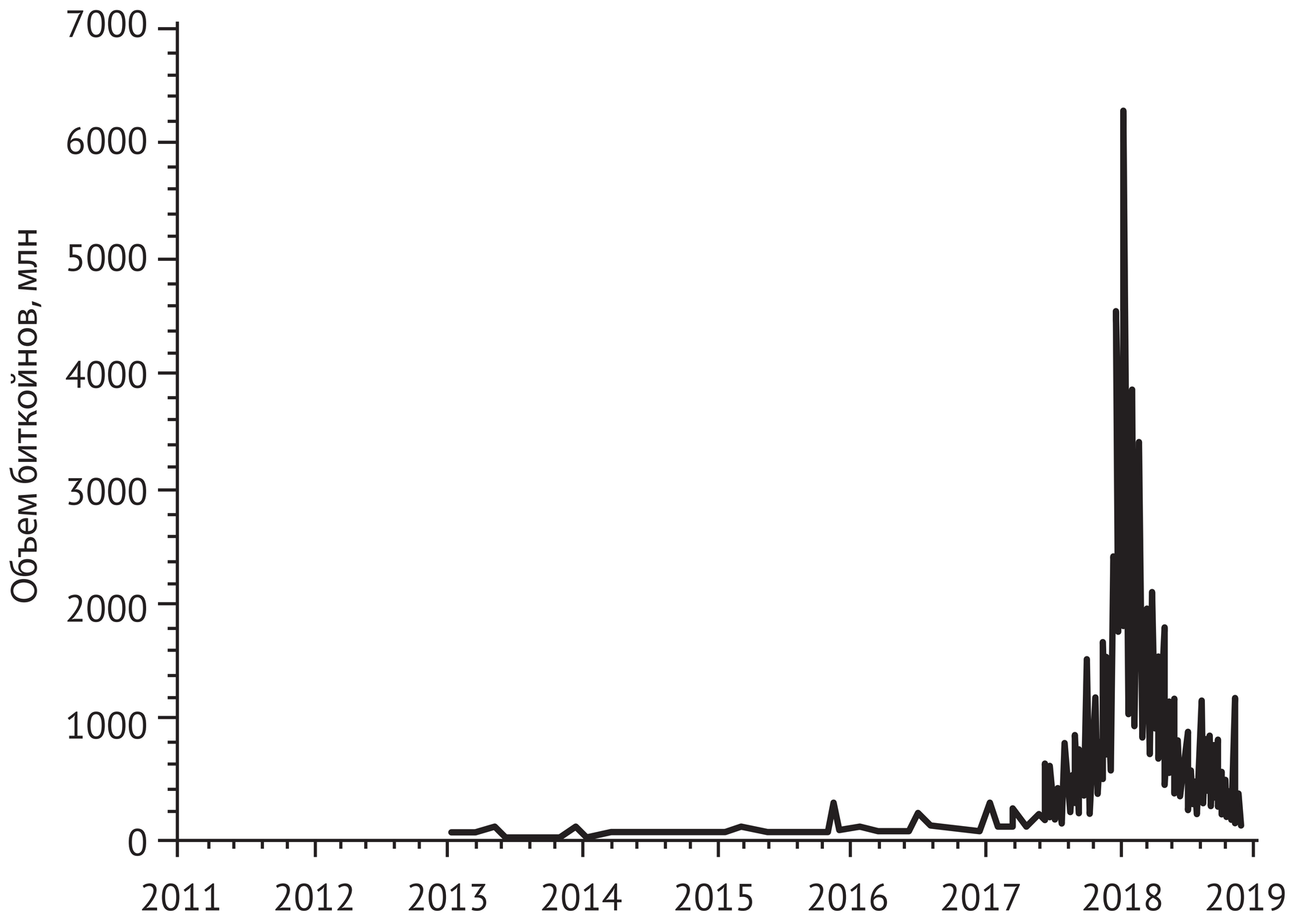

Рис. 7.2. Биткойн-мания

Учитывая, что бóльшую часть изученного Лю и Цывински периода времени биткойн представлял собой спекулятивный пузырь, вполне разумно ожидать сильного моментума в ценах. Однако, как только пузырь лопнет, моментум испарится и больше не будет полезен для предсказания движений цен на биткойн.

Как обсуждалось в главе 6, Лю и Цывински рассмотрели вопрос о том, насколько точно поисковые запросы Google по термину «биткойн» предсказывают цены на биткойн. Они также развернули направление зависимости вспять, чтобы увидеть, насколько точно цены на биткойн предсказывают еженедельные запросы Google по термину «биткойн». Влияние цен на поисковые запросы Google по термину «биткойн» не является особенно полезным, но, по меньшей мере, есть правдоподобное объяснение такой взаимосвязи – в отличие от подавляющего большинства переменных, которые они проанализировали.

Подсчитайте их

В табл. 7.1 подытожены результаты. В целом Лю и Цывински оценили 810 корреляций между ценами на биткойн и различными переменными и обнаружили 63 взаимосвязи, которые были статистически значимыми на уровне 10 %. Это несколько меньше, чем 81 статистически значимая взаимосвязь, которые можно было бы ожидать, если бы они просто откоррелировали (соотнесли) цены биткойна со случайными числами.

Таблица 7.1. Оценочные коэффициенты биткойна

	Количество
Оценочные коэффициенты	810
Коэффициенты, значимые внутривыборочно	63 из 810
Коэффициенты, значимые (и с одинаковыми знаками) внутривыборочно и вневыборочно	7 из 63

Семь из этих 63 корреляций продолжали иметь те же знаки и быть статистически значимыми вневыборочно. Пять из этих семи корреляций были для уравнений, в которых цены на биткойн использовались для предсказывания поисковых запросов по термину «биткойн», и указанные корреляции, похоже, являются наиболее логически правдоподобными взаимосвязями из рассмотренных ими. По иронии судьбы этот вывод является аргументом против проведенной ими глубокой переработки данных и в пользу ограничения внимания логическими взаимосвязями – именно они остаются неизменными.

Что касается сотен других рассмотренных ими взаимосвязей, то менее 10 % были значимыми внутривыборочно на 10%-ном уровне и менее 10 % из этих 10 % продолжали оставаться значимыми вневыборочно. Разумеется, при наличии достаточного количества данных всегда можно найти случайно возникающие закономерности, и по чистой удаче некоторые из этих случайно возникших закономерностей будут продолжаться вне выборки. Следует ли сделать вывод, что, поскольку возвратность биткойна оказала статистически значимое негативное влияние на возвратность акций в индустрии картон-

ных контейнеров и коробок, что было подтверждено вневыборочными данными, была обнаружена полезная, наполненная смыслом взаимосвязь? Или же необходимо заключить, что эти находки соответствуют тому, чего можно было бы ожидать, если бы во всех оцененных уравнениях использовались случайные числа со случайными метками вместо реальных экономических переменных?

Авторы не пытаются объяснить обнаруженные ими взаимосвязи: «Мы не даем объяснений, мы просто документируем это поведение». Закономерности без объяснений – вероломны. Поиск закономерностей в крупных базах данных почти наверняка будет приводить к их обнаружению, а обнаруженные закономерности, скорее всего, будут исчезать, когда результаты будут использованы для составления предсказаний. Какой смысл документировать исчезающие закономерности?

Отчет Чарни

В ходе одного исследования был сделан вывод, что «Теперь у нас есть неопровержимые свидетельства в пользу того, что атмосфера действительно меняется и что мы сами вносим свой вклад в эти изменения. Концентрация углекислого газа в атмосфере неуклонно растет, и эти изменения связаны с использованием человеком ископаемых видов топлива и эксплуатацией земель». Вас можно простить за ваше предположение, что это недавний научный ответ отрицателям изменения климата, но вы ошибаетесь; исследование было опубликовано в 1979 году.

Доклад назывался «Углекислый газ и климат: научная оценка» (Carbon Dioxide and Climate: A Scientific Assessment) и был подготовлен для Национальной академии наук исследовательской группой под руководством Джул Чарни (Jule Charney). В нем содержалось тревожное предсказание: если концентрация CO_2 удвоится, то «более реалистичные результаты моделирования предсказывают глобальное потепление земной поверхности на 1,5–4,5 °C с более высоким увеличением в высоких широтах». В докладе также содержится тревожное предсказание о том, что потепление может быть отложено из-за способности океана удерживать тепло, но достижение предсказываемой температуры теплового равновесия будет неизбежным. Другими словами, даже когда станет ясно, что глобальное потепление является серьезной проблемой и человечество станет углеродно-нейтральным, потепление будет продолжаться десятилетиями.

Скептицизм был понятен сорок лет назад, когда не было самоочевидным, что в мире наблюдается потепление. В то время даже появилось несколько статей, предупреждавших о «глобальном похолодании». Тем не менее большинство климатологов сочли свидетельства убедительными. Сегодня, сорок лет спустя, можно проверить их предсказание и посмотреть, как оно оправдалось. На рис. 7.3 показаны годовые глобальные температуры земной поверхности, начиная с 1880 года, насколько позволяют данные НАСА, и указан год публикации этого исследования. Указанные данные представлены в виде отклонений от средней температуры за тридцатилетний период с 1951 по 1980 год.

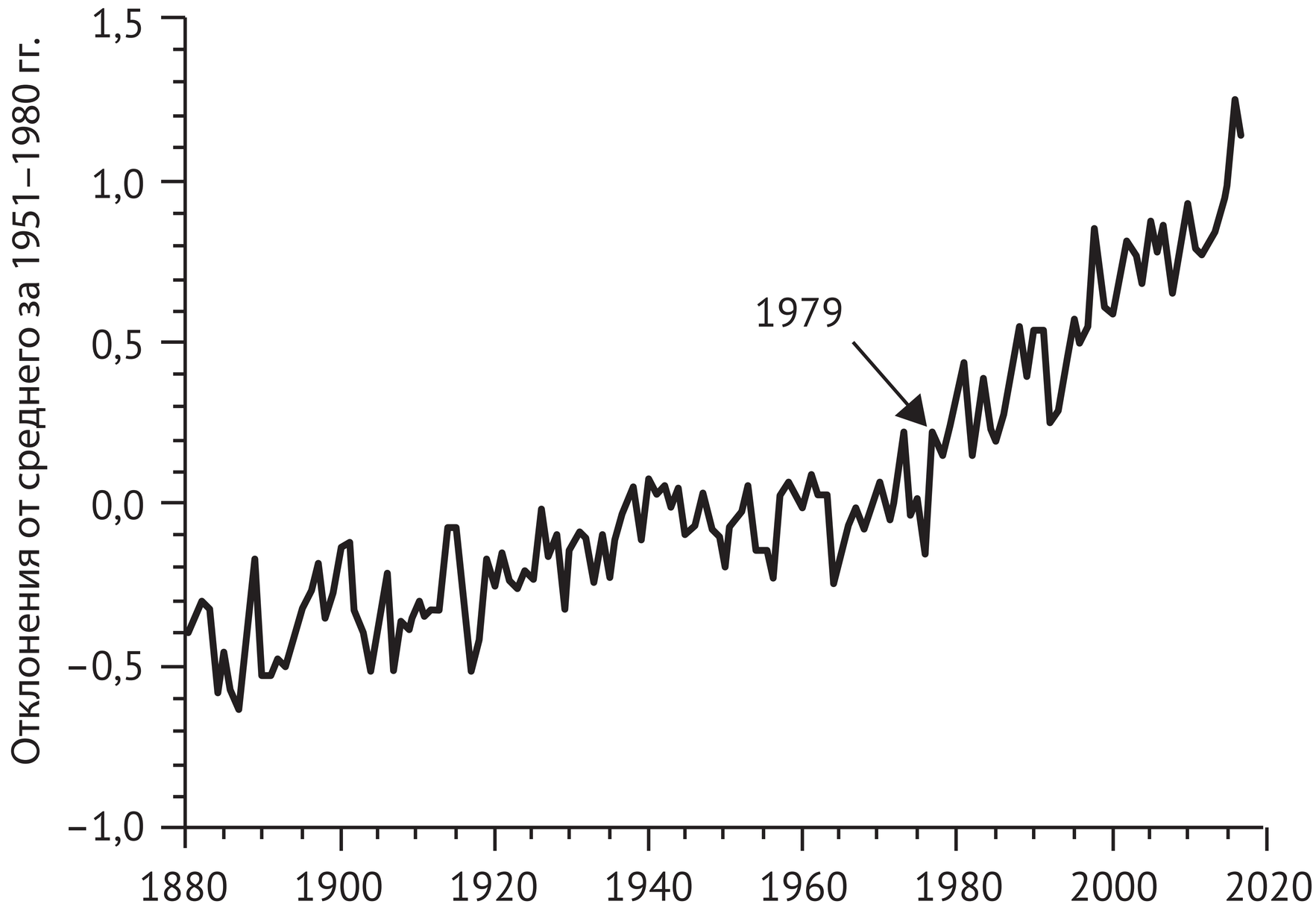

Рис. 7.3. Распознавание тренда до того, как он произойдет

С 1979 года уровень выбросов CO_2 увеличился на 21 %, а средняя глобальная температура на поверхности повысилась на 0,66 °C, что соответствует предсказанному в отчете Чарни повышению глобальной температуры на поверхности на 3 °C, сопровождающемуся 100%-ным увеличением выбросов CO_2 – как раз в середине диапазона, предсказанного в отчете Чарни. Как этим ученым удалось подготовить отчет, в котором содержался надежный сорокалетний прогноз?

Первое, что говорило в пользу группы, – это то, что ее возглавлял эксперт, которому было мало равных как в техническом, так и в научном плане. На самом деле если бы вам нужно было выбрать человека, которого вы хотели бы поблагодарить за то, что он заложил основу для современных успехов в области предсказания погоды, то вы, скорее всего, выбрали бы Джула Чарни. Чарни родился в Сан-Франциско в первый день Нового 1917 года, а в 1946 году защитил докторскую диссертацию в Калифорнийском университете в Лос-Анджелесе. Затем в 1950 году он работал над самыми ранними погодными моделями в Принстоне, где его уравнения произвели революцию в области численного предсказания погоды (в том числе в том, как предсказывать циклоны). В течение двадцати пяти лет работы в Массачусетском технологическом институте он был наставником бесчисленного количества студентов, продолжая проводить новаторские исследования.

Следующее, что группа сделала правильно, – это выдвинула гипотезу и предсказание, основанные на физике и химии, а не на поиске (возможно, ложных) закономерностей в данных. Есть соблазн отыскать тренд и придумать историю, объясняющую причину, по которой он будет продолжаться, но серьезная наука поступает наоборот: сначала теория, а потом данные.

Последним ингредиентом, который способствовал успеху этой вечной статьи, была здоровая доза «не обманывать самих себя». Авторы были осмотрительны в том, чтобы не преувеличивать то, что они могли с уверенностью предсказать, и рассматривали все возможные причины, которые могли бы опровергнуть их предсказания. К сожалению, Чарни умер в 1981 году, так и не дождавшись того момента, когда предсказания его коллектива подтвердились.

Скептицизм против отрицательства

Тот факт, что исследовательские работы иногда содержат неряшливую научную информацию, не означает, что следует отвергать общепринятые научные выводы. Кризис воспроизводимости поощряет скептицизм, но не отрицательство.

Мы должны скептически относиться к отдельным исследованиям, которые по-разному ущербны, но мы также должны скептически относиться к заявлениям, которые преувеличивают научную неопределенность или отвергают целые области исследований в целях продвижения политической или экономической повестки дня. Что касается изменения климата, то во всем мире существует около 200 научных организаций, которые поддерживают точку зрения о том, что изменение климата вызвано действиями человека, и ни одна из них не отвергает ее. (Неудивительно, что Американская ассоциация геологов-нефтяников была последним несогласным, но в 2007 году перешла от противостояния консенсусу к уклончивости.)

Конечно, наука – это не опрос общественного мнения. Книга 1931 года, в которой теория относительности Альберта Эйнштейна (Albert Einstein) оценивалась скептически, называлась «100 авторов против Эйнштейна». Ответ Эйнштейна был таков: «Почему 100? Если бы я был неправ, то одного было бы достаточно». Аргументы в пользу глобального потепления основаны не на результатах опроса, а на совокупности данных, полученных по нескольким направлениям исследований: пыльца, годичные кольца деревьев, керны льда, коралловые рифы, таяние ледниковых и полярных шапок, повышение уровня моря, экологические сдвиги, увеличение содержания углекислого газа, беспрецедентный рост температуры – все это сходится к одному выводу.

Кризис воспроизводимости реален, но он затрагивает не все области в равной степени и не должен использоваться в политических целях. Мы все выиграем, если наука останется плодотворной и продуктивной и будет изолирована от идеологов. Будьте осторожны с группами, которые преувеличивают неопределенность в целях продвижения своих собственных интересов. Табачная промышленность разработала методичку по этому вопросу, и сегодня ей следуют отрицатели изменения климата.

Научное сообщество реагирует на кризис воспроизводимости, устанавливая процедуры для выявления и пресечения различных нарушений, подобных выявленным комитетом по расследованию деятельности Дидерика Стапеля. Многие журналы теперь требуют, чтобы планы исследований представлялись до начала исследования, и настаивают на том, чтобы все используемые в исследовании данные были общедоступными. Научная деятельность – это одно из

направлений человеческой деятельности, в котором долгосрочный прогресс неоспорим, и так должно продолжаться и впредь.

Как не быть введенным в заблуждение призрачными закономерностями

Попытки воспроизвести опубликованные исследования нередко оказываются безуспешными, потому что в основе исследований лежит глубокая переработка данных – поиск закономерностей в данных без каких-либо заранее определенных, связных теорий. Опасности глубокой переработки данных могут усугубляться за счет издевательства над данными – нарезки, разбивки и других способов искажения данных с целью создания закономерностей. Если для какой-либо закономерности нет глубинного основания, то она, скорее всего, исчезнет, когда кто-то попытается повторить исследование.

Большие данные и мощные компьютеры – это не решение, а часть проблемы, поскольку они могут легко выявлять практически неограниченное количество призрачных закономерностей и взаимосвязей, которые исчезают при встрече со свежими данными.

Если исследователю выгодно его заявление, то, скорее всего, оно будет предвзятым, систематически смещенным. Если заявление звучит неправдоподобно, то оно, вероятно, вводит в заблуждение. Если статистические свидетельства кажутся слишком убедительными, чтобы быть правдой, то, вероятно, так оно и есть.

Глава 8

Кто-то вмешался?

В эпоху величия Швеции (1611–1718) несколько амбициозных королей превратили сельскую глушь в военную державу. Шведское огнестрельное оружие, изготовленное из меди и железа, было исключительным. Шведские корабли, изготовленные из первозданной древесины, внушали страх. Шведские солдаты вели себя героически. Швеция правила Прибалтикой.

В 1625 году король Густав II разместил заказы на четыре военных корабля на верфях Хибертссона в Стокгольме. В то время в кораблестроении было мало теоретических знаний. Конструкторы не знали, как определять центр тяжести или устойчивость различных конструкций. Это в буквальном смысле делалось методом проб и ошибок – узнавая, что работает, а что нет, наблюдая, какие корабли годятся для плавания, а какие опрокидываются и тонут. Хенрик Хибертссон (Henrik Hybertsson) был мастером-кораблестроителем на верфях Хибертссона, и он начал строительство традиционного 108-футового корабля Васы для короля без подробных планов или даже грубого эскиза, поскольку он и строители были хорошо знакомы с такими кораблями.

Большинство военных кораблей того времени перевозили сотни солдат и имели единственную орудийную палубу, вооруженную небольшими пушками. Традиционная военная стратегия заключалась в том, чтобы с помощью пушек выводить из строя вражеский корабль, который затем солдаты брали на абордаж. Король Густав предпочел альтернативную стратегию – использовать пушечный огонь для потопления вражеских кораблей, и он изменил заказ на 120-футовую Васу, вооруженную 3000-фунтовыми пушками, стреляющими 24-фунтовыми ядрами.

Узнав, что Дания строит корабль с двумя орудийными палубами, король приказал пристроить к Васе вторую орудийную палубу и увеличить длину киля судна до 135 футов. В дополнение к добавлению пушек высокая орудийная палуба обеспечивала бы бóльшую дальность полета пушечных ядер. 111-футовый киль уже был построен, поэтому Хибертссон расширил проект, по-прежнему работая без подробных планов. Хибертссон никогда не строил корабль с двумя орудийными палубами, но предполагалось, что это будет простой экстраполяцией 108-футового корабля с одной орудийной палубой.

111-футовый киль состоял из трех кусков дуба, соединенных вместе вплотную. К одному концу был добавлен четвертый кусок дуба, чтобы увеличить высоту киля Васы до 135 футов. Одна из проблем заключалась в том, что ширину и

глубину киля тоже следовало увеличить, но делать это было уже слишком поздно, что сделало судно менее устойчивым. Кроме того, для размещения второй орудийной палубы была расширена верхняя часть корабля, но нижнюю часть расширить было невозможно, что опять-таки делало судно менее устойчивым, а также ограничивало пространство для балласта, что могло привести к снижению центра тяжести корабля. Даже если бы они смогли добавить достаточно балласта, чтобы существенно опустить центр тяжести, это привело бы к тому, что первая орудийная палуба оказалась бы ниже уровня моря.

Вместо тридцати двух пушек на одной палубе Васа несла сорок восемь пушек, по двадцать четыре на каждой палубе, что повышало центр тяжести корабля. Король также приказал установить сотни позолоченных и раскрашенных резных фигур из дуба высоко на корабле, чтобы их могли видеть вражеские солдаты; это тоже повысило центр тяжести. Вы знаете, к чему ведет эта история (и этот корабль).

Окончательным эмпирическим злоключением стало то, что один рабочий использовал линейки, калиброванные в шведских футах, которые делятся на двенадцать дюймов, а другие рабочие использовали линейки, калиброванные в амстердамских футах, которые делятся на одиннадцать дюймов, и поэтому для разных рабочих шесть дюймов означали разные вещи. В результате Васа оказалась перекошенной – с левого борта она была тяжелее, чем с правого.

Когда 10 августа 1628 года Васа была спущена на воду в Стокгольмской гавани, ветер был настолько слабым, что команде пришлось убирать паруса вручную, надеясь поймать хотя бы небольшое легкое дуновение, чтобы продвинуть судно вперед. Двадцать минут спустя в 1300 футах от берега внезапный порыв ветра со скоростью восемь узлов опрокинул судно, и оно затонуло.

Дерево пережило под водой сотни лет из-за ледяной, бедной кислородом воды в Балтийском море, и 333 года спустя, в 1961 году, Васа была поднята со дна моря. После семнадцатилетней обработки консервантом она теперь выставлена в собственном музее в Стокгольме.

Васа – это трагический пример важности теории и опасного риска полагаться на закономерности, как это сделал Хайбертссон, когда он взял закономерность, которая подходила для 108-футовых суден, и попытался применить ее к 135-футовому судну. Теперь мы много знаем об устойчивости суден и знаем, что если взять конструкцию судна, подходящую для одного размера, и ее расширить, скажем, увеличив все размеры на 25 %, то это может закончиться катастрофой.

Васа не является аномалией. Призрачные закономерности вводят людей в заблуждение уже очень давно и стали причиной многих неловких ошибок и оплошностей.

Новая кола

Тысячи шарлатанов впаривают тысячи патентованных лекарств, называющихся так потому, что по утверждению их промоутеров секретная формула доказала свою эффективность и была запатентована. Как правило, ни то ни другое не соответствует действительности. Для того чтобы запатентовать лекарство, необходимо раскрыть секретную формулу. А для того чтобы доказать

эффективность, лекарство должно быть протестировано. Большинство эликсиров никогда не тестировались, а только рекламировались.

Патентованные лекарства изготавливались из экзотических трав, специй и других ингредиентов и, как считалось, излечивали практически от всего, включая расстройство желудка, туберкулез и рак. Лекарственные препараты, содержащие змеиное масло, даже стали источником сегодняшнего оскорбления: *торговец змеиным маслом*. Нередко первичным ингредиентом был алкоголь. А иногда там были морфий, опиум или кокаин.

Во второй половине XIX века огромной популярностью пользовалась смесь бордосского вина и листьев коки под названием «Вин Мариани», созданная французским химиком Анджело Мариани. В этом вине кокаин экстрагировался из листьев коки, что позволило получать мощное сочетание вина и кокаина. Вино Мариани употребляли королева Виктория, Томас Эдисон и тысячи выдающихся артистов, политиков и римских пап.

Красочные плакаты и реклама утверждали, что вино Мариани «укрепляет и освежает тело и мозг. Восстанавливает здоровье и жизненную силу».

Папа римский Лев XIII говорил, что он пил вино Мариани, чтобы «подкрепить себя, когда молитв было недостаточно». Он был настолько очарован этим вином из коки, что наградил его специальной медалью, которую Мариани быстро превратил в рекламную возможность.

В Соединенных Штатах бывший полковник армии конфедератов по имени Джон Пембертон (John Pemberton) был ранен во время гражданской войны и пристрастился к морфию. Он владел аптекой (Pemberton›s Eagle Drug and Chemical House) в Джорджии и в 1885 году разработал тонизирующее средство для нервов, названное им французской винной кокой Пембертона (Pemberton’s French Wine Coca), которое он, очевидно, создал по образцу вина Мариани и намеревался не только заменить морфий, но и добиться коммерческого успеха вина Мариани. В дополнение к вину и коке он добавил орехи кола, которые содержали кофеин. В 1886 году местные власти запретили производство, продажу или употребление алкогольных напитков, поэтому Пембертон заменил вино газированной водой и сахаром и назвал свой тоник «кока-кола», что отражает стимулирующее сочетание листьев коки и орехов кола.

Пембертон рекламировал кока-колу как «напиток для поддержания трезвости», который излечивает многие болезни, включая головные боли, зависимость от морфия и – в качестве бонуса – является «самым замечательным средством, укрепляющим половые органы».

Продажи были скромными, и в 1888 году еще один аптекарь, Эйса Григгс Кэндлер (Asa Griggs Candler), приобрел рецепт кока-колы и права на маркетинг, как сообщается, за 2300 долл. Кэндлер основал компанию Coca-Cola Company, которая впоследствии стала доминировать на мировом рынке безалкогольных напитков. Одним из его нововведений стала продажа кока-колы в бутылках. Это позволило людям пить кока-колу на улице, в том числе чернокожим, которым на Юге не разрешалось сидеть у газированных киосков.

Кэндлер внес несколько изменений в оригинальную рецептуру, чтобы улучшить вкус, а также отреагировал на негативную реакцию общественности против кокаина (кульминацией которой стало то, что в 1914 году в Соединенных Штатах кокаин был запрещен), убрав кокаин из напитков компании.

В 1929 году ученые компании Coca-Cola выяснили, как получить экстракт коки, полностью не содержащий кокаина. Управление по борьбе с наркотиками теперь разрешает компании Coca-Cola импортировать листья коки на охраняемый химический завод в Нью-Джерси, где производится жидкий экстракт, не содержащий кокаина, с маркировкой «Товар № 5».

Нынешняя формула кока-колы засекречена и тщательно охраняется, а ее секрет лежит в хранилище и, как говорят, известен только двум сотрудникам компании, которым запрещено путешествовать вместе и чьи личности тоже являются секретом.

В 1985 году, после 99 лет практически неизменного вкуса, компания Coca-Cola решила перейти с сахара на новый кукурузный сироп с высоким содержанием фруктозы, чтобы сделать вкус кока-колы более сладким и нежным – более похожим на вкус ее главного конкурента, пепси. Этому историческому решению предшествовал сверхсекретный опрос 190 000 человек стоимостью 4 млн долл., в ходе которого новая формула превзошла старую на 55 % к сорока пяти. Чего руководители компании Coca-Cola не учли – так это того, что многие из 45 % тех, кто предпочитал старую кока-колу, делали это с энтузиазмом. Пятьдесят пять процентов тех, кто проголосовал за новую колу, возможно, и смогли бы прожить со старой формулой, но, с другой стороны, многие клялись, что не смогут переварить новую колу.

Объявленные компанией Coca-Cola изменения вызвали возмущенные протесты и паническое накопление запасов среди поклонников старой кока-колы. Вскоре компания Coca-Cola пошла на попятную и вернула старую колу под названием «Классическая кола». Некоторые циники предположили, что весь этот сценарий был спланирован компанией Coca-Cola как хитроумный способ получить бесплатную рекламу и создать, по словам старшего вице-президента по маркетингу, «потрясающую связь с нашей публикой». В 1985 году новая кола заняла 15,0 % всего рынка безалкогольных напитков, классическая кола – 5,9 %, а пепси – 18,6 %. В 1986 году доля новой колы упала до 2,3 %, классическая кока-кола выросла до 18,9 %, а пепси прочно удержалась на уровне 18,5 %.

В 1987 году газета The Wall Street Journal заказала опрос 100 случайно отобранных потребителей колы, из которых 52 заранее объявили себя сторонниками пепси, 46 – приверженцами классической колы и двое – новичками в употреблении колы. В ходе проведенного журналом слепого дегустационного теста победителем стала новая кола, набрав 41 голос, за ней последовали пепси с тридцатью девятью и классическая кола с двадцатью. Семьдесят из ста участвовавших в дегустационном тестировании человек ошибочно решили, что выбрали свой любимый бренд; некоторые были возмущены. Любитель классической кока-колы, выбравший пепси, сказал: «Я не опущусь до того, чтобы пить пепси. Это слишком модно. Слишком “да”. Новое поколение – это звучит как нацистская пропаганда. Кока-кола – более непринужденная». Один из выбравший кока-колу поклонников пепси сказал: «Я отношу кока-колу к людям, которые просто принимают существующий статус-кво. Думаю, что Pepsi более бунтарская компания, и во мне есть что-то бунтарское».

В 1990 году компания Coca-Cola выпустила новую колу под новым названием – «Кола II» – и новую банку, окрашенную в традиционный для компании Pepsi синий цвет. Она была не более успешной, чем новая кока-кола.

Руководители компании Coca-Cola и многие другие представители индустрии безалкогольных напитков по-прежнему убеждены, что любители кока-колы предпочитают вкус новой колы, даже несмотря на то, что они по-прежнему беззаветно преданы старым коле и пепси, и эта лояльность объясняется скорее рекламными кампаниями, чем вкусом. Учитывая миллиарды долларов, которые компании тратят на то, чтобы убедить потребителей в том, что имидж колы является важной частью их вкусовых ощущений, слепые дегустационные тесты могут оказаться просто неуместными.

Данные, которые изобилуют ошибками и пропусками или просто не имеют отношения к делу, являются причиной того, что афоризм «мусор на входе, мусор на выходе» является таким известным. В 2016 году IBM подсчитала, что ежегодные расходы на принятие решений, основанных на неверных данных, составили 3,1 трлн долл. Охмурение данными есть следствие введения в заблуждение призрачными закономерностями.

Передача Биллу Клинтону правды

За две недели до выборов в конгресс в 1994 году еженедельник The New Yorker опубликовал статью о шансах демократической партии сохранить контроль над конгрессом. Главные советники президента Билла Клинтона были уверены в непопулярности республиканцев среди избирателей – их уверенность была основана на обширных опросах, свидетельствующих о сильном неодобрении. Но автор статьи заметил, что:

> … опросчики-демократы сформулировали ключевые вопросы таким образом, чтобы получать ответы, которые президент, по-видимому, хотел услышать. Когда опросчики простым, без прикрас, нейтральным языком спрашивали об основных идеях республиканской повестки дня – снижении налогов, сбалансированном бюджете, ограничении сроков полномочий, усилении обороны и т. д., – большинство респондентов их одобряли.

Одна из участниц президентского опроса объяснила, что ей пришлось «сформулировать вопрос очень четко», чтобы получить желаемые ответы. Так, один из заданных президентскими опросчиками вопросов был таким:

> [Кандидат от Республиканской партии X] подписал контракт с республиканским руководством в Вашингтоне, заявив, что он поддержит дальнейшее снижение налогов для богатых и увеличит расходы на оборону, которые будут оплачены за счет сокращения социального обеспечения и медицинской помощи. Именно такую провальную политику мы наблюдали в течение двенадцати лет, помогая богатым и представителям особых интересов за счет среднего класса. Вы ее одобряете?

Когда журналист The New Yorker высказал предположение высокопоставленному советнику Клинтона, что такие сложные вопросы дают президенту искаженное представление о мнении избирателей, он ответил: «Разве не в этом заключается суть опроса?» Еженедельник The New Yorker также процитировал неизвестного «главного политического стратега» демократов, который не входил в ближайшее окружение Клинтона:

> Президент и его политические соратники не понимают, что здесь произошло. Никто из них никогда не выходит из этого здания. Они приходят туда в 7 утра и уходят в 10 вечера, и ни разу не выходят. Они живут как в коконе, в своем собственном диснеевском мире. Они ходят по дому, бледные и изможденные, сжимая в руках свои бумаги, бегают с собрания на собрание и понятия не имеют, что здесь происходит. Я имею в виду, ни малейшего понятия.

Выборы подтвердили догадку этого стратега и опровергли систематически смещенные опросы общественного мнения. Республиканцы получили восемь мест в сенате США и пятьдесят шесть мест в палате представителей, впервые за сорок лет получив контроль над конгрессом. Республиканцы также получили двенадцать губернаторских постов, что дало им в общей сложности тридцать мест (включая семь из восьми крупнейших штатов), и получили более 400 мест в законодательных органах штатов, что дало им большинство в семнадцати штатах, ранее контролировавшихся демократами.

Когда на следующий день после выборов президент провел пресс-конференцию, обозреватель газеты The Washington Post написал, что Клинтон был «больше похож на старого моряка, одержимого и болтливого». Еженедельник The New Yorker сообщил, что «на это было больно смотреть. <…> Заявления о дружбе и извинениях были сведены на нет общим тоном президента, полным непонимания и недоверия».

Вопреки мнению советников президента Клинтона, опросы общественного мнения – это не только постановка сложных вопросов. Клинтону было бы гораздо полезнее проводить опросы с четкими формулировками, которые дали бы ему и его советникам четкое представление о том, чего хотят избиратели. После провала 1994 года Клинтон сменил большинство своих опросчиков и, как только у него появилась более полная информация о том, что, собственно, беспокоит избирателей, одержал убедительную победу на переизбрании в 1996 году. Билл Клинтон был одним из величайших участников предвыборной кампании в истории США, но ему нужны были достоверные данные.

Секреты рынка акций

Рынок акций всегда был источником вдохновения для людей, которые ищут закономерности в данных. Данных много, и на полезной закономерности можно заработать много денег.

Проблема в том, что цены на акции не определяются физическими законами, такими как зависимость между объемом и давлением газа при заданной температуре. Цены на акции определяются готовностью инвесторов покупать и продавать акции.

Если о компании поступают хорошие новости (например, об увеличении ее заработков), то цена акций будет расти до тех пор, пока не будет достигнут баланс между покупателями и продавцами. В этом есть толк, но особенность новостей в том, что они по определению непредсказуемы. Если инвесторы знают, что что-то должно произойти, то это не станет новостью, когда это произойдет. Например, во время президентской избирательной кампании 1988 года было широко распространено мнение, что рынок акций выиграет больше от

президентства Джорджа Буша, чем от президентства Майкла Дукакиса. Однако 20 января 1989 года, в день вступления Джорджа Буша в должность президента, цены на акции немного упали. Инаугурация Буша была старой новостью. Все импульсы, которые президентство Буша дало рынку акций, произошли во время избирательной кампании, когда его победа стала более определенной. С точки зрения рынка акций инаугурация не была долгожданным событием.

Цены на акции также растут и падают из-за заразительных эмоций – инвесторы следуют за стадом и поддаются необоснованному ликованию или отчаянию – «животному духу» Кейнса. Любой, кто пережил рыночные пузыри и крахи, знает, что инвесторами иногда в массовом порядке овладевает безудержный оптимизм или безудержное уныние. Особенность жизнерадостности в том, что она тоже непредсказуема. Никто не знает, когда эмоциональные американские горки рынка акций начнут подниматься вверх или когда они внезапно развернутся и начнут свободное падение.

Нелепо думать, что рынок акций раздает деньги даром. Есть история о двух профессорах в области финансов, которые увидели на тротуаре стодолларовую купюру. Когда один профессор за ней потянулся, другой сказал: «Не беспокойтесь, если бы она была настоящей, ее бы уже кто-нибудь подобрал». Профессора в области финансов любят повторять, что финансовые рынки не оставляют стодолларовые купюры на тротуаре, имея в виду, что если бы существовал простой способ заработать деньги, то кто-нибудь уже давно бы догадался об этом. Но это не останавливает людей искать закономерности. Если отыскать надежную закономерность, то можно заработать огромные деньги. Среди бесчисленных обнаруженных (а затем признанных бесполезными) закономерностей есть следующие:

- рынок акций преуспевает в годы, оканчивающиеся на 5: 2005, 2015 и т. д.;
- рынок акций преуспевает в годы, оканчивающиеся на 8: 2008, 2018 и т. д.;
- рынок акций преуспевает в годы Дракона по китайскому календарю;
- нью-йоркский биржевой брокер выбирал акции, изучая газетные комиксы;
- биржевой брокер из Миннеаполиса разрешил своему золотистому ретриверу выбирать акции;
- индикатор снега в Бостоне (B. S.) основан на показателях снега в Бостоне в канун Рождества;
- суперкубковый предсказатель рынка акций на основан на данных о победителе Суперкубка.

Почему люди верят в такую бестолковщину? Потому что они легко поддаются обману призрачными закономерностями.

Наилучший месяц для покупки акций

Марк Твен предупреждал:

> Октябрь – это один из самых опасных месяцев для инвестирования в акции. Другие опасные месяцы – июль, январь, сентябрь, апрель, ноябрь, май, март, июнь, декабрь, август и февраль.

Если серьезно, то все-таки какие месяцы являются лучшими и худшими для покупки акций? Не теряющие надежды инвесторы изучают прошлые возвратности акций в поисках хороших и плохих месяцев. Веб-сайт Investopedia сообщает, что «средняя возвратность в октябре исторически является положительной, несмотря на рекордные падения на 19,7 и 21,5 % в 1929 и 1987 годах». Январский эффект говорит в пользу покупок в декабре. Ралли Санта-Клауса говорит в пользу покупок в ноябре. Что касается продаж, то некоторые говорят: «Продавай в мае и уходи». Другие считают, что худшие месяцы для акций – это август и сентябрь.

Такие советы выдают желаемое за действительное, когда преисполненные надежд инвесторы ищут способ рассчитать время входа рынок или выхода из него. Неудобная правда состоит в том, что не может быть постоянно лучшего или худшего месяца. Если бы декабрь был самым прибыльным месяцем для акций, то люди покупали бы в ноябре, временно делая ноябрь лучшим месяцем, что заставляло бы людей покупать в октябре, и т. д. Любая регулярная закономерность обречена на самоуничтожение.

Этот вывод противоречит интуиции, поскольку в нашей жизни существуют регулярные ежедневные, еженедельные и ежемесячные циклы. Солнце всходит утром и заходит вечером. Кукурузу сажают весной и собирают осенью. С возрастом люди становятся сильнее, а затем слабее.

Чарльз Доу (Charles Dow), вдохновитель получившей его имя теории, которая была популяризирована Уильямом Гамильтоном (William Hamilton) в его редакционных статьях в Wall Street Journal, считал, что на нашу жизнь влияют регулярные циклы в экономике и других сферах. И действительно, в первом предложении книги Гамильтона «Барометр рынка акций» цитируется теория экономических циклов «солнечных пятен». Если экономика переживает регулярные циклы, а цены на акции отражают состояние экономики, то представляется вероятным, что цены на акции тоже должны проходить предсказуемые циклы и что опытные инвесторы могут извлекать выгоду из распознавания этих циклов.

Однако тот факт, что некоторые компании придерживаются сезонных закономерностей, не означает, что их акции следуют таким же закономерностям. Если перед Рождеством продажи игрушек резко возрастут, будут ли цены на акции Mattel и других производителей игрушек расти каждый год перед Рождеством? А вот и нет. Когда акции Mattel торгуются летом, инвесторы принимают во внимание общеизвестный факт, что продажи игрушек растут в сезон отпусков. Они пытаются предсказать праздничные распродажи и соответственно оценить стоимость акций Mattel. Если их прогнозы окажутся верными, то нет никаких причин для изменения курса акций Mattel на момент публикации отчета о продажах. Акции Mattel вырастут или обвалятся только в том случае, если продажи окажутся неожиданно высокими или разочаровывающими.

Цены на акции растут или падают не потому, что сегодняшний день отличается от вчерашнего, а потому, что сегодняшний день не соответствует ожиданиям рынка – рынок удивлен. Неожиданности по определению предсказывать невозможно, как и краткосрочные колебания цен на акции.

Поскольку нет оснований для ежемесячной закономерности в неожиданностях, нет оснований и для ежемесячной закономерности в ценах на акции.

Закономерности, которые неизбежно обнаруживаются при тщательном всматривании в прошлое, являются не более чем временными случайными совпадениями. В 1990-х годах лучшим месяцем для рынка акций был декабрь. В 2000-х годах лучшим месяцем был апрель, а декабрь не стоил и выеденного яйца (шестой лучший месяц). В 2010-х годах октябрь был лучшим месяцем, а апрель занял лишь седьмое место. Эти расчеты лучших месяцев в прошлом примерно так же полезны, как и расчеты среднего телефонного номера.

Однако это не мешает людям составлять такие таблицы и думать, что их расчеты наполнены смыслом. Отчет исследовательской группы банка J. P. Morgan по рынку акций в Северной Америке за февраль 2019 года был озаглавлен «Сезонность показывает, что сейчас самое время покупать американские игровые акции». Авторы проанализировали ежемесячную возвратность от вложений в игровые акции за январь 2000 года и пришли к выводу, что «на наш взгляд, сейчас самое время покупать. Исторически сложилось так, что март и апрель были лучшими месяцами для приобретения игровых акций в США».

В некоторые месяцы возвратность обязательно будет выше, чем в другие, просто случайно. Такова природа непредсказуемых «американских горок», которые мы называем рынком акций, или фондовым рынком. Выявление месяца, который принес наибольшую возвратность в прошлом, вообще ничего не доказывает.

Мы провели небольшой эксперимент, чтобы продемонстрировать, что месячные закономерности могут быть призрачными. За двадцатилетний период с января 1999 года по декабрь 2018 года среднегодовая возвратность индекса цен на акции S&P 500 составила 6,65 %, а лучшим месяцем оказался март со среднегодовой возвратностью в 22 %. Похоже, это является убедительным свидетельством в пользу того, что март является лучшим месяцем для покупки акций.

Однако предположим, что мы берем 240 ежемесячных возвратов от инвестиций за этот двадцатилетний период и тщательно их перетасовываем по двенадцати категориям, которые мы назовем псевдомесяцами, – псевдоянварь, псевдофевраль и т. д. Каждый фактический возврат за месяц, скажем в марте 2008 года, будет получен в любом из этих псевдомесяцев с одинаковой вероятностью. Возврат в псевдоянваре с такой же вероятностью будет реальным июньским возвратом, как и реальным январским возвратом.

Мы повторили этот эксперимент миллион раз. В среднем самый результативный псевдомесяц показал среднюю возвратность в 23 %, что почти в точности соответствует реальной возвратности в 22 % в марте. В 84 % симуляций был хотя бы один псевдомесяц со средней возвратностью выше 20 %.

Напомним, что это не настоящие месяцы. Это псевдомесяцы. Да, некоторые псевдомесцы имеют более высокую среднюю возвратность, чем другие. Это замечание неизбежно и бесполезно, как и любая рекомендация покупать акции, основываясь на месяце года.

Оптимизация портфеля

Исторически сложилось так, что акции были очень прибыльными инвестициями, приносящими в среднем двузначную годовую возвратность, но цены на

акции также очень изменчивы. На рис. 8.1 и 8.2 показано, что годовая возвратность индекса S&P 500 сильно меняется из года в год, гораздо больше, чем у казначейских облигаций. Индекс S&P 500 сам по себе является среднерыночным показателем и скрывает еще бóльшую нестабильность в отдельных акциях. Акции могут удваиваться либо обесцениться за считанные минуты.

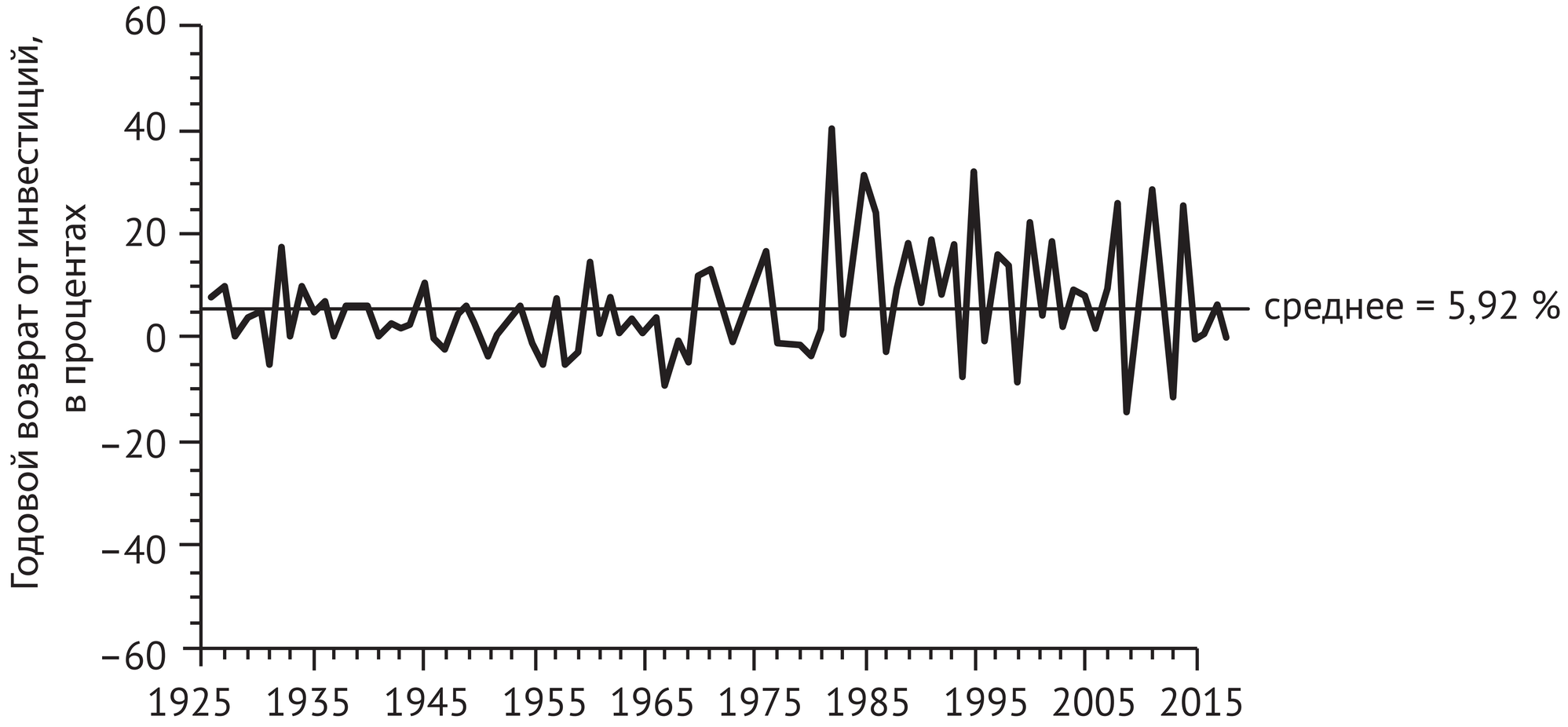

Рис. 8.1. Годовая возвратность долгосрочных казначейских облигаций с 1926 года

Рис. 8.2. Годовая возвратность акций S&P 500 с 1926 года

В 1950-х годах Гарри Марковиц (Harry Markowitz) и Джеймс Тобин (James Tobin) разработали аналитическую систему, позволяющую учитывать как риск, так и возвратность от инвестиций в акции. Опираясь на свои статистические данные, они в качестве правдоподобного приближения предположили, что ин-

весторы основывают свои решения на двух факторах – среднем значении и дисперсии ожидаемой возвратности. Среднее значение – это просто еще одно название средней ожидаемой возвратности. Поскольку это ожидаемое значение игнорирует неопределенность, они предложили использовать статистическую дисперсию в качестве меры риска.

Их подход стал известен как среднедисперсный анализ (Mean-Variance Analysis) или Современная портфельная теория (Modern Portfolio Theory), и он имеет ряд важных следствий. Одно из них заключается в том, что риск инвестора может быть снижен за счет выбора разнообразного (диверсифицированного) портфеля акций. Второй вывод заключается в том, что наиболее эффективно снизить риск можно, выбирая акции, возвратность которых коррелирует слабо, не коррелирует или даже отрицательно коррелирует друг с другом. Третий вывод заключается в том, что истинная мера риска для акции состоит не в том, насколько сильно колеблется ее цена, а в том, насколько эти колебания цен коррелируют с колебаниями цен на другие акции.

Акции, которые показывают плохие результаты, когда другие акции показывают плохие результаты, являются рискованными, поскольку не дают преимуществ диверсификации. Акции, которые движутся независимо от других акций или, что еще лучше, показывают хорошие результаты, когда другие акции показывают плохие результаты, снижают риск портфеля. Следовательно, адекватная мера риска должна учитывать корреляцию акций с другими акциями. Модель ценообразования капитальных активов (The Capital Asset Pricing Model, аббр. CAPM) является ответвлением среднедисперсного анализа, который позволяет это сделать.

Это адекватные и полезные модели, имеющие прочные теоретические корни. Проблема в том, что модели нуждаются в данных об ожиданиях инвесторов относительно будущего, а доступность исторических данных заставляет слишком многих допускать, что будущее повторит прошлое, – допускать, что будущие средние значения, отклонения и корреляции равны прошлым средним значениям, отклонениям и корреляциям. Люди, которые полагаются на исторические данные, неявно принимают допущение о том, что акции, которые показывали хорошую результативность в прошлом, будут показывать хорошую результативность и в будущем и что акции, которые были относительно безопасными в прошлом, будут относительно безопасными и в будущем. Это всего лишь изощренная форма поклонения закономерностям.

Математический аппарат среднедисперсного анализа пугает, а расчеты – трудоемки, поэтому вокруг готовых моделей управления портфелем ценных бумаг развилась целая индустрия. Специалисты по финансовому планированию и управляющие активами приобретают либо подписываются на программные продукты и интернет-службы, в которых для вычисления средних значений, отклонений и корреляций используют исторические данные, и рекомендуют портфели на основе этих исторических данных.

Историческая возвратность может вводить в заблуждение относительно будущего и приводить к созданию несбалансированных портфелей, которые вкладывают значительные средства в небольшое количество акций, показывавших хорошую результативность в прошлом, что противоречит диверсификации, рекомендуемой среднедисперсным анализом.

Например, в 2019 году Гэри обратил внимание на веб-сайт, предлагавший обучать людей работе с финансовыми данными. Веб-сайт был усовершенствован и включал интерактивную демонстрацию программы оптимизации портфеля, которая была представлена на конференции R/Finance для людей, использующих язык программирования R для создания финансовых алгоритмов. Веб-сайт хвастался тем, что «мы можем использовать управляемый данными анализ для оптимизации размещения инвестиций в корзине акций».

На веб-сайте упоминались такие впечатляющие модные термины, как «современная портфельная теория» и «модель ценообразования капитальных активов», и утверждалось, что «решение о размещении инвестиций становится автоматизированным с использованием современной портфельной теории. <...> [Программа] помогает управляющему активами ***принимать более качественные инвестиционные решения, которые будут постоянно улучшать финансовую результативность и, таким образом, удерживать клиентов***» (выделено жирным шрифтом и курсивом в оригинале).

Гэри опробовал эту программу оптимизации портфеля. Единственными требуемыми входными данными являются названия трех акций. Что может быть проще? Управляющим активами не нужно делать предсказания относительно будущей результативности, поскольку программа исходит из допущения, что будущая результативность будет такой же, как и прошлая. Если в прошлом акции-однодневки показывали действительно хорошие результаты, то программа оптимизации рекомендовала инвестировать в них значительные средства.

Гэри протестировал программу с тремя солидными компаниями (без однодневок): Google, IBM и Microsoft. Глядя на результативность этих трех акций за пятилетний период, с 2009 по 2013 год, программа-оптимизатор рекомендовала портфель из 63 % акций IBM, 23 % акций Google и 14 % акций Microsoft. Программа сообщила, что этот портфель опережает S&P 500 на четыре процентных пункта в год при меньшей волатильности.

Было заманчиво думать, что этот портфель позволит «принимать более качественные инвестиционные решения, которые будут постоянно улучшать финансовую результативность и, таким образом, удерживать клиентов», потому что он был отобран с помощью причудливого алгоритма. Однако указанный портфель не был основан на чьем-либо анализе будущего. Это был просто отчет о том, что указанный портфель опережал индекс S&P 500 на четыре процентных пункта в год в период с 2009 по 2013 год. Это называется *тестированием на истории* (бэк-тестированием) и представляет опасность для вашего благосостояния.

После свершившегося факта легко выявлять стратегии, которые были бы результативны в прошлом, но прошлая результативность имеет очень мало общего с будущей результативностью на рынке акций. IBM получила наибольший вес, а Microsoft – наименьший, потому что в период с 2009 по 2013 год IBM показала наилучшую результативность, а Microsoft – наихудшую. Это очень похоже на то, как если бы вы заметили закономерность в подбрасывании монеты и допустили, что эта закономерность продолжит проявляться.

Какую результативность показал этот утяжеленный акциями IBM портфель в течение следующих пяти лет, с 2014 по 2018 год? Он уступил S&P 500 на два

процентных пункта в год, потому что, как показано на рис. 8.3, стоимость Microsoft выросла втрое, в то время как IBM потеряла 28 % своей стоимости. Основываясь на этой новой результативности за пять лет, программа-оптимизатор изменила свои рекомендации с 63 % IBM и 14 % Microsoft на 76 % Microsoft и только 4 % IBM. Увы, это все еще отчет в отношении прошлого, а не надежное предсказание в отношении будущего.

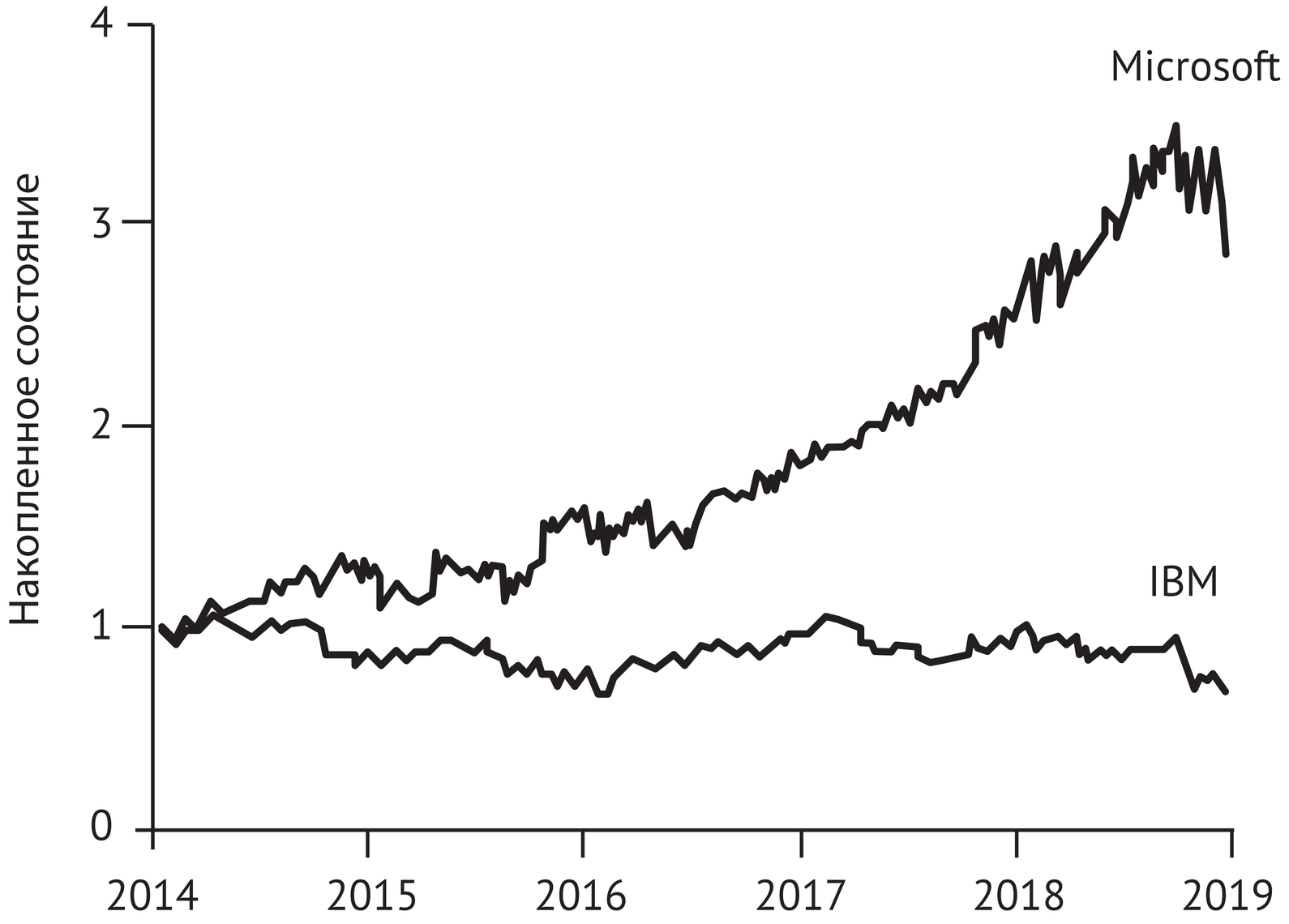

Рис. 8.3. IBM оказалась не лучшим вариантом для инвестирования

Легко выявлять акции, которые показали хорошую результативность, но трудно подбирать акции, которые будут результативными в будущем. Если кто-то допускает, что прошлые результативности акций являются надежным ориентиром для будущей результативности, то он одурачивает себя призрачными закономерностями.

Система месяца

Недавно Гэри получил электронное письмо с предложением получить доступ к сотням алгоритмов автоматической торговли: «Как в магазине приложений, только для алгоритмов». Компания утверждала, что «система месяца» для дневной торговли принесла бы прибыль в размере 82 000 долл. при инвестициях в размере 2600 долл. за предыдущие три с половиной года. Стоимость составляла 70 долл. в месяц, что в сумме составляет 840 долл. в год (42 % от первоначальных инвестиций!) плюс комиссия в размере 7,50 долл. за каждую сделку (а в системах дневной торговли сделок много). Система, должно быть, была невероятно успешной, чтобы преодолеть свои крутые затраты.

Было заявлено, что система полностью автоматизирована, бесконтактная, не вызывает эмоций и протестирована на истории.

И вот чем фишка. Она была основана на обратном тестировании. Системы, протестированные на истории, показывают впечатляющие результаты, потому что всегда можно найти систему, которая показала хорошие результаты в прошлом, если бы о ней знали заранее. Об этом написано мелким шрифтом:

> НЕ ДЕЛАЕТСЯ НИКАКИХ ЗАЯВЛЕНИЙ О ТОМ, ЧТО КАКОЙ-ЛИБО СЧЕТ ПРИНЕСЕТ ИЛИ МОЖЕТ ПРИНЕСТИ ПРИБЫЛЬ ИЛИ УБЫТКИ, АНАЛОГИЧНЫЕ ПОКАЗАННЫМ НА РИСУНКЕ. НА САМОМ ДЕЛЕ ЧАСТО СУЩЕСТВУЮТ РЕЗКИЕ РАЗЛИЧИЯ МЕЖДУ ГИПОТЕТИЧЕСКИМИ РЕЗУЛЬТАТАМИ И ФАКТИЧЕСКИМИ РЕЗУЛЬТАТАМИ, ДОСТИГНУТЫМИ ВПОСЛЕДСТВИИ С ПОМОЩЬЮ КАКОЙ-ЛИБО КОНКРЕТНОЙ ТОРГОВОЙ ПРОГРАММЫ.
>
> ОДНИМ ИЗ ОГРАНИЧЕНИЙ ГИПОТЕТИЧЕСКИХ РЕЗУЛЬТАТОВ ЯВЛЯЕТСЯ ТО, ЧТО ОНИ, КАК ПРАВИЛО, СОСТАВЛЯЮТСЯ С УЧЕТОМ ПРЕИМУЩЕСТВ ПРОШЛОГО ПЕРИОДА.

Кто бы сомневался!

Бросание кубиков на доллары

Некоторые компании продают компьютеризированные системы, чтобы обыграть рынок. Другие привлекают инвесторов, которые хотят, чтобы их деньгами управляли компьютеризированные системы. В обоих случаях результаты обычно разочаровывают.

Биржевой фонд (exchange-traded fund, аббр. ETF) – это пакет ценных бумаг, подобный взаимному фонду, который торгуется на Нью-Йоркской бирже ценных бумаг и других биржах как обычные акции. В настоящее время в 5000 ETF-фондов инвестировано почти 5 трлн долл.

Одной из их привлекательных особенностей является то, что, в отличие от акций обычных взаимных фондов, которые можно покупать или продавать только после закрытия рынков каждый день, акции ETF-фондов могут торговаться непрерывно, пока рынки открыты. Это, очевидно, привлекает инвесторов, которые думают, что могут впрыгивать в рынок и выпрыгивать из него, покупая до того, как цены вырастут, и продавая до того, как цены упадут. К сожалению, дневная торговля занимает много времени, а результаты часто разочаровывают. Возможно, компьютеры способны справиться лучше?

Термин «искусственный интеллект» (ИИ) уже стал дежурным. В 2017 году Ассоциация национальных рекламодателей выбрала «AI» (ИИ) в качестве маркетингового слова года. В октябре того же года компания с крутым названием EquBot запустила AIEQ, который претендовал на звание первого ETF-фонда на базе ИИ и не просто каким-либо ИИ, а ИИ с использованием мощной компьютерной системы IBM под названием Watson, которая победила лучших игроков-людей в американском игровом шоу Jeopardy.

Понятно? AI расшифровывается как artificial intelligence (искусственный интеллект), а EQ – как equity (собственный капитал). Сложите их вместе, и вы получите AIEQ, инструмент для подбора акций с помощью искусственного интеллекта. EquBot гордится тем, что AIEQ – это «инновационное приложение трех форм искусственного интеллекта»: генетических алгоритмов, нечеткой логики и адаптивной настройки. Инвесторы могут и не знать, что означают эти слова, но в этом и заключается привлекательность. Если кто-то говорит

что-то, чего мы не понимаем, естественно думать, что он умнее нас. Однако иногда загадочные слова и зашифрованные фразы предназначены для того, чтобы произвести впечатление, а не информировать.

Генеральный директор и соучредитель компании EquBot Чида Хатуа использовала обычный английский, но по-прежнему не вдавалась в подробности: «Технология искусственного интеллекта EquBot с Watson позволяет имитировать работу армии аналитиков рынка акций, работающих круглосуточно 365 дней в году, исключая при этом человеческие ошибки и предвзятость в процессе работы». Мы помним совет Уоррена Баффета (Warren Buffett) никогда не инвестировать в то, чего вы не понимаете. Мы также признаем, что было бы неплохо исключить человеческие ошибки и предвзятость (есть кто-нибудь, кто поддерживает ошибки и предвзятость?), но компьютеры, которые ищут закономерности, тоже допускают ошибки и имеют предвзятость (систематические смещения).

Компьютерные алгоритмы отбора кандидатов на работу, определения цен автострахования, одобрения заявок на получение кредита и определения сроков тюремного заключения – все они содержали существенные ошибки и систематические смещения, обусловленные не предвзятостями программистов, а природой закономерностей. Алгоритм компании Amazon для оценивания кандидатов на работу допускал дискриминацию в отношении женщин, которые учились в женских колледжах или состояли в женских организациях, поскольку в базе данных алгоритма среди действующих сотрудников было мало женщин. Алгоритм страховой компании Admiral Insurance для определения тарифов автострахования был основан на публикациях соискательниц в Facebook. Одним из приведенных компанией примеров был вопрос о том, нравится ли человеку Майкл Джордан или Леонард Коэн, – который люди распознают как полный ошибок и предвзятостей.

В Admiral Insurance сообщили, что их алгоритм:

> ...постоянно меняется с учетом новых свидетельств, которые мы получаем из данных. Таким образом, наши расчеты отражают то, как водители себя обычно ведут в социальных сетях и насколько их поведение предсказуемо, в отличие от фиксированных допущений о том, как может выглядеть безопасный водитель.

Это заявление было сделано для того, чтобы показать, что их алгоритм является гибким и инновационным. На самом деле оно показывает, что их алгоритм находит не полезные предсказатели, а исторические закономерности. Алгоритм постоянно меняется, поскольку не имеет логической основы и постоянно сталкивается с кратковременными корреляциями.

Основанные на закономерностях алгоритмы по своей природе предрасположены к обнаружению бессмысленных случайных совпадений, которые человеческая мудрость и здравый смысл распознали бы как таковые. Полное отстранение людей от процесса и сокрытие бессмысленного обнаружения закономерностей в непроницаемом черном ящике искусственного интеллекта вряд ли закончится хорошо.

На рис. 8.4 показано, как зарабатывал биржевой фонд AIEQ. Несмотря на обещания генетических алгоритмов, нечеткой логики и адаптивной настройки, фонд AIEQ, похоже, является «скрытным индексатором», отслеживающим

индекс S&P 500, но не достигающим его результативности. С момента создания и по 1 ноября 2019 года совокупная возвратность фонда AIEQ составила 25 %, по сравнению с 32 % для индекса S&P 500. В то же время за эту явно посредственную результативность компания получила годовое управленческое вознаграждение в 0,77 %.

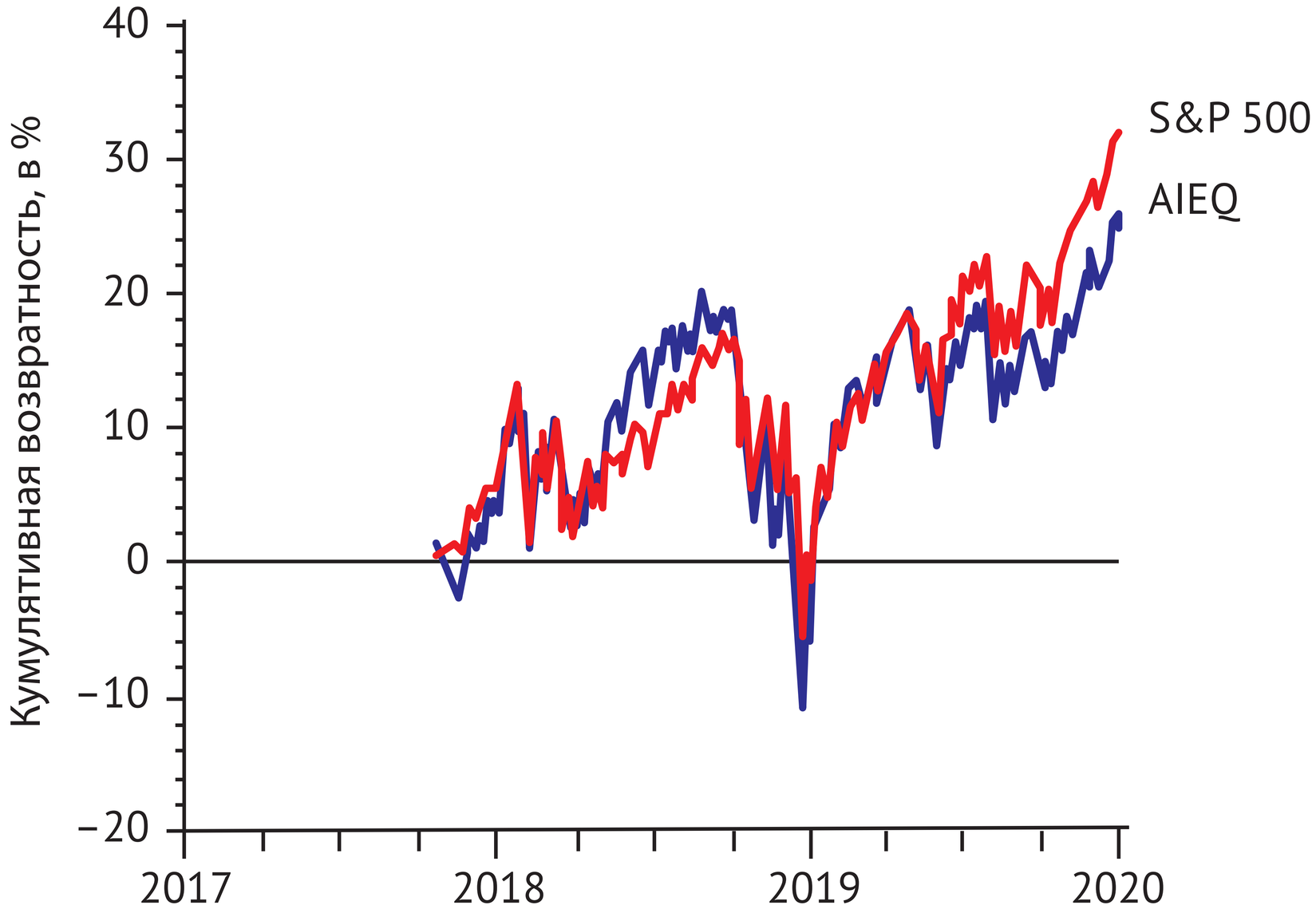

Рис. 8.4. Не вызывающая восторга результативность

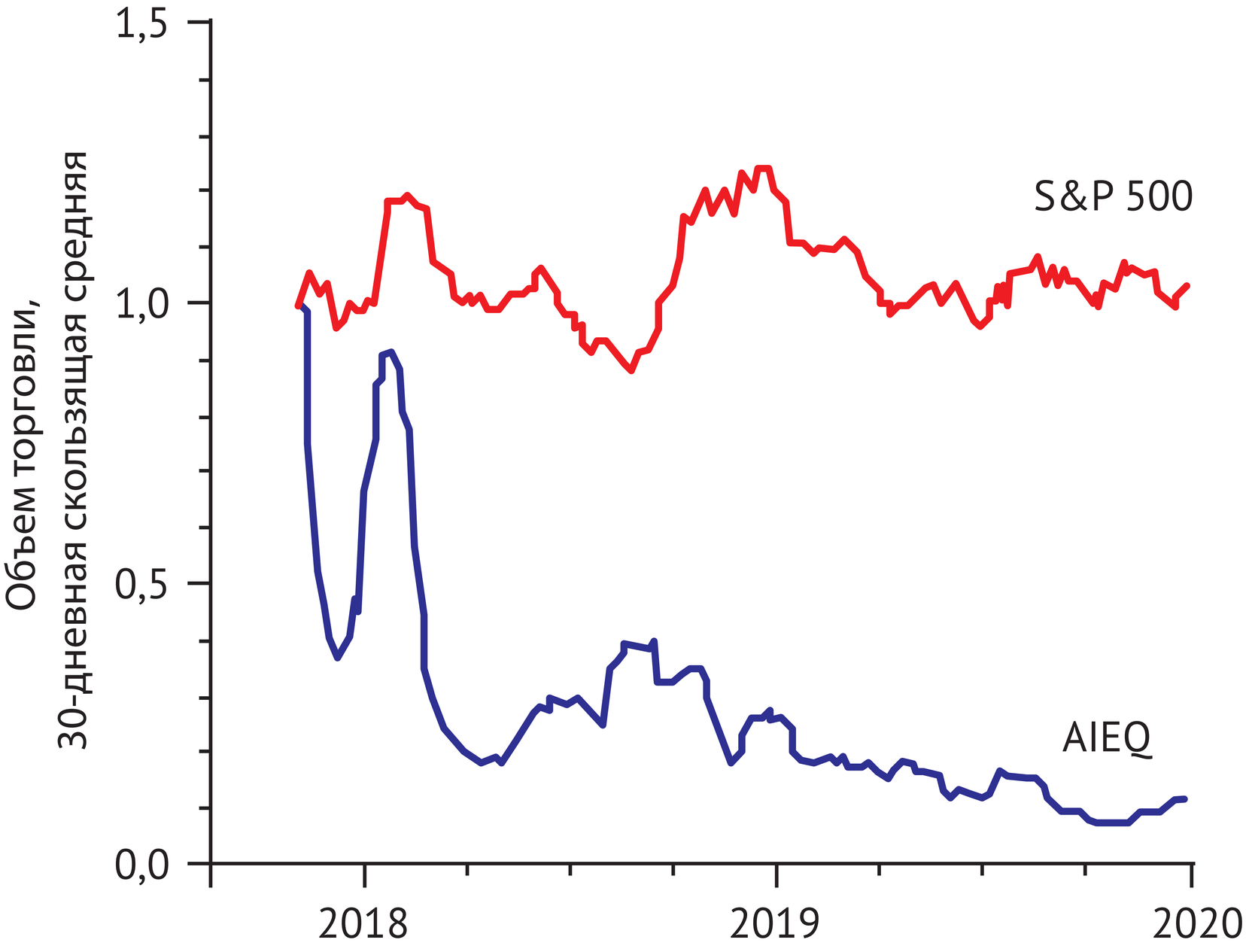

Рис. 8.5. Подавляющее отсутствие интереса

На рис. 8.5 сравнивается объем торгов в акциях фонда AIEQ с объемом торгов в акциях S&P 500, которые на момент запуска фонда AEIQ были равны 1. Как только стали очевидны разочаровывающие результаты, клиенты потеряли интерес.

В инвестировании и других сферах деятельности ярлык «искусственный интеллект» зачастую более эффективен с точки зрения маркетинга, чем с точки зрения результативности.

В чем подвох?

Обнадеженные инвесторы не единственные, кто влюбился в искусственный интеллект. В 2017 году, когда термин «искусственный интеллект» был признан маркетинговым словосочетанием года и был запущен фонд AIEQ, стартап One Concern собрал 20 млн долл., чтобы «обеспечить миру будущее».

Стартап One Concern[11] был вдохновлен опытом аспиранта Стэнфорда (Ахмада Вани), который во время наводнения навещал свою семью в Кашмире. Его семье пришлось семь дней ждать, пока их спасут из затопленной квартиры на третьем этаже, и они видели, как соседи гибли под обрушивающимися домами.

Когда Вани вернулся в Стэнфорд, он вместе с двумя другими студентами, Николь Ху и Тимом Фрэнком, продолжил работать над проектом по предсказанию воздействий, оказываемых на здания землетрясениями, которые в Кремниевой долине случаются чаще, чем наводнения. Они превратили этот проект в компанию One Concern и сообщили, что их алгоритм искусственного интеллекта способен менее чем за пятнадцать минут после начала землетрясения оценить ущерб, нанесенный каждому кварталу, с точностью до 85 %. В интервью журналу Time Вани сказал: «Наша миссия – спасать жизни. Как принимать наилучшие решения, чтобы спасти как можно больше жизней?»

Эта вдохновляющая история получила широкую огласку, и вскоре One Concern получил первые 20 млн долл. финансирования (с тех пор он получил гораздо больше) и несколько клиентов, в том числе в Сан-Франциско и Лос-Анджелесе. С тех пор компания выпустила сопутствующий продукт по борьбе с наводнениями Flood Concern (очевидно, существует не одна озабоченность).

Когда проект получил первоначальное финансирование в размере 20 млн долл., Вани написал, что:

> …наша технология на базе искусственного интеллекта будет назначать уникальный, верифицированный «цифровой отпечаток» каждому природному или рукотворному элементу, от мельчайшего камня до законченных сооружений, мегаполисов и в конечном счете всей планеты. One Concern позволит получать представление обо всем временном горизонте – будь то дни до наводнения, минуты после землетрясения или перспективная политика и планирование.

Скептика можно простить за то, что он думает, что это видение – брехня, за которой ничего нет, и этот скептицизм подкрепляется обещаниями на их отполированном веб-сайте:

[11] В переводе на русский «одна озабоченность». – *Прим. перев.*

> Наше долгосрочное видение заключается в обеспечении устойчивости в планетарном масштабе, когда каждый человек живет в безопасном, устойчивом и справедливом мире.

Госпожа Ху заявила, что «мы действительно считаем, что можем обеспечить будущее без стихийных бедствий».

Успешный алгоритм будет помогать направлять спасателей туда, где они необходимы, но как более эффективные спасательные операции помогут предотвращать стихийные бедствия и обеспечивать, чтобы «все жили в безопасном, устойчивом и справедливом мире»? Наводнения и землетрясения по-прежнему будут приводить к гибели людей, и какова актуальность модных слов «устойчивый и справедливый мир»?

Единственное их заявление, которое можно измерить, – это то, что их алгоритм может предсказывать повреждения от квартала к кварталу с точностью до 85 %. Число 85 выбрано удачно. Оно не настолько низкое, чтобы разочаровывать, но и не настолько высокое, чтобы казаться невероятным. Оно – точнее, чем 80 или 90, что говорит о том, что оно основано на научном исследовании.

Что означает точность в 85 %? Если речь идет о размере ущерба, нанесенного каждому кварталу, то точность в 85 % не имеет смысла. Если речь идет о предсказаниях того, какой квартал будет наиболее поврежден, то как агрегировать травмы, жизни людей и здания в меру ущерба?

Пожалуй, оно означает, что из тех зданий, которые алгоритм идентифицирует как поврежденные, 85 % оказываются поврежденными. Например, предположим, что 85 % зданий в районе повреждены, и алгоритм выбирает случайным образом 100 зданий, из которых 85 были повреждены. Это можно истолковать как 85%-ную точность, даже если алгоритм бесполезен. Или предположим, что 100 из 1000 зданий в районе повреждены, и алгоритм выбирает случайным образом 850 зданий, из которых 85 действительно повреждены. Это тоже можно истолковать как 85%-ную точность, даже если алгоритм бесполезен.

Алгоритм следует сравнить с предсказаниями, сделанными без участия алгоритма – контрольной группой. Работает ли он лучше, чем опытные и знакомые с районом местные спасатели или летающие над районом вертолеты? Когда Вани семь дней ждал спасения в Кашмире, возможно, проблема была не в том, что никто не знал о затоплении его района, а в том, что не было достаточного количества спасательных средств.

Возможно, One Concern – это просто одна брехня, за которой ничего нет.

В августе 2019 года газета New York Times опубликовала статью «Это высокотехнологичное решение для реагирования на стихийные бедствия может оказаться слишком хорошим, чтобы быть правдой», написанную Шери Финк, которая получила Пулитцеровскую премию за журналистские расследования и еще одну – за международные репортажи.

В Сан-Франциско закончился контракт с One Concern, а в Сиэтле возникли сомнения в стоимости и надежности программы. В ходе одной из тестовых симуляций программа проигнорировала крупный склад Costco, поскольку в основном опиралась на данные о жилых домах. Программа также содержала ошибки в оценках прочности зданий и учитывала каждую квартиру в здании как отдельную конструкцию. Когда компания One Concern внесла в програм-

му изменения, включив в нее пропавший склад Costco, таинственным образом исчез Вашингтонский университет. С каждой итерацией предсказания ущерба резко менялись.

One Concern убедила страховую компанию заплатить 250 000 долл. за использование алгоритма в Сиэтле. В свою очередь, страховая компания сможет воспользоваться модельными предсказаниями для «разработки наших страховых продуктов, а также для ценообразования». Это определенно звучит так, будто компания ищет основания для повышения ставок.

Один из бывших сотрудников рассказал Финку, что показатель точности в 85 % вводит в заблуждение и что «один из главных недостатков заключается в том, что он отвлекает внимание от людей, которые действительно нуждаются в помощи». Другой сказал, что его уволили за критику нечестного подхода компании «учись на ошибках и пробуй что-то новое», потому что при ликвидации последствий стихийных бедствий «если вы учитесь на ошибках, то люди умирают».

Ответ мистера Вани был показательным: «Мы ни в коем случае не говорим этим сотрудникам службы экстренного реагирования, что мы заменяем ваше суждение или способность к принятию решений». А разве это не то, что они рекламировали?

Другие отметили, что в рецензируемых журналах не публиковались тесты алгоритма и что ярлык «ИИ» вводит в заблуждение. Настоящие программы с искусственным интеллектом работают на обширных объемах данных, а данных о землетрясениях сопоставимой магнитуды в сопоставимых местах относительно мало. Если алгоритмы занимались экстраполяцией данных землетрясений определенной магнитуды на землетрясения другой магнитуды, землетрясений в определенных географических точках на другие местоположения и определенных типов зданий на другие здания, то они были введены в заблуждение призрачными закономерностями. Насколько адекватными будут данные о повреждениях зданий в результате землетрясения магнитудой 7,3 в Индонезии для предсказания ущерба от землетрясения магнитудой 6,5 в Сан-Франциско?

В конце концов мистер Вани пересмотрел показатель точности с 85 до 78 %, но не стал вдаваться в подробности относительно того, что это означает: «Вы знаете, мы даже не называем это “точностью”; мы называем это “ключевым показателем результативности”. Если бы вам пришлось отправлять службы экстренного реагирования после стихийного бедствия, скажем, для проведения поисково-спасательных работ в городах, то вы бы с точностью до 78 % или выше, или, по меньшей мере, более чем на 78 % с этим справились». Цифры разные, но все равно не имеют смысла.

В ходе проведенной компанией One Concern симуляции было предсказано, что земля под одним из калифорнийских шоссе расплавится в результате сильного землетрясения, но эта информация легкодоступна в штате. И схожим образом симуляции опираются на предсказаниях ущерба, сделанных другой компанией с использованием бесплатных правительственных данных, при этом генеральный директор прямо заявляет: «Это не искусственный интеллект».

Компания One Concern все активнее сотрудничает со страховыми компаниями и кредитно-рейтинговыми агентствами, которые могут использовать

предсказания One Concern для повышения индивидуальных страховых ставок и понижения рейтинга городских облигаций. Вместо спасения жизней их внимание, похоже, переключилось на обоснование более высоких страховых взносов и ставок по кредитам. Основываясь на интервью с бывшим сотрудником, Финк написал:

> Сдвиг в финансовой модели «казался очень лживым» и вызвал у многих сотрудников чувство разочарования, сказала Карин Понсе, бывшая исполнительная ассистентка и управляющая офисом.

В статье, опубликованной в журнале Fast Company в развитие предыдущей позитивной статьи, делается вывод о том, что:

> …как показывает расследование журнала Times, ранний успех стартапа – с венчурным финансированием в размере 55 млн долл. советниками, такими как генерал в отставке, бывший директор ЦРУ Дэвид Петреус, и членами команды, такими как бывший глава FEMA Крейг Фьюгейт, – был основан на вводящих в заблуждение заявлениях и отполированном дизайне. <…> Из-за неисправной технологии, которая, как сообщается, не так точна, как утверждает компания, одной из проблем может быть угроза жизни людей.

Поход по магазинам за доллары

Tesco – британская компания, которая выросла из небольшой сети супермаркетов в один из крупнейших ретейлеров в мире.

Трансформация началась в 1994 году, когда Tesco запустила программу лояльности Clubcard. Покупатели, владельцы клубных карт Clubcard, получают баллы, которые можно обменивать на товары в Tesco и партнерских ресторанах, гостиницах и других местах. Традиционная корпоративная привлекательность программ лояльности заключается в привлечении клиентов. Люди, которые имеют право на получение бесплатного кофе после покупки девяти чашек кофе в Julie›s Java, чаще будут покупать кофе в Julie›s.

Компания Tesco радикально изменила положение дел. Реальная ценность карты Clubcard для Tesco заключается в том, что она предоставляет подробную информацию о своих покупателях – точно так же как сеть магазинов Target смогла использовать покупательские привычки своих клиенток для выявления беременности и предсказания дат родов, чтобы «нацеливать» их на специальные предложения. Tesco заключила контракт с фирмой по глубокой переработке данных dunnhumby (причудливое британское название без заглавных букв, образованное фамилиями жены и мужа, Эдвины Данн и Клайва Хамби, которые основали фирму на своей кухне в 1989 году). Когда в 1994 году компания dunnhumby представила свой первоначальный анализ совету директоров Tesco, председатель совета директоров, насколько известно, сказал: «Что меня пугает в этом всем – так это то, что через три месяца вы будете знать о моих клиентах больше, чем я знаю за 30 лет».

Позже Tesco купила dunnhumby и ушла с головой в работу. Сегодня у половины всех британских домохозяйств есть карты Clubcard. Аналитики Tesco

используют данные карт Clubcard вместе с данными из социальных сетей, чтобы предсказывать предложения, которые будут наиболее привлекательными для отдельных клиентов. Одного покупателя могут привлекать скидки на мыло, но больше заботить качество помидоров; другой покупатель может испытывать прямо противоположные чувства. И поэтому первому клиенту отправляются скидочные купоны на мыло, а второму – скидочные купоны на помидоры.

С 1994 по 2007 год прибыль Tesco выросла в пять раз, а стоимость ее акций выросла в десять раз. В ноябре 2007 года Tesco начала открывать магазины Fresh & Easy в Соединенных Штатах, уверенная, что их качество обработки данных обеспечит им успех. Генеральный директор Терри Лихи похвастался, что «мы можем исследовать характеристики и сконструировать идеальный магазин для американского потребителя в XXI веке. Мы провели все наши исследования, и у нас это хорошо получается».

Исследователи компании Tesco решили, что американским потребителям нужен универсальный магазин по принципу «все в одном месте», где они могли бы приобретать органические продукты по разумным ценам в экологически чистом магазине. Поэтому они построили чистые, хорошо освещенные магазины с органическими продуктами питания и солнечными крышами.

К концу 2012 года на западе Соединенных Штатов было опубликовано более 200 статей о магазинах Fresh & Easy. И все это закончилось не очень хорошо. В апреле 2013 года Tesco объявила, что отказывается от американского рынка. Общая стоимость этого фиаско составила 3 млрд долл.

Одна из проблем с данными заключалась в словах американцев о том, что им нравятся универсальные магазины «все в одном», но, как однажды посоветовал один американский политик, «следите за тем, что мы делаем, а не за тем, что мы говорим». Многие американцы обожают свои автомобили и с удовольствием ездят из одного магазина в другой. Американцам, возможно, нравятся чистые, хорошо освещенные магазины, но они также любят делать покупки в необычных заведениях, таких как Jungle Jim’s, Kalustyan’s, Rouses, Trader Joe’s и Wegmans. Как бы то ни было, информационные гуру компании Tesco совершенно неправильно поняли американский рынок.

На рис. 8.6 и 8.7 показано последующее снижение прибыли Tesco и цены акций. Помимо неудачного вторжения на рынок США, у Tesco были некоторые проблемы с бухгалтерией и жесткая конкуренция со стороны агрессивно настроенных фирм с низкими издержками, таких как немецкая сеть супермаркетов Aldi, однако в статье журнала Harvard Business Review за 2014 год утверждалось, что важной частью проблемы были ошибки в данных: «Менее чем за десятилетие движущая сила и определяющий фактор успеха Tesco скатились до аналитического альбатроса».

Одна из проблем, связанных со скидками по карте Clubcard, которая хорошо знакома пиццериям, заключается в том, что, как только клиенты привыкают к купонам, они больше не покупают без них. Например, домохозяйство, которое обычно готовит пиццу раз в неделю, может купить пиццу во второй день недели, если увидит привлекательную скидку, что увеличивает выручку пиццерии на этой неделе. Это может повториться на второй неделе, а затем и на третьей, но теперь домохозяйство пользуется скидками дважды в неделю, что

сокращает прибыль предприятия. Затем семья начинает считать скидочные купоны новой нормой и возвращается к употреблению пиццы раз в неделю, но только при наличии купона.

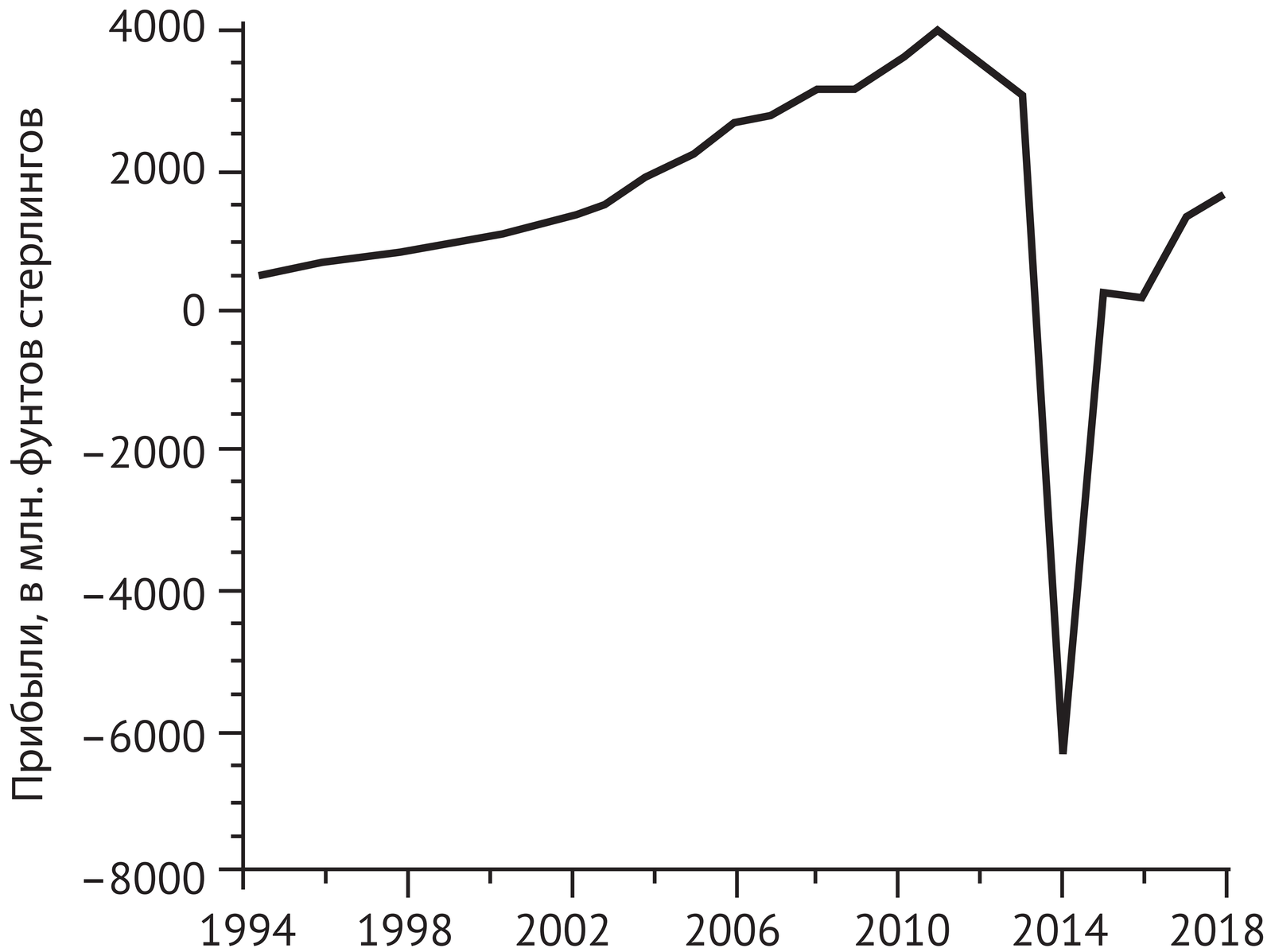

Рис. 8.6. Прибыли стремительно устремляются в пропасть

Рис. 8.7. Американские горки курса акций Tesco

Некоторые домохозяйства считают купоны бесполезными. Вместо того чтобы собирать и выкупать купоны, они предпочитают делать покупки там, где цены всегда низкие. Другие покупатели, разобравшись с системой, начинают в нее играть. Игра похожа на электронный торг. Магазин присылает покупателю скидочный купон на помидоры, но покупатель его не выкупает – возможно, потому что забывает им воспользоваться. В следующий момент магазин присылает еще бóльшую скидку, пытаясь найти приемлемую для покупателя цену. Теперь клиент разбирается в игре и ждет получения третьего или четвертого купона и делает ход только после этого.

Мы видели, как это происходило с Amazon. Покупатель смотрит на продукт в течение нескольких минут, изучая характеристики и читая отзывы, а затем отказывается от покупки. В следующий раз, когда покупатель смотрит на продукт, цена таинственным образом оказывается ниже. Случается и обратное. Клиент заказывает пакетик кофе в зернах несколько месяцев подряд, а затем замечает, что цена таинственным образом выросла. Amazon, очевидно, надеется, что он этого не заметит. Затем клиент просит друга или родственника заказать для него кофе в зернах, и ему предлагают изначальную цену.

С целевыми купонами связана еще одна проблема – каннибализация. Купон на мыло для одного покупателя в одном магазине Tesco может увеличивать продажи в этом магазине за счет продаж в другом магазине Tesco.

Более общая проблема заключается в наличии у Tesco данных только о людях с картами Clubcard, которые совершают покупки в сети магазинов Tesco. Tesco мало что знает о покупателях Tesco, у которых нет карт, и ничего не знает о покупках, которые люди (с картами Clubcard или без) совершают в других магазинах. Например, Tesco может заметить, что в некоторых ее магазинах владельцы карт Clubcard платят повышенные цены за старинные сорта помидор, но не замечает, что большинство покупателей покупают помидоры дешевле в других местах. Если в какой-то район приезжают новые домохозяйства и совершают покупки не в магазинах Tesco, то ее аналитики данных не будут иметь представления о том, что они покупают и зачем.

Как и все алгоритмы разыскивания закономерностей, перемалывание чисел в Tesco также уязвимо для совпадений; например, алгоритм может заметить, что люди, которые купили помидоры в четверг в одном магазине, случайно купили и зубную пасту, и сделать вывод, что это совпадение что-то значит.

Еще одна проблема с данными заключается в том, что, хотя микронацеливание является привлекательным и потенциально выгодным, можно совершить пресловутую ошибку, не увидев лес за деревьями. Алгоритм микронацеливания, который замечает, что отдельные покупатели покупают меньше подслащенных хлопьев, может использовать купоны, чтобы попытаться привлечь этих покупателей к покупке большего количества, не замечая при этом, что клиенты магазина все меньше интересуются подслащенными хлопьями, потому что они становятся старше. Настоящая проблема заключается не в том, чтобы продать больше подслащенных хлопьев пожилым покупателям, а в том, чтобы найти способы привлечь молодых покупателей. Аналогичным образом если отдельные покупатели покупают меньше говядины, то это может быть признаком того, что многие люди сокращают потребление говядины и предпочли бы не больше мясных купонов, а больше альтернатив говядины.

Наконец, когда мы начали осознавать, насколько тщательно отслеживается наше поведение, многие стали этим возмущаться. Марк Цукерберг из Facebook однажды назвал людей, которые доверили ему свои данные, «тупыми ублюдками». Некоторые из этих тупиц просыпаются и больше не принимают такое положение вещей. Проведенное в 2019 году исследование показало, что в период с 2017 по 2019 год число пользователей Facebook в США сократилось на 15 млн, хотя многие из этих беглецов перешли в принадлежащий Facebook Instagram.

Компания Nordstrom прекратила использовать датчики Wi-Fi, чтобы отслеживать покупателей в своих магазинах на следующий день после того, как филиал CBS сообщил о слежке. На компанию Urban Outfitters был подан коллективный иск за ложное сообщение покупателям, расплачивающимся кредитной картой, что они должны сообщить магазину свой почтовый индекс, который затем может быть использован для определения их домашнего адреса. Проведенное в 2013 году исследование показало, что, по сообщениям трети опрошенных, они перестали пользоваться веб-сайтами или совершать покупки у компаний из-за проблем с конфиденциальностью. Некоторые люди даже вернулись к оплате наличными и ничего не покупают через интернет.

В 2018 году Тим Кук из Apple предупредил:

> Наша собственная информация – от повседневной до глубоко личной – используется против нас с военной эффективностью. <…> Каждый день миллиарды долларов переходят из рук в руки, и бесчисленное количество решений принимается на основе наших симпатий и антипатий наших друзей и семей, наших взаимоотношений и разговоров, наших желаний и страхов, наших надежд и мечтаний. <…> Мы не должны приукрашивать последствия. Это слежка. <…> Для того чтобы искусственный интеллект был по-настоящему умным, он должен уважать человеческие ценности, включая неприкосновенность частной жизни. Если мы ошибемся, опасность будет огромной.

Британская газета Telegraph написала о проблемах компании Tesco:

> Судя по переписке читателей газеты Telegraph и разочарованным покупателям, одна из причин, по которой потребители обращаются к [дискаунтерам] Aldi и Lidl, заключается в том, что они считают их простыми и свободными от уловок. Покупатели задаются вопросом, а не являются ли карты лояльности, такие как Clubcard, более полезными для супермаркета, чем для покупателя.

Недовольство клиентов картой Clubcard компании Tesco может также отражать растущее недоверие к сборщикам данных и растущее стремление к конфиденциальности. Некоторые клиенты сети магазинов Target были недовольны алгоритмом-предсказателем беременности Target (поэтому магазины Target начали добавлять в свои объявления случайные товары, такие как газонокосилки, чтобы скрывать, как много они знают). Клиенты магазинов Tesco, возможно, не хотят, чтобы мир знал о том, сколько алкоголя, слабительных, презервативов и виагры они покупают.

В 2015 году Tesco выставила компанию dunnhumby на продажу, но не получила никаких привлекательных предложений.

Как не быть введенным в заблуждение призрачными закономерностями

Данные, несомненно, полезны для получения ответов на многие интересные и важные вопросы, но одних только данных недостаточно. Данные без теории являются источником большого (и растущего) числа промахов, оплошностей и злоключений с данными.

Мы должны сопротивляться искушению поверить, что данные могут давать ответы на все вопросы и что чем больше данных, тем надежнее ответы. Данные могут содержать ошибки и пропуски или не иметь отношение к делу, поэтому охмурение с помощью данных является следствием одурачивания призрачными закономерностями. В дополнение к этому закономерности, обнаруженные в прошлом, исчезнут в будущем, если только для закономерности не будет глубинного основания.

Тестирование на истории моделей на рынке акций представляет особую опасность, потому что очень легко обнаруживать случайно возникающие закономерности, которые оборачиваются дорогостоящими ошибками. Эта эндемичная проблема в настоящее время распространилась далеко и широко ввиду огромного объема данных, которые могут использоваться учеными, предпринимателями и государственными исследователями, чтобы обнаруживать призрачные закономерности.

Глава 9

Видение вещей такими, какие они есть

В январе 1994 года, когда Всемирная паутина (WWW) только начинала набирать обороты, два аспиранта Стэнфорда, Джерри Янг и Дэвид Фило, запустили веб-сайт под названием «Путеводитель Джерри и Дэвида по Всемирной паутине». Их путеводитель представлял собой список веб-страниц, которые они считали интересным. Годом позже они зарегистрировали компанию с сексапильным названием Yahoo!, дополненным восклицательным знаком. Сколько названий компаний содержат восклицательные знаки? Теперь у них был каталог из 10 000 веб-сайтов и 100 000 пользователей Yahoo в день, чему способствовал тот факт, что в популярном браузере Netscape была кнопка «Каталог», которая отправляла людей на Yahoo.com.

По мере развития Всемирной паутины Yahoo нанимала сотни людей для поиска сайтов в интернете, чтобы добавлять их в свой растущий в геометрической прогрессии каталог. Они добавляли графику, новостные сюжеты и рекламу. К 1996 году Yahoo посещало более десяти миллионов человек в день. Yahoo считала себя медиакомпанией, чем-то вроде журнала Fortune, куда люди приходили получать информацию и развлекаться, а рекламодатели платили за возможность поймать блуждающий взгляд читателей. Для создания уникального контента Yahoo наняла собственных авторов. По мере того как Yahoo добавляла больше контента, например спортивных и финансовых страниц, она привлекала целевую рекламу, ориентированную на людей, интересующихся спортом или финансами. Так появились Yahoo Mail, Yahoo Shopping и другие службы.

Затем в 1998 году появилась компания Google с ее революционным алгоритмом поиска. Yahoo не могла самостоятельно угнаться за развитием интернета, поэтому в течение четырех лет использовала поисковую систему Google, пока не разработала свой собственный алгоритм. Тем временем Google зарекомендовал себя как ведущий поисковый веб-сайт и сегодня по-прежнему занимает 90 % поискового рынка.

В 1999 году Yahoo допустила ряд колоссальных ошибок, заплатив 3,6 млрд долл. за GeoCities и 5,7 млрд долл. за Broadcast.com. (Для сравнения: Google заплатила 50 млн долл. за Android в 2005 году, 1,65 млрд долл.

за YouTube в 2006 году и 3,1 млрд долл. за DoubleClick в 2008 году.) Тем не менее люди заходили на страницы Yahoo, потому что там был отличный контент и ввиду инертности пользователей. В январе 2000 года акции Yahoo взлетели до 118,75 долл. В отличие от большинства интернет-магазинов Yahoo была прибыльной. Тем не менее она приносила всего пять центов на акцию. Акции, как правило, считаются опасно дорогими, когда соотношение цена/заработки (price/earnings, аббр. P/E) превышает, скажем, 50. Показатель P/E у Yahoo составил ошеломляющие 2375. По некоторым оценкам, для того чтобы оправдать свою рыночную оценку, Yahoo должна была бы быть такой же прибыльной, как Wal-Mart в 2000 году, в два раза прибыльнее в 2001 году, в три раза прибыльнее в 2002 году и т. д. до бесконечности.

Компания Yahoo была безумно переоценена. Вскоре рынок акций коллективно пришел в себя, и Yahoo рухнула с пресловутого обрыва. За следующие двенадцать месяцев акции Yahoo упали на 90 %.

Обвал акций Yahoo, несомненно, был вызван чрезмерной реакцией (Уолл-стрит этим славится), и в течение следующих нескольких лет акции компании снова росли. За это время в Yahoo сменилась целая карусель руководителей. Никто из них, казалось, ничего не изменил, но им платили баснословные суммы.

Харизма редко является решением проблем компании. Как сказал Уоррен Баффет, «когда управленец с отличной репутацией встречается с компанией с плохой репутацией, именно репутация компании остается нетронутой».

Разочарование чаще бывает особенно сильным, когда на должность генерального директора приглашается посторонний. Во-первых, посторонний человек не знает культуру компании, а также сильные и слабые стороны ее сотрудников. Свои люди знают, что Джек – жалобщик, Джилл – хвастунья, а Джон – бездельник. Посторонние люди не знают ни Джека, ни Джилл, ни Джона. Во-вторых, совет директоров, принимающий решение о приеме на работу, не знает посторонних людей столь хорошо, как своих собственных людей. Чем меньше информации, тем больше разрыв между восприятием и реальностью.

Yahoo усвоила эти уроки на собственном горьком опыте, наняв за пять лет пять генеральных директоров (четверо из них были посторонними людьми) в тщетной попытке спасти медленно тонущий корабль.

Когда акции Yahoo упали до 10,47 долл. за акцию, первоначальный генеральный директор Yahoo Тим Кугл был уволен 7 марта 2001 года, а 17 апреля 2001 года его сменил Терри Семел. Семел проработал в Warner Brothers двадцать пять лет. Он был председателем правления и одним из главных исполнительных директоров и был известен своими успехами в заключении сделок в Голливуде.

К сожалению, его голливудская квалификация не слишком хорошо зарекомендовала себя в Кремниевой долине. Семел хотел создать подразделение Yahoo Entertainment и управлять Yahoo так, как если бы это была киностудия, подобная студии Warner Brothers, в которой ему было комфортно. Он сорвал несколько сделок, в том числе отказался от возможности купить Google и Facebook. Google хотела получить 3 млрд долл.; Семел не согласился на сумму выше одного миллиарда. У Семела была сделка по покупке Facebook за 1 млрд долл., но в последнюю минуту он попытался снизить цену до 800 млн

долл. Сообщается, что у Семела также было соглашение о покупке DoubleClick за 2,2 млрд долл., которое в последнюю минуту было расторгнуто. С другой стороны, Microsoft предложила купить Yahoo, но получила отказ.

Семел фигурирует в нескольких списках худших руководителей в области технологий, однако за шесть лет работы на посту генерального директора ему заплатили почти 500 млн долл. В 2007 году совет директоров Yahoo выразил Семелу вотум доверия, а неделю спустя его уволил.

Затем Yahoo повысила Джерри Янга, одного из соучредителей компании, и, как и соучредители Google Сергей Брин и Ларри Пейдж, Янг получил символическую зарплату в 1 долл. в год (его акции Yahoo стоили миллиарды). Однако бóльшая часть его работы на этом посту, вероятно, не стоила и доллара. Янг продержался чуть больше года, и это был не самый удачный год. Прибыль Yahoo упала на 12 %, а курс ее акций – на 60 %. Он также отклонил предложение компании Microsoft купить Yahoo за 44,6 млрд долл. (на 62 % выше рыночной цены). Помимо участия в создании компании, самое большое достижение Янга произошло еще до того, как он стал генеральным директором, когда в 2005 году он возглавил покупку 40 % акций интернет-стартапа Alibaba за 1 млрд долл. Alibaba стала аналогом Amazon, eBay и Google в Китае. Поскольку акции компании Alibaba еще не торговались на бирже, некоторые инвесторы купили акции Yahoo только для того, чтобы получить инвестиции в Alibaba.

После ухода Янга генеральным директором стала Кэрол Бартц. Она была генеральным директором компании Autodesk, производящей программное обеспечение для дизайна, и объявила, что переходит в Yahoo, чтобы «надрать кому-нибудь задницу». За двадцать месяцев своего правления она уволила сотни сотрудников, но акции Yahoo выросли всего на 7 % в то время, как индекс NASDAQ вырос более чем на 60 %. После того как в 2010 году ей заплатили 47,2 млн долл., компания Glass Lewis, проводившая голосование через представителя, назвала ее самым высокооплачиваемым генеральным директором в стране. В сентябре 2011 года ее уволили по телефону.

Следующим был Скотт Томпсон, который был президентом PayPal. За четыре месяца он уволил еще 2000 сотрудников Yahoo, прежде чем его уволили из-за обвинений в ложном утверждении, что у него есть ученая степень в области теории и практики вычислительных систем. За 130 дней работы на посту генерального директора ему заплатили более 7 млн долл.

Исполнительный вице-президент Yahoo Росс Левинсон согласился исполнять обязанности временного генерального директора и проработал так три месяца. Он ушел, по сути, с выходным пособием в размере 5 млн долл. после того, как 16 июля 2012 года на должность генерального директора была принята Марисса Майер.

Майер было всего 37 лет, но она была умна и энергична, и у нее было потрясающее резюме. Она была танцовщицей, участницей дебатов, волонтером и преподавателем в Стэнфорде, получив при этом степени бакалавра и магистра. У нее было несколько патентов в области искусственного интеллекта, и она была сотрудником номер 20 в Google, где быстро дослужилась до должности вице-президента. В 2012 году компания Yahoo наняла Майер на должность президента и генерального директора, чтобы обратить вспять долгое падение Yahoo (отчасти вызванное стремительным ростом Google).

Майер намеревалась восстановить славу компании Yahoo, сделать ее таким же игроком, как Amazon, Apple, Facebook и Google. Это было бы непросто. Поисковая система Yahoo занимала лишь небольшую долю рынка. Новостной контент Yahoo был ценным, но пользователи не могли легко получать к нему доступ со своих смартфонов. Служба Yahoo Mail ежедневно получала миллиарды электронных писем, но пользователи не могли легко читать их на своих смартфонах.

План Майер состоял в том, чтобы (a) разработать мобильные приложения Yahoo для новостей, электронной почты и обмена фотографиями; (b) создавать привлекательный (мастридный) оригинальный контент, чтобы привлекать читателей во вселенную Yahoo; (c) усовершенствовать алгоритм поиска Yahoo и (d) снова сделать Yahoo веселым и захватывающим местом для работы, местом, которое привлекало бы и воспитывало хорошие таланты.

Она сказала сотрудникам Yahoo, что «наша цель – вдохновлять и радовать наших пользователей, создавать красивые службы, вещи, которыми люди любят пользоваться каждый день, и это наша возможность».

За первые два года работы Майер отказалась от старых продуктов, создала новые и купила 41 стартап – отчасти для привлечения новых талантов и изобретательности. Она утверждала, что «на мой взгляд, приобретения были для Yahoo не выбором, а, скорее, необходимостью». Yahoo пыталась предложить новые интернет-службы, которые заманивали бы пользователей и привлекали рекламодателей, но не смогла эффективно конкурировать с Craigslist, eBay, Facebook и Google. Основной отличительной службой компании Yahoo был премиальный контент, который нуждается в дорогостоящем человеческом труде.

В ноябре 2012 года она переманила Энрике де Кастро из Google и назначила его главным операционным директором (COO) компании Yahoo, сделав его самым высокооплачиваемым исполнительным директором в стране. Говорили, что Де Кастро был раздражительным как в Google, так и в Yahoo. Доходы от рекламы после его найма падали ежеквартально, и в январе 2014 года он был уволен с выплатой выходного пособия в размере 60 млн долл. Общая сумма компенсации за пятнадцать разочаровывающих месяцев составила 109 млн долл.

Ни одна из целей Майер не была достигнута. Приложения были тусклыми, цифровые журналы непопулярными, а поисковая система еще больше отставала от Google.

Гэри был поклонником Yahoo из-за отличного контента. В отличие от первоначального веб-сайта Google, который был просто поисковой системой, на Yahoo были ссылки на важные новостные материалы из New York Times, Wall Street Journal, BBC и других авторитетных источников. В них также были хорошо организованные разделы, посвященные спорту и финансам, с хорошо организованными данными и интересными историями, некоторые из которых были написаны собственными сотрудниками Yahoo. Гэри мог просматривать любые интересующие его акции и находить в них массу интересной информации.

Затем кто-то решил, что Yahoo может увеличить свой доход от рекламы, размещая приманку для кликов, т. е. рекламу, замаскированную под новостные сюжеты, например такие, как «Нынешний вид этой актрисы-подростка заста-

вит отвиснуть вашу челюсть!», «Отшельница, ставшая величайшим коллекционером акций» и «Ужасающие засекреченные фотографии Вудстока».

Пользователь, который переходил по ссылке на Вудсток, попадал на слайд-шоу с надписью «Забытый Вудсток: невиданные ранее кадры величайшего рок-концерта всех времен!». Это слайд-шоу окружено рекламой и другими «историями» с названиями вроде «Невероятно неловкие семейные фотографии», «В происходившее трудно было поверить» и «Она понятия не имела, почему другие политики так на нее пялились» (в сопровождении фотографии женщины с большой грудью). При нажатии на любую из этих историй открывалось еще одно слайд-шоу, окруженное большим количеством рекламных объявлений и «историй». Трудно перемещаться по любому слайд-шоу, не нажимая кнопку **Далее** или стрелку, которая приводит пользователя к рекламе. Слайд-шоу крайне скучны. (Гэри продрался только через четыре слайда Вудстока.)

Легитимные истории, которые остаются на главной странице Yahoo, – это в основном сплетни о знаменитостях: две кинозвезды расстаются; двух знаменитостей видят вместе; у школьной учительницы роман с ученицей; в больнице произошла путаница с ребенком.

В то самое время, когда компания Yahoo превращалась в комбинацию желтой прессы National Enquirer и Yellow Pages, компания Google пополнила страницу новостей реальными новостями из CNN, New York Times, Wall Street Journal и Chicago Tribune о темах, которые были важнее сплетен о знаменитостях. Там могли появляться сообщения о наводнении в Париже, о тысячах людей в Гонконге, празднующих массовую резню на площади Тяньаньмэнь, и сплетни о том, что Facebook отслеживает разговоры со смартфонов.

Теперь веб-страница Google News является для Гэри дежурной.

В 2016 году Майер объявила, что Yahoo уволит 15 % из 11 000 своих сотрудников и закроет половину своих цифровых изданий: Yahoo Autos, Yahoo Food, Yahoo Health, Yahoo Makers, Yahoo Parenting, Yahoo Real Estate и Yahoo Travel. Сообщалось, что сотрудники Yahoo были деморализованы переходом от качественного контента к низкопробным дешевым развлечениям на «дерьмовой домашней странице». Распространенная жалоба от авторов была такой: «Вы конкурируете с задницей Ким Кардашьян».

Курс акций Yahoo продолжал падать, а ключевые сотрудники бежали.

В 2014 году Yahoo получила денежные вливания в размере 9,4 млрд долл., продав 27 % своих акций Alibaba. По имеющимся сведениям, в 2016 году оставшиеся 383 млн акций Alibaba (15 % Alibaba) были единственной ценностью, которая оставалась у Yahoo. Сторонние аналитики подсчитали, что рыночная стоимость Yahoo была меньше, чем стоимость ее активов в холдинге Alibaba, – бизнес Yahoo имел отрицательную рыночную стоимость! Инвесторы предложили Yahoo разделить компанию на две части: одна часть – ценные акции Alibaba, другая – практически бесполезный бизнес Yahoo. Крах Yahoo, казалось бы, был полным: компания стоимостью 128 млрд долл. превратилась в ничто. Летом 2016 года Yahoo продала свой стержневой бизнес, за исключением Alibaba, компании Verizon за 4,8 млрд долл., что больше нуля, но ненамного больше.

По оценкам, за руководство этим тонущим кораблем Майер было выплачено более 200 млн долл., включая выходное пособие в размере 57 млн долл.

На рис. 9.1 показана история курса акций Yahoo, скорректированная с учетом дробления акций, а также временные рамки пяти генеральных директоров, которые привлекались для спасения Yahoo после падения курса акций. (Трехмесячная промежуточная позиция Росса Левинсона опущена.)

Рис. 9.1. Взлет и падение компании Yahoo!

Пятеро высокооплачиваемых генеральных директоров в совокупности не смогли сделать для Yahoo ничего, кроме удачной покупки акций Alibaba, инициатором которой был Джерри Янг в то время, когда он не был генеральным директором. За это им заплатили почти миллиард долларов. В каждом случае Yahoo платила немалые деньги за то, чтобы найти суперзвездного генерального директора, который восстановил бы славу Yahoo. В каждом случае разрыв между надеждой и реальностью оказывался колоссальным.

Эта печальная история является примером того, что призрачные закономерности не обязательно должны быть числовыми. Каждый генеральный директор, который входил во вращающуюся дверь компании Yahoo, в прошлом добивался успеха. Правление компании Yahoo исходило из того, что прошлые примеры успеха являются надежным предсказателем будущих успехов. Однако генеральный директор редко является причиной успеха компании. Компанию великой делают как раз тысячи или десятки тысяч сотрудников, приходящих на работу каждый день. Здесь тоже может быть большая доля удачи. Некоторые компании преуспевают, потому что их продукт оказывается гораздо успешнее, чем кто-либо мог предположить. Кто знал, что фрисби, хула-хупы и кубики Рубика будут так популярны? Кто знал, насколько успешными окажутся самоклеящиеся листочки для заметок Post-it и автомобиль Toyota Corolla? Немного. И уж точно не генеральные директора.

Иногда компания преуспевает вопреки своему генеральному директору. Нет необходимости называть имена. Суть в том, что простая закономерность – Крис Харизма становится генеральным директором, компания преуспевает – не является надежным предсказателем того, насколько успешно будет работать следующая компания, которая наймет Харизму.

Принцип Питера

Это касается не только руководителей компаний. Обычно сотрудников повышают в должности за их выдающиеся достижения на текущей работе. Однако должности, на которые они переводятся, могут требовать совершенно других навыков; например, повышение инженера до руководителя или продавца до менеджера по продажам. Вот почему люди, которых продвигают по службе не на основе способностей, необходимых им для достижения успеха на новой работе, а на основе их текущей работы, зачастую разочаровывают.

Получившие повышение сотрудники, которые и в самом деле добиваются успеха на своей новой работе, скорее всего, будут повышены до еще более высокого уровня в иерархии. И так будет продолжаться до тех пор, пока они не достигнут должности, на которой они не будут успешны, и перестанут продвигаться дальше по цепочке.

Продвижение по службе людей, которые проявляют себя хорошо в том, что они делали, до тех пор, пока они не перестанут проявлять себя хорошо в том, что делают сейчас, – это принцип Питера, предложенный Лоуренсом Дж. Питером (Laurence J. Peter): «Управленцы поднимаются до уровня своей некомпетентности».

Из принципа Питера следует, что большинство управленцев неэффективны, а те сотрудники, которые эффективны, эффективны только до тех пор, пока их не переведут на работу, где они неэффективны. Принцип Питера – циничная, но слишком распространенная реальность.

Google

Коллегу из последней компании Джея спросили, как они могут добиться успеха в качестве интернет-компании, когда Google – крупная рыба в море. Он ответил: «Google не крупная рыба в море, они и есть море». Итак, как компания Google стала морем?

Ларри Пейдж окончил Мичиганский университет; Сергей Брин родился в Москве и окончил Мэрилендский университет. Оба они отказались от получения степени доктора философии в Стэнфорде по теории и практике вычислительных систем (компьютерным наукам), чтобы основать Google. Первоначально они полагали, что о важности веб-страницы можно судить по количеству ссылок на нее. Вместо того чтобы платить людям за то, чтобы они угадывали, какие страницы могут заинтересовать искателя информации, они использовали коллективное голосование интернет-пользователей, чтобы определить наиболее релевантные и полезные страницы. Для того чтобы собрать эти данные, они создали мощный веб-обходчик, который мог блуждать

по Сети, подсчитывая ссылки. Этот поисковый алгоритм, именуемый PageRank, трансформировался в другие, более сложные алгоритмы, позволяющие Google доминировать на поисковом рынке.

Роботы Google (именуемые *пауками* или *обходчиками*) постоянно блуждают по Сети, собирая данные для своих алгоритмов соотнесения. В 2019 году было подсчитано, что Google обрабатывал 63 000 запросов в секунду, или 5,5 млрд запросов в день, что соответствует данным с 60 млрд веб-страниц. Огромная база данных искателей информации и результатов поиска позволяет Google постоянно совершенствовать свои алгоритмы и оставаться далеко впереди потенциальных конкурентов. Богатые становятся еще богаче.

Компания Google делает это не как публичная служба. Она использует эту персонализированную информацию для размещения целенаправленной рекламы и соотнесения компаний с людьми, которых они пытаются охватить и на которых они пытаются повлиять. Как говорится, когда вы пользуетесь поисковой системой, вы не клиент, вы – продукт.

В 2018 году выручка компании Google составила 136 млрд долл., из которых 116 млрд пришлось на рекламу. Выручка компании Google от рекламы генерирует огромный денежный поток, который она инвестирует в другие проекты, включая Gmail; Google Chrome (который вытеснил Internet Explorer компании Microsoft в качестве самого популярного браузера); Google Документы, Google Таблицы и Google Слайды (которые угрожают Microsoft Word, Excel и PowerPoint); Google Карты и, конечно же, Google Авто.

Успех компании Google зависит от двух составляющих. Первая – это способность Google определять, что именно люди ищут, когда они вводят таинственные фразы с перепутанными словами и ошибками в написании. Вторая – это способность Google определять веб-страницы, которые наиболее релевантны для намеченного поискового запроса.

Одним из преимуществ компании Google является то, что она может предоставлять адаптированные, персонализированные поисковые запросы, используя данные предыдущих поисковых запросов пользователя, а также информацию из сообщений Gmail и профилей Google+, чтобы отбирать наиболее релевантные веб-сайты. Однако у нее есть еще одно не столь секретное оружие: широкое использование A/B-тестов.

Понятно, что компания Google не публикует подробности своих алгоритмов, но мы знаем, что она проводит множество A/B-тестов. Компания даже предлагает клиентам бесплатный инструмент для A/B-тестирования под названием Google Оптимизация.

A/B-тест похож на лабораторный эксперимент. Создаются две версии страницы – как правило, текущая страница и предлагаемая модификация, возможно, с другим баннером, заголовком, описанием продукта, отзывом или надписью на кнопке. Для определения того, какую страницу увидит пользователь, используется генератор случайных чисел. На языке медицинского теста те, кто отправляется на текущую страницу, являются *контрольной группой*, а те, кто отправляется на измененную страницу, – *процедурной группой*. Для определения страницы-победителя используется предопределенная метрика, например количество покупок или подписок на рассылку. Вариантов может быть более двух (A/B/C/D-тест и т. д.), и между страницами может быть множество различий.

Первый A/B-тест Google, проведенный в 2000 году, был использован для определения оптимального количества результатов поиска, отображаемых на странице. Они решили, что лучше всего использовать десять, и это число используется до сих пор.

Дэн Сирокер покинул Google в 2008 году, чтобы стать директором по аналитике президентской кампании Барака Обамы в 2008 году. Один из его проектов касался дизайна приветственной страницы на главном веб-сайте президентской кампании. На рис. 9.2 показан первоначальный дизайн до того, как Сирокер провел A/B-тесты.

Рис. 9.2. Изначальная приветственная страница президентской кампании «Обама 2008». Дэн Сирокер

Интересующей метрикой была частота, с которой люди нажимали на кнопку **SIGN UP** (Зарегистрироваться), дававшая президентской кампании адрес электронной почты, который можно было использовать для будущих сообщений и запросов. Сирокер провел A/B-тесты на четырех разных кнопках и шести разных фотографиях. Они обнаружили, что замена изначальной кнопки на кнопку **LEARN MORE** (Узнать больше) увеличила количество подписчиков на 18,6 %, а замена изначальной цветной фотографии Обамы на черно-белую семейную фотографию увеличила количество подписчиков на 13,1 %. Эти два изменения в совокупности, показанные на рис. 9.3, увеличили число подписчиков на 40,6 %, с 8,26 до 11,60, что составило три миллиона дополнительных адресов электронной почты.

Компания Google использует A/B-тесты, чтобы помогать себе превращать поисковые запросы в наполненные смыслом слова и фразы, определять, что именно пользователи ищут, выявлять веб-страницы, которые искатели информации сочтут наиболее полезными, и определять способ вывода результатов на экран.

Рис. 9.3. Новая и улучшенная приветственная страница президентской кампании «Обама 2008».

Optimizely

Google ежегодно проводит десятки тысяч, а может быть, и сотни тысяч A/B-тестов в замкнутом цикле, в котором искатели информации используются в качестве бесплатной рабочей силы с целью улучшения результатов поиска и их закрепления в качестве пользователей Google.

В следующий раз, когда вы будете выполнять поиск в Google, вы вполне можете принять участие в A/B-тестировании.

Контролируемые эксперименты

Франклин Д. Рузвельт был президентом Соединенных Штатов в течение двенадцати лет, с марта 1933 по апрель 1945 года – с самого начала Великой депрессии до разгрома нацистской Германии во Второй мировой войне. Его вдохновляющие речи, инновационные социальные программы и ободряющие радиопередачи «Беседы у камина» помогли создать представление о том, что федеральное правительство в целом и президент США в частности несут ответственность за благополучие страны.

Несмотря на то что Рузвельт излучал силу и оптимизм, в 1921 году, когда ему было 39 лет, он был поражен болезнью, которая навсегда парализовала его ниже пояса. Его публичные выступления тщательно контролировались, чтобы публика не видела его инвалидного кресла. Во время публичных выступлений он держался прямо, опираясь на помощников или крепко держась за прочную кафедру.

Болезнь Рузвельта была диагностирована как полиомиелит, который в то время был рецидивирующей эпидемией, хотя сейчас считается, что более вероятно, у него был синдром Гийена–Барре. В 1938 году Рузвельт основал Национальный фонд борьбы с детским параличом, ныне известный как «Марш десятицентовиков», для поддержки исследований и образования в области полиомиелита. Вполне уместно, что после смерти Рузвельта федеральное правительство заменило монету с крылатой головой Свободы на монету Рузвельта 30 января 1946 года, когда Рузвельту исполнилось бы 64 года.

Полиомиелит – это острое вирусное заболевание, которое приводит к параличу, мышечной атрофии, необратимым деформациям и даже смерти. Это заболевание отличается особенной жестокостью, поскольку большинство его жертв – дети. В 1916 году в США разразилась первая эпидемия полиомиелита, и в течение следующих сорока лет от нее пострадали сотни тысяч американцев. В начале 1950-х годов ежегодно регистрировалось более 30 000 случаев острого полиомиелита.

Исследователи, многие из которых получили поддержку фонда «Марш десятицентовиков», обнаружили, что большинство взрослых в течение своей жизни переносили легкую форму полиомиелита, и их организм вырабатывал антитела, которые не только защищали от инфекции, но и делали их невосприимчивыми к другим атакам. Подобным же образом они обнаружили, что полиомиелит встречается реже всего в обществах с самыми плохими санитарными условиями. Это объясняется тем, что почти все дети в таких обществах подвергались воздействию заразного вируса, будучи еще достаточно маленькими, чтобы быть защищенными материнскими антителами, и поэтому они вырабатывали собственные антитела, даже не страдая от этой болезни.

Вследствие этого ученые работали над созданием безопасной вакцины, которая стимулировала бы выработку антител в организме, не вызывая паралича или чего-то худшего. В 1930-х годах было опробовано несколько вакцин, но от них отказались, потому что они иногда вызывали заболевание, для предотвращения которого были разработаны. К 1950-м годам в результате обширных лабораторных исследований было разработано несколько многообещающих вакцин, которые, как казалось, производили безопасные антитела против полиомиелита.

В 1954 году Служба общественного здравоохранения организовала общенациональное тестирование вакцины Джонаса Солка от полиомиелита, в котором приняли участие два миллиона школьников. Поскольку на протяжении 1940-х и начала 1950-х годов эпидемии полиомиелита сильно варьировались от места к месту и из года в год, Служба общественного здравоохранения решила не предлагать вакцину всем детям ни по всей стране, ни в конкретном городе. В противном случае не было бы контрольной группы, и служба здравоохранения не смогла бы определить, были ли изменения в заболеваемости полиомиелитом вызваны вакциной либо капризами эпидемий. Вместо этого было намечено, что вакцина будет предложена всем второклассникам в выбранных школах. Затем опыт этих детей (процедурную группу) можно было сравнить с опытом первоклассников и третьеклассников (контрольной группой), которым вакцина не предлагалась. В идеале процедурная группа и контрольная группа были похожи во всех отношениях, если бы не тот факт, что процедурная группа получила вакцину, а контрольная группа – нет.

У этого предложенного эксперимента было две проблемы. Во-первых, участие было добровольным, и были опасения, что те, кто согласился пройти вакцинацию, как правило, будут иметь более высокие доходы и более высокую гигиену и, как объяснялось ранее, будут более подвержены полиомиелиту. Во-вторых, школьным врачам было поручено выявлять как острые, так и легкие случаи полиомиелита, а диагностировать легкие случаи было нелегко. Если бы врачи знали, что многие второклассники были вакцинированы, а первоклассники и третьеклассницы – нет, то это знание могло бы повлиять на диагноз. Врач, который надеялся, что вакцина будет успешной, мог бы с большей вероятностью заметить симптомы полиомиелита у невакцинированных, чем у вакцинированных.

Второе предложение состояло в проведении двойного слепого теста, в ходе которого только половине детей-добровольцев была бы введена вакцина Солка, и ни дети, ни врачи не знали бы, что ребенок получил: вакцину или плацебо из раствора соли и воды. Жидкость для инъекций каждому ребенку выбирается случайным образом из коробки и записывается серийный номер. Только после того, как будет установлен диагноз полиомиелита, станет известно, что именно ребенок получал: вакцину или плацебо. Первостепенным возражением против этого предложения было неудобство обращения к родителям с просьбой поддержать программу, в рамках которой вероятность того, что их дети будут вакцинированы, составляла всего 50 %.

Как выяснилось, около половины школ решили сделать вакцину всем добровольцам из второго класса и использовать первоклассников и третьеклассников в качестве контрольной группы. Оставшаяся половина согласилась на проведение двойного слепого теста с использованием детей, получавших плацебо, в качестве контрольной группы. Результаты представлены в табл. 9.1. Вакцина Солка снизила заболеваемость полиомиелитом в обоих случаях. Число диагностированных случаев полиомиелита снизилось примерно на 54 % в первом подходе и на 61 % в контрольной группе, получавшей плацебо. Если бы не был проведен двойной слепой эксперимент, то доводы в пользу вакцины Солка были бы менее убедительными, поскольку снижение было меньшим и поскольку скептики могли бы объяснить это снижение подсознательным желанием врачей, чтобы вакцина сработала.

Таблица 9.1. Результаты общенационального испытания вакцины против полиомиелита в 1954 году

	Контрольная группа из перво- и третьеклассников		Двойная слепая с плацебо	
	Дети	**Полио на 100 тыс.**	**Дети**	**Полио на 100 тыс.**
Процедурная группа	221 998	25	200 745	28
Контрольная группа	725 173	54	201 229	71
Отсутствие согласия	123 605	44	338 778	46

Проведенные в 1954 году испытания стали важным национальным экспериментом в области общественного здравоохранения, который убедительно

доказал ценность вакцины против полиомиелита. В конечном итоге вакцина Солка была заменена еще более безопасным и эффективным препаратом Сэйбина. На рис. 9.4 показано резкое снижение заболеваемости полиомиелитом после начала национальных программ иммунизации в 1950-х годах. Сегодня около 92 % всех детей в США получают прививки от полиомиелита в возрасте от 19 до 36 месяцев, и с 1979 года в США не было зарегистрировано ни одного случая заболевания полиомиелитом.

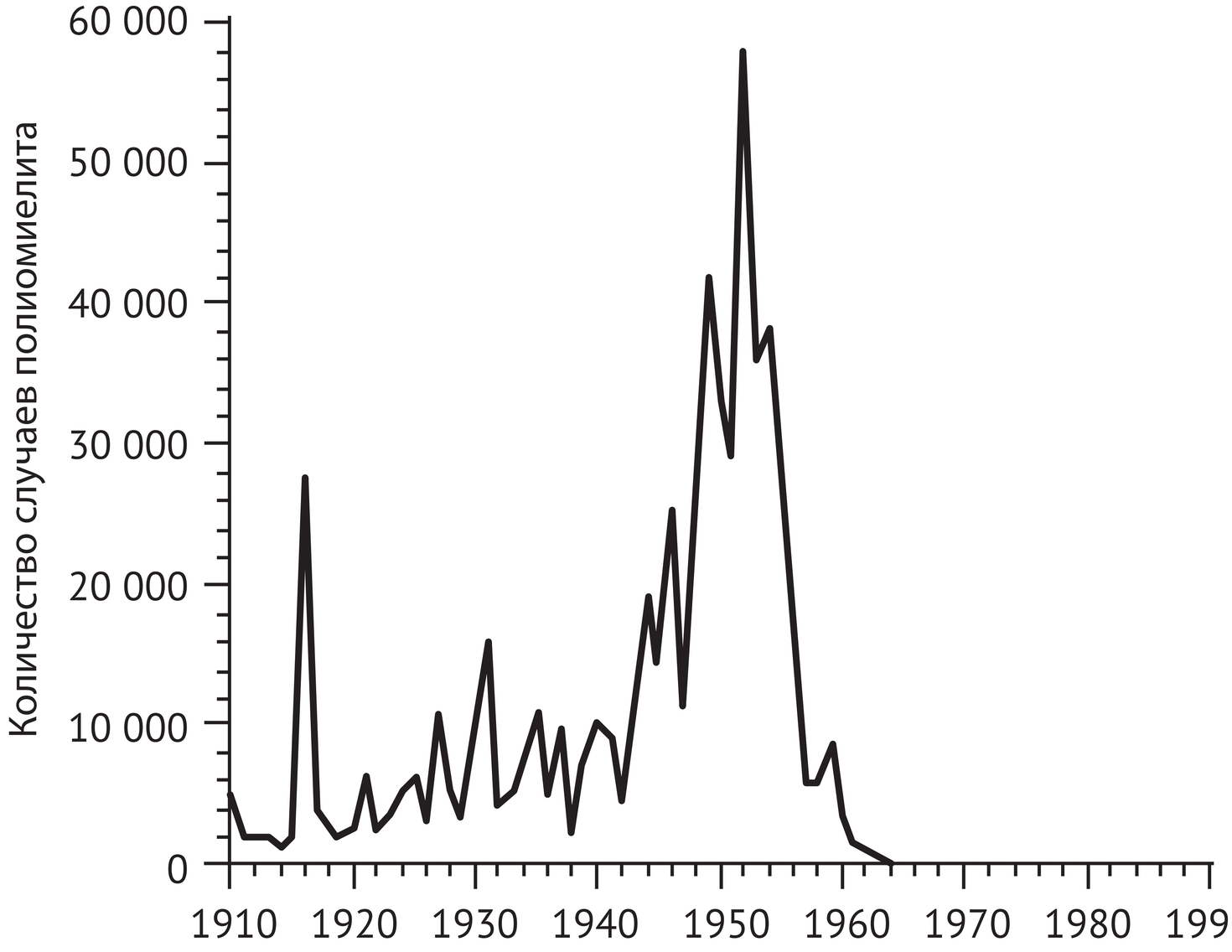

Рис. 9.4. Исчезновение полиомиелита

Естественные эксперименты

Иногда бывает трудно делать выводы на основе данных, полученных в результате наблюдений (т. е. того, что мы наблюдаем, в отличие от контролируемых экспериментов). Например, люди, которые служат в вооруженных силах США, как правило, имеют более низкие доходы после увольнения из армии по сравнению со своими сверстниками, которые не служили в армии. Это может быть связано с тем, что подготовка, которую люди получают в армии, менее полезна для гражданской работы, чем образование и опыт работы, которые они могли бы получить, если бы не поступили на военную службу. Как вариант, это может быть связано со систематическим смещением за счет самоотбора[12], которое часто встречается в данных наблюдений. Разные исходы для людей, которые выбирают тот или иной вид деятельности, и для тех, кто этого не делает, могут быть обусловлены различиями между людьми, которые делают такой выбор. Люди с высшим образованием могут отличаться от людей без высшего образования не из-за того, чему они учатся в колледже, а потому что люди, которые

[12] Систематическое смещение за счет самоотбора (self-selection bias) возникает, когда решение об участии в исследовании принимается исключительно самими людьми. – *Прим. перев.*

решили поступить в колледж и получить высшее образование, отличаются от тех, кто этого не делает. Женатые люди могут отличаться от неженатых не из-за того, что делает с ними брак, а из-за различий между теми, кто решил вступить в брак, и теми, кто решил этого не делать. Люди, решившие поступить на военную службу, впоследствии могут получать относительно низкую зарплату, поскольку у них были ограниченные перспективы трудоустройства, и они не были заинтересованы в поступлении в колледж, когда они решали поступить на военную службу.

Для того чтобы обойти проблемы, связанные с данными наблюдений, можно провести рандомизированное контролируемое исследование, но мы не можем заставить случайно выбранных людей жениться или не жениться, поступать в колледж или не поступать. Однако мы можем заставить людей пойти на военную службу. Естественный эксперимент произошел во время войны во Вьетнаме, когда система селективной службы в США использовала лотерею для случайного отбора людей, которые призывались либо не призывались в армию.

Первая лотерея была проведена 1 декабря 1969 года и определила призывной приказ из 850 000 мужчин, родившихся в период с 1 января 1944 года по 31 декабря 1950 года. 366 дней года (включая 29 февраля) были написаны на листочках бумаги и помещены в синие пластиковые капсулы, которые были высыпаны в большой пластиковый контейнер, похожий на нижнюю половину бутылки из-под воды.

Обеспокоенным мужчинам (и их семьям) было сказано, что первые 125 выбранных дат рождения, скорее всего, будут призваны в армию; номера 126–250 могут не быть призваны, а номера выше 250 считаются безопасными.

На транслировавшейся по национальному телевидению церемонии главный республиканец в комитете палаты представителей по вооруженным силам Александр Пирни вытащил капсулу с датой 14 сентября, и эта дата была помещена на доску рядом с номером 001, что указывает на то, что мужчины, родившиеся 14 сентября, будут призваны на службу первыми.

Остальные капсулы выбирались группой молодых людей, чтобы продемонстрировать, что молодые люди были частью этого процесса. Студент из Род-Айленда Пол Мюррей выбрал вторую капсулу (24 апреля), и так продолжалось до тех пор, пока последнему выбранному дню (8 июня) не был присвоен номер 366. Как оказалось, все даты с номерами 195 и ниже должны были явиться на службу.

Это было драматично и вызывало тревогу, но указанная лотерея действительно стала естественным экспериментом, который позволил избежать систематического смещения за счет самоотбора. На военную службы были призваны не все, у кого было низкое число (у некоторых были проблемы со здоровьем, такие как костные шпоры), но самая большая разница между теми, у кого были первые 195 дат рождения, и теми, у кого были более поздние даты рождения, заключается в том, что первая группа попадала под призыв.

Последующее исследование показало, что до лотереи не было разницы в среднем заработке попадавших под призыв белых мужчин и не попадавших под призыв белых мужчин, но после лотереи разница была существенной. Участие в призыве играло роль, и не в лучшую сторону. У тех, кто получил низкие числа призыва, доходы в среднем были ниже, чем у тех, кто избежал

отбора. Разница в доходах, естественно, была наибольшей во время службы попавших под призыв мужчины в армии, но сохранялась и после того, как они выполнили свои воинские обязанности. Через десять лет после окончания службы ветераны зарабатывали в среднем на 15 % меньше, чем те, кто не был ветеранами. Военный опыт был плохой заменой гражданскому образованию и опыту работы. Различия среди цветных мужчин были менее выражены, но все ветераны, как правило, жертвовали частью своего будущего финансового благополучия, а также своим временем.

Обойденные вниманием женщины доктора Спока

В 1946 году педиатр по имени Бенджамин Спок (Benjamin Spock) опубликовал книгу «Книга здравого смысла о младенцах и уходе за ними» (The Common Sense Book of Baby and Child Care), которая разошлась тиражом более 50 млн экземпляров по всему миру. Когда он умер в 1998 году в возрасте 96 лет, в некрологе журнала Time говорилось, что доктор Спок «в одиночку изменил приемы воспитания родителями своих детей».

До Спока считалось, что детям нужна строгая дисциплина – определенное время приема пищи и сна, а также порка за плохое поведение. Спок советовал родителям «доверять собственному здравому смыслу». Если ребенок не голоден, не заставляйте его есть. Если ребенок не устал, не заставляйте его спать. Прежде всего детям нужна любовь, а не правила и наказания. Некоторые эту гибкость высмеивали, но многие ее одобряли. В разговорах родителей нередко звучала фраза: «По словам доктора Спока…». Даже в 2019 году, спустя 73 года после публикации книги «Младенец и уход за детьми», в статье газеты The Washington Post о президентстве Дональда Трампа говорилось: «И, по словам доктора Спока…».

Спок был ярым противником войны во Вьетнаме, и некоторые его обвиняли в том, что он давал советы по воспитанию детей мятежному юноше в поношенной одежде и с растрепанными волосами, который протестовал против войны. В 1968 году Спок и четверо других обвиняемых предстали перед федеральным судом штата Массачусетс за сговор с целью нарушения Закона о военной службе 1967 года путем консультирования, оказания помощи и подстрекательства к сопротивлению призыву. Спок и трое других обвиняемых были осуждены, но подали апелляцию.

Секретарь суда засвидетельствовал, что он отбирал потенциальных присяжных заседателей, случайным образом тыча пальцем в список взрослых жителей округа. Отобранным им людям были разосланы анкеты, и после того, как те, кто был дисквалифицирован по закону, были отсеяны, группы по 300 человек были приглашены в центральную ложу жюри, из которых была отобрана группа из 100 человек (именуемая «судебным приказом»[13]). Как оказалось, несмотря на то что более половины жителей округа составляли женщины, только девять из 100 перечисленных в судебным приказе по делу доктора Спока человек были женщинами.

---

[13] Судебный приказ о сформировании состава присяжных (venire). – *Прим. перев.*

Процесс отбора присяжных мог быть предвзятым (систематически смещенным). Вслепую проведенный по списку палец почти наверняка окажется рядом с несколькими именами, и секретарь суда, возможно, отдавал предпочтение мужчинам, когда решал, какие имена использовать. При перекрестном допросе секретарь признал, что, возможно, так оно и было:

Ответ: …Я указываю пальцем на нужное место и имя на странице, а затем ставлю рядом с ним отметку ручкой.

Вопрос: Вы делаете это, не глядя на страницу?

Ответ: Мне нужно достаточно внимательно посмотреть на него, чтобы понять, где он находится по отношению к моему пальцу.

Вопрос: Да.

Ответ: Я не собираюсь внимательно изучать имя…

Вопрос: Я предполагаю, что в какой-то момент вам придется посмотреть на имя, чтобы разослать анкету?

Ответ: Правильно.

Вопрос: Есть ли у вас какое-либо объяснение этому [различию между количеством анкет, отправленных мужчинам и женщинам], кроме того, что вы, возможно, увидели имя, распознали в нем женское имя и решили, что будет немного эффективнее не отправлять слишком много анкет женщинам?

Ответ: Это единственное возможное объяснение, кроме чистой случайности…

В судебных протоколах не указано, каким образом была отобрана коллегия из 100 человек из центральной коллегии присяжных, но это тоже могло быть вызвано субъективными факторами.

Это были данные наблюдений, но анализ был основан на правдоподобной теории о том, что женщины, возможно, подвергались дискриминации, – в отличие от анализа данных 100 различными способами и выявления необычного количества левшей, чьи фамилии начинались на букву «S». Возможность проведения рандомизированного контролируемого испытания отсутствовала, но существовала естественная контрольная группа.

Через несколько месяцев после окончания судебного процесса защита получила данные, свидетельствующие о том, что из 598 присяжных заседателей, недавно назначенных судьей первой инстанции, который вел дело доктора Спока, женщинами были только 87 (14,6 %), тогда как из 2378 присяжных заседателей, назначенных другими шестью судьями округа, женщинами были 29 %. Для того чтобы предположить, как это систематическое смещение могло повлиять на шансы Спока на оправдательный приговор, в апелляции приводился опрос Гэллапа 1968 года, в котором 50 % мужчин назвали себя ястребами, а 33 % – голубями во Вьетнаме, по сравнению с 32 % ястребов и 49 % голубей среди женщин.

Возможно также, что женщины, воспитывавшие своих детей «по советам доктора Спока», могли симпатизировать его антивоенной деятельности. В апелляции утверждалось, что вероятность того, что среди случайно отобранных присяжных будет непропорционально много мужчин, мала и что «таким образом, практически неизбежен вывод о том, что секретарь, должно быть, вы-

бирал кандидатуры для судьи первой инстанции из центральной ложи присяжных так, что это каким-то образом систематически уменьшало долю женщин в составе присяжных заседателей». (Вероятность такого большого гендерного неравенства между присяжными заседателями, которых использовал этот судья, и присяжными заседателями, которых использовали остальные шесть судей, составляет примерно один к двадцати восьми триллионам).

В конечном счете обвинительный приговор Споку был отменен, хотя и на основании первой поправки, а не из-за неправильного отбора присяжных. Однако принятый в 1969 году закон предписывал проводить случайный отбор присяжных в федеральных судах с использованием статистически приемлемых методов, таких как генераторы случайных чисел. Впоследствии мужчины и женщины в федеральных судах уже представлялись в равной степени.

Сначала теория, потом данные

Многие легкомысленные, даже нелепые стратегии обыгрывания рынка акций были обнаружены путем рыскания по историческим данным в поисках закономерностей. Закономерности рынка акций особенно соблазнительны тем, что они таят в себе соблазн легких денег – на самом деле денег ни за что. Упомянутые в начале этой главы глупые системы являются лишь небольшой выборкой. Финансовые астрологи изучают расположение планет и звезд. Рейтинг купальников журнала Sports Illustrated основан на том, не из США ли модель на обложке, а рейтинг систем от головной боли основан на продажах аспирина. Система BB основана на производстве сливочного масла в Бангладеш.

Жизнеспособные стратегии начинаются с правдоподобной теории – логической основы, – а не с неограниченного разведывания данных. Дэниел Канеман и Амос Тверски задокументировали вид ошибочных рассуждений, который они назвали «законом малых чисел». Когда происходит что-то необычное, мы склонны бросаться к поспешному выводу о том, что это, скорее всего, повторится снова и что событие является типичным, а не необычным. Это может стать основой для чрезмерной реакции на тривиальные события. Мы можем подумать, что баскетболист, который делает сложный бросок, – хороший стрелок. Мы можем подумать, что комментатор, который делает правильные политические предсказания, – проницателен. Мы можем подумать, что человек, который рассказывает смешные анекдоты, – прирожденный комик. Канеман и Тверски собрали самые разные формальные экспериментальные свидетельства, которые подтвердили их гипотезу о том, что люди склонны придавать излишний вес новой информации.

Что касается рынка акций, то Кейнс заметил, что «ежедневные колебания прибыли от существующих инвестиций, которые, очевидно, носят эфемерный и незначительный характер, как правило, оказывают совершенно чрезмерное и даже абсурдное влияние на рынок [акций]». Если это правда, то такая чрезмерная реакция может стать основой для памятного совета Уоррена Баффета: «Бойтесь, когда другие жадничают, и будьте жадными, когда другие боятся». Если инвесторы часто слишком остро реагируют, вызывая чрезмерные колебания цен на акции, то может оказаться выгодным сделать ставку на то, что за значительными колебаниями цен последуют ценовые развороты.

Одним из примеров (из, по сути, бесконечного списка) стало выступление производителя программного обеспечения Oracle 9 декабря 1997 года. Аналитики ожидали, что продажи компании Oracle во втором квартале будут на 35 % выше, чем годом ранее, а прибыль – на 25 %. После закрытия рынка 8 декабря Oracle сообщила, что ее продажи во втором квартале выросли всего на 23 % по сравнению с предыдущим годом, а прибыль – всего на 4 %. На следующий день было продано 171,8 млн акций Oracle, что составляло более одной шестой всех находящихся в обращении акций Oracle, и цена акций упала на 29 %, что привело к снижению общей рыночной стоимости компании Oracle более чем на 9 млрд долл.

Как это часто бывает, рынок отреагировал слишком бурно. Годовая возвратность акций Oracle в течение следующих 21 одного года, до 31 декабря 2018 года, составила 15,8 % по сравнению с 4,1 % у индекса S&P 500. Инвестиции в Oracle в размере 10 000 долл. на следующий день после ее краха в 1997 году выросли бы до 133 000 долл. по сравнению с 39 000 долл. у индекса S&P 500.

Совсем недавно, 24 января 2013 года, компания Apple сообщила о рекордной квартальной прибыли в размере 13,1 млрд долл., продав на 28 % больше смартфонов iPhone и на 48 % больше планшетов iPad, чем годом ранее, но акции компании упали более чем на 12 %, снизив ее рыночную стоимость на 50 млрд долл. Компания Apple продала рекордные 47,8 млн смартфонов iPhone, но это оказалось меньше консенсусного прогноза в пятьдесят миллионов. Заработки в расчете на акцию оказались выше предсказаний (13,81 против 13,44 долл.), как и выручка (54,7 против 54,5 млрд), но инвесторы привыкли к тому, что Apple не оправдывает предсказаний. Здесь есть небольшой парадокс. Если аналитики ожидали, что компания Apple превзойдет их предсказания, почему они их не повысили? В любом случае инвесторы были напуганы, а рынок отреагировал слишком бурно. На рис. 9.5 показано, что с 24 января 2013 года по 31 декабря 2019 года индекс S&P 500 вырос примерно на 100 %, а Apple – на 400 %.

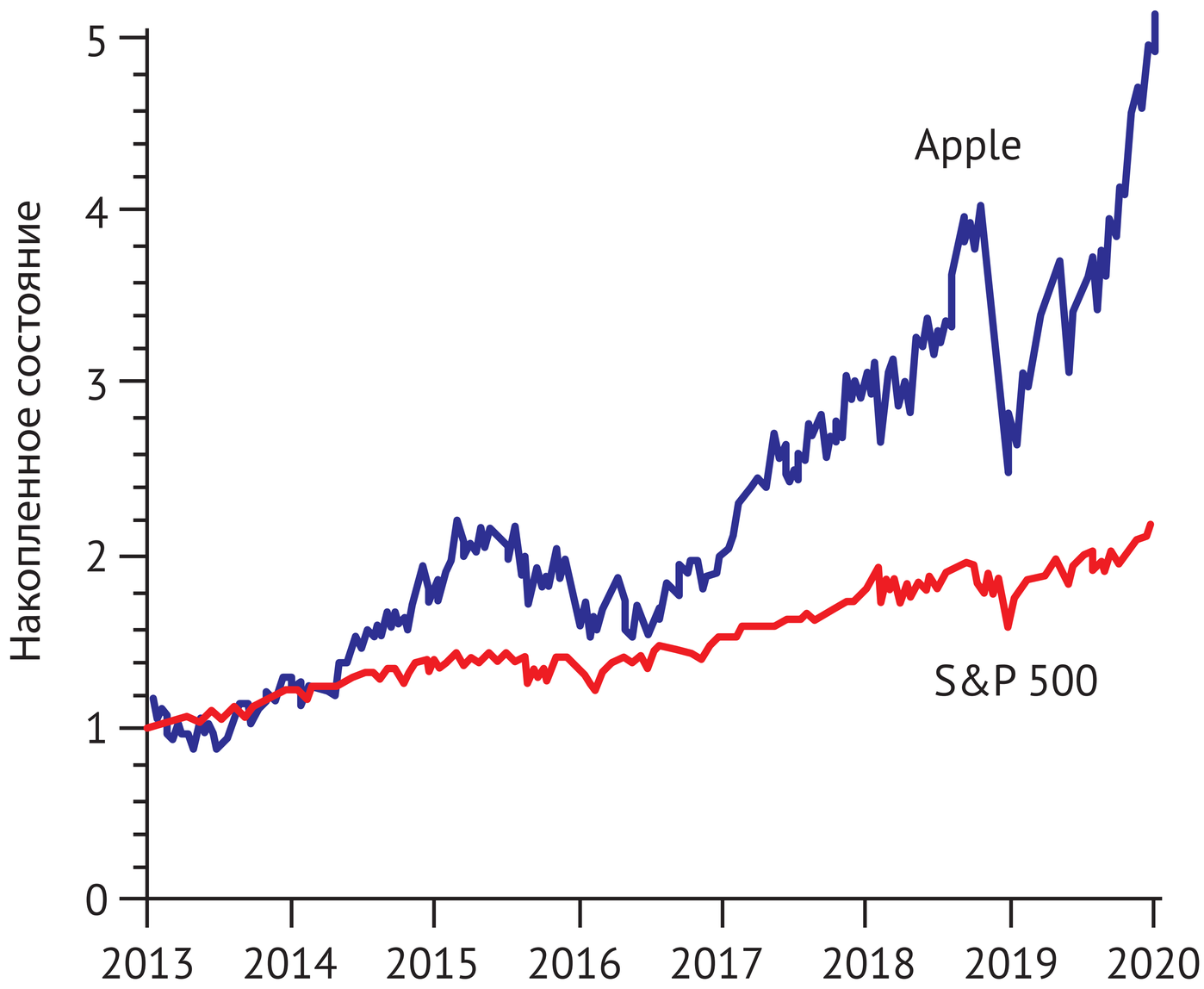

Рис. 9.5. Apple устремляется ввысь

Являются ли эти примеры избирательным припоминанием или свидетельством обычной чрезмерной реакции на рынке акций? Для того чтобы это выяснить, Гэри проанализировал ежедневную возвратность акций индекса Доу Джонса для акций промышленных компаний с 1 октября 1928 года, когда индекс Доу был расширен с двадцати до тридцати акций, по 31 декабря 2015 года, в общей сложности за 22 965 торговых дней.

Каждый день скорректированная дневная возвратность каждой акции рассчитывалась относительно средней возвратности остальных двадцати девяти акций Доу за тот же день. Использование скорректированной возвратности предназначено для того, чтобы привлечь внимание к акциям, которые поддерживаются или снижаются из-за раздражающих (идиосинкратических) новостей или мнений, характерных для отдельной компании, в отличие от общих рыночных скачков или обвалов, вызываемых макроэкономическими новостями или эмоциями.

Большим днем для акций Доу считался день, когда их скорректированная возвратность была выше пяти процентов или ниже отрицательных пяти процентов. Для проверки надежности Гэри также пересмотрел расчеты с отсечками в большие дни, равные шести, семи, восьми, девяти или десяти процентам. Эти альтернативные ограничения не повлияли на выводы.

Результативность каждой акции, которая пережила большой день, отслеживалась в течение десяти дней после этого. В табл. 9.2 показано, что большинство акций упали после того, как у них был день большого роста, и выросли после того, как у них был день большого падения. Это различие было в высшей степени статистически значимым.

Таблица 9.2. Процент акций с положительной скорректированной возвратностью на следующий день после большого дня

Отсечка, в процентах	После дня большого роста	После дня большого падения
5	44,34	53,60
6	44,34	52,45
7	43,95	50,90
8	42,87	53,83
9	45,35	53,22
10	46,10	53,12

На рис. 9.6 и 9.7 показана средняя скорректированная дневная возвратность и средняя скорректированная совокупная возвратность за десять дней, следующих за положительными и отрицательными большими днями. Средняя возвратность была отрицательной в течение первых девяти дней после положительного большого дня и положительной в течение девяти из десяти дней после отрицательного большого дня. Различия между совокупной возвратностью после положительных и отрицательных больших дней были существенными и весьма статистически значимыми. На десятый день кумулятивная потеря после положительного большого дня составил в среднем 0,59 %, а кумулятивная прибыль после отрицательного большого дня – в среднем 0,87 % (что соответствует очень высокой возвратности в годовом исчислении).

Рис. 9.6. Средняя дневная скорректированная возвратность после пятипроцентного большого дня

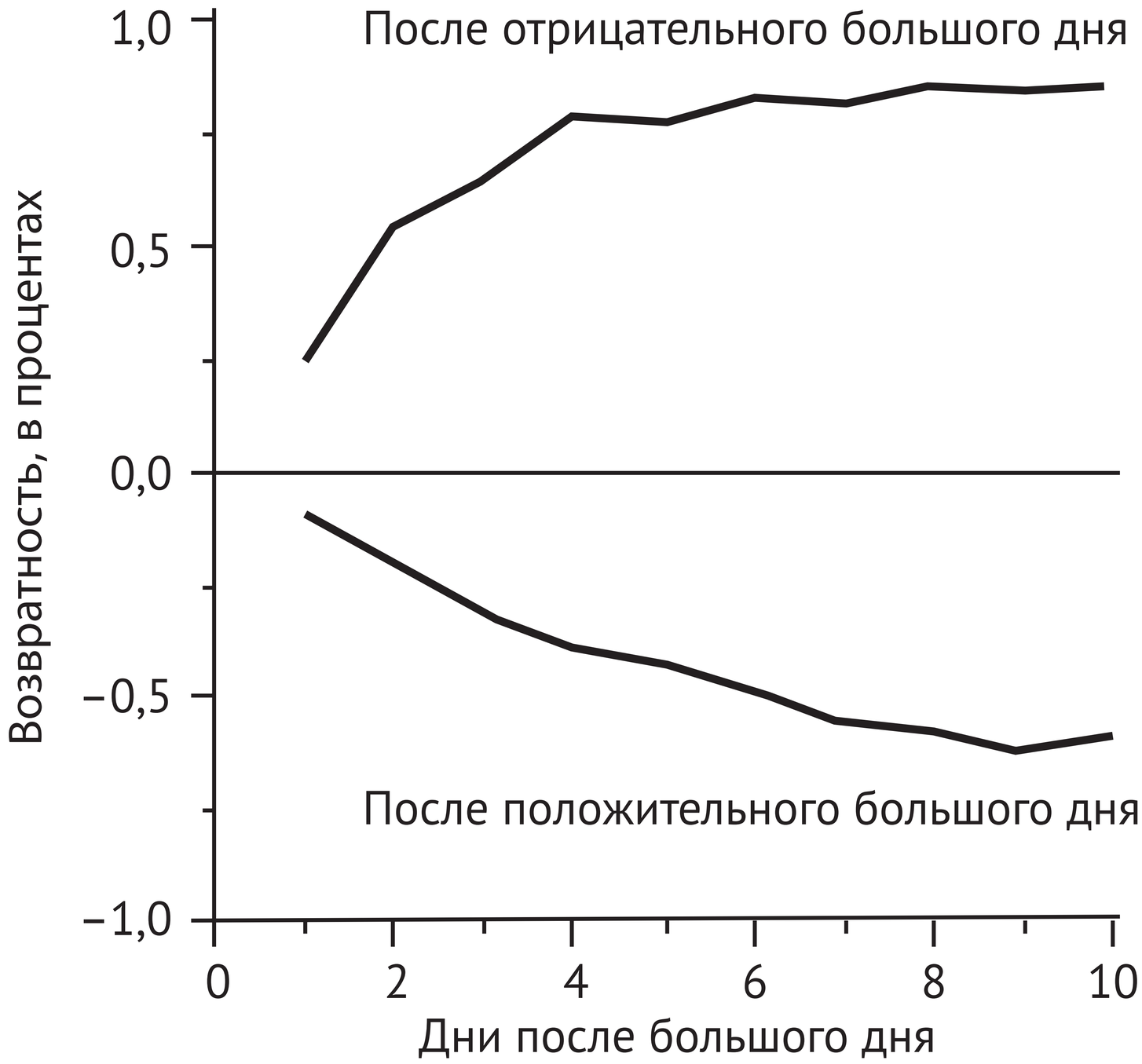

Рис. 9.7. Кумулятивная среднесуточная скорректированная возвратность после пятипроцентного большого дня

Чрезмерная реакция на негативные события кажется более чрезмерной в том смысле, что ценовые развороты в обратную сторону происходят чаще в дни большого спада, чем в дни большого подъема. Возможно, это как-то связано с описанным Канеманом и Тверски неприятием потерь. Чрезмерная реакция более выражена на плохие новости, чем на хорошие, потому что инвесторы более подвержены панике, чем жадности.

Это исследование дало интересные, правдоподобные и потенциально полезные результаты, потому что оно было основано на правдоподобной теории о том, что инвесторы чрезмерно реагируют на новостные события, а не на использовании алгоритма глубокой переработки данных для выявления загадочной и мимолетной закономерности, такой как тенденции роста цен на акции в течение четырнадцати дней после того, как они падали два дня подряд, а затем росли два дня подряд.

Чрезмерные ожидания

Облигации приносят проценты, а акции – дивиденды, вот почему Уоррен Баффет называет акции «замаскированными облигациями». Вплоть до конца 1950-х годов многие инвесторы оценивали акции и облигации, сравнивая процентные ставки по облигациям с дивидендной отдачей по акциям. Облигация стоимостью 100 долл., по которой выплачиваются проценты в размере 3 долл. в год, имеет трехпроцентную годовую возвратность. Акция стоимостью 100 долл., по которой выплачиваются дивиденды в размере 3 долл. в год, имеет трехпроцентную дивидендную отдачу. Поскольку акции являются более рискованными инструментами, чем облигации, инвесторы привыкли считать, что дивидендная отдача должна быть выше процентных ставок.

На рис. 9.8 показано, что вплоть до 1958 года средняя дивидендная отдача по индексу цен на акции S&P 500 почти всегда превышала процентную ставку по казначейским облигациям. В 1950 году средняя дивидендная отдача составляла почти девять процентов, в то время как процентная ставка по казначейским облигациям составляла всего два процента.

Это сравнение было неправильным. И действительно, Баффет утверждал, что между процентами и дивидендами существует важная разница, поэтому он редко инвестирует в облигации. Процентные выплаты по облигациям фиксированы, но дивиденды по акциям со временем могут (и обычно увеличиваются) наряду с корпоративными доходами. Если дивидендная отдача по акциям составляет 9 %, а дивиденды растут на 5 % в год, то уровень возвратности (норма прибыли) в годовом исчислении фактически составляет 14, а не 9 %.

Поскольку в 1950-х и 1960-х годах ценность экономического роста все больше осознавалась, растущие цены на акции привели к тому, что дивидендная отдача оказалась ниже отдачи облигаций. К началу 1970-х годов инвесторы, казалось, были заинтересованы только в росте, в особенности в акциях ведущих компаний с маркировкой Nifty 50. Среди этих немногих избранных были IBM, Xerox, Disney, McDonald’s, Avon, Polaroid и Schlumberger.

Рис. 9.8 Дивидендная отдача и процентные ставки

Как это часто бывает, инвесторы бросались из одной крайности в другую, платя, оглядываясь назад, нелепые цены. Они перешли от игнорирования экономического роста к фиксации на нем. Среднее соотношение цены к заработкам (P/E) для S&P 500 с 1871 года составляло около 15. В 1972 году у 16 из 50 акций Nifty это соотношение было выше 50, а у некоторых даже превышало 100.

Дэвид Дреман рассказал о мечтах об упущенных возможностях, которые возникали в головах инвесторов:

> Если бы кто-то вложил 10 000 долл. в компанию Haloid-Xerox в 1960 году, когда был представлен первый обычный копировальный аппарат 914, то десять лет спустя инвестиции составили бы 16,5 млн долл. За период с 1961 по 1966 год заработки компании McDonald's выросли в несколько тысяч раз, а затем – более скромно – снова выросли в четыре раза к 1971 году, когда был выпущен восьмимиллиардный гамбургер. Любой достаточно сообразительный человек, купивший акции McDonald's в 1965 году, когда они стали публичными, за следующие семь лет заработал бы в сорок раз больше. Инвестор, вложивший в 1914 году 2750 долл. в компанию Thomas J. Watson's Computing and Tabulating Company, к началу 1970-х годов владел акциями IBM на сумму более 20 млн долл.

Печальным последствием такой зацикленности на цене Nifty 50 стало убеждение, что никогда не бывает плохого времени для покупки растущих акций и что за них не приходится платить слишком высокую цену. Увлеченный растущими акциями финансовый управляющий написал, что «сейчас самое время покупать растущие акции. Все предназначение таких инвестиций – участие в будущих более крупных заработках, поэтому, по сути, любая задержка

с принятием решения ведет к поражению». Он не упомянул цену – упущение, которое неизменно является грехом инвестора.

Последующие результаты многих из Nifty 50 были разочаровывающими. В 1973 году Avon продавала свою компанию по 140 долл. за акцию, что в шестьдесят раз превышало заработки. В 1974 году цена упала до 19 долл.; двенадцать лет спустя, в 1986 году, она продавалась по 25 долл. за акцию. В 1972 году Polaroid был продан за 150 долл. (в 115 раз больше заработков!), в 1974 году – за 14 долл., а в 1986 году – за 42 долл., прежде чем обанкротился в 2001 году.

Это были не единичные случаи. Многие гламурные акции, которые в 1970-х годах достигали невероятных уровней соотношения цены к заработкам (P/E), в течение следующих нескольких десятилетий демонстрировали значительно худшие результаты, чем рынок. Инвестор, выбравший крайне неудачный момент и купивший акции Nifty 50 в конце 1972 года с соотношениями P/E выше 50, в течение следующих сорока пяти лет (до декабря 2018 года) получил бы среднегодовую возвратность в размере отрицательных 14 %, по сравнению со средней возвратностью более 10 % у S&P 500.

Как не быть введенным в заблуждение призрачными закономерностями

Закономерности не обязательно должны быть комбинациями цифр. Например, сотрудники – от клерков до генеральных директоров, – которые превосходно выполняют свою работу, нередко оказываются менее успешными, когда их переводят на новые должности, и это разочарование увековечено принципом Питера: «Управленцы достигают уровня своей некомпетентности».

Закономерности в данных наблюдений могут вводить в заблуждение из-за систематического смещения за счет самоотбора, поскольку наблюдаемые различия между людьми, принимающими разные решения, могут быть обусловлены типом людей, принимающих такие решения. Рандомизированные контролируемые испытания и A/B-тесты, когда это практически возможно, являются эффективными способами определения причины и следствия. При вынужденном использовании данных наблюдений важно, чтобы перед тем, как обращаться к данным, были детализированы проверенные теории. В противном случае мы, скорее всего, мы будем одурачены призрачными закономерностями.

Эпилог

Все дело в Байесе

В XVIII веке пресвитерианский священник по имени Томас Байес (Thomas Bayes) бился над сложным вопросом – какова вероятность существования Бога. Преподобный Байес был не только священником, он изучал логику и, скорее всего, математику в Эдинбургском университете (пресвитерианам не разрешалось посещать английские университеты).

Многие интеллектуалы в то время спорили о том, что до сих пор остается логической загадкой. Если вселенную создал благожелательный Бог и неусыпно следит за нами, почему в мире так много страданий и зла? Неужели заботливый Бог допустил бы, чтобы хорошие люди голодали, верующие умирали от ужасных болезней, а невинные дети гибли от рук солдат?

Первой попыткой Байеса вступить в эту дискуссию была возвышенная, но не очень убедительная брошюра, опубликованная под псевдонимом Джеймс Нун, с кажущимся бесконечным названием «*Божественная благосклонность, или Попытка доказать, что главной целью божественного провидения и руководства является счастье его созданий. Это ответ на брошюру, озаглавленную Божественная праведность [Джона Бэлгуя]. <…> С опровержением выдвинутых в ней представлений о красоте и порядке, причинах наказания и необходимости испытания, предшествующего совершенному счастью*».

Обоснованно неудовлетворенный взаимными доводами, Байес решил привести математический аргумент – расчет вероятности существования Бога!

В то время математики быстро разрабатывали теоремы и доказательства для расчета вероятности различных азартных игр. Например, если десять раз подбросить честную монету, то какова вероятность выпадения пяти орлов и пяти решек? (Ответ равен 0,246.)

Байес задал обратный вопрос, т. е. *инверсную* вероятность: если монету подбросить десять раз и она выпадет пять раз орлом и пять раз решкой, то какова вероятность того, что это честная монета? Если бы он смог ответить на такой вопрос, то, возможно, смог бы вычислить вероятность существования Бога.

Вместо того чтобы пытаться определить эту вероятность:

> Если Бог существует, то какова вероятность того, что мир будет таким, какой он есть?

Байес хотел рассчитать инверсную вероятность:

> Если мир таков, каков он есть, то какова вероятность того, что Бог существует?

Байесу не удалось вычислить вероятность существования Бога, но его работа о том, как перейти от одной вероятности к инверсной, оказалась невероятно ценной и в настоящее время является основой для байесовского подхода к теории вероятностей и статистике.

Теорема Байеса

При расчете инверсных вероятностей для различных азартных игр Байес обнаружил, что для расчета требуется априорная вероятность (до того, как результат будет наблюдаться), которая затем пересматривается в апостериорную вероятность (после того, как результат будет наблюдаться).

Давайте разберем простой пример, чтобы увидеть, как априорные вероятности пересматриваются в апостериорные. Предположим, что в черный мешочек помещены двадцать монет; девятнадцать – обычные монеты, а у одной монеты орлы с обеих сторон. Из мешочка случайным образом выбирается монета, на одной стороне которой изображен орел. Какова вероятность того, что это монета с двумя орлами? Как вы думаете? Вероятность того, что это монета с двумя орлами, составляет десять процентов? Тридцать процентов? Пятьдесят процентов? Семьдесят процентов?

В табл. Е.1 показан простой способ организации нашего мышления. Предположим, что эксперимент проводится сорок раз. Поскольку вероятность того, что монета с двумя орлами будет выбрана с вероятностью один к двадцати, можно ожидать, что эта монета будет выбрана дважды, а обычная монета – тридцать восемь раз. В тех двух случаях, когда выбирается монета с двумя орлами, всегда выпадает орел. В тридцати восьми случаях, когда выбирается обычная монета, можно ожидать, что девятнадцать раз выпадет орел, а девятнадцать раз – решка.

Таблица Е.1. Инверсная вероятность

	Выпали орлы	Выпали решки	Итого
Выбрана монета с двумя орлами	2	0	2
Выбрана нормальная монета	19	19	38
Итого	21	19	40

Таким образом, из двадцати одного случая, когда выпавшей стороной является орел, вероятность того, что это монета с двумя орлами, составляет 9,5 %, 2/21 = 0,095.

Перед выбором монеты априорная вероятность выбора монеты с двумя орлами равна 1/20. После того как монета выбрана и одна сторона показывает орла, апостериорная вероятность выпадения монеты с двумя орлами почти удваивается с 1/20 до 2/21.

Это понимание имеет решающее значение для современного использования правила Байеса. Апостериорная вероятность пересматривает априорную вероятность на основе данных, которые стали доступны. В этом суть усваивания знаний из данных – в использовании новой информации для пересмотра нашего мнения.

Джону Мейнарду Кейнсу принадлежит замечательный ответ на критику в адрес того, что он изменил свои политические рекомендации: «Когда меняется моя информация, я меняю свои выводы. А вы что делаете, сэр?» Было бы интеллектуально нечестно не менять свое мнение в свете новой информации. И правило Байеса направляет нас при таких пересмотрах.

Давайте применим это рассуждение к медицинскому диагнозу.

Медицинские диагнозы

Некоторые медицинские диагнозы – прямолинейны. Да, вы подвернули лодыжку. Да, вы ушибли большой палец молотком. Другие диагнозы – менее однозначны. Предположим, что во время обычного медицинского осмотра врач обнаруживает подозрительное образование в груди пациентки. Сказать с уверенностью, является ли опухоль злокачественной либо доброкачественной, просто невозможно. Не исключено, что врач знает статистику, что среди женщин с аналогичной историей болезни опухоль оказывается злокачественной в одном случае из 100. Для того чтобы собрать больше информации, врач назначает маммографию. Известно, что в тех случаях, когда опухоль является злокачественной, существует вероятность 0,80 того, что рентген даст положительный результат, а когда опухоль является доброкачественной, существует вероятность 0,90 того, что рентген даст отрицательный результат.

Априорная вероятность того, что опухоль является злокачественной, равна 0,01; апостериорная вероятность – это пересмотренная вероятность, основанная на результатах рентгена. Предположим, что рентгеновский снимок оказался положительным. Какова, по вашему мнению, апостериорная вероятность того, что опухоль является злокачественной? Восемьдесят процентов? Девяносто процентов?

Используя правило Байеса, получается, что апостериорная вероятность равна 7,5 %. Это особенно интересный пример, потому что он является очень яркой иллюстрацией того, как вероятности и инверсные вероятности бывают совершенно разными. Даже если злокачественная опухоль будет правильно идентифицирована как таковая в 80 % случаев, вероятность того, что положительный рентгеновский снимок отражает злокачественную опухоль, составляет всего 7,5 %.

Это также показывает, как правило Байеса изменяет вероятности в свете новых данных, – здесь с априорной вероятности 0,01 (до рентгеновского снимка) до апостериорной вероятности 0,075 (после рентгеновского снимка). Если бы рентгеновский снимок оказался отрицательным, то апостериорная вероятность была бы пересмотрена в сторону понижения до 0,002.

Правило Байеса можно использовать снова и снова, чтобы продолжать пересматривать вероятность, основываясь на результатах дополнительных рентгеновских снимков. Предположим, что первый рентгеновский снимок оказался положительным, и врач назначает повторный рентген. Теперь правило Байеса говорит, что вероятность 0,075 после первого теста либо увеличивается до 0,393, либо уменьшается до 0,018, в зависимости от показаний второго рентгеновского снимка.

Неудивительно, что компьютеры широко опираются на правило Байеса, когда используют данные для пересмотра вероятности того, что вам понравится определенный фильм, гаджет или политическая позиция. Например, есть шансы, что если вам понравилось телешоу «Во все тяжкие», то вам больше понравятся фильмы «Лучше звоните Солу» и «Эль Камино». Для количественной оценки этих вероятностей можно использовать правило Байеса. Кроме того, правило Байеса может учитывать другую релевантную информацию, которая может повлиять на ваши предпочтения, такую как ваш возраст, пол, род занятий и доход, и может обновлять свою оценку вероятности по мере того, как узнает о вас больше, – точно так же как правило Байеса учитывало результаты повторного рентгена при оценивании вероятности злокачественного новообразования.

При использовании правила Байеса компьютеры имеют явное преимущество перед людьми, потому что мы склонны совершать две систематические ошибки. Во-первых, мы не в полной мере осознаем разницу между вероятностями и инверсными вероятностями, и, во-вторых, мы не пересматриваем вероятности так, как того требует правило Байеса.

В примере с маммографией какова, по вашим прикидкам, была вероятность того, что опухоль является злокачественной, после того как первый рентген дал положительный результат? Она была ближе к 7,5 либо к 80 %? Один и тот же вопрос был задан сотне врачей, и девяносто пять из них дали вероятность около 75 %, хотя правильная вероятность составляет одну десятую от этой! По словам проводившего это исследование ученого, «ошибавшиеся врачи обычно сообщают, что, по их предположениям, вероятность рака при условии, что у пациента был положительный рентгеновский снимок… была приблизительно равна вероятности положительного рентгеновского снимка у пациента с раком». Неправильное толкование медиками инверсных вероятностей может иметь трагические последствия.

Помимо того, что люди путают вероятности и инверсные вероятности, они не очень хорошо умеют пересматривать вероятности на основе новых данных. Помните пример с двадцатью монетами, одна из которых с двумя орлами? Какой была ваша догадка о вероятности того, что выбранная монета имела орлов с обеих сторон? Была ли она десять процентов (правильный ответ) или ближе к тридцати, пятидесяти или даже семидесяти процентам? Не смущайтесь. Когда речь идет о пересмотре вероятностей в свете новой информации, несовершенство человеческой интуиции, как известно, проявляет себя во всей красе.

Этот недостаток настолько распространен, что у него даже есть название: *логическая ошибка неучета базовой частоты*[14]. Когда у людей есть общая информация и конкретные данные, они склонны выделять конкретные данные и пренебрегать общей информацией (или информацией о базовой частоте). Предположим, что базовая частота такова, что менее четверти процента женщин в стране являются юристами, а конкретные данные таковы, что Джилл носит деловой костюм и носит портфель. Какова вероятность того, что Джилл юрист? Большинство людей зациклены на том, что юристы часто носят дело-

[14] Логическая ошибка неучета базовой частоты (base rate fallacy) – это когнитивное искажение, при котором люди склонны игнорировать базовую, статистическую информацию о частоте какого-либо события в пользу более конкретных данных о частном случае. – *Прим. перев.*

вые костюмы и портфели, и не обращают должного внимания на базовую частоту, равную 0,25 %.

Это еще один из путей, которым мы подвержены обману со стороны призрачных закономерностей. Мы склонны придавать большое значение малой порции данных – запоминающейся или необычной закономерности, – хотя не должны этого делать.

Правило Байеса в зале суда

Суды присяжных тоже дают убедительные свидетельства того, что люди не понимают инверсных вероятностей и допускают ошибки при пересмотре вероятностей. Например, в статье журнала Law Review приводится пример того, как присяжные заседатели нередко ошибаются в оценке ценности ассоциативных свидетельств, таких как отпечатки пальцев или образцы крови, которые совпадают с образцами крови обвиняемого.

Представьте, что вы являетесь присяжным заседателем в суде по делу об убийстве и, основываясь на свидетельствах, которые вы уже услышали, считаете, что вероятность виновности подсудимого составляет пятьдесят на пятьдесят. Обвинение теперь представляет свидетельства, убедительно доказывающие, что, судя по углу нанесения смертельного удара, нападавший был правшой. Вы смотрите на стол защиты и видите, что обвиняемый делает пометки правой рукой. Какова, по вашему мнению, вероятность того, что обвиняемый виновен?

Согласно правилу Байеса, апостериорная вероятность виновности составляет 0,53, что лишь незначительно превышает априорную вероятность, равную 0,50. Показания правши имеют малую доказательную ценность, поскольку праворукость очень распространена среди населения. Они стоят немногим больше, чем доказательство того, что у нападавшего есть нюх. Тем не менее авторы статьи в Law Review утверждают, что, по их опыту, присяжные нередко придают большое значение свидетельствам, которые практически не имеют ценности.

Ситуация была бы иной, если бы обвиняемый оказался левшой и были представлены убедительные свидетельства того, что нападавший является левшой, поскольку леворукость встречается редко. В этой ситуации апостериорная вероятность возрастает до 0,91.

Как ни странно, хотя присяжные заседатели зачастую принимают во внимание тривиальные свидетельства, они, как правило, недооценивают весомые свидетельства. Исследователи провели инсценированные судебные процессы, в ходе которых присяжные-добровольцы заслушивали свидетельства и им поручалось записывать свои личные оценки вероятности вины обвиняемого. В ходе одного эксперимента 144 присяжным-добровольцам сообщили, что винный магазин был ограблен мужчиной в лыжной маске. Возле магазина полиция арестовала подозреваемого, рост, вес и одежда которого соответствовали описанию, полученному от продавца. Лыжная маска и деньги были найдены в соседнем мусорном баке.

После заслушивания этих показаний присяжных попросили записать их оценку вероятности того, что арестованный «действительно это совершил». Средняя вероятность составила 0,25.

Затем судебно-медицинский эксперт засвидетельствовал, что образцы волос подозреваемого совпадают с волосами, найденными внутри лыжной маски, и что только у двух процентов населения волосы совпадают с волосами, найденными в лыжной маске. Половине присяжных также сказали, что в городе с населением в миллион человек эта двухпроцентная вероятность означает, что волосы совпадут примерно у 20 000 человек.

Девятнадцать присяжных дали апостериорную вероятность 0,98, очевидно, полагая, что если такой тип волос только у двух процентов невиновных людей, то вероятность невиновности обвиняемого составляет всего два процента. Это называется *логической ошибкой прокурора*, потому что такой аргумент используется прокурорами для вынесения обвинительного приговора. Это логическая ошибка, потому что присяжные заседатели путают вероятности и инверсные вероятности – в частности, они не проводят различия между вероятностью того, что процент невиновных имеет такой тип волос, и инверсной вероятностью того, что человек с таким типом волос невиновен.

Шесть из семидесяти двух присяжных, которым была дана только двухпроцентная вероятность, и двенадцать из семидесяти двух присяжных, которым были даны как двухпроцентная вероятность, так и число в 20 000, вообще не пересматривали свои вероятности, очевидно, полагая, что совпадение волос является бесполезной информацией, потому что очень многие люди имеют такой тип волос. Это называется *логической ошибкой адвоката защиты*, потому что такой аргумент используется адвокатами защиты. Это логическая ошибка, потому что обвиняемый не был выбран случайным образом из 20 000 человек с таким типом волос. Другие свидетельства, касающиеся описания подозреваемого, а также выброшенной лыжной маски и денег, уже убедили присяжных в том, что существует значительная вероятность того, что обвиняемый – виновен.

При априорной вероятности, равной 0,25, из правила Байеса следует, что апостериорная вероятность после учета совпадения волос равна 0,94. Три четверти присяжных заседателей не попались на удочку обвинения или адвоката защиты; однако их пересмотренные вероятности были неизменно более консервативными, чем байесовские апостериорные вероятности, вытекавшие из их априорных вероятностей. Средняя пересмотренная вероятность составила всего 0,63.

В экспериментах с участием присяжных заседателей этот консерватизм проявлялся снова и снова. В десяти таких экспериментах, где средние байесовские апостериорные вероятности находились в диапазоне от 0,80 до 0,997, средние пересмотренные вероятности присяжных были неизменно ниже и составляли от 0,28 до 0,75.

Присяжные заседатели, которые, в конце концов, тоже люди, склонны полагать, что важны тривиальные данные, а важные данные неубедительны. Они считают, что случайная закономерность важна, в то время как уйма данных – нет.

Одним из крайних примеров был «процесс века» в отношении О. Джея Симпсона, в ходе которого присяжные, по-видимому, пришли к выводу, что указывающие на вину свидетельства ДНК были менее важны, чем тот факт, что найденные на месте преступления пропитанные кровью кожаные перчат-

ки не подходили Симпсону по размеру: «Если они не подходят, то вы должны его оправдать». Было много возможных объяснений, почему они не подходили: Симпсон был в латексных перчатках, когда примерял кожаные перчатки; возможно, перчатки стали жесткими из-за того, что их пропитали кровью, несколько раз замораживали и размораживали; Руки Симпсона распухли из-за того, что он перестал принимать лекарства от артрита две недели назад, и Симпсон притворился, что ему трудно надевать перчатки.

Присяжные предпочли тривиальное серьезному, возможно, потому что им было трудно оценить важность каждого из них, или, возможно, они увидели то, что хотели увидеть.

Ирония заключается в том, что байесовский анализ, который предназначен для того, чтобы позволить людям выражать и пересматривать свои личные вероятности, выполняется компьютерами лучше, чем людьми. С другой стороны, компьютеры по-прежнему сталкиваются с непреодолимой проблемой: оценкой того, *следует* ли использовать конкретные данные для пересмотра вероятностей, т. е. различением хороших данных от плохих. Компьютеры в этом отношении ужасны.

Когда речь идет об оценке качества данных, у людей есть явное преимущество. Подумайте еще раз об инсценированном судебном процессе с участием грабителя в лыжной маске и показании эксперта о том, что обнаруженные в лыжной маске волосы встречаются только у двух процентов населения. Возможно, люди, которые не изменили свою вероятность вины или изменили ее не так сильно, как это вытекает из правила Байеса, не доверяли показанию эксперта. Они вполне могли не учитывать двухпроцентную цифру, поскольку предполагали, что обвинение будет использовать платного эксперта, который дал наиболее благоприятное показание. Как вариант, они, возможно, рассматривали возможность того, что полиция подложила волосы в лыжную маску. Их пересмотренные вероятности, возможно, были для них правильными субъективными вероятностями.

Недавно Джей был вызван в качестве присяжного заседателя и просидел на галерее целых четыре дня, в то время как потенциальных присяжных разбивали на группы по двенадцать человек и допрашивали адвокаты обвинения и защиты в надежде найти двенадцать подходящих присяжных заседателей. Вначале было восемьдесят потенциальных присяжных, и адвокаты проверили семьдесят пять из них, после чего остановились на двенадцати присяжных. Примечательно, что Джей был в числе пяти, которых не допрашивали.

Он испытал некоторое облегчение, потому что между ним и адвокатом обвиняемого, который неоднократно инструктировал присяжных о том, как им следует обращаться со «свидетельством склонности», неизбежно должна была состояться оживленная дискуссия. В некоторых ситуациях судьи могут разрешать обвинению представлять свидетельства совершения подсудимым аналогичных преступлений в прошлом. Адвокат сказал присяжным, что «если свидетельства *по этому делу* не будут доказаны вне всяких разумных сомнений, вы не сможете потом сказать: “Ну он уже совершал нечто подобное в прошлом. Нет дыма без огня”, – и затем признать его виновным. Люди все равно должны доказать каждое обвинение *в этом деле* вне всяких разумных сомнений».

Свидетельств прошлых преступлений само по себе недостаточно, чтобы доказать, что подсудимый виновен в текущем преступлении, но адвокат, по-видимому, утверждал, что свидетельства прошлых преступлений никогда не должны убеждать присяжных изменять свое заключение с «невиновен» на «виновен». Если бы это было правдой, то свидетельствам склонности подсудимого не придавалось бы никакого значения, и судьи никогда не должны были позволять обвинению представлять свидетельства совершения обвиняемым аналогичных преступлений.

Предположим, что присяжный заседатель считает 95%-ную вероятность вины «не вызывающей разумных сомнений», а все доказательства, относящиеся к текущему преступлению, указывают на вероятность вины в 94 %. Наличие свидетельства совершения аналогичных преступлений в прошлом, несомненно, повысило бы вероятность выше 95 %. Если это так, то свидетельства по текущему делу не обязательно должны быть «вне разумных сомнений», чтобы присяжный пришел к выводу, что все свидетельства вины не вызывают разумных сомнений. Излишне говорить, что у многих потенциальных присяжных возникли проблемы с пониманием аргументации адвоката, и их исключили из состава присяжных, потому что казалось, что на них повлияют прошлые преступления, – хотя прошлые преступления должны были повлиять на них!

Джей слышал, что в еще одном деле потенциального присяжного спросили: «Как вы думаете, больше ли шансов, что обвиняемый будет признан виновным только потому, что его арестовали?» Присяжный ответил: «Конечно! Если только вы не арестовываете людей наугад». Это было сказано истинным байесовцем.

Закономерности и апостериорные вероятности

Та же логика применима к оценке закономерностей, выявленных с помощью глубокой переработки данных. Предположим, что мы используем алгоритм глубокой переработки данных для отыскания закономерности, которая может быть использована для предсказания цен на акции, лечения болезни или для достижения какой-либо другой стоящей цели. Мы используем надежный статистический тест, который в 95 % случаев правильно выявляет реальную закономерность как действительно полезную, а случайно возникшую – как действительно бесполезную.

Мы знаем, что существует много бесполезных закономерностей по сравнению с действительно полезными. Поэтому давайте предположим, что одна из каждых 1000 закономерностей, которые можно найти с помощью глубокой переработки данных, полезна, а остальные 999 бесполезны. Наша априорная вероятность, до того как мы найдем закономерность и проведем статистический тест, равна 1/1000. После того как мы нашли закономерность и определили, что она статистически значима, апостериорная вероятность того, что она полезна, оказывается меньше 1/500. Это больше, чем 1/1000, но вряд ли убедительно. И все же гораздо ближе к тому, что была обнаружена закономерность, которая на самом деле бесполезна.

Это яркий пример того, как вероятности и инверсные вероятности бывают совершенно разными. Вероятность того, что игрок английской премьер-лиги является мужчиной, составляет 100 %, но вероятность того, что случайно выбранный мужчина играет в английской премьер-лиге, близка к 0. Здесь, несмотря на то что шанс статистической значимости бесполезной закономерности составляет всего 5 %, более 98 % статистически значимых закономерностей бесполезны.

Априорная вероятность 1/1000, безусловно, является слишком оптимистичной при глубокой переработке больших данных. В табл. E.2 приведены апостериорные вероятности для других значений априорной вероятности.

Таблица E.2. Вероятность полезности обнаруженной закономерности

Априорная вероятность	Апостериорная вероятность при статистической значимости
0,001	0,018664
0,0001	0,001897
0,00001	0,000190
0,000001	0,000019

Мы не знаем точно, сколько бесполезных закономерностей ожидает своего обнаружения, но мы знаем, что с большими данными и мощными компьютерами это очень большое число, и с каждым днем оно становится все больше. В главе 5 рассказывается об истории студента, которого послали по ложному следу предсказать обменный курс турецкой лиры к доллару США. Без особых усилий он просмотрел более 75 млн бесполезных закономерностей.

Наша точка зрения является предостерегающей. Наши далекие предки выжили и процветали, потому что распознавали полезные закономерности в окружающей их среде. Сегодня резкий рост числа измеряемых и регистрируемых факторов может предоставлять нам полезную информацию, но он также невероятно увеличил количество случайных совпадений и ложных статистических зависимостей, которые только и ждут, чтобы ввести нас в заблуждение.

В настоящее время объем данных составляет около сорока триллионов гигабайт, что в сорок раз больше, чем количество звезд в известной Вселенной, и процесс создания данных не замедляется. Девяносто процентов всех данных, которые когда-либо были собраны, были созданы за последние два года.

Количество возможных закономерностей практически не ограничено, а возможности больших компьютеров по их отысканию поражают. Поскольку полезных закономерностей относительно немного, бóльшая часть найденного – это мусор, мнимые корреляции, которые могут нас позабавить, но не должны восприниматься всерьез.

И пусть вас не вводят в заблуждение призрачные закономерности.

Библиография

1. Соревнования всех звезд НБА 2010–2019 годов по трехочковым броскам. 2010–2019 NBA All Star 3-Point Contests. Accessed September 29, 2019. **https://www.justallstar.com/contests/3point/3p-2019-2018-2017-2016-2015-2014-2013-2012-2011-2010/#1548011618477-20160cea-ddf95d9f-c82f**.
2. Abney, Wesley. Прямой эфир из Вашингтона. Лотерейная ночь 1969 года. Live from Washington, It's lottery night 1969!, Vietnam, History Net. 2010. **https://www.historynet.com/live-from-dc-its-lottery-night-1969.htm**.
3. Anderson, Chris. Конец теории: лавина данных делает научный метод устаревшим, The end of theory: The data deluge makes the scientific method obsolete, Wired, June 23, 2008. **https://www.wired.com/2008/06/pb-theory/**.
4. Anderson, Christopher, J., et al. Ответ на комментарий к статье «Оценивание воспроизводимости в психологической науке», Response to comment on 'Estimating the reproducibility of psychological science', Science 351 (2016): 1037.
5. Anderson, Jeff, Smith, Gary. Отличная компания может быть отличной инвестицией, A great company can be a great investment. Financial Analysts Journal 62, no. 4 (2006): 86–93.
6. Angrist, Joshua, D. Заработок за всю жизнь и лотерея призыва на военную службу времен войны во Вьетнаме: данные из административных записей социального обеспечения, Lifetime earnings and the Vietnam era draft lottery: Evidence from Social Security administrative records, The American Economic Review 80, no. 3 (1990): 313–36.
7. Asmelash, Leah, Muaddi, Nadeem. Ребенок родился в день 7-Eleven в 19:11, весом 7 фунтов и 11 унций, Baby born on 7-Eleven Day at 7:11 p.m., weighing 7 pounds and 11 ounces. CNN online. July 13, 2019. **https://edition.cnn.com/2019/07/13/us/711-baby-trnd/index.html**.
8. Athalye, Anish, Engstrom, Logan, Ilyas, Andrew, Kwok, Kevin. Одурачивание нейронных сетей в физическом мире с помощью противоборствующих 3D-объектов, Fooling neural networks in the physical world with 3D adversarial objects. labsix. October 31, 2017. **https://www.labsix.org/physical-objects-that-fool-neural-nets/**.
9. Athey, Susan. Влияние машинного обучения на экономику, The Impact of Machine Learning on Economics, in The Economics of Artificial Intelligence: An Agenda, edited by Ajay Agrawal, Joshua Gans, Avi Goldfarb, 507–52. Chicago, IL: University of Chicago Press, 2018.
10. Aubrey, Allison, Может ли диета беременной женщины повлиять на пол ребенка? Can a pregnant woman's diet affect baby's sex? Morning Edition, NPR. January 15, 2009. **https://www.npr.org/templates/story/story.php?storyId=99346281&t=1580054221847**.
11. Baker, Monya, 1500 ученых приоткрывают завесу над воспроизводимостью, 1,500 scientists lift the lid on reproducibility. Nature 533, no. 7604 (2017): 452–4.
12. Bayer, Dave, and Diaconis, Persi. Отслеживание ласточкиного хвоста до ее гнезда, Trailing the dovetail shuffle to its lair. Annals of Applied Probability 2, no. 2 (1992): 294–313.
13. Begoli, Edmonn, Horey, James. Конструктивные принципы для эффективного обнаружения знаний из больших данных, Design principles for effective knowledge

discovery from big data. Paper presented at Software Architecture (WICSA) and European Conference on Software Architecture (ECSA), 2012 Joint Working IEEE/IFIP Conference, Helsinki, August 2012.

14. Bem, Daryl, J. Чувство будущего: экспериментальные свидетельства аномального ретроактивного влияния на познание и эмоции, Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect. Journal of Personality and Social Psychology 100, no. 3 (2011): 407–25.
15. Bhattacharjee, Yudhijit. The mind of a con man. The New York Times Magazine, April 24, 2013.
16. Bode, Johann Elert. Ясное наставление для познания звездного неба, Deutliche Anleitung zur Kenntniß des gestirnten Himmels [Clear Instruction for the Knowledge of the Starry Heavens]. Hamburg: Dietrich Anton Harmsen, 1772.
17. Bohannon, J. Попытка повторить эксперимент провоцирует похвалу – и обвинения в «запугивании», Replication effort provokes praise – and "bullying" charges. Science 344 (2014): 788–9.
18. Bolen, Johan, Mao, Huina, Zeng, Xiaojun. Настроение в Twitter предсказывает рынок акций, Twitter mood predicts the stock market. Journal of Computational Science 2, no. 1 (2011): 1–8.
19. Borjas, G. J. Этнический капитал и межпоколенческая мобильность, Ethnic capital and intergenerational mobility. Quarterly Journal of Economics 107, no. 1 (1992): 123–50.
20. Borjas, G. J. Межпоколенческая мобильность иммигрантов, The intergenerational mobility of immigrants, Journal of Labor Economics 11, no. 1 (1993): 113–35.
21. Brown, Nicholas, J. L., Coyne, James C. Может ли язык Twitter надежно предсказывать сердечное заболевание? Does Twitter language reliably predict heart disease? A commentary on Eichstaedt et al. (2015a). PeerJ 6 (2018): e5656. **https://doi.org/10.7717/peerj.5656**.
22. Camerer, Colin, F., et al. Оценивание воспроизводимости лабораторных экспериментов в экономике, Evaluating replicability of laboratory experiments in economics. Science 351 (2016): 1433–6.
23. Camerer, Colin, F., et al. Оценивание воспроизводимости экспериментов в области социальных наук в журналах Nature и Science в период с 2010 по 2015 год, Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. Nature Human Behaviour 2, no. 9 (2018): 637–44.
24. Carliner, G. Заработная плата, заработки и часы работы американских мужчин первого, второго и третьего поколения, Wages, earnings, and hours of first, second, and third generation American males, Economic Inquiry 18, no. 1 (1980): 87–102.
25. Carrier, David, R., Kapoor, A. K., Tasuku Kimura, Nickels, Martin K., Scott, Eugenie C., So, Joseph K., Trinkaus, Erik. Энергетический парадокс человеческого бега и эволюции гоминидов [и комментарии и ответы], The energetic paradox of human running and hominid evolution [and comments and reply]. Current Anthropology 25, no. 4 (1984): 483–95. **https://doi.org/10.1086/203165**.
26. CBS News online. Трампа освистали на Мировой серии и скандирование: «Заприте его в клетке!» Trump booed at World Series and chants of 'Lock him up!' break out. October 29, 2019. **https://www.cbsnews.com/news/trump-booed-world-series-lock-him-up-chant-break-out-during-game-5-washington-nationals-stadium-pa/**.
27. Chambers, Sam. Единственное, что алгоритм больших данных Tesco никогда не предсказывал, The one thing Tesco's big data algorithm never predicted. Bloomberg online, October 20, 2015.

28. Chappell, Bill. Неравенство доходов в США ухудшается, образуя новый разрыв, U.S. income inequality worsens, widening to a new gap. NPR. September 26, 2019. **https://www.npr.org/2019/09/26/764654623/u-s-income-inequality-worsens-widening-to-a-new-gap**.

29. Chatterjee, Rhitu. Новые подсказки обнаружены в затонувшем многовековом шведском корабле, New clues emerge in centuries-old Swedish shipwreck. PRI, February 23, 2012. **https://www.pri.org/stories/2012-02-23/new-clues-emerge-centuries-old-swedish-shipwreck**.

30. Chiswick, B. R. 1977. Сыновья иммигрантов: находятся ли они в невыгодном положении с точки зрения заработка?, Sons of immigrants: are they at an earnings disadvantage? American Economic Review: Papers and Proceedings, 67(1), 376–80.

31. Choi, Hyuntoung. 2012. Предсказывание настоящего с помощью Google Trends, Predicting the Present with Google Trends, The Economic Record, 88, special issue, 2–9.

32. Chordia, Tarun, Goyal, Amit, and Saretto, Alessio. Манипуляция p-значением: свидетельства из двух миллионов торговых стратегий, p-hacking: Evidence from two million trading strategies. Swiss Finance Institute Research Paper No. 17–37. Zurich: Swiss Financial Institute, 2017.

33. Christian, Brian. A/B-тест: внутри технологии, которая меняет правила бизнеса, The A/B Test: Inside the technology that's changing the rules of business. Wired, April 25, 2012.

34. Cios, Krzysztof, J., Pedrycz, Witold, Swiniarski, Roman W., Kurgan, Lukasz Andrzej. Глубокая переработка данных: подход к обнаружению знаний, Data Mining: A Knowledge Discovery Approach, New York: Springer, 2007.

35. Climate Research Board. Углекислый газ и климат, Carbon Dioxide and Climate. Washington, D.C.: National Academies Press, 1979. **https://doi.org/10.17226/19856**.

36. Coase, Ronald. Как экономистам следует делать выбор?, How should economists choose? in Ideas, Their Origins and Their Consequences: Lectures to Commemorate the Life and Work of G. Warren Nutter, Washington, D.C: American Enterprise Institute for Public Policy Research, 1988.

37. Courtland, Rachel. Гордон Мур: человек, чье имя означает прогресс. Инженер-визионер размышляет о 50-летии закона Мура, Gordon Moore: The man whose name means progress. The visionary engineer reflects on 50 years of Moore's Law. IEEE Spectrum online. March 30, 2015. **https://spectrum.ieee.org/computing/hardware/gordon-moore-the-man-whose-name-means-progress**.

38. Crane, Burton. Изощренный инвестор, The Sophisticated Investor. New York: Simon and Schuster, 1959, 56.

39. Docquier, Frédéric. Распределение дохода, невыпуклости и связь между рождаемостью и доходом, Income distribution, non-convexities and the fertility-income relationship. Economica 71, no. 282 (2004): 261–73.

40. Dunn, Peter, M. Джеймс Линд (1716–94) из Эдинбурга и лечение цинги, M. James Lind (1716–94) of Edinburgh and the treatment of scurvy. Archives of Disease in Childhood 76 (1997): F64–5.

41. Egami, Naoki, Fong, Christian J., Grimmers, Justin, Roberts, Margaret E., Stewart, Brandon M. Как делать причинно-следственные выводы с помощью текста, How to make causal inferences using text. 2018. arXiv:1802.02163v12016.

42. Engber, Daniel. Дэрил Бем доказал реальность экстрасенсорного восприятия: это означает, что наука нарушена, Daryl Bem proved ESP is real: Which means science is broken. Slate, May 17, 2017.

43. Enserink, Martin. Заключительный отчет: дело Стапеля указывает на более серьезные проблемы в социальной психологии, Final report: Stapel affair points to bigger problems in social psychology. Science, November 28, 2012.
44. Fairley, Richard, E. Почему затонуло судно «Васа»: 10 извлеченных уроков, Why the Vasa sank: 10 Lessons Learned. IEEE Software online 20, no. 2 (2003): 18–25.
45. Fayyad, Usama, Piatetsky-Shapiro, Gregory, Smyth, Padhraic. От глубокой переработки данных к обнаружению знаний в базах данных, From data mining to knowledge discovery in databases. AI Magazine 17, no. 3 (1996): 37–54.
46. Fink, Sheri. Это высокотехнологичное решение для ликвидации последствий стихийных бедствий может оказаться слишком хорошим, чтобы быть правдой, This high-tech solution to disaster response may be too good to be true. The New York Times, August 9, 2019.
47. Fisher, Philip, A. Обыкновенные акции и необыкновенные прибыли, Common Stocks and Uncommon Profits. Hoboken, NJ: John Wiley & Sons, 1958. Reprinted and updated 2003.
48. Garen, Brenton. На NYSE дебютирует ETF-фонд на базе искусственного интеллекта, Artificial intelligence powered ETF debuts on NYSE. ETF Trends, October 18, 2017.
49. Gelsi, Steve, Ветеринарное IPO залаяло на рынке, Veterinary IPO barking in market. MarketWatch online, November 8, 2001. **https://www.marketwatch.com/story/veterinary-firm-to-trade-under-woof-ticker**.
50. Gilbert, Daniel, T., King, Gary, Pettigrew, Stephen, Wilson, Timothy. D. Комментарий к статье «Оценивание воспроизводимости в психологической науке», Comment on 'Estimating the reproducibility of psychological science'. Science, 351 (2016): 1037.
51. Glaeser, Edward, Huang, Wei, Ma, Yueran, and Shleifer, Andrei. Бум на рынке недвижимости с китайской спецификой, A real estate boom with Chinese characteristics. Journal of Economic Perspectives 31, no. 1 (2017): 93–116.
52. GlobalPropertyGuide, Тренды цен на дома, Home price trends. **https://www.globalpropertyguide.com/home-price-trends/China. 2019**.
53. Gray, Jim. eScience – преобразованный научный метод, eScience – A Transformed Scientific Method. n.d. **http://research.microsoft.com/en-us/um/people/gray/talks/NRC-CSTB_eScience.ppt**.
54. Gray, Paul, E. Человек, который любил детей: доктор Бенджамин Спок (1903–1998), The man who loved children: Dr. Benjamin Spock (1903–1998). Time 151, no. 12. March 30, 1998. **http://content.time.com/time/magazine/article/0,9171,988061,00.html**.
55. Harish, Ajay. Почему затонул шведский корабль «Васа», Why the Swedish Vasa ship sank, SimScale (blog), March 21, 2019. **https://www.simscale.com/blog/2017/12/vasa-ship-sank/**.
56. Heinonen, O. P., Shapiro, S., Tuominen, L., Turunen, M. I. Применение резерпина в связи с раком молочной железы, Reserpine use in relation to breast cancer. Lancet 2 (1974): 675–7.
57. Hennessey, Raymond. Дебют VCA Antech оказался неудачным: компания вышла на биржу во второй раз, VCA Antech has a dismal debut as it makes second public run. Dow Jones Newswires online. November 23, 2001. **https://www.wsj.com/articles/SB10063724332089572 00**.
58. Hernandez, Ben. Том Лайдон заявил в Yahoo! Finance LIVE, что без ИИ вы окажетесь «в невыгодном положении», You'll 'be at a disadvantage' without AI, says Tom Lydon on Yahoo! Finance LIVE. ETF Trends online. March 25, 2019.

59. Herndon, Thomas, Ash, Michael, Pollin, Robert. Является ли высокий государственный долг постоянным сдерживающим фактором экономического роста? Does high public debt consistently stifle economic growth? A critique of Reinhart and Rogoff. Cambridge Journal of Economics 38, no. 2 (2014): 257–79.

60. Kahneman, Daniel, Tversky, Amos. Теория перспектив: анализ решений в условиях риска, Prospect theory: An analysis of decision under risk. Econometrica 47, no. 2 (1979): 263–92.

61. Inskeep, Steve. Ребенок, родившийся в Миссури 11 июля, известный как день 7–11, Missouri baby born on July 11, which is known as 7–11 Day. Morning Edition: Strange News, NPR. July 15, 2019.

62. Ioannidis, John, A. Противоречивые и изначально более сильные эффекты в часто цитируемых клинических исследованиях, Contradicted and initially stronger effects in highly cited clinical research. Journal of the American Medical Association 294, no. 2 (2005): 218–28.

63. Jaki, Stanley, L. Ранняя история закона Тициуса-Боде, The Early History of the Titius-Bode Law, American Journal of Physics 40, no. 7 (1972): 1014. **https://doi-org.ccl.idm.oclc.org/10.1119/1.1986734**.

64. John, Leslie, K., Loewenstein, George, Prelec, Drazen. Измерение распространенности сомнительных исследовательских практик с помощью стимулов для высказывания правды, Measuring the prevalence of questionable research practices with incentives for truth-telling. Psychological Science 23, no. 5 (2012): 524–32.

65. Jones Larry, E., Schoonbroodt Alice, Tertilt, Michèle. Теории фертильности: могут ли они объяснить отрицательную зависимость между фертильностью и доходом? Fertility theories: Can they explain the negative fertility-income relationship? NBER Working Paper 14266. Cambridge, MA: National Bureau of Economic Research, 2008.

66. Jordan, Brigitte Jordan. Распознавание закономерностей в эволюции человека и его значение для этнографии, антропологии и общества, Pattern Recognition in Human Evolution and Why It Matters for Ethnography, Anthropology, and Society, in Advancing Ethnography in Corporate Environments: Challenges and Emerging Opportunities, edited by Brigitte Jordan, 193–214. Abingdon: Routledge, 2016.

67. Karelitz, Samuel, Fisichelli, Vincent R., Costa, Joan, Kavelitz, Ruth, Rosenfeld, Laura. Связь плача в раннем младенчестве с речевым и интеллектуальным развитием в возрасте трех лет, Relation of crying in early infancy to speech and intellectual development at age three years. Child Development 35 (1964): 769–77.

68. Kecman Vojislav. Предисловие к книге «Глубокая переработка данных: подход к обнаружению знаний», Foreword, in Data Mining: A Knowledge Discovery Approach, edited by Krzysztof J. Cios, Witold Pedrycz, Roman W. Swiniarski, Lukasz Andrzej Kurgan, xi. New York, Springer, 2007.

69. Кенийцы выслеживают и отлавливают убивающих коз гепардов, Kenyans chase down and catch goat-killing cheetahs. BBC online. November 15, 2013. **https://www.bbc.com/news/world-africa-24953910**.

70. Knapp, Gunnar. Почему вас никогда не выберут конгрессменом от Бронкса, если ваша фамилия не начинается на букву «Б», Why You'll Never Get Elected Congressman from the Bronx Unless Your Last Name Starts with B. unpublished thesis, Yale University, December 7, 1973.

71. LeClaire, Jennifer. Закону Мура исполняется 50 лет. Создатель говорит, что он не будет действовать вечно, Moore's Law Turns 50, Creator Says It Won't Go on Forever. Newsfactor, May 12, 2015.

72. Lee, Jane, J. Откуда петух знает, что нужно кукарекнуть на рассвете, How a rooster knows to crow at dawn. National Geographic News, March 19, 2013. **https://www.nationalgeographic.com/news/2013/3/130318-rooster-crow-circadian-clock-science/**.

73. Levelt Committee, Noort Committee, Drenth Committee. Ущербная наука: мошеннические методы исследования социального психолога Дидерика Стапеля, Flawed science: The fraudulent research practices of social psychologist Diederik Stapel. November 28, 2012. **https://poolux.psychopool.tu-dresden.de/mdcfiles/gwp/Reale%20F%C3%A4lle/Stapel%20-%20Final%20Report.pdf**.

74. Levy, Steven. Эксклюзив: как алгоритм Google управляет интернетом, Exclusive: How Google's algorithm rules the web. Wired, February 22, 2010.

75. Lévi-Strauss, Claude. Грустные тропики, Tristes Tropiques. Translated by John Weightman and Doreen Weightman. New York: Penguin Books, 1992.

76. Lewis, Tanya. Великолепный упадок: вторая гибель шведского военного корабля «Васа», Gorgeous Decay: The Second Death of the Swedish Warship Vasa. Wired, September 3, 2012. **https://www.wired.com/2012/09/swedish-warship-vasa/**.

77. Liebenberg, Louis. Значение охоты на изнурение для эволюции человека, The relevance of persistence hunting to human evolution, Journal of Human Evolution 55, no. 6 (2008): 1156–9.

78. Liu, Yukun, Tsyvinski, Aleh. Риски и возвратность криптовалюты, Risks and returns of cryptocurrency. NBER Working Paper No. w24877. August 6, 2018. **https://ssrn.com/abstract=3226952**.

79. Lo, Bruce, M., Visintainer, Catherine M., Best, Heidi A., Beydoun, Hind A. Разоблачение мифа: использование экстренных служб в пятницу 13-го, Answering the myth: Use of emergency services on Friday the 13th. The American Journal of Emergency Medicine 30, no. 6 (2012): 886–9.

80. Macaskill, Sandy. Топ-10: футбольные суеверия, способные посоперничать с Коло Туре из «Арсенала», Top 10: Football superstitions to rival Arsenal's Kolo Toure. The Telegraph February 25, 2009.

81. Man, Joyce Yanyun, Zheng, Siqi, Ren, Rongrong. Жилищная политика и рынки жилья: тренды, закономерности и доступность, Housing Policy and Housing Markets: Trends, Patterns, and Affordability, in China's Housing Reform and Outcomes, edited by Joyce Yanyun Man, 3–18. Cambridge, MA: Lincoln Institute of Land Policy, 2011.

82. Mann, Brad. Сколько раз следует тасовать колоду карт? How many times should you shuffle a deck of cards? UMAP 15, no. 4 (1994): 303–32.

83. Markowitz, Harry. Выбор портфеля: эффективная диверсификация инвестиций, Portfolio Selection: Efficient Diversification of Investments. New York: John Wiley & Sons, 1959.

84. Marr, Bernard. Большие данные в Tesco: аналитика в реальном времени в британском ретейловом гиганте продуктовов питания, Big Data at Tesco: Real Time Analytics at The UK Grocery Retail Giant. Forbes, November 17, 2016. **https://www.forbes.com/sites/bernardmarr/2016/11/17/big-data-at-tesco-real-time-analytics-at-the-uk-grocery-retail-giant/#7196f50061cf**.

85. Mathews, Fiona, Johnson, Paul J., Neil, Andrew. Вы – то, что ест ваша мать: доказательства влияния предгравидарного рациона питания матери на пол будущего ребенка у людей, You are what your mother eats: Evidence for maternal preconception diet influencing foetal sex in humans. The Royal Society: Proceedings: Biological Sciences 275, no. 1643 (2008): 1661–8.

86. Mattson, Mark, P. Превосходная обработка закономерностей – это суть развитого человеческого мозга, Superior pattern processing is the essence of the evolved human brain. Frontiers in Neuroscience 8 (2014): 265. doi:10.3389/fnins.2014.00265.

87. Mayer, Jürgen, Khairy, Khaled, Howard, Jonathon. Начертание слона с использованием четырех сложных параметров, Drawing an elephant with four complex parameters. American Journal of Physics 78 (2010): 648. DOI:10.1119/1.3254017.

88. McMahon, Dinny. Великая стена долга Китая: теневые банки, города-призраки, огромные кредиты и конец китайского чуда, China's Great Wall of Debt: Shadow Banks, Ghost Cities, Massive Loans, and the End of the Chinese Miracle. Boston, MA: Houghton Mifflin Harcourt, 2018.

89. McRae, Mike. Отчет показывает, что «кризис репликации» науки затронул даже самые уважаемые журналы, Science & apos;s 'replication crisis' has reached even the most respectable journals, report shows. ScienceAlert. August 27, 2018. **https://www.sciencealert.com/replication-results-reproducibility-crisis-science-nature-journals**.

90. Meier, Paul. Крупнейший эксперимент в области общественного здравоохранения: полевые испытания вакцины Солка против полиомиелита в 1954 году, The Biggest Public Health Experiment Ever: The 1954 Field Trials of the Salk Poliomyelitis Vaccine, in Statistics: A Guide to the Unknown, edited by Judith Tanur, Frederick Mosteller, William H. Kruskal, Richard F. Link, Richard S. Pieters, Gerald R. Rising, 2–13. San Francisco: Holden-Day, 1972.

91. Milbank, Dana. Президент Трамп вступает в свои ужасные два года, President Trump is entering his terrible twos. The Washington Post, January 7, 2019.

92. Mitchell, Robert, L. 12 глюков в предсказательной аналитике, 12 predictive analytics screw-ups. Computerworld online, July 24, 2013. **https://www.computerworld.com/article/2484224/12-predictive-analytics-screw-ups.html**.

93. Moore, Geoffrey. Втискивание большего количества компонентов в чипы интегральных схем, Cramming more components onto integrated circuits. Electronics 38, no. 8, April 19, 1965: 114.

94. Mullen, Jethro, Stevens, Andrew. Миллиардер: китайская недвижимость – «крупнейший пузырь в истории», Billionaire: Chinese real estate is "biggest bubble in history". CNN Business, CNN online. September 29, 2016. **https://money.cnn.com/2016/09/28/investing/china-wang-jianlin-real-estate-bubble/**.

95. Myers, D. Вертикальная мобильность в пространстве и времени: уроки иммиграции, Upward mobility in space and time: lessons from immigration, in America's Demographic Tapestry, edited by James W. Hughes, Joseph J. Seneca, 135–57. New Brunswick, NJ: Rutgers University Press, 1999,

96. Neill, Tess. Набег Tesco и его неудача в США, Tesco's foray, and failure, in the U.S., AdAge October 4, 2013s.

97. Nelson, Leland, R., Furst, Merrick L. Объективное исследование влияния ожиданий на конкурентоспособность, An objective study of the effects of expectation on competitive performance. The Journal of Psychology 81 (1972): 69–72.

98. Nicholls, Neville. Доклад Чарни: 40 лет назад ученые точно предсказали изменение климата, The Charney Report: 40 years ago, scientists accurately predicted climate change. Phys.org online. July 23, 2019. **https://phys.org/news/2019-07-charney-years-scientists-accurately-climate.html**.

99. Nisbet, Robert, Elder, John, Miner, Gary. Справочник по приложениям статистического анализа и глубокой переработки данных, Handbook of Statistical Analysis & Data Mining Applications. Burlington, MA: Academic Press, 2009.

100. Nunlist, Tom. Рынок жилья Китая: железный пузырь, China's housing market: The iron bubble. Cheung Kong Graduate School of Business Knowledge online. July 31, 2017.
101. Open Science Collaboration. Оценивание воспроизводимости психологической науки, Estimating the reproducibility of psychological science. Science 349, no. 6251 (2015). DOI: 10.1126/science.aac4716.
102. Palca, Joe. Ставка на искусственный интеллект для управления реагированием на землетрясения. Все учтено, Betting on artificial intelligence to guide earthquake response. All Things Considered, NPR, April 20, 2018.
103. Pashler, Harold, Wagenmakers, Eric. Введение редактора к специальному разделу «Воспроизводимость в психологической науке: кризис доверия?» Editors' Introduction to the special section on 'Replicability in psychological science: A crisis of confidence?'. Perspectives on Psychological Science 7, no. 6 (2012): 528–30.
104. Peirce, Charles Sanders, Ketner, Kenneth Laine. Рассуждение и логика вещей: лекции на Кембриджской конференции 1898 года, Reasoning and the logic of things: The Cambridge conferences lectures of 1898. Cambridge, MA: Harvard University Press, 1992, 194–6.
105. Piatetsky-Shapiro, G. Обнаружение знаний в реальных базах данных: отчет о семинаре IJCAI-89, Knowledge discovery in real databases: A report on the IJCAI-89 workshop. AI Magazine 11, no. 5 (1991): 68–70.
106. Powell, Adam, Люди – разгоряченные, потные, прирожденные бегуны, Humans hot, sweaty, natural-born runners. The Harvard Gazette April 19, 2007.
107. Preis, Tobias, Moat, Helen Susannah, Stanley, H. Eugene. Количественное оценивание торгового поведения на финансовых рынках с использованием Google Trends, Quantifying trading behavior in financial markets using Google trends. Scientific Reports 3 (2013): 1684.
108. Rashes, Michael. Инвесторы, находящиеся в глубоком замешательстве, принимают явно невежественные решения, Massively confused investors making conspicuously ignorant choices. The Journal of Finance 56 (2001): 1911–27.
109. Redman, Thomas, C. Плохие данные обходятся США в 3 триллиона долларов в год, Bad data costs the U.S. $3 trillion per year. Harvard Business Review, September 22, 2016.
110. Roman, John, Chalfin, Aaron. Существует ли волна интернет-преступлений? Is there an iCrime Wave? Washington, D.C.: Urban Institute, 2007.
111. Sadrô, J., Janudi, Izzat, Sinha, Pawan. Роль бровей в распознавании лиц, The role of eyebrows in face recognition. Perception 32, no. 3 (2003): 285–93.
112. Sagiroglu, S, Sinanc, D. Большие данные: обзор, Big data: A review. Paper presented at The 2013 International Conference on Collaboration Technologies and Systems (CTS 2013), San Diego, CA, May 2013.
113. Scanlon, Thomas, J., Luben, Robert, N., Scanlon, F. L., Singleton, Nicola. Вредна ли пятница 13-е для здоровья?, Is Friday the 13th bad for your health? British Medical Journal 307, no. 6919 (1993): 1584–6.
114. Schrage, Michael. 2014. Падение Tesco – предупреждение для ретейлеров, ориентированных на данные, 2014. Tesco's Downfall Is a Warning to Data-Driven Retailers, Harvard Business Review, October.
115. Schuld, Jochen, Slotta, Jan E., Schuld, Simone, Kollmar, Otto, Schilling, Martin K., Richter, Sven. Распространенное мнение встречается с хирургической реальностью: влияние фаз Луны, пятницы 13-го и знаков зодиака на экстренные операции и интраоперационную кровопотерю, Popular belief meets surgical reality: Impact of lunar

phases, Friday the 13th and zodiac signs on emergency operations and intraoperative blood loss. World Journal of Surgery 35, no. 9 (2011): 1945–9.

116. Schwab, Katherine. Помощь при стихийных бедствиях подорвана. Сможет ли ИИ это исправить? Disaster relief is dangerously broken. Can AI fix it?. Fast Company online, November 15, 2018.
117. Schwab, Katherine. Этот знаменитый стартап поклялся спасать жизни с помощью ИИ. Теперь он стал поучительной историей, This celebrated startup vowed to save lives with AI. Now, it's a cautionary tale. Fast Company online, August 13, 2019.
118. Shanks, David, R., Newell, Ben R., Lee, Eun Hee, Balikrishnan, Divya, Ekelund, Lisa, Cenac, Zarus, Kavvadia, Fragkiski, Moore, Christopher. Подготовка разумного поведения: неуловимый феномен, Priming intelligent behavior: An elusive phenomenon. PLOS One 2013. **https://doi.org/10.1371/journal.pone.0056515**.
119. Shapiro, Samuel. Взгляд в XXI век: извлекли ли мы уроки из своих ошибок или обречены усугублять их?, Looking to the 21st century: Have we learned from our mistakes, or are we doomed to compound them?. Pharmacoepidemiology and Drug Safety 13, no. 4 (2004): 257–65.
120. Shapiro, S., Slone, D. Исследование методом случай-контроль: консенсус и противоречия, Case-control study: Consensus and controversy. Comment, Journal of Chronic Disease 32 (1979): 105–7.
121. Sharpe, William F. Возвратность паевых инвестиционных фондов, Mutual fund performance. Journal of Business January (1966): 119–38.
122. Simmons, Joseph P., Nelson, Leif D., Simonsohn, Uri. Ложноположительная психология: нераскрытая гибкость в сборе и анализе данных позволяет представлять что угодно как значимое, False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. Psychological Science 22 (2011): 1359–66.
123. Simonite, Tom, Закон Мура мертв. Что теперь? Moore's law is dead. Now what? MIT Technology Review, May 13, 2016. **https://www.technologyreview.com/s/601441/moores-law-is-dead-now-what/**.
124. Simonsohn, Uri. Просто опубликуйте: урок, извлеченный из двух случаев фальсификации данных, обнаруженных только с помощью статистики, Just post it: The lesson from two cases of fabricated data detected by statistics alone. SSRN 2012. **https://doi.org/10.2139/ssrn.2114571**.
125. Smith, Gary. Отличная компания, отличная инвестиция, пересмотренная версия, Great company, great investment revisited. Journal of Wealth Management 19, no. 1 (2016): 34–9.
126. Smith, Gary. Иллюзия искусственного интеллекта, The AI Delusion. Oxford: Oxford University Press, 2018.
127. Smith, Gary, Levere, Michael, Kurtzman, Robert. Поведение игроков в покер после крупных выигрышей и крупных проигрышей, Poker player behavior after big wins and big losses. Management Science 55, no. 9 (2009): 1547–55.
128. Smith, Gary. Парадокс больших данных, The Paradox of Big Data. Unpublished, 2019.
129. Smith, Gary, Cordes, Jay. 9 подводных камней науки о данных, The 9 Pitfalls of Data Science. Oxford: Oxford University Press, 2019.
130. Smith, Margaret Hwang, and Smith, Gary. Пузырь, пузырь, где же пузырь на рынке жилья? Bubble, bubble, where's the housing bubble? Brookings Papers on Economic Activity 2006, no. 1 (2006): 1–50.

131. Smith, Margaret Hwang, Smith, Gary. Какая мать, такая и дочь? Экономическое сравнение матерей-иммигранток и их дочерей, Like mother, like daughter? An economic comparison of immigrant mothers and their daughters. International Migration 51 (2013): 181–90.

132. Stein, Gregory, M. Что будет делать Китай, когда истечет срок действия прав на землепользование? What will china do when land use rights begin to expire? Vanderbilt Journal of Ttransnational Law 50 (2017): 625–72.

133. Su, Francis, E., et al. Семь перетасовок. Интересные математические факты, Seven shuffles. Math Fun Facts. **https://math.hmc.edu/funfacts/seven-shuffles/. n.d.**

134. Thelwell, Richard. Собака, которая не лаяла: почему аналитика клиентов в Tesco дала такие серьезные ошибки? The dog that didn't bark: How did customer analytics go so badly wrong at Tesco? **https://www.matillion.com/insights/how-did-customer-analytics-go-so-badly-wrong-at-tesco/**.

135. Tullock, Gordon. Комментарий к статье Дэниела Кляйна «Обращение к экономистам, выступающим за свободу», A comment on Daniel Klein's 'A plea to economists who favor liberty'. Eastern Economic Journal 27, no. 2 (2001): 203–7.

136. Turner, Annie. Поучительная история Tesco: большие данные и «аналитический альбатрос», The cautionary tale of Tesco: Big data and the 'analytic albatross'. Inform online. January 2015. **https://inform.tmforum.org/features-and-analysis/2015/01/cautionary-tale-tesco-big-data-analytic-albatross/**.

137. Tversky, Amos, Kahneman, Daniel. О психологии предсказания, On the psychology of prediction. Psychological Review 80 (1973): 237–51.

138. Tversky, Amos, Kahneman, Daniel. Суждение в условиях неопределенности: эвристика и систематические смещения, Judgement under uncertainty: Heuristics and biases. Science 185 (1974): 1124–31.

139. Ullman, S. B. Математическая формулировка закона Боде, A mathematical formulation of Bode's law. Popular Astronomy, 57 (1949): 197.

140. Urban Institute. Плеер iPod: громоотвод для преступников?, The iPod: Lightning rod for criminals? Washington, D.C.: Urban Institute, September 27, 2007.

141. Vogelstein, Fred. Как Yahoo! все испортил, How Yahoo! blew it, Wired, February 1, 2007.

142. Wei, Shang-Jin, Zhang, Xiaobo, Liu, Yin. Владение жильем как конкуренция за статус: некоторые теории и доказательства, Home ownership as status competition: Some theory and evidence. Journal of Development Economics 127 (2017): 169–86.

143. Weinberg, Robert, S., Gould, Daniel, Yukelson, David, and Jackson, Allen. Влияние изначальной и искусственно созданной самоэффективности на соревновательную задачу на мышечную выносливость, The effect of preexisting and manipulated self-efficacy on a competitive muscular endurance task. Journal of Sport and Exercise Psychology 3, no. 4 (1981): 345–54.

144. Wicherts, Jelte, M., Bakker, Marjan, Molenaar, Dylan. Готовность делиться исследовательскими данными зависит от убедительности свидетельств и качества представления статистических результатов, Willingness to share research data is related to the strength of the evidence and the quality of reporting of statistical results. PLoS ONE 6, no. 11 (2011). **https://doi.org/10.1371/journal.pone.0026828**.

145. Worland, Justin, Как предотвратить гибель людей от землетрясений еще до того, как они произойдут, How we can stop earthquakes from killing people before they even hit. Time September 28, 2017.

146. Wu, Jing, Deng, Yongheng, Liu, Hongyu. Индекс цен на жилье на зарождающемся рынке жилья: случай Китая, House price index construction in the nascent housing

market: The case of China. The Journal of Real Estate Finance and Economics 48, no. 3 (2014): 522–45.

147. Young, Shalise Manza. НФЛ и Ассоциация игроков НФЛ заявляют, что нет никаких доказательств того, что Эрик Рид из «Пантерс» подвергался тестированию на наркотики, NFL, NFLPA say there's 'no evidence' Panthers' Eric Reid was targeted for drug testing. Yahoo Sports, January 9, 2019. **https://sports.yahoo.com/nfl-nflpa-say-theres-no-evidence-panthers-eric-reid-targeted-drug-testing-150426603.html**.

148. Young, Stanley, S., Bang, Heejung, Oktay, Kutluk. Гендерный отбор, обуславливаемый зерновыми? Вероятнее всего, множественный ложноположительный результат тестирования, Cereal-induced gender selection? Most likely a multiple testing false positive. The Royal Society: Proceedings: Biological Sciences 276, no. 1660 (2009): 1211–12.

149. Zeisel, Hans. Доктор Спок и дело об исчезновении женщин-присяжных, Dr. Spock and the case of the vanishing women jurors. University of Chicago Law Review 37, no. 1 (1969): 1–18.

Предметный указатель

Нумерованный

666, число 59

А

Абракадабра 51
Анализ среднедисперсный 157
Аутизм и вакцина MMR 138

Б

Байер, Дэйв 87
Байес, Томас 195
Балл индекса академической успеваемости (API) 92
Бартц, Кэрол 174, 177
Баскетбол
  промежутки, когда игрок в ударе 80
  соревнование по трехочковым броскам 81
Баффет, Уоррен 161, 173, 188, 192
Бем, Дэрил 132
Бентли, Дональд 61
Беременность
  и хлопья на завтрак 137
  предсказатель Target 170
Биткойн 123, 139
Блейн, Гилберт 33
Блуждание случайное 77
Блэк-джек, дробление карт 55
Боде, Иоганн 68
Борхас, Джордж Хесус 107
Браге, Тихо 67
Бридж, карточная игра 86
Брин, Сергей 174, 178
Булочка монахини 23
Бэктестирование. См. Тестирование на истории

В

Вакцина MMR и аутизм 138
Вани, Ахмад 163
Вариативность малой выборки 92
Васа, судно 147
Вероятность апостериорная 196, 200, 202
Вероятность априорная 196, 200, 202
Витамин С 32
Война во Вьетнаме
  лотерея призыва в армию 185
  суд над Споком 186
Выдвижение гипотез после того, как результаты известны (HARKing) 39
Вытеснение истинных переменных 103

Г

Галилей 67
Гамильтон, Уильям 154
Гельман, Эндрю 135
Гипертермия 14
Гипотеза об эффективности рынка 139
Гипотермия 14
Гиппократ 52
Гравитация 20
График открытие-максимум-минимум-закрытие 78
График цен на золото 78
Грейнджер, Клайв 38
Густав II, король Швеции 147

Д

Данные вневыборочные 102
Данные внутриборочные 102
Данн, Эдвина 166
Дева Мария на сырном сэндвиче 23
Джобс, Стив 10
Джуманджи 124
Диаконис, Перси 87
Доу, Чарльз 154
Доход и военный опыт 185
Дух животный 140

З

Закон Боде 68
Закон малых чисел 188
Закон Мура 71

Закономерность 202
Закономерность на букву «М» 54
Заполняемость гостиниц 30
Зук, Джон 114

И

Изменение климата 143
Имена женские и ураганы 139
Интеллект искусственный (ИИ) 99, 160, 165, 170
Использование Google Trends для отслеживания цен на биткойн 123
Испытание контроллируемое рандомизированное (РКИ) 35
Исследования зомбические 138

К

Канеман, Дэниел 109, 136, 188, 192
Кастро, Энрике де 175
Кейнс, Джон Мейнард 140, 153, 188, 197
Кеплер, Иоганн 67
Клинтон, Билл 151
Кнапп, Гуннар 55
Кока-кола 148
Кола новая 148
Коллинз, Джим 39
Конфиденциальность 170
Коперник, Николай 66
Космические совпадения 65
Коуз, Рональд 101
Кризис воспроизводимости 132
Кругман, Пол 107
Кубик Рубика 177
Кугл, Тим 173
Кукареканье петухов 38
Кук, Тим 170
Купюра стодолларовая на тротуаре 153
Кэндлер, Эйса Григгс 149

Л

Ландау, Сара 43
Ланкастер, Джеймс 32
Левинсон, Росс 174, 177
Леви-Стросс, Клод 27
Либерман, Дэниел 13
Линд, Джеймс 33
Лира турецкая 112
Лихи, Терри 167
Логическая ошибка «после этого, следовательно, из-за этого» 37
Лотерея призыва в армию 185
Лучшие места для работы, список 43
Лю, Юкун 123, 139
Лян, Уэсли 128

М

Майер, Марисса 174, 177
Малярия 51
Марковиц, Гарри 156
Марш десятицентовиков, фонд борьбы с детским параличом 182
Мать-иммигрантка 106
Мать Тереза 23
Метод научный 53, 74, 100
Миазмы 52
Мировая серия по бейсболу 94
Мобильность иммигрантов межпоколенческая 107
Мобильность экономическая мигрантов 106
Модель ценообразования капитальных активов 157
Мур, Гордон 71
Мэнсон, Патрик 52
Мюррей, Пол 185

Н

Наука неряшливая 133
Нейман, Джон фон 70
Новорожденный в день 7-Eleven 24
Новые Гебриды 51
Номера лотерейные и мечты 60
Носек, Брайан 137
Нумерология 63
Ньютон, Исаак 20

О

Обама, Барак 180
Ожидание самореализующееся 140
Окленд Эйс, команда 95
Отбор естественный 15
Отрицательство 145
От хорошего к отличному 39
Отчет Чарни 143
Охота на изнурение 13

Ошибка логическая адвоката защиты 200
Ошибка логическая неучета базовой частоты 198
Ошибка логическая прокурора 200

П

Парадокс больших данных 99, 102, 105, 116
Пейдж, Ларри 174, 178
Пеле 57
Пембертон, Джон 149
Пепси-кола 150
Переработка данных глубокая 100, 202
Переработка твитов Трампа глубокая 114
Пирни, Александр 185
Питер, Лоуренс Дж. 178
Пифагор 62
Планета карликовая 71
Плач и коэффициент интеллекта 19
Повреждение от землетрясения 163
Подвергание данных пыткам 101
Поле электромагнитное 91
Полиомиелит 182
Помощь финансовая колледжа и зачисление 31
Поселенцы Катана 88
Правило Байеса в зале суда 199
Предсказане пола ребенка по диете роженицы 137
Предсказание обменного курса валют 112
Предсказание покупки страхового полиса 106
Предсказание степени расшатанности игрока в покер 108
Предсказатель беременности Target 170
Предсказывание безработицы 126
Предсказывание цен на акции с помощью астрологических знаков 121
Предсказывание цен на акции с помощью службы Google Trends 118
Предсказывание цен на биткойн с помощью службы Google Trends 123
Принцип Питера 178
Проект по воспроизводимости 139
Проект по репликации экспериментальной экономики (Experimental Economics Replication Project) 137
Промежуток, когда игрок в ударе 80
Пророчество самоисполняющееся 57
Процесс века над О. Джеем Симпсоном 200
Пузырь китайский жилищный 128
Пятница, тринадцатое число 58, 60

Р

Разлив Нила 29
Рак груди 197
Распознавание закономерностей 16
Распознавание лиц 23
Рассуждение «после этого, следовательно, из-за этого» 37
Регрессия множественная 104
Резерпин 136
Рид, Эрик 85
Росс, Рональд 52
Роузбек, Оуэн 43
Рузвельт, Франклин Д. 181

С

Сад с расходящимися дорожками 135
Связь причинно-следственная 27
Связь причинно-следственная Грейнджера 38
Семел, Терри 173, 177
Серия 75
Символ тикерный 45
Симуляция методом Монте-Карло 104
Сириус 29
Сирокер, Дэн 180
Система геоцентрическая 66
Система гелиоцентрическая 66
Скептицизм 145
Скопление онкологических заболеваний 89
Слон четырехпараметрический 70
Смещение систематическое при верификации 134, 135
Сначала теория, потом данные 188
Солк, Джонас 182
Соломон, Майкл 48
Софтостроитель 101
Список Fortune наиболее уважаемых компаний 41
Сплетни 41
Спок, Бенджамин 186
Стапель, Дидерик 132, 133, 135, 136, 139, 145
Суеверие 53

Суперкубок 22
Сэндвич сырный с обликом Девы Марии 23

Т

Тасование карт 87
Твиттер 114
Твиты Трампа 114
Теорема Байеса 196
Теория современная портфельная 157
Тестирование на истории 158
Тест на наркотики в НФЛ 85
Тикер продуманный 45
Титиус, Иоганн 68
Тобин, Джеймс 156
Томпсон, Скотт 174, 177

У

Ульман, С. Б. 70
Упражнения физические и интеллект 14
Ураганы и женские имена 139
Уровень безработицы 126
Уровень преступности 93

Ф

Фаллоу, Сара Стэнли 40
Фило, Дэвид 172
Фишер, Филип 41
Фонд биржевой (ETF) 160
Фрэнк, Тим 163

Х

Хамби, Клайв 166
Харрингтон, Майкл 106
Харрис, Джо 81
Хатуа, Чида 161
Хибертссон, Хенрик 147
Хлопья на завтрак и новорожденные-мальчики 137
Ходжес, Крейг 83
Холдем техасский (Texas Hold 'Em), игра в покер 109
Ху, Николь 163

Ц

Цена на авокадо и поисковые запросы в отношении знака Дева 30
Цена на акции
  астрологические знаки 121
  животный дух 140
  наилучший месяц для покупки 153
  поисковые запросы по термину «долг» 119
  система месяца 159
  теория Доу 154
  чрезмерная реакция 188
  Nifty 50 192
Цинга 32
Цукерберг, Марк 170
Цывински, Алех 123, 139

Ч

Чарни, Джул 143
Четыре и смерть, в японском языке 58
Число 47 колледжа Помона 61
Чудо-Метс 96

Ш

Шорт, Первис 80
Шутка самая смешная 18

Э

Эйнштейн, Альберт 145
Эксперимент контролируемый 181
Эпоха величия Швеции 147
Эффект затухания 136

Я

Янг, Джерри 172, 174, 177
Янг, Стэнли 138

A

A/B-тест 36, 179
Admiral Insurance, страховая компания 161
Affluent Market Institute 40
AIEQ, фонд 160
Aldi, немецкая сеть супермаркетов 167
Alibaba 174, 176
Android 172
Apple 189

B

Broadcast.com 172

C

Clubcard 166
Club Med, курорты 37

D

DoubleClick 173
dunnhumby, компания 166

E

EquBot, компания 160

F

Facebook 170, 173
Flood Concern 163
Fresh & Easy, сеть магазинов 167
Full Tilt Poker (Покер во весь опор),
онлайновая покерная комната 109

G

GeoCities 172
Glassdoor, компания 43
Google 172, 173, 178
Google Trends
уровень безработицы 126
цены на акции 118
цены на биткойн 123

I

iPod, плеер 10
iTunes, плеер 77

M

Microsoft 174, 179

N

Netscape, браузер 172
Nifty 50 192
Nordstrom, компания 170

O

One Concern, стартап 163
Oracle 189

P

PayPal 174
Post-it, самоклеящиеся листочки 177

T

Tesco, компания 166
Toyota Corolla, автомобиль 177
Tversky, Amos 109, 188, 192

U

Urban Institute, аналитический центр 11
Urban Outfitters, компания 170

V

VCA Antech 48

W

Wal-Mart 173

Y

Yahoo! 172
YouTube 173

Книги издательства «ДМК Пресс» можно купить оптом и в розницу
на складе издательства по адресу:
г. Москва, ул. Электродная, д. 2, стр. 12, офис 7,
тел. +7 (499) 322-19-38,
а также заказать на сайте **www.dmkpress.com**
с доставкой в любой регион РФ

Гэри Смит и Джей Кордес

Ловушки призрачных закономерностей в анализе данных

Главный редактор *Мовчан Д. А.*
Зам. главного редактора *Яценков В. С.*
editor@dmkpress.com

Перевод с английского *Логунов А. В.*
Корректор *Абросимова Л. А.*
Верстка *Паранская Н. В.*
Дизайн обложки *Трофимова С. В.*

Формат 70×100 $^{1}/_{16}$.
Печать цифровая. Усл. печ. л. 17.88.
Тираж 100 экз.

Веб-сайт издательства: **www.dmkpress.com**

В современном мире на нас обрушивается лавина данных и стремительно растет число закономерностей, которые мы в них обнаруживаем. От природы человеку свойственно верить, что каждая закономерность наполнена смыслом. Но это не так! Шансы, что найденная закономерность окажется полезной, быстро приближаются к нулю.

Эта увлекательная книга благодаря бесчисленным примерам поможет читателю не поддаваться дурману данных, не прибегать к бесполезным инвестиционным стратегиям и не бояться делать прививки.

Основные темы книги:

- парадокс больших данных;
- кризис воспроизводимости;
- теорема Байеса;
- контролируемые и естественные эксперименты;
- примеры призрачных закономерностей.

Издание адресовано всем, кто использует данные в работе, – ученым, аналитикам, инженерам, консультантам, студентам и предпринимателям, а также любознательным читателям, которые хотят научиться правильно оценивать связь между событиями.

Гэри Смит – профессор экономики имени Флетчера Джонса в колледже Помона. Он получил степень доктора философии по экономике в Йельском университете и в течение семи лет проработал там ассистентом профессора. Получил две премии в области преподавания и написал более восьмидесяти научных статей и тринадцать книг.

Джей Кордес – исследователь данных, которому нравится решать сложные задачи и помогать будущим исследователям данных избегать распространенных ошибок, с которыми он сталкивался в корпоративном мире. Получил степень по математике в колледже Помона, окончил магистерскую программу Калифорнийского университета в Беркли по информатике и обработке данных (MIDS).

OXFORD

ДМК ПРЕСС
издательство

www.дмк.рф

ISBN 978-5-93700-433-8