

Генезис статистики



Источники и части

Уже в Библии упоминается о переписях населения. Чем не повод для создания статистики? Еще раньше, видимо, было важно считать и систематизировать звезды. Тоже подходящий повод. Люди, конечно, всегда стремились найти объяснение того, что они наблюдали, однако разумное объяснение найти удавалось не всегда. Тогда построили модель мира, в которой все события, все факты, все наблюдения были жестко детерминированы. В этой модели нет места случайности. А где нет случайности, там нет и статистики.

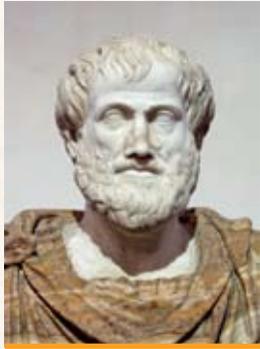
Но очень давно люди начали замечать, что есть случаи, когда удобнее предположить, что бывают события, причины которых теряются где-то во мраке неопределенности. Еще Аристотель говорил: «Уничтожение случая влечет за собой нелепые последствия. Есть многое, что совершается не по необходимости, а случайно». Примерно через тысячу лет на Среднем Востоке Аль-Фараби писал: «...случайны явления, причины которых неизвестны». Умение понимать и учитывать случайные явления рассматривалось как сакральное знание.

И только новое время с его рационализмом оформило статистику как науку об изучении сложных социально-экономических явлений и математическую статистику как ветвь теории вероятности. Между статистикой и математической статистикой начался флирт, продолжающийся по сей день. Науки возникают из практики. Так, принято считать, что теория вероятностей порождена знаменитой перепиской П. Ферма и Б. Паскаля по поводу так называемых задач кавалера де Мере, которая началась в 1654 г. Если толчком к созданию теории вероятности послужили азартные игры, то статистика нашла себе применение во многих областях, таких как демография, астрономия, а затем и эмпирические науки.

Адлер Юрий Павлович — президент Международной гильдии профессионалов качества, главный редактор журнала «Менеджмент качества», кандидат технических наук, профессор.

$$P(AB) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B)$$

$$P(B/A) = \frac{P(AB)}{P(A)}$$

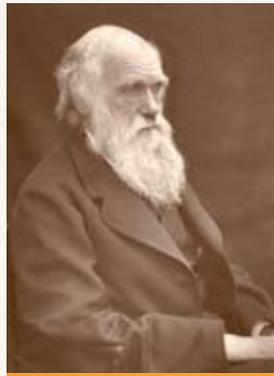


Аристотель

Поскольку статистика долгое время не входила в обязательный набор университетских дисциплин, она оставалась уделом отдельных творческих людей. Один из ярких примеров такого рода — судьба знаменитого британского контр-адмирала Р. Фицроя, создателя современной гидрометеорологии, систем прогноза погоды и штормовых предупреждений. Он был назначен главным статистиком-метеорологом Британии, не получив, видимо, никакого специального образования. Вообще, прогноз погоды — пример хорошей статистической задачи, решение которой все еще впереди. То время богато на статистические прорывы. Достаточно вспомнить работы Ч. Дарвина, которого как раз Р. Фицрой возил в кругосветное путешествие. Забавно, что сводку основных результатов, полученных к этому времени в вероятностных представлениях, сделал человек, который был убежденным детерминистом. Его звали Пьер Симон де Лаплас. Эта сводка была опубликована в 1812 г. Свои мысли он выражал точно и недвусмысленно: «Все явления, которые, по-видимому, не подчинены никаким естественным законам, точно так же естественно необходимы, как вращение солнца вокруг оси... Случая не существует. Случай есть не более чем выражение нашего неведения». Бельгиец А. Кетле слушал лекции Лапласа и находился под его сильным влиянием. Тем не менее, он существенно продвинул «политическую арифметику», введя, например, понятие «средний человек». Тем временем Ч. Дарвин плавал в поисках те-

ории эволюции (1831–1836), а его двоюродный брат Френсис Гальтон, родившийся в 1822 г., подрастал и готовился к череде важных открытий, которые ему суждено было совершить.

Ф. Гальтона заинтересовала связь между талантами людей и талантами их предков. Он решил проверить гипотезу об этой связи, в результате чего появилась знаменитая книга «Наследственный талант». Увлечшись этой темой, он обнаружил, что надо как-то собирать и обрабатывать данные. А поскольку он не был статистиком, ему пришлось самому разработать «вариационную статистику», основанную на идее распре-



Ч. Дарвин

деления признаков у людей, а затем и у других биологических видов. Концепция variability пришла Гальтону из работы Чарльза Дарвина о происхождении видов и прошла красной нитью через все, что делал Гальтон, — от исследования отпечатков пальцев (антропометрия), широко используемого теперь в криминалистике, до корреляционного и регрессионного анализа. Среди последователей Гальтона выделялся выдающийся статистик Карл Пирсон, который к тому же оказался гораздо ближе к математике. Именно на его долю выпала честь создания того варианта прикладной статистики, который сохранился до наших дней.

В 1901 г. Пирсон вместе с Гальтоном и морским зоологом Р. Уелдоном, пионером прикладной биометрии, учредили первый в Евро-

пе да и в мире журнал *Biometrika*, от начала которого вполне можно отсчитывать время современного периода развития статистики. Пирсон полагал, что журнал будет способствовать экспериментальной проверке гипотезы Дарвина о происхождении видов с помощью построения кривых распределения (кривых Пирсона) и проверки гипотезы о том, что они не противоречат результатам наблюдений в соответствии с критерием «хи-квадрат» (критерием Пирсона). Для этого предлагалось по экспериментальным данным вычислять первые четыре момента кривых распределения, а именно: среднее, дисперсию, асимметрию и эксцесс. Биометрия — детище Пирсона — была не единственной областью его интересов. Главным делом его жизни стала «Грамматика науки». Для нас важна его мысль о том, что статистика — это прежде всего то, что задает грамматическую структуру всего, что претендует на название «наука».

Австрийский монах Грегор Мендель (1822–1884) вместо того, чтобы искать «вариабельность в неслучайном», показал, что «совокупность случайностей может вести к неслучайному результату». Так возникла генетика — важная область приложения статистических методов.

Современный период

На современный период развития статистики приходится целый XX в. Историю статистики в XX в. удобнее всего проследить по прекрасной книге Дэйвида Солсберга «Леди, дегустирующая чай: как



Р. Фицрой

$$P(AB) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B)$$

$$P(A/B) = \frac{P(AB)}{P(B)}$$

статистика революционизировала науку в двадцатом веке». Конечно, строить всякие границы и рубежи — дело сомнительное. Но, видимо, людям удобнее структурировать события и описывать их по кластерам. Это позволяет нам проследить процесс рождения новой парадигмы в рамках старой.

Итак, в наследство от Гальтона нам достались: корреляция, регрессия и основы вариационной статистики. Пирсон добавил к этому теорию распределений и инструменты проверки гипотез. С этим мир вступил в прошлый век. А начало века ознаменовалось прорывными работами Уильяма Госсета.

Случая не существует. Случай есть не более чем выражение нашего неведения.

Он не только осознал проблемы малых выборок и ввел распределение и критерий, получившие имя Стьюдента, но и предвосхитил идеи планирования эксперимента и контрольных карт.

Распределение Стьюдента было первым звонком, оно подрывало представление об универсальности нормального распределения, введенного К.-Ф. Гауссом. Говорят, что Гаусс по просьбе астрономов нашел математическую форму кривой, которая хорошо описывала поведение отклонений результатов траекторных наблюдений. А пользователи решили, что Гаусс нашел некоторый универсальный закон, которому подчиняется поведение отклонений. Так его и стали воспринимать практики. Это ощущение усиливалось еще и так называемой центральной предельной теоремой теории вероятности. Она формулирует условия, при которых среднее как угодно распределенных данных с ростом числа наблюдений стремится именно к нормальному распределению.

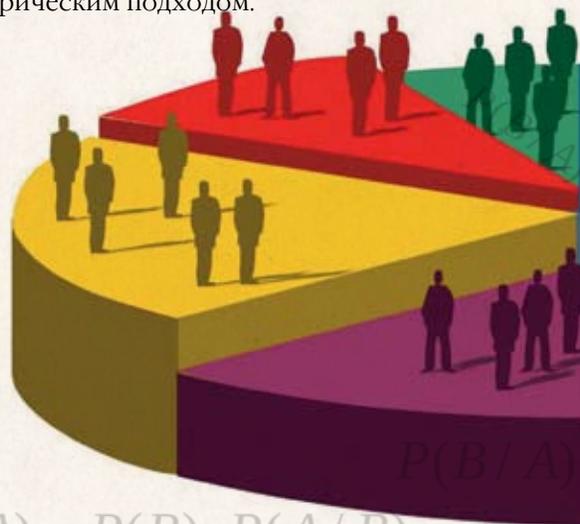
Конечно, было известно, что есть дискретные распределения, например биномиальное распределение или распределение Пуассона. Они не вызывали особого беспокойства, видимо, потому, что имели «другую природу», а с практической точки зрения почти не отличались от нормального. Но к непрерывным и ненормальным распределениям пришлось привыкать. Жизнь вызвала потребность в логарифмически нормальном распределении, востребованном, например, в минералогии и обогащении полезных ископаемых, в семействе распределений Вейбулла, используемых, в частности, в многочисленных задачах надежности.

Более того, оказалось, что при практическом использовании некоторой модели закона распределения возникает неопределенность, связанная как с обычно малым объемом выборки, так и с известным произволом при выборе самой модели. И если одной из главных проблем прикладной статистики в XX в. была малость выборок, то теперь начала бурно развиваться ветвь прикладной статистики, у которой главная проблема — огромность выборок. В связи с малостью выборок Рихард фон Мизес, младший брат Людвиг фон Мизеса, занимаясь частотной интерпретацией теории вероятности, выяснил, что число 30 наблюдений для объема выборки может рассматриваться уже как «большое». Это стало на практике удобным ориентиром, оставив Стьюденту совсем маленькие выборки.

Что же касается произвола при выборе модели, то дело было не только в том, что лишь в нескольких простых ситуациях выбор вида модели был очевиден. Так,

было известно, что распад радиоактивного изотопа хорошо описывается моделью распределения Пуассона, а подбрасывание правильной монеты — биномиальным распределением. В большинстве случаев наших априорных знаний оказывается совершенно недостаточно для идентификации распределения. Поэтому принцип «максимума правдоподобия», предложенный Р. Фишером, не получил очень широкого распространения.

Практика показывала, что с распределениями связаны и другие неприятности. Например, «выбросы» или «дикие наблюдения» неизменно подстерегают исследователя. Всегда возможны человеческие ошибки, описки и т.п. Поэтому неудивительно, что в 1945 г. Франк Уилкоксон предложил новый подход, получивший название непараметрического [1]. Он сразу начал стремительно развиваться и теперь вполне конкурентоспособен при сравнении с классическим подходом. Дилемма очень проста. Если вы не допускаете мысли о том, что распределение не может быть иным, чем то, на котором вы настаиваете, то, конечно, надо пользоваться параметрическим подходом. Но если у вас есть сомнения относительно закона распределения, то, чем больше сомнения, тем естественнее воспользоваться непараметрическим подходом.

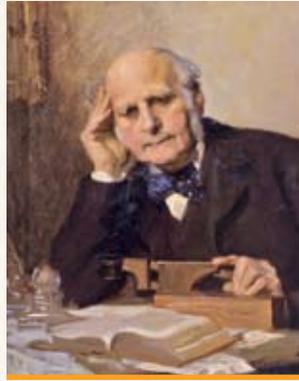


$$P(AB) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B)$$

$$P(B/A) = \frac{P(AB)}{P(A)}$$

Все модели ошибочны, но...

Еще в XVIII в. Томас Байес в Англии в свободное от своих пасторских обязанностей в англиканской церкви и от разведения кроликов время занимался условными вероятностями и вывел формулу, получившую впоследствии его имя, — формулу условных вероятностей Байеса. Она широко использовалась в стандартных вероятностных расчетах, но только в XX в. Джон Мейнард Кейнс заметил в «Трактате о вероятности» (1921), что «все вероятности условные». Так, наша оценка вероятности некоторого события до начала его исследования (априори или до) не будет совпадать с нашей же оценкой вероятности этого события после исследования или эксперимента (апостериори или после). Можно выдвинуть произвольную гипотезу о виде закона распределения и, накапливая данные, систематически проверять выдвинутую гипотезу о нормальности распределения до тех пор, пока не накопится достаточно данных, чтобы уверенно отбросить ее. Тогда нам придется выдвинуть новую гипотезу, учитывая информацию о том, что, например, нормальный закон противоречит накопленным данным. Так действует механизм, открытый в этом же веке несколько позже Карлом Поппером как механизм верификации и фальсификации гипотез и удачно обобщенный Джорджем Боксом в афо-



Ф. Гальтон

ризме: «Все модели ошибочны, но некоторые из них полезны».

Здесь представляется уместным напомнить, что статистика всегда была открыта флирту со смежными направлениями. Наиболее яркие примеры — это теория игр, теория полезности и теория принятия решений. Игровые модели органичны статистическим представлениям и не случайно возникли в работах Эмиля Бореля во Франции в 1920-е годы, в несколько более поздних работах Джона фон Неймана, но до появления ставшей классической работы фон Неймана и Оскара Моргенштерна «Теория игр и экономическое поведение» не были в центре внимания.

Статистику голыми руками не возьмешь

Всередине прошлого века повсеместно чувствовалось, что со статистикой все труднее справиться голыми руками. Как раз появились мощные компьютеры, которым было суждено изменить лицо статистики. Этому предшествовали интенсивные поиски программных средств и создание компаний, которые в свою очередь создали соответствующие программные системы. К таким компаниям, например, относятся SAS, SPSS, Statistika, Statgraphics, Minitab и др.

Компьютер сумел преодолеть и предотвратить многие трудности во многих статистических задачах и среди них — в задачах регрессионного анализа. Созданный, как мы помним, еще Ф. Гальтоном,

он начал быстро развиваться, но с ростом числа независимых переменных быстро достиг пределов возможностей «невооруженного» человека. Делались попытки прорваться за счет использования техники ортогональных полиномов Чебышева, но только после того, как М. Эфроимсон предложил алгоритм шагового метода, дело постепенно пошло на лад. Н. Дрейпер и Г. Смит в своей книге «Прикладной регрессионный анализ» последовательно в трех изданиях подробно рассказывают о развитии вычислительных процедур регрессии.

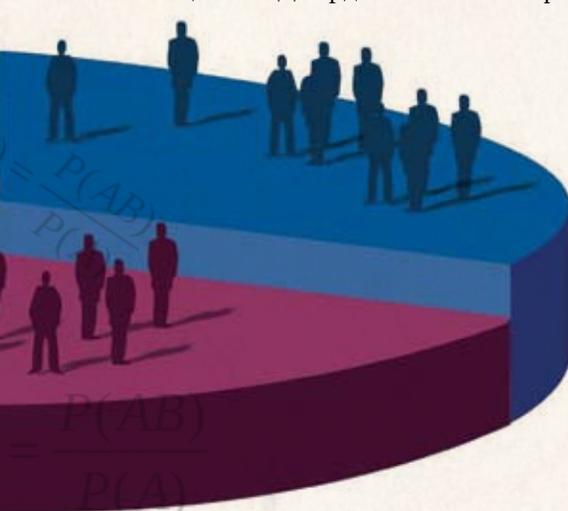
Регрессионная модель оказалась наиболее глубоко проработанной. Более того, при желании можно всю статистику рассмотреть сквозь призму этой модели. Тем более что практически любую статистическую модель можно представить как многомерную. Все подобные обобщения немедленно приводят к компьютерам и программным продуктам. Каждый метод имеет свою специфику, но алгоритмы оказываются очень похожими. Даже такие внешне далекие методы, как дисперсионный анализ и факторный анализ, возникший в связи с задачами психологии, технически различаются не слишком сильно.

Возможность быстро и дешево проводить вычисления большого объема породила целую гамму новых методов, называемых для краткости методами «бутстрепа». За этим обобщающим названием скрываются такие методы, как метод «складного ножа», метод «перепроверки». В окончательном виде они были сформулированы Брэдли Эфроном в начале 60-х годов прошлого века.

Вместе с тем, иногда нормальному распределению следуют не сами интересующие нас величины, а, скажем, их логарифмы. Можно ли найти такое преобразование этих величин, которое превратит их распределение в наперед заданное? Мы думаем, что практи-

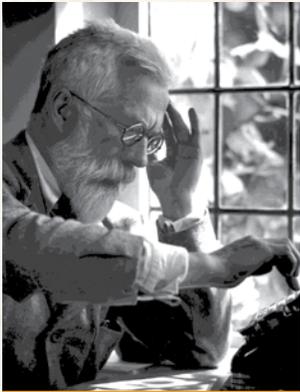
$$= \frac{P(AB)}{P(A)}$$

$$= \frac{P(AB)}{P(A)}$$



$$P(AB) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B)$$

$$P(A/B) = \frac{P(AB)}{P(B)}$$



Р. Фишер

чески всегда ответ на этот вопрос положителен. Трудность заключается лишь в том, что преобразования меняют метрику пространства и требуют интерпретации результатов уже в новой метрике. А придумать интерпретацию в новой метрике значит создать новую содержательную теорию, как это сделал, например, Сванте Аррениус в 1903 г. в теории химической кинетики, за что был удостоен Нобелевской премии по химии. В XX в. усилиями Джорджа Бокса и его коллег была построена систематическая теория преобразований.

Стоит отметить, что на статистику кроме вышеназванных влияли и влияют еще многие науки. Часто эти влияния бывают взаимными и плодотворными. Один из важнейших примеров — кибернетика. Тезис кибернетики о том, что «оптимальное управление сложными системами возможно при неполном знании их механизмов», оказалось для статистики как нельзя более кстати. Важную роль сыграла и модель «черного ящика». В середине XX в. возникли целые системы наук, имеющие хорошие перспективы военного приложения и, как обычно бывает, получившие благодаря конкуренции и мощной поддержке государств быстрое и эффективное развитие. К таким наукам относятся исследование операций, теория надежности систем, собственно теория систем, системотехника, теория очередей, теория массового обслуживания, теория распознавания образов и др. Все

эти науки существенным образом использовали статистику.

Переписи и обследования принадлежат, видимо, к самым древним задачам, решение которых трудно себе представить без использования статистических методов. Ушедший век, тем не менее, внес в их организацию и проведение много нового. Тщательному анализу подверглась идея выборочной оценки. Возникли два ключевых понятия: репрезентативность (представительность) выборки и рандомизация при отборе элементов выборки. Идея рандомизации принадлежит Рональду Фишеру. Она была высказана им в связи с планированием эксперимента и принадлежит к числу важнейших из многочисленных идей Фишера [2]. Хотя статистика — неотъемлемая часть планирования эксперимента, случилось так, что она как-то дистанцировалась, чувствуя, что в планировании эксперимента есть что-то чужеродное, нестатистическое. Видимо, именно с этим связано то, что планирование до сих пор не заняло того места в арсенале статистики, которого, несомненно, заслуживает и которого ждут многочисленные задачи практически во всех областях человеческой деятельности. Что ж, ждем этого от наступившего века.

Очертания сложившейся парадигмы

Естественная экспансия статистики привела к тому, что в ее сферу стали попадать объекты, традиционно не рассматриваемые в рамках этой науки. Это направление стало называться статистикой объектов нечисловой природы [3]. Другим расширением стала статистика случайных процессов, временных рядов и вообще объектов, имеющих пространственно-временную структуру. Прежде всего, это оказалось востребованным теорией коммуникаций.

Статистика многими своими достижениями обязана математи-

ке вообще и теории вероятности в частности. Иногда их вообще трудно разделить и различить. Интересно, что вектор развития математики в сторону «утраты определенности» совпадает с вектором развития статистики в сторону учета неопределенности, ее оценки.

Можно представить себе, например, следующую систему. Основа статистики — теория вероятности. Статистика — математическая дисциплина с богатыми и разнообразными приложениями. Будучи частью математики, статистика может и должна добиваться некоторого приемлемого уровня математической строгости в духе представлений Анри Лебега или знаменитого проекта Николая Бурбаки. Вот замечание А. Н. Колмогорова, сделанное им в предисловии к вышеупомянутой книге Лебега: «...у математиков существует склонность, уже владея законченной математической теорией, стыдиться ее происхождения...»

Сам по себе разрыв между теоретическими моделями и практикой их использования не страшен, но он порождает быстро усложняющуюся проблему выбора. Ответственность за адекватность выбора, естественно, лежит на том, кто собирается применять выбранную модель. А как обеспечить эту адекватность? Выбор невелик. Мы либо должны *знать* заранее, что выбрали самую лучшую модель именно для нашей задачи, либо должны *верить*, что это наша модель, либо, наконец, нам придется *проверять*, действительно ли ее предпосылки хорошо согласуются с реальными обстоятельствами нашей задачи.

Конечно, первая ситуация — самая желанная, но на практике она встречается очень редко. Сочетание редкости и желанности рождает мифы. Ключевая проблема выбора подходящей модели заключается в том, что даже в случае «знаю» есть риски, которые трудно поддаются оценке. Вера в «пра-

$$P(AB) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B)$$

$$P(B/A) = \frac{P(AB)}{P(A)}$$

вильность» модели обычно опирается на моду. К сожалению, нет убедительных доказательств того, что искренность и усердность молитвы повышает шансы того, что данная модель окажется удачной в данном случае. Поэтому мы скептически относимся в такой возможности выбора модели.

Что же остается? Проверка пригодности модели? Но это же очень затратно! Конечно, но дело не только в этом. Методы проверки предпосылок моделей можно разделить на два класса: прямые и косвенные методы. В прямых методах главная проблема — неэффективное использование ресурсов и проблема взаимопонимания. В результате возникает лингвистическая пропасть, преодолеть которую может не каждый.

Косвенные методы проверки гораздо мягче. Один из самых распространенных — это воспользоваться какой-нибудь модной моделью, и если результат представляется приемлемым, то по умолчанию и постулаты модели считаются приемлемыми. Другие косвенные методы основаны на различных технических приемах вроде анализа остатков в регрессионной модели.

Статистика — основной генератор моделей. Но чтобы от модели была польза, ее надо наполнить конкретным содержанием. Для этого есть несколько путей. Можно, например, собрать те данные, кото-

рые уже есть. Можно организовать сбор данных специально для выбранной модели. Можно, наконец, спланировать экспериментальное исследование и реализовать его. При выборе метода сбора данных возникает коллизия. Опыт показывает, что планирование эксперимента, развитое в XX в. усилиями сэра Рональда Фишера и его многочисленных коллег, во многих отношениях более эффективно, чем анализ кое-как собранных случайных данных. Однако с планированием эксперимента, как правило, связаны дополнительные расходы на организацию

К новой статистической парадигме

Как мы представляем себе парадигму, в которой сейчас живем? Видимо, первым «возмутителем спокойствия» был Госсет. Он одним из первых усомнился в незыблемости нормального закона. Свой вклад в разрушение основ внесли и Р. Фишер, и многие другие исследователи. Но самые плодотворные результаты, направленные на разрушение старой парадигмы, возникли от взаимодействия статистики с контролем качества. Все началось в 1922 г. в Лабораториях Белла. Руковод-

Переписи и обследования принадлежат к самым древним задачам, решение которых трудно себе представить без использования статистических методов.

эксперимента и достижение независимости варьирования выбранных факторов.

Пусть данные, так или иначе, собраны. Что дальше? Теперь самое время выяснить, в каких отношениях находятся собранные данные и выбранная модель. Для статистика это новое поле деятельности: проверка статистических гипотез, прежде всего гипотезы об адекватности выбранной модели. Но тут предпосылки снова заявляют о себе, они нужны для корректной статистической проверки гипотез. В сущности, эта проверка мало что проясняет, поскольку ограничения на ресурсы и время, с одной стороны, и естественная вариабельность систем — с другой создают условия, в которых потенциальная мощность критериев оказывается часто слишком слабой. И тогда возникает соблазн выдать то, что интуитивно кажется правильным, за результат, освященный авторитетом статистики. Такая фальсификация данных может проходить неосознанно.

дитель департамента статистики доктор Джонс поручил молодому инженеру-физику Уолтеру Шухарту разобраться в причинах отказов усилительных подстанций телефонных сетей в произвольные моменты времени, а не через определенное число часов, указанных в технической документации. Через несколько месяцев Шухарт пришел к нескольким важным выводам.

1. Все системы, как естественные, так и искусственные, постоянно подвержены вариациям.
2. Если в жизнь системы не вмешиваться, то ее собственные вариации остаются статистически неизменными.
3. Если в жизнь системы вмешиваются случайные воздействия, то она выходит из управляемого состояния и становится непредсказуемой.
4. Для выяснения того, в каком именно состоянии — управляемом или неуправляемом — нахо-



А. Колмогоров

$$P(A/B) = \frac{P(AB)}{P(B)}$$

$$P(AB) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B)$$

дится наша система, нужна система мониторинга ее состояния, для чего Шухарт предложил строить график, названный позднее контрольной картой Шухарта.

Это послужило основой для построения процедуры принятия решений в различных практических ситуациях. Так, например, выяснилось, что на практике менеджмент имеет дело с обычной таблицей с двумя входами. Одним входом служит состояние системы, определенное в процессе мониторинга: стабильное или нестабильное. Вторым входом служит удовлетворенность или неудовлетворенность работой системы с точки зрения потребителей и владельцев данного бизнеса. Первый квадрант: система стабильна, и все довольны. Это идеальная ситуация. Ни во что вмешиваться не надо, важно только следить. А рано или поздно она непременно нарушится. Тогда мы попадаем во второй квадрант: система нестабильна, но все довольны. Казалось бы, тоже можно пока ничего не делать, но Шухарт, а потом и Э. Деминг думали иначе. Здесь надо немедленно вмешиваться в систему, причем, только снизу, то есть вмешиваться должны люди, работающие в системе, а вовсе не менеджмент, поскольку люди в системе имеют гораздо больше шансов выяснить природу «особой» вариации, которую продемонстрировала карта Шухарта, а значит, и найти средства и методы ее элиминирования.

В третьем квадранте имеем ситуацию: система стабильна, но результаты не устраивают клиентов или владельцев. В этом случае, конечно, нужно вмешательство, только на этот раз сверху, то есть со стороны менеджмента. В последнем квадранте нет ни стабильности, ни удовлетворенности. Надо начать со стабильности, то есть снизу, и только потом заняться всем остальным. Так сложилась концепция, которую теперь принято называть статистическим мышлением.

Для разработки инструментов статистического мышления нужны профессиональные знания статистики, однако использование этих инструментов на практике вовсе не требует знаний статистики. Еще нужна практика под руководством квалифицированного инструктора.



У. Шухарт

Хотя контрольными картами Шухарта относительно легко пользоваться, сами они оказались необычными статистическими объектами. В классической статистике принято считать, что с ростом числа наблюдений среднее асимптотически стремится к некоторому математическому ожиданию. Но для контрольной карты Шухарта, очевидно, нет никакого математического ожидания. Среднее будет флуктуировать со временем и только. Да и границы карты были выбраны эмпирически, скорее, из экономических, чем из статистических соображений. Причем в работе карты есть несколько важных особенностей. Так, например, карта, как мы знаем, обладает предсказательной силой. Почему это столь важно? Деминг говорил, что никакое управление невозможно без предсказания. А предсказание — это, по существу, вывод из прошлого к будущему. То есть, это *индуктивный* вывод. А мы знаем, что индуктивные выводы опасны. Что может сделать индуктивный вывод надежным? Оказывается, что в данном случае — стабильность контрольной карты при сохранении статус-кво. Это теория.

Надежное предсказание производственного будущего

Для получения надежного предсказания будущего критически важно иметь стабильную в статистическом смысле систему, неотягченную никакими привходящими возмущениями, более существенными, чем достигнутый уровень стабильности, и не способными его поменять. А всегда ли это нам нужно? Отвечая на этот вопрос, Деминг предлагает разделить все статистические исследования на два класса: численные и аналитические. Несколько упрощая, можно сказать, что Деминг делит статистические исследования на интерполяционные, где мы не задаемся вопросами «почему?» и «что делать?», а просто хотим получить конкретную информацию, которая нас интересует. Однако нас всегда начинают заботить условия, обеспечивающие надежность индуктивного вывода. Выходит, что интерполяция — задача гораздо более простая, чем экстраполяция.

Как же мы отличаем статистически стабильное состояние системы от нестабильного? Технически это выглядит очень просто. Контрольная карта — это график, по горизонтальной оси которого отложено время, а по вертикальной — некоторый показатель, интересующий наших клиентов. Результаты измерений наносятся на этот график в виде точек. Теперь остается нанести на карту верхний и нижний пределы. Их физический смысл понятен — это границы, задаваемые естественной вариацией нашей системы. Шухарт предложил строить их на расстоянии трех квадратичных отклонений выше и ниже центральной линии. Этот выбор не связан непосредственно с моделями данных, и многие восприняли его как выбор модели нормального распределения. Для статистического анализа это очень удобно. Но, к сожалению, на практике это совершенно неприемлемо. Такая модель данных делает диагностику непредсказуемой.

$$P(B/A) = \frac{P(AB)}{P(A)}$$

Поэтому Шухарт, а вслед за ним и Деминг, предложил отказаться от постулатов статистической модели данных и заменить их *операциональными определениями*. Тогда границы карты на уровне трех сигм становятся эмпирическим решением, основанным на балансировке двух возможных и ведущих к опасным последствиям ошибок. Одна ошибка состоит в том, чтобы принять стабильный процесс за нестабильный, а вторая, напротив, в том, чтобы принять нестабильный процесс за стабильный. Каждая такая ошибка чревата большими и трудноизмеримыми потерями. Избежать их раз и навсегда не в нашей власти, но можно попытаться минимизировать суммарные потери. Из таких соображений и были выбраны Шухартом пределы карты. Теперь появление любой одной точки за пределами контрольных границ — не важно, выше или ниже — становится операциональным определением выхода системы из состояния статистической управляемости. И это решение не связано со статистической моделью данных. Мы можем считать, что это непараметрическая или робастная ситуация, но на принимаемые решения все подобные рассуждения не оказывают существенного влияния.

Таким образом, если наш диагноз гласит, что система статистически стабильна в данный момент, то мы спешим воспользоваться этим благоприятным обстоятельством для прогнозирования ее развития на ближайшее будущее. А что делать,

если карты говорят иное? Тогда, как мы уже выяснили, надо немедленно вмешиваться в систему, дабы попытаться вернуть ее в стабильное состояние. Причем, это должно быть вмешательством снизу. Но как люди, работающие в системе, узнают, что именно им следует делать? В Японии сразу после окончания Второй мировой войны на-

казалось бы, у такого набора методов очень мало шансов на практический успех, но жизнь показала, что с их помощью решается около 95% всех реальных задач. Интересно понять, что же дало такой потрясающий эффект? Дело, видимо, в том, что авторы семи инструментов нарушили одно из самых священных табу:

Статистика многими своими достижениями обязана математике вообще и теории вероятности в частности.

чались работы над новым подходом. Это направление возглавил профессор Каору Исикава.

Великолепная семерка качества

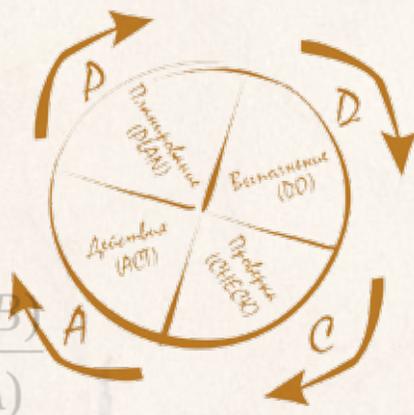
То, что сделали Исикава и его сотрудники, обычно теперь называют семью простыми инструментами статистического контроля качества. Эти методы предназначены для двух разных целей. Первая из них — поиск обстоятельности, действий, воздействий, которые вызвали или способствовали вызову некоторой особой причины, нарушившей стабильность системы, и выработка действий, направленных на устранение нестабильности, или хотя бы смягчение ее последствий. Вторая цель — решение локальных проблем в ходе стабильной работы системы в рамках механизма непрерывного совершенствования, который японцы называют словом «кайдзен».

Первая задача заключалась в том, чтобы выбрать несколько (например, семь) очень простых инструментов, которые должны изначально обладать, по крайней мере, следующими свойствами.

1. Иметь как можно более простую и многомерную графическую интерпретацию.
2. Охватывать по возможности все типичные структуры данных.
3. Иметь простое и доступное математическое обоснование.

они отказались от требования математической строгости. Проще говоря, предлагая вам для какого-то конкретного случая воспользоваться одним из семи инструментов, они вовсе не гарантируют вам успех. А чтобы не возник конфуз, они предусмотрели два вида «защиты». Первый из них заключается в том, что в структуре самого инструмента заложен некоторый простой критерий, позволяющий распознавать те случаи, когда данный инструмент может не сработать. Например, строя гистограмму для изучения распределения некоторой случайной величины, они предлагают вам набор стандартизованных картинок и краткую инструкцию по интерпретации каждой из них.

Второй защитный механизм связан с использованием подхода, который мы теперь называем циклом Шухарта — Деминга: планируй — делай — исследуй — действуй. Исикава понял это так: планируй — делай — проверяй — действуй. Этот цикл он и включил в методологию применения семи инструментов. Теперь отдельная неудача, хотя она и замедляла прогресс, все-таки не имеет решающего значения, поскольку мы имеем возможность исправить ее в следующем цикле. Жизнь предприятия, таким образом, погружается в среду непрерывного совершенствования стабильных про-



$$P(AB) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B)$$

$$P(A/B) = \frac{P(AB)}{P(B)}$$

цессов и систем с перерывами на поиск и устранение особых источников вариации. При этом используется единый набор инструментов — семь простых.

Необходимым условием работы этих инструментов является создание специальных команд часто на постоянной основе, называемых обычно кружками качества, хотя командный стиль работы был мало распространен среди статистиков.

Технологии качества

В послевоенные годы сложилась одна из самых оригинальных систем представлений о роли, месте и возможностях применения статистических методов в системах менеджмента качества, да и в общем менеджменте. На Западе их стали называть методами Тагути по имени их создателя профессора Гениги Тагути.

Для того, чем занимается Тагути, он придумал название Quality Engineering, которое часто переводят на русский как «инжиниринг качества». Мне кажется, что это недоразумение. Поэтому я предлагаю использовать естественный термин «технологии качества».

Первый из методов Тагути — функция потерь по Тагути. Одна из вечных проблем исследователей — это постоянная потребность в одновременном исследовании нескольких разных показателей качества или откликов. Изучать каждый отклик отдельно часто не по карману. Поэтому люди издавна занимались поисками подходов к свертыванию многомерных откликов в скалярные показатели. Об этом написаны горы книг. Но проблема постоянно требует априорной информации о возможном характере свертки. Г. Тагути удалось построить очень простую нелинейную свертку, опираясь на представления о требованиях потребителей и о границах допусков на эти требования, согласованных между продавцом и потребителем.



Э. Деминг

Следующий момент в концепции Тагути относится к связи между разработкой, проектированием, конструированием, пуском и наладкой, с одной стороны, и эксплуатацией действующей системы (например, производства) — с другой. Тагути не просто обратил внимание на эту связь, но и предложил технологию систематического решения задач разработки и внедрения продукции и оборудования в производство.

Главное новшество в планировании экспериментов по Тагути — концепция «робастности». Интересно, что этот термин, который, как мы знаем, ввел Дж. Бокс, Тагути использовал хотя и в близком, но все-таки ином смысле. Он предложил классифицировать факторы следующим образом: факторы, влияющие на номинал отклика («сигнал»); факторы, влияющие на вариацию, разброс значений отклика; факторы, влияющие и на то, и на другое, и, наконец, факторы, не влияющие ни на то, ни на другое, то есть несущественные для исследования данного объекта.

Задача планирования эксперимента как раз и состоит в том, чтобы расклассифицировать факторы в данной задаче. Тогда можно будет отбросить за ненадобностью факторы, попавшие в последнюю группу, оставить на потом факторы третьей группы в надежде на то, что до них не дойдет дело, поскольку удастся получить удовлетворительное решение с помощью только факторов первых двух групп, а для факторов первой и второй групп построить по полученным экспериментальным данным отдельные модели,

чтобы с их помощью провести независимую оптимизацию.

Таким образом, робастные технологии позволяют построить такой объект или такую систему, которая будет обеспечивать наилучший возможный сигнал, одновременно стремясь к минимально возможному шуму. Значит, система становится устойчивой по отношению к факторам, создающим шум. И это достигается не за счет дополнительных капиталовложений, а только за счет «разума», за счет выбора наилучшей конфигурации системы. При этом Тагути предлагает провести планирование эксперимента уже в пространстве допусков на факторы, чтобы решить, какие из них имеет смысл ужесточить, чтобы добиться желаемого результата с наименьшими затратами.

От статанализа к анализу данных

В 1962 г. вышла в свет большая статья Джона Тьюки, в которой автору понадобилось новое словосочетание «анализ данных» (раскопки, добыча данных) вместо привычного «статистический анализ». Почему? Мы думаем, что это явный признак предчувствия смены парадигм. Тьюки не привел точного определения термина «анализ данных», он лишь заметил, что этот термин гораздо шире, чем статистика, и целиком включает ее в себя. Тем не менее, из его работы ясно, что он чувствует «узость» статистического парадигмы и стремится расширить ее за счет эффективных приемов визуализации, позже названных когнитивной графикой, за счет интенсивного использования нарождающегося нового поколения «дружелюбных» компьютеров.

Настаивая на визуализации, Тьюки подчеркивал, что числовые свертки данных делаем акцент на ожидаемом, а графические представления — на неожиданном. Развивая свой подход, Тьюки пришел к идее разведочного анализа. Анализ данных во всех вариантах вы-

$$P(B/A) = \frac{P(AB)}{P(A)}$$

звал большой интерес, но не стал сразу повальным увлечением. Требовалось время для освоения возможностей компьютеров. Тьюки даже ввел несколько терминов, которые теперь общеприняты, например бит.

Можно еще упомянуть концепцию «Шесть сигм». Она представляется нам американской реакцией на методы Тагути и интересна, прежде всего, с психологической точки зрения.

Нам неизвестно о каких-либо целенаправленных усилиях в этом направлении, предпринятых в нашей стране. Можно отметить только работы профессора В.В. Налимова, который упорно разрабатывал концепцию вероятностного мышления. В прошлом столетии толчком к такому пониманию послужила революция в физике, приведшая к квантовой механике с ее вероятностной структурой микромира.

Осенью 2009 г. в Токио профессор Хироэ Цубаки в частной беседе рассказал мне о своей идее, касающейся смены парадигм в статистике. Он заметил, что существующая парадигма, связанная еще с именем Карла Пирсона, себя исчерпала, и ей на смену, как он полагал, может прийти концепция, основанная на идеях профессора Тагути. Некоторые из своих соображений он отразил в только что вышедшей тогда книге *The Grammar of Technology Development* [4]. Эта идея показалась мне весьма заманчивой, и я стал ее обдумывать. Определенное стимулирующее влияние на ход моих мыслей оказало



К. Исикава

появление книг Талеба, особенно его «Черного лебедя». В результате появилась работа, которую вы сейчас читаете.

Моменты новых взглядов

Не знаю, можно ли построить модель новой парадигмы, опираясь только на концепцию Тагути, думаю, что нет, но мне ясно, что в ней есть ряд моментов, заслуживающих рассмотрения. Вот как я вижу сейчас основные моменты новых взглядов.

Главная причина того, что мы обсуждаем, ясна: *мир стал вероятностным*. И сейчас мы думаем, следуя В.В. Налимову, что это невероятностная интерпретация детерминированного мира, на которую наложена неопределенность, порожденная незнанием. Напротив, мы полагаем, что мир и в самом деле вероятностен, что это — его внутреннее имманентное свойство.

Если встать на эту точку зрения, то мы будем вынуждены отказаться от многих привычных вещей, прежде всего, от фундаментальных принципов детерминизма, ведущих к концепции причинности. Впрочем, это не слишком большая новость. В восточных мировоззренческих системах принципы индетерминизма сформулированы давно.

Нам не суждено узнать, справедлива ли гипотеза о вероятностном мире, но в ее пользу накопилось гораздо больше подтверждений, чем в пользу гипотезы детерминизма. Именно поэтому и происходит смена парадигм:

1. Вероятностный мир вместо детерминированного и причинного.
2. Вместо отдельных наук — междисциплинарные конгломераты.
3. Вместо теоретических моделей модели типа «черного ящика» с большим числом факторов и со сверткой многомерных откликов.
4. Вместо аксиом операциональные определения.
5. Вместо формализации визуализация.

6. Вместо статистического анализа моделирование в широком диапазоне.

7. Вместо управления системами по наитию изучение статистической устойчивости систем и действия в соответствии с диагнозом.

8. Использование планирования экспериментов как обязательного элемента научного исследования.

9. Вместо хаотичной обработки данных использование компьютерных раскопок (добычи) данных.

10. Вместо индивидуальных прорывов работа в смешанных командах над реальными проектами. Форма работы — непрерывный диалог.

11. Вместо случайного обучения — непрерывное обучение всех членов команды с поддержанием эффективного диалога.

12. Вместо детерминированного принятия решений учет элементов неопределенности.

13. Статистика — генератор моделей науки.

Конечно, все вышеперечисленные пункты нуждаются в тщательном детальном обсуждении. Но, какие бы коррективы ни были внесены в эти данные, все равно остается центральный вопрос: что делать, чтобы наилучшим образом подготовиться к грядущим переменам? Перемены не ждут. ■

ПЭС 12166/15.11.2012

Литература

1. Холлендер М., Вулф Д. Непараметрические методы статистики / Пер. с англ. Д.С. Шмерлинга; Науч. ред. Ю. П. Адлер и Ю.Н. Тюрин; Предисловие Ю.П. Адлера, Ю.Н. Тюрина и Д.С. Шмерлинга. М.: Финансы и статистика, 1983. 518 с.
2. Fisher R.A. *The Design of Experiments*. Edinburgh, London: Oliver and Boyd, 1937. 260 p.
3. Орлов А.И. *Нечисловая статистика*. М.: МЗ-Пресс, 2004.
4. Tsubaki H., Nishina K., Yamada S. (Editors). *The Grammar of Technology Development*. Japan: Springer, 2008. 221 p.

$$P(AB) = P(A) \cdot P(B/A) = P(B) \cdot P(A/B)$$

$$P(A/B) = \frac{P(AB)}{P(B)}$$