

Вводные лекции по статистике для ИММ (курс Китаевой А.В.)

«Исчислите всё общество сынов Израилевых по родам их, по семействам их, по числу имен, всех мужского пола поголовно...»

Четвертая книга Моисеева.
Числа. Ветхий завет, гл. 1,2.

У истоков статистической науки стояли две школы – **немецкая описательная и английская школа политических арифметиков.**

Представители описательной школы считали, что задачей статистики является описание достопримечательностей государства: территории, населения, климата, вероисповедания, ведения хозяйства и т. п. только в словесной форме, без цифр и вне динамики, т. е. без отражения особенностей развития государств в те или иные периоды, а только лишь на момент наблюдения. Видными представителями описательной школы были **Г. Конринг** (1606-1661), **Г. Ахенваль** (1719-1772), **А. Бюшинг** (1724-1793) и др.

Политические арифметики ставили целью изучать общественные явления с помощью числовых характеристик. Это был принципиально новый этап развития статистической науки по сравнению со школой государственоведения, так как от описания явлений и процессов статистика перешла к их измерению, исследованию, оценке и выработке вероятных гипотез будущего развития. Политические арифметики видели основное назначение статистики в изучении массовых общественных явлений. Виднейшим представителем и основателем этого направления был **В. Петти** (1623-1687). Именно школа политических арифметиков стала основообразующей в развитии современной статистики.

Математическое направление в статистике развивалось в работах англичан – сэра **Фрэнсиса Гальтона** (1822-1911) и **Карла Пирсона** (1857-1936), **Рональда Фишера**.

Прогрессу статистической методологии способствовали труды российских статистиков – **А.А. Чупрова** (1874-1926), **В.С. Немчинова** (1894-1964), **С.Г. Струмилина** (1877 - 1974), **В.Н. Старовского** (1905-1975) и др.

В настоящее время термин «статистика» употребляется в трех значениях:

Во-первых, под статистикой понимают **отрасль практической деятельности**, которая имеет своей целью **сбор, обработку, анализ и публикацию массовых данных о самых различных явлениях общественной жизни**. Осуществляется сбор данных в каждом регионе и по стране в целом о численности и составе населения, ведется подсчет предприятий и организаций, собираются данные об объемах производства и объемах продаж и т.д. Эту деятельность на профессиональном уровне осуществляет **Федеральная служба государственной статистики (Госкомстат РФ)** и система ее учреждений, организованных, по административно-территориальному признаку.

Во-вторых, статистикой называют **цифровые материалы**, служащие для характеристики какой-либо области общественных явлений или территориального распределения какого-то показателя, публикуемые в периодической прессе, справочниках, сборниках. Например, динамика цены на бензин в Томской области, представленная за второе полугодие 2008 года.

В-третьих, статистикой называется **отрасль знания, особая научная дисциплина**, которая в широком понимании **разрабатывает методы сбора, систематизации, анализа, интерпретации и отображения результатов наблюдений массовых случайных явлений и процессов с целью выявления существующих в них закономерностей**. Например, исследования взаимосвязи между качеством трудовых ресурсов и экономическим ростом в регионах РФ.

Основные этапы статистического анализа:

- 1. Планирование исследования** включает составление подробного плана сбора данных, возможно, с использованием случайной выборки из генеральной совокупности.
- 2. Предварительное исследование данных** включает рассмотрение набора данных с разных точек зрения, описание и обобщение данных. Выполнение этого этапа помогает убедиться, что запланированный анализ адекватен данным, а при необходимости позволяет внести в процесс анализа определенные коррективы.
- 3. Оценивание неизвестной величины** дает наиболее обоснованное возможное предположение о значении, основанное на исходных данных. Кроме того, есть возможность вычислить величину ошибки, которая возникает при использовании оценки вместо фактического неизвестного значения.
- 4. Проверка статистических гипотез** заключается в использовании данных для выбора одной из двух (или больше) различных возможностей при решении вопроса в неопределенной ситуации. Такая проверка позволяет убедиться, действительно ли данные обладают определенным интересным свойством, или мы имеем дело с "чистой случайностью", которая не представляет интереса.

Вероятность, исходя из предположений об изучаемой ситуации, показывает возможность или шанс наступления в будущем каждого из нескольких потенциальных событий. **Вероятность** – это понятие, в некотором смысле обратное статистике: вероятность показывает, какие данные Вы скорее всего получите, если известна характеристика ситуации, а статистика помогает охарактеризовать ситуацию в результате анализа и обобщения данных.

Основные термины

- **Статистика (statistics)**
- **Планирование исследования (designing the study)**
- **Предварительное исследование данных (exploring the data)**
- **Оценка неизвестной величины (estimating an unknown quantity)**
- **Проверка статистических гипотез (hypothesis testing)**
- **Вероятность (probability)**

Пример. Данные в менеджменте

Данные постоянно используют в менеджменте. Ниже приведен краткий перечень видов ежедневно используемой в менеджменте информации (перечень используемых данных).

1. Финансовые отчеты (и другие виды бухгалтерской отчетности).
2. Курсы и объемы ценных бумаг, процентные ставки (и другая информация, относящаяся к инвестициям).
3. Состояние бюджета (и другие сообщения правительства).
4. Отчеты о продажах (и другие внутренние отчеты).
5. Результаты обзора состояния рынка (и другие маркетинговые отчеты).
6. Данные о качестве продукции (и другие производственные отчеты).
7. Отчеты о производительности рабочих (и другие внутренние данные фирмы).
8. Цена и объем проданной продукции (и другие данные о продажах).
9. Расходы на рекламу и результаты рекламной компании (и другая рекламная информация).

Подумайте об этом. Вероятно, большая часть Вашей деятельности будет зависеть, по крайней мере, косвенно от данных. Вероятно, кто-то будет работать для Вас и консультировать Вас по этим вопросам, Вы редко будете видеть фактические данные. Время от времени Вы могли бы попросить дать Вам "сырые данные", чтобы иметь представление о перспективе. Просмотрев данные и задав вопросы, можно получить неожиданные результаты: можно обнаружить, что качество данных не так высоко, как Вы думали (и у Вас появится мысль: на чем же мы строим наши прогнозы?), или, наоборот, Вы получите обнадеживающие результаты и обретете уверенность. Другими словами, данные заслуживают внимания.

Общий совет

Статистические результаты должны допускать простое непосредственное объяснение (даже если соответствующая теория намного сложнее). Ниже приведено несколько общих советов.

1. Доверяйте своим суждениям, учитывайте здравый смысл.
2. Сохраняйте здоровый скептицизм.
3. Не дайте себя ввести в заблуждение с помощью на первый взгляд оригинального статистического анализа. Он может опираться на нереальные или неподходящие предположения.

Из-за большой гибкости, доступной аналитику на каждой стадии статистического анализа, один из самых важных факторов, который надо принять во внимание при оценке результатов статистического анализа, звучит так: *Кто это финансировал?* Помните, что аналитик много раз делает выбор – при определении проблемы, при планировании сбора данных, при выборе структуры или модели для анализа, при интерпретации результатов.

Набор **данных** содержит одно или несколько значений для каждого из отдельных объектов, называемых **элементарными единицами**. В качестве таких объектов могут выступать люди, домохозяйства, города, телевизионные приемники или что угодно, что представляет интерес для изучения. Для каждого из объектов регистрируют один и тот же признак (или признаки). Признак, который регистрируют для каждого из объектов (например, стоимость), называется переменной.

Существуют **три основных способа классификации наборов данных**: по **количеству переменных** (одномерный, двумерный и многомерный); по **типу** представленной каждой из переменных **информации** (числа или категории) и в зависимости от того, является ли набор данных **временным рядом**, или это **данные об одном временном срезе**.

Одномерные наборы данных (одна переменная) содержат информацию только об одном признаке, зарегистрированную для каждого объекта. Одномерный набор данных позволяет определить типичное значение и характеристику изменчивости данных, а также выделить специфические особенности или проблемы в данных.

Двумерные наборы данных (две переменные) содержат два признака, значения которых регистрируются для каждого объекта. Двумерные данные в дополнение к информации о каждой переменной как наборе одномерных данных позволяют изучить **связь между двумя переменными** и предсказать значение одной переменной на основе значения другой.

Многомерные наборы данных (много переменных) содержат три или больше признаков, значения которых регистрируются для каждого объекта. Многомерные данные в дополнение к информации о каждой переменной как наборе одномерных данных дают возможность изучить связь между переменными и предсказать значение одной переменной на основе значения других.

Значения переменных, которые регистрируются как числа, имеющие содержательный смысл, называют **количественными данными**. Дискретная количественная переменная может принимать значения только из некоторого списка конкретных чисел (таких как, например, 0 или 1, или перечень чисел 0, 1, 2, 3, ...). Любую количественную переменную, которая не является **дискретной**, будем называть **непрерывной**. Значения непрерывной переменной не ограничены простым перечнем возможных значений.

Если переменная содержит информацию о том, какой из нескольких нечисловых категорий принадлежит объект, то она называется **качественной переменной**. Если категории можно естественным образом и содержательно осмысленно упорядочить, то речь идет о **порядковой качественной переменной**. Если такой порядок отсутствует, то речь идет о **номинальной качественной переменной**. Несмотря на то, что часто значения качественной переменной можно записать с помощью чисел, такая переменная все равно остается качественной, а не количественной, поскольку эти числа не имеют какой-либо интерпретации, содержательно присущей этой переменной.

К количественным данным можно применять те же операции, что и к обычным числам: подсчет частоты, ранжирование, арифметические действия. С порядковыми данными можно выполнять только подсчет частоты и ранжирование, с номинальными данными – только подсчет частоты.

Если последовательность записи данных имеет содержательный смысл, то соответствующий набор данных представляет собой **временной ряд**. Если последовательность записи данных не важна, то соответствующий набор содержит **данные об одном временном срезе**. Анализ временных рядов сложнее анализа данных об одном временном срезе.

Если вы самостоятельно планируете сбор данных (даже если собственно сбор данных делают другие), то получаете **первичные данные**. Если вы используете данные,

предварительно собранные другими людьми и для других целей, то вы имеете дело с **вторичными данными**. Получение первичных данных часто обходится дорого и занимает много времени, но вы получаете то, что вам необходимо. Вторичные данные можно получить дешевле (или даже бесплатно), но вы можете найти, а можете и не найти то, что вам необходимо.

Основные термины

- **Набор данных (data set)**
- **Элементарные единицы (elementary units)**
- **Переменная (variable)**
- **Одномерный (univariate)**
- **Двумерный (bivariate)**
- **Многомерный (multivariate)**
- **Количественная (quantitative)**
- **Дискретная (discrete)**
- **Непрерывная (continuous)**
- **Качественная (qualitative)**
- **Порядковая или ординальная (ordinal)**
- **Номинальная (nominal)**
- **Временные ряды (time series)**
- **Об одном временном срезе (cross-sectional)**
- **Первичные данные (primary data)**
- **Вторичные данные (secondary data)**

Определения статистики

Статистика – это научно упорядоченное изображение действительности, вызванное необходимостью совершенствовать искусство управления.

1954 С моей точки зрения, собственно статистику можно, пожалуй, определить как искусство оперирования с неопределенностью и различиями в поведении людей в ситуациях принятия решений.
(L. J. Savage. The Foundations of Statistics).

1954 Статистика как наука является одним из разделов прикладной математики, и ее можно рассматривать как математику, применяемую при разработке результатов массового наблюдения... Статистику можно рассматривать как: 1) учение о совокупностях, 2) учение о вариации и 3) учение о методах приведения данных в компактной форме.
(Р. А. Фишер. Статистические методы для исследователей. 1958).

1959 В наше время принято считать, что статистика есть наука, изучающая теорию принятия решений в условиях неопределенности. Это определение статистики выкристаллизовывалось в результате многих лет ее развития. Достоинство этого определения состоит в том, что оно в сжатой и ясной форме излагает научное существо статистики.
(Г. Чернов, Л. Мозес. Элементарная теория статистических решений. 1962).

1963 Статистику иногда определяют как искусство и науку количественной обработки наблюдений, подверженных изменениям.
(E. V. Lewis. Statistical Analysis. Ideas and Methods).

1964 Значение этого слова (статистика) за последние два столетия претерпело значительные изменения. Слово "статистика" имеет один корень со словом "государство" (state) и первоначально оно означало искусство и науку управления; первые преподаватели статистики университетов Германии XVIII в. сегодня назывались бы специалистами по общественным (политическим) наукам. Поскольку решения правительства до некоторой степени основываются на данных о населении, промышленности (ремеслах), сельском хозяйстве и т. д., статистики, естественно, стали интересоваться такими данными, и постепенно слово "статистика" стало означать сбор данных о

государстве, а затем вообще сбор и обработку данных. Это значение слова все еще широко употребляемо, но наблюдается и дальнейшее изменение значения. Нет смысла собирать данные, если из этого не извлекается какая-то польза, и статистики, естественно, начинают заниматься интерпретацией данных. Современный статистик изучает методы, при помощи которых можно сделать выводы о популяции на основе данных, которые обычно получают из выборки популяции.
(J. L. Hodges, Jr, E. L. Lehmann. Basic Concepts of Probability and Statistics).

1965 В основе статистического вывода лежит индуктивное рассуждение, приводящее к утверждениям, верным лишь с определенной "степенью достоверности". Строгий смысл "степени достоверности" обеспечивается математическими методами и вероятностным подходом, образующими основы современной статистической теории
(С. Р. Рао. Линейные статистические методы и их применение. М., 1968).

1968 Математическая статистика – это ветвь теории вероятностей. В ней рассматриваются задачи, связанные с оперативными характеристиками правил индуктивного поведения, основанных на случайных экспериментах.
(Ю. Нейман. Вводный курс теории вероятностей и математической статистики).

1968 Статистика — это математическая теория того, как узнать нечто о мире через опыт.
(W. Thompson. The Future of Statistics).

Основные тенденции современной статистики



1 октября 2008 г. с лекцией, посвященной уровню общественного доверия к статистике, а также проблемам изучения благосостояния, в ГУ-ВШЭ выступил профессор Энрико Джованини, директор статистического департамента Организации экономического сотрудничества и развития (ОЭСР), автор множества научных трудов, один из крупнейших авторитетов мировой статистики.

«Вы знаете, что означает слово «статистика»— наука о государстве. А я пытаюсь внедрить термин «социотистика», то есть наука об обществе. Изначально ведь статистика была создана для того, чтобы политические руководители могли принимать решения. Однако сегодня статистику для принятия своих решений нужно знать всем, и мы пытаемся довести ее до людей, до населения. Я считаю, что эта тенденция носит глобальный характер, и это очень серьезное изменение».

Сейчас ученые, по мнению профессора, пытаются выйти за пределы экономического измерения благосостояния. Долгие десятилетия все ссылки делались на показатель ВВП, служивший главным критерием, по которому определялось, благоприятна ли ситуация в обществе, или нет. «Но теперь люди начинают интересоваться в этой связи и другими вопросами — скажем, состоянием окружающей среды. Ведь общественные потребности выходят за чисто экономические рамки существования государства».

Во многих странах сейчас наблюдается откровенное недоверие к национальным правительствам. И скепсис по отношению к официальной статистике часто является отражением этого самого недоверия к власти. Более того, в обществе появляется все больше различных неправительственных организаций, ассоциаций, предоставляющих свои цифры. Раньше только государственные статистические службы могли предоставлять цифры. Сейчас, похоже, собирать, обобщать и представлять статистику могут очень многие.... В нашем распоряжении сейчас больше информации, чем когда-либо ранее. И мы просто теряемся, мы запутались в этих данных».

Информационная асимметрия, указал Э. Джованини, является очень важной частью объяснения того, как работают рынки. Это относится и к политическим моделям. Например, ситуация выборов правительства. Предстоит голосовать за какого-то кандидата, однако, чем он занимается, результатов его деятельности людям по существу не видно. Спрашивается, как определить, хорош этот политик или плох? Ведь ни достоверной информацией, ни необходимыми индикаторами оценки эффективности проводимой им политики избиратели не располагают. «Вот тут-то и нужна статистика, призванная объективно измерять то, что происходит в стране». Если граждане, могут наблюдать за индикаторами, фиксирующими эффективность проводимой политики, то и решения будут приниматься более осознанно, а значит, на более высоком уровне будут действовать экономические рынки и демократия».

«У нас много статистики, но качество этих данных не всегда является хорошим», — сказал Э. Джованини, приведя данные некоторых опросов. К примеру, экономистам в разных странах был задан такой вопрос: считаете ли вы, что общественные дебаты в вашей стране основаны на надежной информации? Среди западных экономистов 62% ответили положительно, 20% затруднились с ответом, а 18% ответили отрицательно. В странах СНГ 43% экономистов считают, что статистика, используемая для общественных дискуссий, хороша. В Африке ситуация практически та же самая. Но помимо специалистов в мире живут простые граждане. «Мы спросили граждан в европейских странах о том, знают ли они, как, почему растут ВВП, инфляция, безработица? В Италии, если говорить о росте ВВП, 15% сказали «да», 70% — «мы не знаем», а еще 15% отвечали наугад». Вслед за этим последовал вопрос: а хотите ли вы получать такого рода информацию? И 42% сказали, что нет, не надо, что «мы лучше будем темными, не перегружайте нас цифрами».

В мире есть страны, где население доверяет официальной статистике, но есть государства, где высок уровень недоверия к ней. Так, например, в северных странах Европы — Дании, Исландии, Финляндии, Швеции — доверие к политикам, использующим статистические данные, высоко. Потому что цифры, которыми они оперируют, соответствуют тому, что знает народ. А вот во Франции и Великобритании доверие к статистике находится на низком уровне, лишь чуть более 30% французов и англичан доверяют статистике. «Почему? Причин несколько. Одна из них заключается в том, что статистка показывает не то, что люди могут сопоставить со своей собственной жизнью. Статистика подается так, что возникает некая удаленность этих цифр от того, что люди испытывают и чувствуют, того, во что они верят», — заметил в этой связи профессор.

Кроме того, на статистику оказывают влияние политики и в этом смысле показателен опыт Франции. «Почему французы не доверяют статистике до такой степени? Дело в том, что во время президентской кампании в апреле 2007 года статистическая служба должна была опубликовать данные по безработице. Но в этот период французские статистики изменили методологию сбора данных (никогда, тут же посоветовал студентам в зале профессор, не меняйте методологию во время избирательной кампании, если в дальнейшем будете отвечать за статистику). Цифры были получены, но выяснилось, что с методологией есть некоторые проблемы, и эти данные было решено не публиковать. По закону же цифры должны были направляться в Евростат. Естественно, что многие продумали, будто какие-то данные скрываются по политическим соображениям. Это сразу возбудило острую дискуссию, и масла в огонь, разумеется, подлили средства массовой информации».

В прошлом году, продолжал итальянский ученый, в Стамбуле по инициативе ООН, Европейской комиссии, Организации исламских конференций и Всемирного банка прошел второй мировой форум по статистике, знаниям и политике. Речь на форуме, в

котором приняли участие 1200 человек из 130 стран, шла об измерении прогресса в обществе. По результатам работы стамбульского форума была подготовлена так называемая Стамбульская декларация. В ней, в частности, отмечалось, что **культура принятия решений, основанных на фактической информации, должна улучшаться на всех уровнях, потому что от этого зависит благосостояние обществ.**

«Мы заявили о своей приверженности к измерению прогресса в обществах во всех аспектах, не только через показатель ВВП, но и с учетом, к примеру, уровня охраны окружающей среды. Революционный характер носило и заявление о том, что статистики должны работать не в изоляции, а вместе с общественностью, с частными организациями, с гражданами. Потому, что если принимать статистику как фактор нормального функционирования общества, необходимо вовлекать в эту работу представителей этого самого общества. Людей нужно слушать и слышать, обсуждать с ними проблемы измерения того, что для них важно».

Конечно, продолжал профессор, сегодня есть Интернет, и у каждого свое мнение. И многие полагают, будто разные мнения имеют одинаковый вес. Однако это не так. «Если вы чувствуете себя бедным, это не значит, что все в России тоже влачат жалкое существование. И если вы считаете, что счастливы, то это не значит, что все в России всем довольны. **Статистику изобрели для того, чтобы получать объективную информацию, которую сами граждане получить не могут».**

Статистика, по словам Джованини, подвергается критике во многих странах – говорят, что статистики-де не представляют мнение населения и, мол, существует серьезный конфликт между мнением народа и статистикой. Стамбульская декларация провозглашает, что **«необходимо стимулировать сообщества к обсуждению того, что означает для них прогресс».**

Сейчас ученые пытаются измерить прогресс различными способами. Этим занимается ОЭСР, различные национальные инициативы, действуют Международное общество изучения качества жизни, в Давосе создан Глобальный форум по эталонным показателям прогресса общества. Президент Франции образовал комиссию по измерению экономических результатов прогресса, в работе которой участвует пять нобелевских лауреатов! В апреле 2009 года будет представлен ее первый доклад.

«Мы, статистики, естественно, говорим о том, как измерить этот прогресс. Существует система национального счета, которую используют для измерения инвестиций, ВВП, потребления и т.д. Это серьезная аналитическая система, и ее можно расширить, чтобы охватить социальный и экологический компоненты. Этим занимаются статистики. Но как же трудно выразить проблемы, связанные, скажем, с загрязнением окружающей среды, в финансовых терминах! Ведь ВВП выражается в денежных

единицах, и если «отталкиваться» от ВВП, нужно, соответственно, использовать финансовую терминологию. А как, к примеру, в таком случае измерить потери биологической массы при вырубке лесов?»?

Вторым способом измерения может быть так называемый **композитный показатель, определяемый по ряду индикаторов**. Например, берется ожидаемый срок жизни при рождении, ВВП, образование, каждому из этих показателей придается вес и рассчитывается композитный показатель. **Есть, например, индекс человеческого развития. Существует более 180 показателей такого рода.**

Третий подход— это **использование ключевых показателей**. «Мы признаем, что все в одной цифре выразить нельзя, что реальность многогранна и сложна, и это надлежит осознавать, а посему и рассматривать следует разные показатели. Кто и как может выбрать эти индикаторы? Это процесс политический. **Сейчас идет большой спор между статистиками по поводу объективных и субъективных индикаторов, таких, например, как индекс счастья. Следует ли статистикам «считать» субъективные индикаторы, или нет? Есть движение в сторону субъективных индикаторов, но оно очень медленное и с большим количеством вопросов. Мы поддерживаем развитие ключевых индикаторов. Но расчет индикаторов недостаточен, если эти измерения не вызывают доверия у общества. Только если мы сможем сформировать консенсус в отношении того, каковы те ключевые индикаторы, которые все должны знать,**— только тогда мы сможем сделать шаг вперед».

Сейчас ОЭСР с целым рядом других международных организаций начала реализацию глобального проекта по измерению прогресса общества. «Самая сумасшедшая идея нашего проекта,— заметил профессор,— это «Википедия прогресса». Откуда возникла эта идея? Статистика, как таковая, людей не интересует, их интересуют какие-то конкретные вопросы, а статистику люди хотят использовать для обсуждения этих вопросов». В чем «изюминка» этой Википедии? «В том, что люди могут дополнять имеющиеся сведения своими статьями, которые, если так можно сказать, полностью интегрируются с теми статьями, которые написаны другими. Таким образом возникает некий коллективный разум... надо полагать, **люди будут с достаточно высокой степенью доверия относиться к «Википедии прогресса», ведь они сами смогут пополнять эту «копилку»** ». Во всяком случае, по словам Э. Джованини, в этой базе имеются, к примеру, данные о потреблении пива на пивном празднике в Мюнхене, или график количества поцелуев, которыми обмениваются юноши и девушки в США. «Не знаю,— заметил профессор,— насколько эти цифры надежны, но в Википедии можно найти и такого рода информацию...»

«Включайтесь в статистическую деятельность, привносите в нее что-то новое, пусть сначала и кажущееся странным, даже глупым. Но в ходе этого процесса придет умение отличать полезную информацию от пустой, и статистка станет интересной, потому

что придет ощущение возможности изменить мир», — на такой оптимистической ноте закончил Э. Джованини свою лекцию.

«К сожалению, в 1990-ые годы качество подготовки квалифицированных кадров для отечественной статистики существенно снизилось. Поэтому руководство ГУ-ВШЭ решило открыть в 2008/2009 учебном году отделение статистики, анализа данных и демографии. «Мы надеемся, что студенты, которые закончат это отделение, с успехом пополнят ряды профессионалов-статистиков России и будут востребованы и наукой, и бизнесом». (первый проректор ГУ-ВШЭ Леонид Гохберг).

«Цель — подготовка статистиков-аналитиков. Дело в том, что с каждым годом в нашей стране и мире растет потребность в экономистах, обладающих аналитическим мышлением, основанном на глубоком знании экономики, математики, статистики и компьютерных технологий. Значительная потребность в статистиках-аналитиках имеется на сегодня на микроэкономическом уровне у государственных и частных предприятий, учреждений и фирм.

В XXI веке, по мнению американских экспертов, **системный аналитик входит в четверку самых востребованных специальностей в мире наряду с финансовым менеджером, менеджером вебсайтов и специалистом по информационным технологиям.** Ведущие университеты мира, занимающие первые места в мировых и национальных рейтингах (Гарвард, Стенфорд, Оксфорд, Кембридж и др.), имеют в своей структуре факультеты (кафедры, отделения, школы) статистики.

Уже сегодня многие министерства в России, а также крупные отечественные и зарубежные компании имеют аналитические подразделения, занимающиеся анализом статистических данных для определения направлений и перспектив своего развития.

На сегодняшний день факультета или отделения статистики нет ни в одном вузе Москвы и Санкт-Петербурга. **Это можно объяснить недооценкой роли измерения в экономике и социальной сфере, где значение каждого показателя является итогом деятельности и жизни сотен, тысяч людей и без методов статистики получить адекватную оценку таких показателей невозможно.** К сожалению, у нас в стране часто путают понятия "статистик" и "статист" (человек, который по определению должен молчать). Я видел объявления — "Требуется младший статист. Оклад 24 000 рублей"». (профессор Владимир Мхитарян, декан отделения статистики)).

Индекс счастья

Политики приходят к выводу, что меряться ВВП — это прошлый век экономической науки. Последняя «фишка» — всевозможные индексы счастья, которые учитывают не только материальный достаток, но продолжительность жизни, экологическое состояние страны и характеристики социальной сферы. Год назад о разработке такой методики заявил Китай, а в начале января президент Николя Саркози попросил экс-руководителя ВБ, нобелевского лауреата по экономике Джозефа Стиглица сконструировать индекс счастья для Франции.

На сегодняшний день самый известный индекс счастья подсчитывается английским исследовательским центром New Economics Foundation. На первых местах там жаркие и экологически чистые (промышленности-то в них почти нет) страны вроде Вануату, Колумбии и Доминики, а вот в развитых государствах со счастьем плохо. Китай занимает 31-е место, Индия на 62-м, Франция на 129-м, США — 150-м, а Россия и вовсе на 172-м

месте из 178 стран.

Индекс счастья – впереди 9 стран Латинской Америки.



Данные организации New Economics Foundation (Великобритания) основаны на официальных показателях (например, продолжительность жизни и т.д.) и собственных исследованиях, проведенных по всему миру. Индекс счастья показывает насколько их жители удовлетворены своей жизнью, начиная с максимально «счастливой» страны и заканчивая совсем «несчастливой». Согласно исследованиям 2006 года десять самых

«счастливейших» стран:

Вануату (государство в Меланезии)

Колумбия

Коста-Рика

Доминика

Панама

Куба

Гондурас

Гватемала

Сальвадор

Сент-Винсент и Гренадины

Индекс счастья, Наука и Жизнь №4, 2006.

В 1972 году король Бутана заявил, что благосостояние страны должно измеряться не внутренним валовым продуктом (ВВП), а внутренним валовым счастьем (ВВС) что включает: обеспечение справедливого и социально-экономического развития, сохранение и развитие традиционных культурных ценностей, охрана природы и правильное управление страной.

ВВП или совокупный общественный продукт, не учитывают многие ценности, производимые в стране или, наоборот, теряемые ею - стоимость неоплачиваемой работы добровольцев, стоимость здоровья, которое люди накапливают во время правильно проведенного отдыха, экономические потери, связанные с деградацией окружающей среды. Счастливый, довольный жизнью человек и трудится лучше несчастного, так что неэкономический показатель явно влияет на экономику.

Американские психологи Эд Динер и Мартин Зелигман: основной целью политиков должно быть повышение благосостояния граждан, а измерять успехи в этой области надо по трем показателям: по ВВП, уровню образования и здравоохранения в стране и по такому субъективному показателю, как уровень удовлетворенности жизнью. Как подчеркивают эти специалисты, с 1945 года американский ВВП на душу населения утроился, но социологические опросы показывают, что «уровень счастья» населения остался примерно прежним, скорее даже немного упал. Так же обстоит дело в других странах западного мира. Однако, например, в Дании за последние 30 лет число довольных жизнью людей выросло, и причины этого неясны.

Голландский психолог Руут Венховен, редактор международного «Журнала исследований счастья», вывел обобщенный показатель удовлетворенности жизнью в той или иной стране. Его показатель «Годы счастливой жизни» комбинирует данные о продолжительности жизни и о степени довольства жизнью. Так, в Канаде средняя продолжительность жизни составляет 78,6 года, а средний уровень удовлетворенности жизнью (вполне субъективный показатель, измеряемый при опросах по условной шкале) — 0,763 балла. Венховен перемножает их, получается 60 «счастливых лет».

В 2003 году пр-во Англии рекомендовало при выборе путей реформы в здравоохранении и образовании останавливаться на том варианте, который даст наибольшее увеличение этого показателя.

Неясна связь удовлетворенности жизни с уровнем благосостояния (например, Япония и Южная Корея менее довольны, чем следовало бы по уровню их доходов).

В западных странах с их общепринятым индивидуализмом эти чувства часто рассматриваются как мера личного успеха. Быть несчастным означает, что вы неудачник, вы не сумели правильно распорядиться своей жизнью и теми возможностями, которые предоставляет окружающий мир.

В странах, где больше ценится коллективизм (Китае, Японии, Южной Корее) люди относятся к счастью с большой долей фатализма. Там принято считать, что счастье ниспосылают небеса. Это освобождает людей от чувства неполноценности или вины за то, что они не очень счастливы. Если счастье дают боги, то вы вполне можете быть достойной и замечательной во всех отношениях личностью, вам просто пока не везет.

Больше всех довольны

Насколько счастливы на работе британцы – таким вопросом ежегодно задаются эксперты консалтинговой компании Cuty & Guilds. «Индекс счастья 2008», представленный публике 4 июня (было опрошено 2000 британцев), показал, что больше всех довольны жизнью работники салонов красоты. Из них каждый третий оценил удовлетворенность работой на 10 баллов, каждый второй чувствует себя счастливее, чем в прошлом году,

82% готовы рекомендовать свою компанию в качестве работодателя (см. рейтинг). Исследователей также интересовало, что именно делает людей счастливыми на их работе.

Эксперты пришли к выводу, что деньги – не самое главное. 95% опрошенных согласились с тем, что интерес к работе имеет большое значение, а больше половины назвали это обстоятельство ключевым. Следующими по важности факторами стали финансовая стабильность, хорошие отношения с коллегами, баланс между работой и личной жизнью. 44% респондентов полагают деньги ключевым фактором.

Самые счастливые	
МЕСТО В РЕЙТИНГЕ	ПРОФЕССИИ
1	Работники салонов красоты
2	Парикмахеры
3	Военные
4	Повара
5	Продавцы
6	Учителя
	Маркетологи и пиарщики
	Бухгалтеры
9	Секретари
	Водопроводчики
	Инженеры
	Архитекторы
13	Журналисты
	Автомеханики
	HR
16	Сотрудники колл-центров
17	IT-специалисты
	Медсестры
	Банковские сотрудники
	Строители

ИСТОЧНИК: CUTY & GUILDS

Ссылки на статистические сайты мира

Статистические службы международных организаций

Отдел статистики ООН

[Statistics Division, United Nations](#)

Статистический отдел Европейской экономической комиссии ООН (ЕЭК ООН)

[Statistical Division, United Nations Economic Commission for Europe](#)

Отдел статистики Экономической и социальной комиссии ООН для стран Азии и Тихого Океана (ЭСКАТО ООН)

[Statistics Division, UN Economic and Social Commission for Asia and the Pacific \(ESCAP\)](#)

Бюро статистики Международной организации труда (МОТ)

[Bureau of Statistics, International Labour Office \(ILO\)](#)

Отдел статистики ФАО (Организация ООН по вопросам продовольствия и сельского хозяйства)

[Statistics Division, FAO](#)

Отдел статистики ЮНЕСКО

[Institute for Statistics, UNESCO](#)

Статистический департамент Международного валютного фонда (МВФ)

[Statistics Department, International Monetary Fund](#)

Департамент развития экономики Мирового банка

[Development Economics Department, The World Bank](#)

Статистическое бюро Европейского Союза (Евростат)

[Statistical Office of the European Communities \(Eurostat\)](#)

Статистический директорат организации экономического сотрудничества и развития (ОЭСР)

[Statistics Directorate, Organisation for Economic Co-operation and Development \(OECD\)](#)

Постоянное бюро Международного статистического института (МСИ)

[Permanent office, International Statistical Institute \(ISI\)](#)

Консорциум "Партнерство в области статистики для развития в 21 столетии" (PARIS21)

[Partnership In Statistics for development in the 21st century](#)

Национальные статистические службы СНГ

ГОСУДАРСТВЕННЫЙ КОМИТЕТ ПО СТАТИСТИКЕ

АЗЕРБАЙДЖАНСКОЙ РЕСПУБЛИКИ

370136, г. Баку-136, проспект Иншаатчилар

Тел: (994-12) 38-11-71

Факс: (994-12) 438-64-98/ 438-24-42

Е-mail: gks@baku-az.net, ssc@mednet.az

<http://www.azeri.com/goskomstat>

НАЦИОНАЛЬНАЯ СТАТИСТИЧЕСКАЯ СЛУЖБА

РЕСПУБЛИКИ АРМЕНИЯ

375010, г. Ереван, Площадь Республики, Дом Правительства,3

Тел: (374-1) 52-42-13

Факс: (374-1) 52-19-21

Е-mail: armstat@sci.am

<http://www.armstat.am>

МИНИСТЕРСТВО СТАТИСТИКИ И АНАЛИЗА

РЕСПУБЛИКИ БЕЛАРУСЬ

220658, г. Минск, Партизанский проспект, 12

Тел: (375-17-2) 49-52-00

Факс: (375-17-2) 49-22-04

Е-mail: minstat@mail.belpak.by

<http://www.belstat.gov.by>

ГОСУДАРСТВЕННЫЙ ДЕПАРТАМЕНТ ПО СТАТИСТИКЕ

МИНИСТЕРСТВА ЭКОНОМИЧЕСКОГО РАЗВИТИЯ ГРУЗИИ

380015, г. Тбилиси, проспект К. Гамсахурдия, 4
Тел: (995-32) 33-14-50
Факс: (995-32) 33-14-93
E-mail: info@statistics.gov.ge
<http://www.statistics.ge>

**АГЕНТСТВО РЕСПУБЛИКИ КАЗАХСТАН
ПО СТАТИСТИКЕ**

480008, г. Алматы 8, проспект Абая, 125
Тел: (732-72) 42-13-89, 61-13-23
Факс: (732-72) 42-08-24
E-mail: stat@mail.online.kz
<http://www.stat.kz>

**НАЦИОНАЛЬНЫЙ СТАТИСТИЧЕСКИЙ КОМИТЕТ
КЫРГЫЗСКОЙ РЕСПУБЛИКИ**

720033, г. Бишкек, ул. Фрунзе, 374
Тел: (996-312) 66-40-38
Факс: (996-312) 66-01-38
E-mail: nsc_mail@nsc.bishkek.su
<http://www.stat-gvc.bishkek.su>

**ДЕПАРТАМЕНТ СТАТИСТИКИ И СОЦИОЛОГИИ
РЕСПУБЛИКИ МОЛДОВА**

2028, г. Кишинев, Хынчешское шоссе, 53
Тел: (373-2) 73-37-74
Факс: (373-2) 22-61-46
E-mail: dass@moldova.md
<http://www.statistica.md>

**ФЕДЕРАЛЬНАЯ СЛУЖБА ГОСУДАРСТВЕННОЙ СТАТИСТИКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

107450, г. Москва, ул. Мясницкая, 39, строение 1
Тел: (7-095) 207-49-02
Факс: (7-095) 207-46-40
E-mail: gks@gks.ru
<http://www.gks.ru/>

**ГОСУДАРСТВЕННЫЙ КОМИТЕТ СТАТИСТИКИ
РЕСПУБЛИКИ ТАДЖИКИСТАН**

734025, г. Душанбе, ул. Бохтар, 17
Тел: (992-372) 23-25-51
Факс: (992-372) 21-43-75
E-mail: stat@tojikiston.com
tad@stat.tajik.net

**НАЦИОНАЛЬНЫЙ ИНСТИТУТ
ГОСУДАРСТВЕННОЙ СТАТИСТИКИ И ИНФОРМАЦИИ
ТУРКМЕНИСТАНА**

744000, г. Ашгабат, проспект Магтымгулы, 72
Тел: (993-12) 39-42-65 / 35-35-96
Факс: (993-12) 35-22-79
E-mail: office@natstat.gov.tm
staff@natstat.gov.tm

Государственный комитет Республики Узбекистан по статистике
700077, г. Ташкент-77, ул. Буюк Ипак Юли, 63
Тел: (998-71-2) 67-25-51
Факс: (998-71-2) 67-78-16 / 67-25-09
E-mail: gds@uzstat.org.uz
<http://www.stat.uz>

**ГОСУДАРСТВЕННЫЙ КОМИТЕТ СТАТИСТИКИ
УКРАИНЫ**

01023, г. Киев, ул. Шота Руставели, д.3
Тел: (380-44) 226-20-21 / 287-24-33
Факс: (380-44) 287-66-11 / 235-37-39
E-mail: minstat@minstat.kiev.ua
petr@minstat.kiev.ua
[http:// www.ukrstat.gov.ua](http://www.ukrstat.gov.ua)

Российская государственная статистика

Госкомстат РФ www.gks.ru
Комитет государственной статистики Республики Башкортостан www.bashstat.ru
Комитет государственной статистики Республики Бурятия www.stat.burnet.ru
Комитет государственной статистики Республики Карелия www.goskomstat.karelia.ru
Комитет государственной статистики Республики Коми www.komistat.ru
Комитет государственной статистики Республики Марий Эл www.gkstat.mari.ru
Комитет государственной статистики Республики Мордовия www.komstat.moris.ru
Комитет государственной статистики Республики Саха (Якутия) www.stat.sakha.ru
Комитет государственной статистики Республики Татарстан www.tatstat.ru
Комитет государственной статистики Удмуртской Республики www.gksur.udmnet.ru
Комитет государственной статистики Чувашской Республики www.cap.ru

Региональные органы статистики

Алтайский край www.amitel.ru/~komstat
Амурская область www.amur.ru/~oblstat
Астраханский облкомстат www.oblstat.astranet.ru/
Архангельская область www.dvinaland.ru/
Башкортостан www.bashstat.ru/
Белгородская область www.oblstat.belnet.ru/
Брянская область www.komstat.bryansk.ru
Бурятия www.stat.burnet.ru/
Волгоградская область www.avtlg.ru/~vcomstat
Вологодский облкомстат www.stat.volcom.ru/
Воронежская область www.web.vrn.ru/obstat
Владимирская область www.vtsnet.ru/vlgks
Комитет государственной статистики Еврейской автономной области www.stat.jar.ru
Иркутская область www.east-siberia.ru/statistic
Калининградский облкомстат www.baltnet.ru/~kokgstat/
Калужский область www.users.kaluga.ru/oblstat
Карелия www.sampo.karelia.ru/~stat/
Кемеровская область www.rosemis.ru/gks/
Кировский облкомстат www.kks.kirov.ru
Коми www.komistat.ru/
Краснодарский край www.krsndstat.boom.ru/
Красноярский край www.krasnet.ru/~statis
Курганская область www.kurganstat.ru
Магаданская область www.sferacom.ru/~nauka
Марий Эл www.mari.ru:8101/gks/
Мосгоркомстат www.mosstat.ru/
Московская область www.mosobl.ru/oblstat/
Мурманская область www.murmanweb.ru/stat/
Нижегородская область www.nizstat.sinn.ru
Мордовский облкомстат www.komstat.moris.ru
Новгородский облкомстат www.novgorod.net/~stat/
Новосибирский облкомстат www.sibstat.gcom.ru/
Орловская область www.aha.ru/~statarel/
Оренбургская область www.esoo.ru/~comstat
Пермский облкомстат www.oblstat.perm.ru/
Пензенский облкомстат www.sura.com.ru/~statt
Приморский край www.personal.primorye.ru/vlstat/
Псковская область www.ellink.ru/clients/oblstat/

Ростовская область www.rdstat.aanet.ru/
Самарская область www.samtel.ru/~samrstat/
г. Санкт-Петербург www.gostat.spb.ru
Сахалинская область www.ns.adm.sakhalin.ru/oblstat/
Свердловская область www.ersds.e-burg.ru/
Ставропольский край www.stavropol.ru/stv/kraykomstat
Тамбовская область www.statkom.tambov.ru
Тверская область www.tvcom.ru/usp/tvstat/
Томская область www.oblstat.tomsk.gov.ru/
Тульский облкомстат www.tula.ru/stat
Тюменский облкомстат www.oblstat.tmn.ru/
Ульяновская область www.statcom.mv.ru
Хабаровский край www.kht.ru/homepage/~statistica/
Челябинская область www.chelstat.ru
Читинская область www.chita.ru/public_html/stat

Зарубежная государственная статистика

Австралия www.abs.gov.au/
Австрия www.oestat.gv.at/
Азербайджан www.azeri.com/goscomstat/
Аргентина www.indec.mecon.ar/
Аруба www.indec.mecon.ar/
Бразилия www.ibge.gov.br/
Великобритания www.ons.gov.uk/
Венгрия www.ksh.hu/
Венесуэла www.ocei.gov.ve/
Германия www.statistik-bund.de/
Греция www.statistics.gr/
Дания www.dst.dk/
Индонезия www.bps.go.id/
Ирландия www.cso.ie/
Испания www.ine.es/
Исландия www.statice.is/
Казахстан www.kazstat.asdc.kz/
Канада www.statcan.ca/
Канада www.canada4you.ru/
Кыргызстан www.stat-gvc.bishkek.su/
Латвия www.csb.lv/
Литва www.std.lt/
Мозамбик www.ine.gov.mz/
Монголия www.statis.pmis.gov.mn/
Нидерланды www.cbs.nl/
Новая Зеландия www.stats.govt.nz/statsweb.nsf
Норвегия www.ssb.no/
Палестина www.pcbs.org/
Португалия www.ine.pt/
Сингапур www.singstat.gov.sg/
Словакия www.statistics.sk/
Словения www.sigov.si/zrs/
США http://www.first.gov/Topics/Reference_Shelf/Data.shtml
Узбекистан www.gov.uz/mms100fr.html
Финляндия www.stat.fi/
Франция www.insee.fr/
Хорватия www.dzs.hr/
Швеция www.scb.se/
Эквадор www.4.inec.gov.ec
Эстония www.stat.ee/
Югославия www.szs.sv.gov.yu/
Южная Африка www.statssa.gov.za/
Япония www.stat.go.jp/

Впервые изучением связей между вероятностью события и качеством исходной информации занялся второй из старших Бернулли — Якоб (1654-1705), дядя известного Даниила Бернулли. Он был еще ребенком, когда Паскаль и Ферма высказали свои замечательные математические идеи, и умер, когда его племяннику Даниилу едва исполнилось пять лет. Талантливый, как все Бернулли, он был современником Исаака Ньютона и, обладая свойственным всем Бернулли сложным и самолюбивым характером, считал себя соперником великого английского ученого. Сама по себе постановка Якобом обсуждаемого вопроса, даже если отвлечься от предложенных им ответов, была научным подвигом. По его признанию, он размышлял над этой проблемой двадцать лет и окончил посвященный ей труд незадолго до смерти, последовавшей в 1705 году.

Якоб был самым мрачным из Бернулли, особенно к концу жизни, несмотря на то что он жил в веселые и легкомысленные времена, наступившие в Англии после реставрации монархии в 1660 году и восшествия на престол Карла II. (Ему была свойственна своеобразная поэтичность, сказавшаяся, к примеру, в пожелании, чтобы на его могильном камне высекали прекрасную спираль Фибоначчи, поскольку ее свойство расширяться, не изменяя формы, является «символом стойкости и неизменности посреди хаоса и напастей, а в конечном итоге — даже нашего воскрешения во плоти»). Под спиралью он потребовал выбить эпитафию: «*Eadem Mutata rurgo*» («Неизменная в вечном движении»).

Де Муавр родился в 1667 году — через 13 лет после Якоба Бернулли — в протестантской семье во Франции, в обстановке возрастающей враждебности ко всем не католикам. В 1685 году, когда ему было 18 лет, король Людовик XIV отменил Нантский эдикт, провозглашенный в 1598 году родившимся в протестантской вере королем Генрихом IV и предоставивший протестантам, называемым гугенотами, равные политические права с католиками. После отмены эдикта исповедование реформатской религии было запрещено, дети гугенотов должны были воспитываться в католической вере, эмиграцию запретили. Де Муавр свыше двух лет провел в тюрьме за свои религиозные убеждения. Ненавидя Францию и все с нею связанное, он в 1688 году бежал в Лондон, где Славная революция как раз покончила с остатками государственного католицизма. На родину он так и не вернулся.

В Англии де Муавр вел печальную и неустроенную жизнь. Несмотря на все усилия, ему не удалось добиться приличной академической должности. Он зарабатывал на жизнь уроками математики и консультациями по применению теории вероятностей для игроков и страховых брокеров. С этой целью он держал неофициальную приемную в кофейне Слайтера, что на улице Святого Мартина, где большей частью и проводил остаток дня по окончании занятий с учениками. Хотя он был другом Ньютона и стал членом Королевского общества уже в тридцать лет, он так и остался едким, ушедшим в себя, асоциальным человеком. Умер он в 1754 году в бедности и слепоте в возрасте 87-ми лет.

Первая прямо посвященная теории вероятностей работа де Муавра озаглавлена «*De Mensura Sortis*» (буквально «Об измерении случайных величин»). Работа была впервые опубликована в 1711 году в журнале Королевского общества «*Philosophical Transactions*». В 1718 году де Муавр предпринял значительно расширенное издание этой работы на английском языке, озаглавленное «Теория случайностей» («*The Doctrine of Chances*»), с посвящением своему близкому другу Исааку Ньютону. Книга имела огромный успех и выдержала еще два издания в 1738-м и 1756 годах.

Де Муавр так подытожил свои исследования: *«Случай порождает Отклонения от закономерности, однако бесконечно велики Шансы, что с течением Времени эти Отклонения окажутся пренебрежимо ничтожными относительно повторяемости того Порядка, который естественным образом является результатом БОЖЕСТВЕННОГО ПРЕДНАЧЕРТАНИЯ».*

Вкладом де Муавра в математику был инструмент, который сделал возможной оценку вероятности того, что заданное число наблюдений попадет в некоторую область вокруг истинного отношения. Этот результат нашел широкое практическое применение.

Например, все производители опасаются того, что результатом сборки может оказаться бракованная продукция, которая дойдет до потребителей. Стопроцентное качество в большинстве случаев практически невозможно — наш мир, похоже, непоправимо враждебен совершенству.

Представьте себе директора булавочной фабрики, который старается добиться, чтобы бракованные булавки встречались не чаще, чем в 10 случаях из 100000, то есть чтобы брак составлял не более 0,01% от объема производства. Для контроля дел он проводит обследование произвольной выборки из 100 000 сошедших с конвейера булавок и выясняет, что у 12 нет головок — на 2 больше, чем он надеялся получить в среднем по всей производимой продукции. Насколько значима эта разница? Какова вероятность найти 12 бракованных булавок из выборки объемом в 100000, если *средний* процент брака составляет 10 бракованных булавок на каждый 1 000 000? Нормальное распределение и среднее квадратичное отклонение де Муавра дают ответ на этот вопрос.

Но обычно вопрос ставится по-иному. Чаще никто точно не знает, сколько именно бракованных изделий *в среднем* выпускает фабрика. Вопреки благим намерениям действительная доля брака может оказаться в среднем выше, чем 10 из 100000. Что скажет выборка из 100000 булавок о вероятности того, что для всей выпускаемой продукции брак в среднем составляет 0,01%? Насколько более точные сведения можно получить из выборки объемом в 200 000 булавок? Какова вероятность того, что процент брака окажется в пределах от 0,009% до 0,011%? А в пределах от 0,007% до 0,013%? Какова вероятность того, что одна наугад взятая булавка окажется бракованной?

Здесь исходными данными являются 10 булавок, 12 булавок, 1 булавка, а вероятность оказывается искомой величиной. В такой постановке задача сводится к вычислению так называемой *обратной вероятности*: какова вероятность того, что по всей произведенной продукции брак составляет в среднем 0,01%, если в выборке из 100000 булавок оказалось 12 бракованных?

Одно из наиболее эффективных решений этой задачи было предложено пастором Томасом Байесом, который родился в 1701 году и жил в Кенте.

Пример практического применения Байесова подхода к статистическим задачам

Обратимся вновь к булавочной компании. Компания имеет две фабрики, причем старая выпускает 40% продукции. Это означает, что взятая наугад булавка, бракованная или нет, с вероятностью 40% выпущена на старой фабрике; это исходная вероятность. Известно, что на старой фабрике процент брака вдвое больше, чем на новой. Если клиент звонит и сообщает о купленной им бракованной булавке, на какую из двух фабрик должен звонить менеджер по сбыту?

Исходная вероятность побуждает утверждать, что, скорее всего, бракованная булавка сделана на новой фабрике, выпускающей 60% продукции компании. С другой стороны, частота появления брака на этой фабрике вдвое меньше, чем на старой. Пересмотрев исходную вероятность с учетом этой дополнительной информации, получаем, что вероятность выпуска бракованной булавки новой фабрикой равна только 42,8%; это значит, что с вероятностью 57,2% виновата старая фабрика. Эта новая оценка становится апостериорной вероятностью.

Пример смешения разнородных совокупностей

	Городские жители		Сельские жители	
	Лечившиеся	Не лечившиеся	Лечившиеся	Не лечившиеся
Выздоровевшие	1 000	50	95	5 000
Не выздоровевшие	9 000	950	5	5 000

A = (выздоровевшие), B = (лечившиеся), C = (городские жители)

	B	B^c	B	B^c
A	1 095	5 050	95	5 000
A^c	9 005	5 950	5	5 000

$$P(A/B) = 1095/10100 \approx 10\%, \quad P(A/B^c) = 5050/11000 \approx 45.9\%.$$

$$P(A/BC) = 1000/10000 = 10\%, \quad P(A/B^cC) = 50/1000 = 5\%, \\ P(A/BC^c) = 95/100 = 95\%, \quad P(A/B^cC^c) = 5000/10000 = 50\%.$$

$$P(A/B) = P(AB)/P(B) = (P(ABC) + P(ABC^c))/P(B) = \\ = P(ABC)/P(BC) \times P(BC)/P(B) + P(ABC^c)/P(BC^c) \times P(BC^c)/P(B) = \\ = P(A/BC) P(C/B) + P(A/BC^c) P(C^c/B) = \\ = 1000/10000 \times 10000/10100 + 95/100 \times 100/10100.$$

$$P(A/B^c) = P(A/B^cC) P(C/B^c) + P(A/B^cC^c) P(C^c/B^c) = \\ = 50/1000 \times 1000/11000 + 5000/10000 \times 10000/11000.$$