

Даниэль Канеман. Шум. Несовершенство человеческих суждений

Два одинаково уважаемых врача могут поставить пациенту совершенно разные диагнозы. Два одинаково честных судьи – вынести абсолютно разные вердикты по одному делу. Два одинаково опытных специалиста по подбору персонала – выбрать на одну и ту же должность разных соискателей... Почему это происходит? От чего зависит? Могут ли на такие важные решения влиять время суток или день недели?

Даниэль Канеман вместе с Оливье Сибони и Кассом Р. Санстейном раскроют секреты шума – посторонних влияний на наши суждения – во многих областях: от медицины до криминалистики, от экономического прогнозирования до юриспруденции, и, что еще важнее, научат, как его уменьшить, а значит, начать находить лучшие решения.

Даниэль Канеман, Касс Р. Санстейн, Оливье Сибони. Шум. Несовершенство человеческих суждений. – М.: АСТ, 2021. – 544 с.



Купить цифровую книгу в [ЛитРес](#), бумажную книгу в [Ozon](#) или [Лабиринте](#)

Введение. Два вида ошибок

Представьте, что в стрелковый тир пришли четыре компании друзей. В компаниях по пять человек; они получают по винтовке на команду, и каждый участник производит один выстрел. В идеале все выстрелы попали бы точно в «десятку».

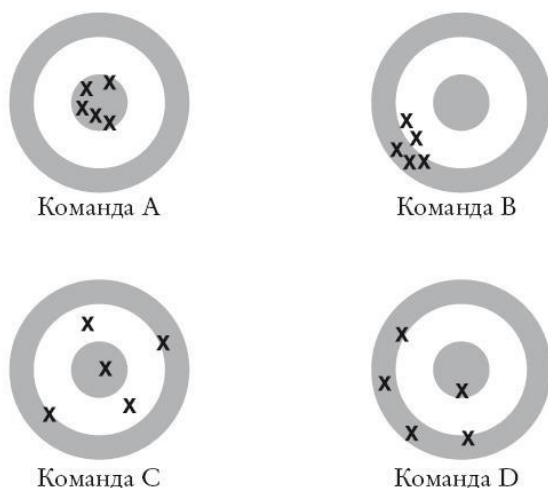


Рис. 1. Четыре команды

У команды А почти все попадания сгруппировались возле цели. Результаты команды В можно назвать *смещенными*. Результаты команды С мы назовем *шумными*. У команды D результаты смещенные и шумные одновременно. Стрельба по мишени – образное представление сбоев в процессе вынесения людьми суждений.

Представьте, что вы заглядываете на обратную сторону мишеней, по которым стреляли наши команды, где «яблочко» не обозначено (рис. 2). С этой стороны вы не определите, кто оказался ближе к «десятке»: команда А или команда В. Зато сразу же понятно, что шумный результат именно у команд С и D. Шум можно обнаружить и измерить, даже ничего не зная ни о цели, ни о смещении, – это одно из его основных свойств.



Рис. 2. Обратная сторона мишеней

Ко многим заключениям мы пришли, изучая суждения, о справедливости которых мы не имеем ни малейшего понятия. Понять природу ошибочных суждений невозможно без понимания сущности смещения и шума. Основную проблему представляет собой именно шум. Однако ему редко уделяется должное внимание. Главную роль всегда отводят смещению. О смещении написаны тысячи научных работ и десятки бестселлеров, а шум при этом едва упоминается. С помощью этой книги мы попытаемся установить равновесие.

Часть I. Как обнаружить шум

Глава 1. Преступление и шумное наказание

Во всем мире судьи имеют возможность свободно решать, насколько суровым должен быть приговор. Эксперты многих стран приветствуют эту свободу действий как справедливую и гуманную. Считается, что при вынесении уголовных приговоров нужно учитывать множество факторов, среди которых не только само преступление, но и личность обвиняемого, а также сопутствующие обстоятельства. Многим казалось, что сама идея надлежащего судебного разбирательства требует неограниченной свободы усмотрения для судей.

В 1973 году внимание к проблеме привлек судья Марвин Франкел. Максимальный тюремный срок за ограбление федерального банка равнялся 25 годам. На практике приговор мог быть любым, в диапазоне от 0 до 25 лет. Приговор одному и тому же обвиняемому по одному и тому же делу мог кардинально отличаться в зависимости от того, какой судья рассматривал дело. Франкел призвал конгресс покончить с «дискриминацией» – так он называл подобный произвол. Под «дискриминацией» в основном имелся в виду шум – необъяснимые расхождения при вынесении приговоров.

Чтобы подвергнуть идею шума испытанию, несколько исследователей сразу же бросились изучать уровень шума в сфере вынесения приговоров. Пятидесяти судьям из разных округов предложили гипотетические дела, по которым требовалось назначить наказание. По всем делам были подготовлены идентичные характеристики подсудимых. Основным итогом стало «повсеместное отсутствие консенсуса» и «поразительный» разброс в выборе наказания. В зависимости от судьи наркотореклер мог получить от года до десяти лет тюрьмы. Наказание за ограбление банка могло составить от пяти до восемнадцати лет лишения свободы.

Многое свидетельствует о том, что несущественная информация – небольшие и, казалось бы, случайные факторы – способны значительно изменить результат рассмотрения дела. Например, как выяснилось, судьи с большей вероятностью предоставляют условно досрочное освобождение

с утра или после обеденного перерыва. Голодные судьи более суровы. Исследование тысяч решений суда по делам несовершеннолетних выявило, что, когда в выходные местная футбольная команда проигрывает, в понедельник судьи выносят более строгие приговоры.

В 1984 году, приняв во внимание доказательства необоснованных расхождений при вынесении приговоров, конгресс США принял Закон о реформе системы назначения наказаний. В следующем году, после анализа 10 000 дел, комиссия разработала рекомендации. При назначении наказания судье следует учитывать два фактора: само преступление и уголовное прошлое подсудимого. На основе сложения этих двух факторов рекомендации предлагают относительно узкий диапазон наказаний. Верхний предел этого диапазона может превышать нижний лишь на шесть месяцев или на 25%. Отклониться от этого диапазона судьи могут при наличииотячающих или смягчающих обстоятельств, обосновав свое решение в апелляционном суде.

Несколько исследований, проведенных с применением различных методов и охватывавших различные исторические периоды, пришли к одному и тому же заключению: рекомендации снижают уровень шума.

Обязательные рекомендации снижают не только шум, но и смещение, вызванное предвзятостью. После того как Верховный суд сделал рекомендации необязательными, значительно увеличилось расхождение в приговорах афроамериканцам и белым правонарушителям, осужденным за схожие преступления. В то же время судьи женщины стали чаще, чем судьи мужчины, действовать на свое усмотрение и проявлять больше снисхождения. То же самое можно сказать и о судьях, назначенных президентами демократами.

История борьбы судьи Франкела за введение рекомендаций по вынесению приговоров дает представление о важнейших моментах, которые мы рассмотрим в этой книге. Во-первых, выносить суждения нелегко, потому что реальный мир сложен и непредсказуем. Во-вторых, масштабы этих разногласий гораздо шире, чем мы ожидаем. Мало кто возражает против принципа судебного усмотрения, но при этом почти все порицают возникающие в ходе его применения огромные расхождения. В-третьих, уровень шума можно снизить. Соблюдение правил и рекомендаций – это один из методов, успешно сокращающих шум. Другие приемы лучше приспособлены для иных видов суждений. При этом некоторые методы снижения уровня шума могут одновременно уменьшать и смещение. В-четвертых, усилия, направленные на сокращение шума, часто встречают сопротивление и наталкиваются на серьезные препятствия.

Глава 2. Шумная система

В крупных компаниях многие специалисты наделены правом выносить суждения, которые связывают эти компании обязательствами. Например, в страховых организациях работает множество андеррайтеров, назначающих размеры взносов за страхование. Там также задействовано много оценщиков страховых убытков, которые прогнозируют расходы в связи с будущими претензиями и ведут переговоры с заявителями в случае возникновения споров.

Размер ставки премии имеет для организации большое значение. Компании выгодно, когда ставка высокая и клиент с ней соглашается, однако завышенная котировка увеличивает риск лишиться сделки. Если же ставка низкая, клиент принимает ее с большей готовностью, но тогда уже компания упускает выгоду. Для любого риска существует размер ставки, отражающий золотую середину. При этом высоки шансы, что в среднем оценки большой группы специалистов не будут сильно от нее отклоняться. Ставки выше или ниже этой величины окажутся убыточными – именно таким образом разброс шумных оценок способен отрицательно повлиять на финансовые показатели компании.

В процессе ревизии одно и то же дело оценивается большим количеством специалистов, и разброс в их заключениях становится очевидным. Это особенно хорошо работает в случае с андеррайтерами и страховыми оценщиками, чьи решения основываются на информации, представленной в письменном виде. Для проведения ревизии шума руководство компании подготовило детальные описания пяти типичных кейсов для каждой из групп специалистов (андеррайтеров и оценщиков). Каждый специалист должен был самостоятельно оценить два-три кейса; при этом испытуемые не знали, что целью исследования была оценка разброса в их суждениях.

До исследования мы задали вопрос руководителям, какова будет разница между двумя оценками в процентах от их среднего показателя? Медианной оценкой был разброс в 10%. Результаты исследования выявили медианную разницу у андеррайтеров в 55%, а у оценщиков – 43%. Т.е., сумма, которую клиент заплатит за страховку, довольно сильно зависит от результатов лотереи выбора специалиста, отвечающего за сделку. Как правило, имея дело с организациями, люди ожидают от системы последовательных суждений, но никак не системного шума.

Одна из определяющих черт системного шума – его *нежелательность*.

Подчеркнуть, что разброс в суждениях нежелателен отнюдь не всегда. Вариативность мнений ожидаема и желательна в условиях конкуренции – когда поощряются наилучшие решения. Когда трейдеры по-разному оценивают стоимость акций, одни на этом заработают, а другие нет. Рынок строится на разности мнений. Однако если случайно выбранному трейдеру поручат подобную оценку от имени компании, и мы обнаружим, что оценки его коллег из той же компании будут кардинально отличаться, тогда налицо проблема системного шума.

Нежелательному разбросу часто не придают значения, полагая, что случайные ошибки взаимно уничтожаются. Однако в шумных системах не принимается множество решений по одному и тому же вопросу – в них принимаются шумные решения по разным вопросам. Если двух преступников, которым полагается по пять лет тюрьмы, приговаривают к трем и семи годам соответственно, справедливость в целом не торжествует. В шумных системах ошибки не компенсируют друг друга – они накапливаются.

Как мнения специалистов, выполняющих одинаковые функции в компании, могли так сильно отличаться и как это могло остаться незамеченным? Ревизия позволила предположить, что авторитетные специалисты и их компании работодатели поддерживали лишь иллюзию согласия и при этом фактически изо дня в день расходились в профессиональных суждениях. Там, где есть место суждению, найдется и шум – и его намного больше, чем вы думаете.

Глава 3. Уникальные решения

Долгое время уникальные решения рассматривались отдельно от типовых – тех, что принимаются взаимозаменяемыми специалистами в крупных организациях. Типовые решения – предмет изучения социологов, тогда как ответственные уникальные решения остаются в ведении историков и гуру менеджмента. Типовые решения рассматриваются в статистическом ключе. Подход к уникальным решениям обычно носит причинно-следственный характер: они обсуждаются в ретроспективе, и основной объект внимания – это причины произошедшего.

Мы не можем измерить уровень шума в уникальных решениях, но мы наверняка знаем о его присутствии. Вспомним, как разные страны отреагировали на пандемию COVID-19. Хотя она началась для всех примерно в одно и то же время и развивалась похожим образом, ответные меры существенно отличались от региона к региону. Такая вариативность четко свидетельствует о шуме в решениях, принимаемых разными государствами. Но что, если бы эпидемия затронула только одну страну? Тогда бы мы просто не смогли заметить никакого разброса. Однако от того, что разброс незаметен, принятое решение не становится менее шумным.

Руководствуясь соображениями уменьшения шума, мы должны относиться к уникальным решениям как к типовым решениям, принимаемым лишь единожды. Выносите ли вы суждение один или сотню раз, ваша цель – снизить при этом как масштаб смещения, так и уровень шума.

Часть II. Ваш разум – измерительный прибор

Глава 4. Субъективные суждения

Обычно люди убеждены, что при вынесении проверяемых суждений стремятся к совпадению прогноза с реальным результатом. На деле же, вне зависимости от степени проверяемости прогноза, они ждут внутреннего сигнала о готовности суждения, а он, в свою очередь, возникает, когда между условиями задачи и сделанным выводом не остается противоречий. Однако гораздо правильнее стремиться к тому, чтобы в процессе вынесения суждений было найдено оптимальное решение для целой совокупности похожих случаев. Люди, которых коснулись последствия оценочных суждений, полагают, что такие суждения отражают решение, принятое системой, а не мнение отдельно взятого судьи. Системный шум – это непоследовательность, а непоследовательность подрывает авторитет системы.

Глава 5. Погрешность измерения

Гаусс предложил правило для оценки вклада индивидуальных ошибок в общую погрешность. Его мера общей погрешности, называемая среднеквадратической ошибкой (MSE), – это среднее значение квадратов индивидуальных погрешностей измерения.

Роль смещения и шума в возникновении погрешностей легко обобщить двумя выражениями, которые мы назовем уравнениями расчета погрешности. Первое из этих уравнений раскладывает погрешность однократного измерения на две составляющие: смещение, или среднюю погрешность, и остаточную «шумную погрешность». Шумная погрешность имеет положительное значение, если погрешность больше, чем смещение, и отрицательное, если меньше. Среднее значение шумных погрешностей равняется нулю.

Погрешность в однократном измерении = Смещение + Шумная погрешность

Второе уравнение расчета погрешности – это разложение на составные части среднеквадратической ошибки, которая может быть представлена как сумма квадратов смещения и шума.

Общая погрешность (MSE) = Смещение² + Шум²

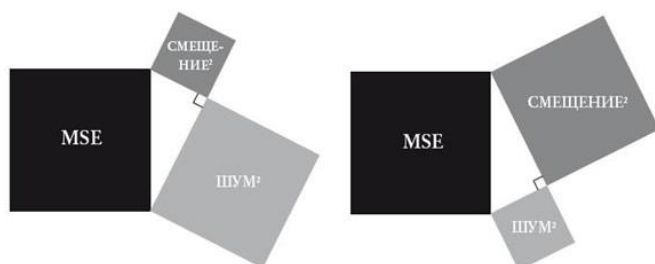


Рис. 3. Два разложения MSE на составляющие

В уравнении расчета погрешности смещение и шум взаимозаменяемы, поэтому независимо от того, какой из этих двух показателей мы уменьшим, снижение общей погрешности будет одинаковым.

Глава 6. Анализ шума

Любой адвокат подтвердит, что у каждого судьи своя репутация: есть судьи «кровожадные», чьи приговоры всегда строже среднего, а есть «сердобольные», чьи приговоры, как правило, мягче. Мы будем называть подобные отклонения *межэкспертным шумом*.

Но существует и *внутриэкспертный шум*. Он отражает сложный внутренний характер отношения конкретного судьи к каждому делу. Один судья, например, может быть в среднем более строгим, но проявлять некоторую снисходительность к обвиняемым, совершившим ненасилиственные преступления в экономической сфере. Другой не отличается особенной суровостью, но менее терпим к рецидивистам. Строгость третьего судьи может быть ближе к средним значениям, при этом он выказывает больше сочувствия к соучастникам преступления, но выносит более жесткие приговоры, если жертва преступления – пожилой человек.

Разложение системного шума на межэкспертный и внутриэкспертный компоненты описывается уравнением:

Системный шум² = Межэкспертный шум² + Внутриэкспертный шум²

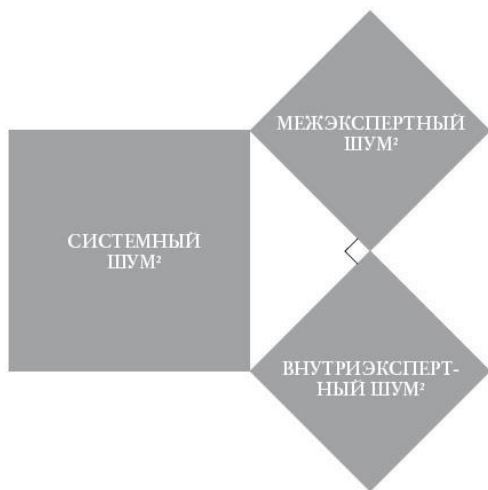


Рис. 4. Разложение системного шума на составляющие

Глава 7. Ситуативный шум

Никого не удивляет разброс в результатах штрафных бросков одного и того же баскетболиста или в показателях других физических процессов. Несколько сложнее наблюдать вариативность наших умственных процессов. Довод, который мы находили неудачным и расплывчатым, вдруг становится ясным и принципиально важным. Опытные консультанты по программному обеспечению могут предложить весьма отличные друг от друга оценки сроков выполнения одного и того же задания, если их спросить об этом дважды в разные дни.

Мы уже сравнивали процесс назначения андеррайтера, судьи или врача с лотереей, создающей системный шум. Ситуативный шум – продукт второй лотереи. Результаты этой лотереи зависят от конкретного момента вынесения суждения экспертом, от его расположения духа, от набора похожих случаев, которые еще свежи в его памяти, а также от бесчисленных других особенностей ситуации. Нам известно лишь то, что окончательное суждение было выбрано из множества возможных. Ситуативный шум – это вариативность среди этих неявных возможностей.

Исследователи задавали испытуемым один и тот же вопрос дважды, вуалируя этот факт. Ситуативный шум уменьшается, если усреднить два ответа. Другие исследователи акцентировали внимание на том, что задавали вопрос повторно, но просили дать ответ, который будет максимально отличаться от первого, не теряя при этом правдоподобия. Они также усреднили значения двух полученных ответов. Исследователи назвали этот метод *диалектическим бутстрэппингом*. Результаты его применения оказались гораздо точнее, чем результаты первого эксперимента. Поскольку испытуемым пришлось посмотреть на задачу в новом свете, они как бы задали тот же вопрос другой версии себя – таким образом, над задачей думали два человека из «внутренней толпы», чьи мнения несколько отличались.

Если есть возможность получить независимую оценку других экспертов, воспользуйтесь ею: настоящая «мудрость толпы» имеет высокие шансы улучшить результат. Если вы не можете этого сделать, подумайте над вторым ответом сами, создавая эффект «внутренней толпы». Такие исследования подтверждают важную идею о процессе вынесения суждений: «Испытуемый скорее выбирает ответ из распределения вероятностей, сложившегося у него в голове, чем отвечает, основываясь на совокупности своих знаний».

Один источник ситуативного шума знаком каждому: это настроение. Эксперименты с эффектом настроения подчеркивают важную истину: *вы не всегда являетесь одной и той же личностью*. Ситуативный шум при вынесении суждений возникает также под воздействием стресса и усталости. Мы считаем, что масштаб ситуативного шума ниже, чем масштабы постоянных межэкспертных и внутриэкспертных различий в принятии решений.

Глава 8. Как группы усиливают шум

Возникновение шума в суждениях одного человека – серьезная проблема. Однако, когда суждения выносятся коллегиально, положение значительно усложняется. Групповые решения могут быть непредсказуемыми и отчасти зависеть от факторов, которые никакого значения иметь не должны. Кто начинает обсуждение, кто его завершает, кто говорит уверенно, кто одет в черное,

кто занимает определенное место за столом, кто улыбается, хмурится или делает нужный жест в подходящий момент – все эти и многие другие обстоятельства могут повлиять на исход дела.

Если участник группы сразу же высказался в защиту определенного решения, другие участники имеют основания последовать его примеру.

Мы уже упоминали феномен «мудрости толпы»: если задать вопрос большой группе людей, высоки шансы того, что усредненный ответ будет близок к истине. Однако «мудрость толпы» работает, только когда участники группы выносят суждения независимо друг от друга. Если они принимают решения не самостоятельно и вместо этого полагаются на чужое мнение, возможно поляризация (например, в вопросе [выбора игроков](#) менеджерами Fantasy Premier League).

Люди выслушивают мнения окружающих, и если эксперты, выступающие первыми, высказываются в пользу какого-то решения, то остальные могут с ними согласиться – по крайней мере в отсутствие причин не доверять выступающим или сомневаться в их правоте. При групповой поляризации мнения участников дискуссии после обсуждения проблемы обычно все дальше сдвигаются в направлении изначально выбранного полюса.

Часть III. Шум в прогнозировании

Чтобы сравнить точность прогнозов мы должны выразить качество прогнозов в цифрах. Единицей измерения является процент согласия (ПС). При идеальной точности прогнозов ПС составит 100%. А если прогнозы совершенно бесполезны, совпадение окажется таким же, как при случайном выборе. ПС составит 50%. Однако специалисты не используют ПС в качестве стандартной единицы измерения. Они применяют коэффициент корреляции r , который находится в диапазоне от 0 до 1. Эти единицы измерения связаны друг с другом:

<i>Коэффициент корреляции</i>	<i>Процент согласия (ПС)</i>
0,00	50%
0,10	53%
0,20	56%
0,30	60%
0,40	63%
0,60	71%
0,80	79%
1,00	100%

Рис. 5. Коэффициент корреляции и процент согласия (ПС)

Глава 9. Оценочные суждения и модели

Люди во многих отношениях слабее статистических моделей. И одна из критических слабостей состоит в том, что суждения людей подвержены влиянию шума. Люди уверены, что они используют комплексный подход и проявляют изобретательность, вынося свои оценки. Однако сложность и изобретательность большей частью напрасная трата сил – обычно они не помогают превзойти точность простых моделей.

Даже по прошествии более шестидесяти лет после выхода книги Пола Мила (Paul E. Meehl, *Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence*) нас по-прежнему шокирует сама мысль о том, что механическое прогнозирование одерживает верх над человеком. В оценках столько шума, что «бесшумная» модель добивается более точных прогнозов, чем эксперт.

Глава 11. Объективная неосведомленность

Истинное незнание (то, чего никак не узнать) и неполная информация (можно было выяснить, но этого не произошло) – делают идеальный прогноз невозможным. Обе эти неизвестные величины – не проблемы смещения и шума в ваших оценках; это объективные характеристики задачи. Объективная неосведомленность о важных, но неизвестных вам вещах жестко ограничивает точность. Люди, которые занимаются прогнозированием, недооценивают свою объективную неосведомленность. Излишняя самоуверенность – одно из самых хорошо документированных когнитивных искажений.

Психолог Филип Тетлок и его жена Барбара Меллерс [изучали](#), насколько хорошо люди справляются с предсказанием мировых событий на относительно короткий период – как правило, меньше года. Они обнаружили, что давать краткосрочные прогнозы трудно, но вполне возможно,

и некоторые люди, названные Тетлоком и Меллерс суперпредсказателями, показывают лучшие результаты, чем другие, – включая работающих в разведке профессионалов. Однако, что чем дальше мы заглядываем в будущее, тем выше объективная неосведомленность.

Когда вы доверяете интуиции, прислушиваясь к внутреннему голосу, вместо того чтобы анализировать факты, вы отрицаете объективную неосведомленность. Модели более совершенны, чем люди; однако их превосходство незначительно. Чаще всего мы сталкиваемся с экспертными оценками средней руки и лишь ненамного лучшими моделями. И все же надо стремиться к лучшему, а лучше все-таки модели.

Глава 12. Долина правдоподобности

Несмотря на то, что корреляция не подразумевает причинной связи, причинная связь подразумевает корреляцию. Там, где есть причинно-следственная зависимость, мы способны прогнозировать – причем корреляция, точность нашего прогноза, является мерой того, насколько верно мы понимаем причинно-следственную зависимость. Коэффициент корреляции является признаком того, насколько много – или насколько мало – мы понимаем.

Если вы задумались об общих категориях: уязвимые семьи как социальная группа и статистические данные, которые эту группу описывают, включая средние значения, статистические отклонения, корреляцию и тому подобное, вы обрели *статистическое мышление*. Другой способ мышления, более естественный для нашего разума, мы назовем здесь *каузальным мышлением*, то есть выражающим связь причины и следствия. Каузальное мышление изобретает сюжеты, в которых отдельные события, люди и объекты оказывают друг на друга воздействие.

Результаты мы воспринимаем, как завершение цепочки событий, неизбежную развязку predetermined трагедии. Когда мы проникаемся чувством неизбежности, то упускаем из виду, что обстоятельства легко могли сложиться иначе – подобно тому, как на каждой развилке дороги фортуна могла выбрать иной путь. Альтернативные сценарии так же правдоподобны, как и основной – при условии, что исход известен. Как только он реализовался, каузальное мышление представляет его вполне объяснимым – а значит, и предсказуемым.

Когда вы объясняете неожиданные, однако правдоподобные события таким образом, predetermined, которое в конечном счете свершилось, всегда приобретает смысл. Это мы и подразумеваем под пониманием сюжета; оно-то и заставляет реальность казаться предсказуемой – задним числом. И вследствие того, что событие, когда оно происходит, объясняет само себя, мы становимся жертвой иллюзии и верим – его можно было предугадать. Как известно из классических исследований феномена, получившего название «Я так и знал!», даже когда субъективная неопределенность какое то время существует, воспоминания о ней почти полностью стираются, едва эта неопределенность заканчивается.

Мы сопоставили два способа мышления: статистическое и каузальное. Каузальный режим избавляет нас от трудоемкого мыслительного процесса, в реальном времени сортируя события по категориям: нормальные и аномальные. Каузальное мышление избегает ненужных усилий, зато помогает сохранить бдительность, необходимую для того, чтобы фиксировать аномальные события. А вот статистическое мышление очень затратно. Оно требует ресурсов внимания.

Источник ошибок в прогнозировании состоит в том, что при рассмотрении единичного случая мы полагаемся на каузальное мышление. И лишь взглянув на события со статистической точки зрения, которую мы будем называть еще *взглядом со стороны*, можно постараться избежать этих ошибок.

Каузальный режим представляется нам более естественным. Даже объяснения, которые по сути должны относиться к статистическим, легко можно превратить в сказки о причинах и следствиях.

Часть IV. Как возникает шум

Глава 13. Эвристика, искажения и шум

Эта книга вобрала в себя столетия исследовательской работы в области интуитивных человеческих суждений – так называемую программу изучения эвристики и искажений.

Результаты первых четырех десятилетий исследований были освещены в книге [Думай медленно...](#)

[решай быстро](#), которая раскрыла психологический механизм, объясняющий как чудеса интуитивного мышления, так и присущие ему недостатки. Центральной идеей программы стало то, что люди, отвечая на сложные вопросы, используют упрощающие операции, названные эвристическими методами (*эвристикой*). Вообще то эвристика, как продукт быстрого, интуитивного мышления, или Системы 1, довольно полезна и помогает подобрать адекватные ответы. Однако порой она ведет к искажениям, которые мы описали как систематические, предсказуемые ошибки.

Когда мы заменяем один вопрос другим, более простым, ошибки неизбежны. К примеру, мы не учтем априорную вероятность, если будем оценивать похожесть вместо вероятности. Разного рода предвзятость побуждает людей представлять факты в искаженном свете в угоду своему первому впечатлению. Мы быстро поддаемся впечатлению и держимся за него, даже узнав противоречащую информацию. Это свойство называется *чрезмерной когерентностью*.

Глава 14. Сравнительные прогнозы

Количество категорий, которые мы можем распознать на шкале интенсивности, – семь плюс минус два.

Мы оба оценили фильм как очень хороший, хотя вы, как мне кажется, получили от просмотра намного меньше удовольствия, чем я. Мы использовали одни и те же слова, а вот шкалы у нас, похоже, разные. Мы предполагали, что второй сезон сериала станет не менее захватывающим, чем первый. Вот вам и сравнительный прогноз! Он подвел нас! Трудно оценить каждое сочинение само по себе. Не попробовать ли вам разложить сочинения по порядку от лучшего к худшему?

Глава 17. Источники шума

Итак, MSE разлагается на квадрат искажений и квадраты трех компонентов шума:

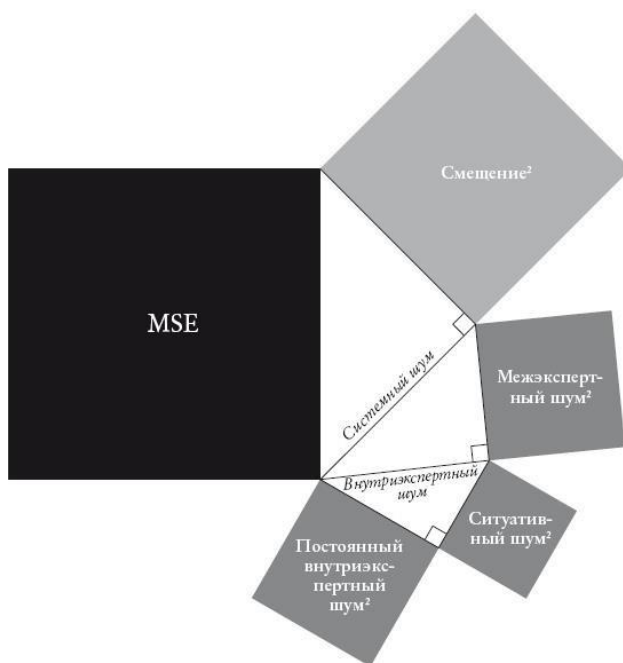


Рис. 6. Ошибки, искажения и компоненты шума

Часть V. Улучшение качества суждений

Глава 18. Лучший судья – лучшее суждение

Как определить наиболее квалифицированных кандидатов, которые будут выносить суждение? Отметим три значимых фактора. Верное суждение зависит от багажа накопленных знаний, эффективности мыслительных процессов и стиля мышления.

Метод измерения уровня интеллекта – «общий уровень умственных способностей» (GMA), заменил тест на IQ. Если работодателю требуются специалисты для вынесения профессионального суждения, имеет смысл набирать такие кадры из кандидатов с высоким уровнем интеллекта.

Независимо от уровня умственных возможностей, люди различаются свойственным им *когнитивным стилем* (подходом к решению задач, требующих вынесения суждения). В целях определения когнитивного стиля создано немало инструментов. Один из них – «Проверка когнитивной рефлексии» (CRT). Данная методика прославилась за счет известного вопроса о мяче и бите: «Бейсбольная бита и мяч вместе стоят 1,1 доллара. Бита ровно на доллар дороже мяча. Сколько стоит мяч?» Каждый из вопросов призван выяснить, насколько человек способен подавить интуитивный и, увы, неверный ответ («десять центов»). Низкий итоговый балл определяет вашу склонность попадаться в «ловушки сознания», а также говорит о мистическом складе ума – вере в привидения, астрологию и телепатию. Оценка по тесту предскажет, купитесь ли вы на фейковую новость. Балл говорит даже о степени вашей зависимости от смартфона.

CRT рассматривают как инструмент измерения склонности к рефлексии либо к импульсивным суждениям. Проще говоря, одни люди предпочитают тщательно обдумать свое решение, тогда как другие, столкнувшись со сходной проблемой, доверяют интуиции.

Отметим также тест Дайаны Халперн «Оценка критического мышления», который изучает умения, включающие природную склонность к рациональному мышлению и комплекс навыков, формирующихся в процессе научения.

Профессор психологии Джонатан Бэйрон разработал шкалу оценки, определяющую задатки «активного открытого мышления». Обладание «открытым умом» означает способность к активному поиску информации, противоречащей вашим же собственным гипотезам. Подобная информация включает в себя противоположные мнения оппонентов, которые следует подвергнуть тщательному анализу на предмет сопоставления новых фактов с уже имеющимися убеждениями.

Подобная способность свидетельствует о гибкости: человек сознает, что суждение – это непрерывный рабочий процесс, и приветствует оппонента, который стремится исправить его потенциальные ошибки.

Глава 19. Устраняем искажения Гигиена принятия решений

Мы рекомендуем выявлять искажения в режиме реального времени (не до и не после принятия решения). Человек часто с готовностью распознает чужие системные ошибки, со своими же все не так просто. Подобный недостаток осознания известен как *мертвая зона когнитивных искажений*.

Знаете ли вы, с каким именно искажением боретесь, в каком направлении оно может повлиять на результат? Если нет, возможно, ваш проект столкнулся с несколькими видами искажений, и неясно, какое из них одержит верх. Прежде чем начать обсуждение решения, давайте выберем наблюдателя за процессом. Мы четко соблюдали гигиенические меры при обсуждении данного решения, значит, высока вероятность, что решение принято оптимальное.

Глава 20. Управление информационным потоком в криминалистике

Там, где есть место суждению, найдется и шум, – это относится и к анализу отпечатков пальцев. У нас есть дополнительная информация о расследовании, однако не будем рассказывать криминалистам все сразу, пока они не определятся, иначе лишние сведения могут исказить их суждение. Расскажем только то, что им совершенно необходимо знать. Второе мнение не будет независимым, если второй эксперт знает мнение первого. С третьим экспертом дело обстоит еще хуже; это и есть каскад наводящей информации.

Глава 21. Стратегии отбора и агрегирования в прогнозировании

Аналитики, как правило, склонны к чрезмерной уверенности. Попросите их сформулировать прогноз в виде доверительного интервала, а не конкретного значения, и они наверняка выберут куда более узкий интервал, чем следовало бы. Исследователи предлагают различные варианты стратегий снижения шума и смещения. О первом из них – принципе «лучший судья – лучшее решение» – мы уже рассказывали в главе 18. Второй метод является одной из универсальных стратегий гигиены принятия решений: агрегирование (обобщение) множества независимых суждений.

Три стратегии улучшения суждений:

1. Обучение: несколько прогнозистов прошли курс обучения, направленный на улучшение их способностей, в ходе которого им преподавали основы вероятностных рассуждений.

Прогнозисты узнали о различных типах искажений (в том числе о пренебрежении априорной вероятностью, чрезмерной уверенности и заданности восприятия), усвоили важный принцип усреднения прогнозов, поступивших из различных источников, и обучились учету сопоставимой информации.

2. Командная работа (одна из форм обобщения): нескольких прогнозистов попросили работать в группах, предоставив им возможность прислушиваться к прогнозам коллег и обсуждать их. Командная работа имеет свойство повышать точность прогноза, поощряя участников группы анализировать мнения оппонентов и стимулируя развитие активного открытого мышления.
3. Отбор: каждый из прогнозистов получал балл за точность прогноза, и через год исследователи выбрали 2% лучших суперпредсказателей. В течение следующего года они уже работали совместно в особых элитных группах.

В итоге выяснилось, что каждая из трех стратегий по-своему результативна. Как минимум участники всех трех групп продвинулись по [шкале Брайера](#).

Глава 23. Определение шкалы оценки профессиональной эффективности

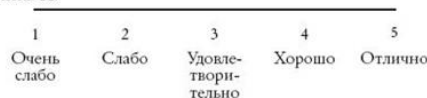
Многих работников нельзя оценить комплексно лишь на основании объективных показателей производительности труда. Именно поэтому сегодня широко распространены аттестации с применением методов суждения.

По результатам опыта служебных аттестаций опубликованы тысячи аналитических статей, и большинство авторов приходит к выводу об огромной доле шума в оценках. В основном тревожные сигналы касаются [метода 360 градусов](#), когда одного и того же работника по самым разным параметрам оценивают многочисленные эксперты. Не станем утверждать, что итоги подобных аттестаций обнадеживают. Исследования нередко говорят о том, что истинная дисперсия (то есть диапазон реальных оценок деловых качеств конкретного человека) составляет лишь от 20 до 30% разброса оценок. Остальные 70–80% являются системным шумом. Наверное, любая оценка деловых качеств куда меньше отражает реальную эффективность работника, чем нам хотелось бы.

Эффективным решением проблемы инфляции рейтингов может стать *принудительное ранжирование*. Относительные суждения содержат меньше шума, чем абсолютные, и данная закономерность также прослеживается в оценке деловых качеств работников.

Секция А

Качество работы сотрудника А:



Секция Б

Оцените ваших подчиненных по признаку надежности. Под надежностью в данном случае понимается четкое следование установленным правилам и инструкциям; поведение, не создающее излишних рисков; очевидные признаки осведомленности и понимания основ безопасности рабочего процесса.



Рис. 7. Примеры абсолютной и относительной шкалы оценки

Ранжирование носит принудительный характер в случаях, когда распределение оценок строится на обязательном условии. Например, стандарты компании предусматривают, что не более 20% работников может попасть в высшую категорию и не менее 15% – в низшую.

Критики системы часто нападают на сам принцип, который осуждают как жестокий, бесчеловечный и приводящий к обратному результату. В ходе одного исследования выяснилось, что 90% боссов, сотрудников и глав кадровых служб считают, что процесс управления эффективностью не приносит тех результатов, на которые рассчитывают компании.

Сторонники революции в процессе управления эффективностью больше сосредоточились на развивающих и работающих на будущее методах обратной связи, нежели на технологиях ретроспективной оценки.

Многие компании пользуются шкалой, привязанной к поведенческим показателям. Однако поведенческая шкала не устраняет шум в достаточной степени. Снижение уровня шума – та еще задача, и простыми техническими методами с ней не справиться.

Специалист по работе с клиентами:
Предупредителен и вежлив с каждым клиентом; надлежащим образом применяет знания о каждом товаре, помогая клиенту сделать выбор; прислушивается к каждому слову, старается выглядеть жизнерадостным, позитивным, полезным.

Выдающийся уровень	5	⇔ При просьбах, выходящих за рамки обычных, отношение к клиенту остается позитивным, сохраняется должная степень вежливости
	4	⇔ Охотно предлагает соответствующую замену, если необходимого клиенту товара нет в наличии
Удовлетворительный уровень	3	⇔ Четко помнит каждую просьбу клиента, не заставляя того повторяться
	2	⇔ Принимая заказ от клиента, прибегает к неприветливо звучащей формуле «Что вам угодно?»
Слабый уровень	1	⇔ Не перестает болтать с другими работниками, заставляя нескольких клиентов ожидать в очереди

Рис. 8. Пример поведенческой шкалы оценки

Глава 24. Структурирование решений по найму персонала

Если вы решили выяснить, какой из кандидатов преуспеет, а какой разочарует, стандартное собеседование (его еще называют неструктурированным) – не слишком информативный метод. Выразимся более определенно: часто он бесполезен.

Традиционные собеседования дают ошибочный прогноз дальнейшей эффективности сотрудника. Некоторые из таких ошибок обусловлены явлением, которое мы назвали «объективной неосведомленностью» (см. главу 11). Эффективность работы зависит от многих факторов, в том числе от того, как быстро нанятый вами человек приспосабливается к новой должности, как влияют на его деятельность различные события в личной жизни. Подавляющее большинство этих факторов в момент приема на работу спрогнозировать невозможно. Неопределенность ограничивает прогностическую ценность интервью; то же самое можно сказать о любом методе подбора кадров.

В одном эксперименте исследователи поручили студентам сыграть роли интервьюера и интервьюируемого, причем беседа должна была состоять из общих вопросов, требующих ответа «да» или «нет». Некоторых «претендентов» втайне попросили отвечать наобум (например, первая буква заданного вопроса сигнализировала о необходимости положительного ответа). Впоследствии экспериментаторы иронически заметили: «Некоторые из «претендентов» беспокоились, что такое интервью обречено – собеседник воспримет их ответы как полную чушь. Однако ничего подобного не случилось, интервью прошли в обычном порядке».

Ни единый интервьюер не осознал, что кандидаты отвечают как попало. Хуже того, когда их попросили подтвердить, что за несколько минут интервью удалось достаточно много узнать о собеседнике, «эксперты», принявшие участие в фарсе, ответили утвердительно, как и их коллеги из контрольной группы, которым «претенденты» давали обдуманные ответы. Такова наша способность к когерентности: мы частенько обнаруживаем воображаемую структуру в данных случайного характера, различаем в облаках знакомые фигуры и находим логику в абсолютно бессмысленных ответах.

Приведем другой пример. Одному из авторов данной книги как-то довелось провести собеседование с кандидатом, который на прежней работе занимал должность финансового директора в компании средних размеров. Интервьюер подметил, что собеседник уволился уже

через несколько месяцев работы, и попросил пояснения. Кандидат рассказал, что подал заявление из-за «стратегических разногласий с руководителем». Один из коллег, проводивший аналогичное интервью с тем же человеком, задал ему тот же вопрос и получил идентичный ответ. В последовавшей по итогам собеседования дискуссии интервьюеры радикально разошлись в оценках. Первый, получивший положительные впечатления от кандидата, усмотрел в его решении уволиться признаки целостности натуры и мужества. Второй же, чьи первые впечатления стали отрицательными, трактовал тот же факт как свидетельство отсутствия гибкости и даже как признак незрелости. Эта история показывает: как бы мы ни были уверены, что выносим суждение о кандидате, опираясь на факты, интерпретация этих фактов всегда окрашена нашим первым восприятием.

Google – хороший пример компании, модернизированной своей практикой подбора персонала и отчитавшейся о достигнутых результатах. Ласло Бок, исполнявший должность старшего вице-президента по персоналу, рассказал об этом в своей книге «Работа рулит!». В Google с одним и тем же кандидатом по очереди беседуют 4 эксперта. Каждый интервьюер обязан беседовать с кандидатом отдельно от коллег, и лишь потом эксперты общаются между собой. Метод обобщения работает, но лишь в том случае, когда суждения независимы друг от друга!

Google структурировал комплексные суждения на основе трех принципов: вычленение, независимость и отсрочка окончательного суждения. *Вычленение* предполагает разъединение решения на промежуточные оценки. Вычленение подобно дорожной карте, определяющей, в каких именно данных возникнет необходимость. Не относящаяся к делу информация отсеивается.

Независимость означает, что информация для формирования каждой из оценок собирается по отдельности. В ходе традиционных интервью каждый из этих компонентов не анализируется независимо. Каждая оценка влияет на следующую, что приводит к возникновению шума. Во время структурированного интервью не стоит задача решить, устраивает ли кандидат в целом: интервьюеру следует собрать данные о каждой оценке в общей структуре и присвоить кандидату балл по каждому из компонентов. Для выполнения поставленной задачи интервьюер обязан задавать предварительно сформулированные вопросы относительно поведения кандидата в определенных ситуациях в прошлом. Ответы регистрируются и оцениваются по заранее утвержденной шкале, использующей унифицированные критерии. Критерии включают в себя опорные примеры средних, хороших и блестящих ответов на каждый вопрос.

Структурированные интервью отличаются большей достоверностью предсказания дальнейшей деятельности кандидата, нежели традиционные, неструктурированные собеседования. Используя метод измерения ПС, скажем, что при проведении структурированного интервью ваши шансы выбрать лучшего претендента оцениваются в 65–69%. Такое значение, безусловно, выше, чем диапазон 56–61% при традиционных собеседованиях.

Отсрочка окончательного суждения – не стоит исключать интуицию, однако использовать ее следует лишь на последнем этапе. В Google итоговая рекомендация о приеме на работу формируется коллегиально на заседании комитета по найму, который анализирует все рейтинги, набранные кандидатом по каждому параметру и в ходе каждого интервью. Таким образом ограничивается склонность интервьюера (и любого из членов комитета) к составлению быстрого, интуитивного впечатления о кандидате и поспешным выводам.

Глава 25. Протокол промежуточных оценок

Основные стадии протокола промежуточных оценок

1. Приступая к процессу, следует разложить решение на факторы для промежуточной оценки (для типовых решений данный шаг выполняется единожды).
2. При формировании перечня желательно не забывать о «взгляде со стороны» (для типовых решений необходимо использовать относительное суждение, по возможности с применением соответствующей шкалы).
3. На стадии анализа следует приложить максимальные усилия для обеспечения независимости одной оценки от другой.
4. На решающем совещании каждая оценка должна рассматриваться отдельно.
5. Каждый участник совещания должен выносить индивидуальное суждение; целесообразно использовать алгоритм «оценка – обсуждение – повторная оценка».

6. Интуитивную составляющую следует отложить до финального этапа обсуждения, однако нецелесообразно запрещать ее вовсе.

Приложение 3. Корректировка прогнозов

Сравнительные прогнозы содержат ошибку, связанную с излишним доверием к процессу интуитивного сопоставления. Обычно мы делаем сравнительный прогноз, полагаясь на имеющуюся в наличии информацию, и ведем себя так, будто она является идеальным (или очень сильным) предиктивным индикатором.

Джули научилась бегло читать в четыре года. Каков у нее средний балл успеваемости в колледже? (Максимальный 4). Дав прогноз в районе 3,8, вы интуитивно рассудили, что в части умения читать четырехлетняя Джули входила в лучшие 10 % своей возрастной группы (и все же не в лучшие 3–5%). Значит, вы невольно предположили, что по успеваемости она войдет в число лучших учеников своей возрастной группы в колледже – это как раз средний балл в районе 3,7–3,8.

Прогноз статистически некорректен, поскольку вы переоценили прогностическую значимость имеющейся информации. Раннее развитие ребенка не всегда предполагает выдающиеся результаты в учебе (и, к счастью, дети, которым чтение в дошкольном возрасте давалось с трудом, необязательно будут находиться в неуспевающей части своего класса).

Чаще всего выдающиеся способности в детстве далее начинают приближаться к среднему уровню. Нет смысла рассуждать о причинах, ибо мы имеем дело со статистическим феноменом. Крайние проявления со временем сглаживаются только потому, что показатели, зарегистрированные в прошлом, далеко не идеально коррелируют с результатами в будущем. Данная тенденция получила название регрессии к среднему значению (соответственно, и сравнительный прогноз мы называем нерегрессионным, поскольку он эту особенность не учитывает).

Суждение, вынесенное вами по поводу Джули, окажется верным лишь в случае, если возраст, в котором проявилась способность к чтению, является четким прогностическим индикатором дальнейшей успеваемости. Иными словами, между двумя упомянутыми факторами прослеживалась бы корреляция. Мы с вами понимаем, что это не так.

Существуют статистические способы, помогающие вынести более точное суждение. Ими невозможно воспользоваться интуитивно, более того, даже человек, имеющий определенные знания статистики, затруднится применить подобный подход. Необходимую процедуру мы покажем на рисунке 9, описывающем пример с Джули.

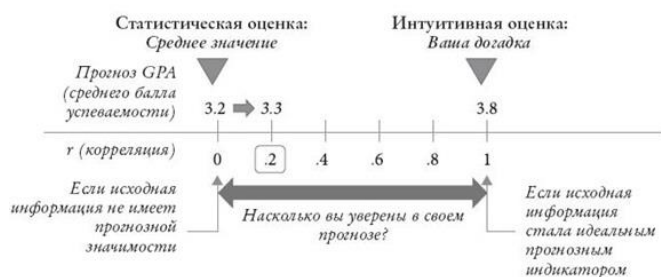


Рис. 9. Адаптация интуитивного прогноза к регрессии к среднему значению

Используем интуитивный подход. Не следует считать вашу интуитивную догадку о будущих успехах Джули бесполезной. То же самое верно в отношении любого прогноза, если вы располагаете исходной информацией. Ваша система 1 (быстрое мышление) легко находит на прогнозной шкале место для исходной информации, позволяя сделать соответствующий вывод о будущей успеваемости Джули. Данная догадка и станет прогнозом, если информация, которой вы располагаете, содержит сильный предиктивный индикатор. Запишите ваш первый вывод.

Ищем среднее значение. Возвращаемся в исходную точку и забываем о том, что мы сейчас знаем о Джули. Что вы скажете о будущей успеваемости Джули, если не будете знать о девочке вообще ничего? Ответ напрашивается сам собой: при полном отсутствии информации лучшим прогнозом станет средняя успеваемость в ее классе – допустим, 3,2. Подобный взгляд продиктован применением более широкого понятия, которое мы уже обсуждали выше, – взгляда со стороны.

Прибегая к подобному подходу, мы рассматриваем анализируемый случай в контексте аналогичных случаев и рассуждаем статистически.

Оцениваем прогностическую ценность исходной информации. Данный этап – самый сложный. Вам необходимо спросить себя: «Какова ценность доступной мне информации для вынесения прогноза?» Мы уже понимаем, почему этот вопрос крайне важен. Если вы знаете лишь размер обуви Джули, то будете совершенно правы, присвоив этой информации оценку «0»; для прогноза придется воспользоваться средним значением балла успеваемости. Если же вы располагаете сведениями об оценках Джули по каждому предмету в школе – это идеальный прогностический индикатор. В таком случае мы просто используем их среднее значение. Между этими двумя крайними случаями лежит область неопределенности. Зная о выдающихся успехах Джули в старших классах школы, мы, безусловно, оценим прогностическую ценность подобной информации куда выше, чем сведения о возрасте, в котором она научилась читать.

Наша задача сводится к количественному определению прогностической ценности исходных данных, которая выражается в виде корреляции с событиями, которые требуется спрогнозировать. За исключением редких случаев, данное значение станет довольно упрощенным. Социология говорит, что корреляция выше 0,5 встречается крайне редко. Коэффициент корреляции в области 0,2 нам уже о чем-то говорит. В примере с Джули, скорее всего, это значение и станет верхним пределом.

Движемся от взгляда со стороны в направлении вашей интуитивной догадки, приближаясь к отметке, которая подаст нам сигнал о прогностической ценности исходной информации. Последний этап представляет собой простое арифметическое сочетание трех цифр, которые у вас уже есть. Вам следует скорректировать среднее значение в сторону вашего интуитивного предположения с учетом величины корреляции, которую вы держите в уме.

Данный этап всего лишь детализирует то наблюдение, которое мы сделали чуть выше: если значение корреляции равно нулю – выбираем среднюю величину; если значение корреляции равно единице – спокойно игнорируем среднюю и получаем идеальный сравнительный прогноз. Что касается Джули, то лучшим прогнозом в условиях имеющейся информации станет предположение, что ее средний балл составит около 3,3 (диапазон между средним баллом класса – 3,2 и вашей интуитивной оценкой – 3,8 составляет 0,6; итоговую оценку мы получили, продвинувшись от среднего значения к вашей догадке не более чем на 20% от величины диапазона).

Подобный метод можно применять ко многим проблемам, требующим суждения.