



управляем
предприятием



УПРАВЛЕНЧЕСКАЯ ТЕОРИЯ ИЗМЕРЕНИЙ

Часть 7

Рекомендации по использованию
в системе показателей



Андрей Мицкевич

К. э. н., доцент Высшей школы финансов и менеджмента РАНХиГС и Научно-исследовательского университета ВШЭ, руководитель консультационного бюро Института экономических стратегий, преподаватель ведущих бизнес-школ Москвы. Консультант, автор четырех книг и более чем 80-и статей,

вышедших в последние годы, по контроллингу, мотивации, управленческому учету и сбалансированной системе показателей.

Незнание теории не освобождает от ответственности

Что такое качество измерений? Это способность передать эмпирическую систему отношений числами. Видим (пусть субъективно), Иванов работает лучше Петрова, а это значит, что и зарплата у него должна быть больше. Все понятно, но как же сложно... Например, в компаниях используется маркетинговая информация. Качество, то есть надежность и достоверность маркетинговой информации, определить весьма не просто. Еще большее значение качество играет при измерении показателей компании. Ведь плохое качество измерения показателя дискредитирует всю систему показателей компании. В этой части статьи мы поговорим об информативности измерений, проанализируем приведенную в одной из предыдущих частей систему оценок в фигурном катании, а также обсудим использование теории измерений в системе показателей компании.



Информативность измерений

Информативность — это то новое, что вы узнаете по результатам измерения.

Информативностью измерений мы будем называть порядковую или количественную меру новой информации, заключенную в измерении.

Надежность системы оценок приходится оценивать по интуиции. Нет критериев оценки самих критериев надежности и достоверности. Каждый раз придется творить.

Количественная мера информативности сама по себе представляет интересный объект теории информации (см. врезку «Мера Шеннона»). Но вот беда: опять теория о своем, а практика совсем иная. Напрямую в системе показателей компании мера Шеннона практически неприменима. Но для практического использования в системе показателей компании это не важно. Гораздо важнее то, что мы уже отмечали: для большей достоверности выводов следует использовать менее информативные шкалы, соблюдая нужный баланс информативности и достоверности.

Информативность измерений изменяется в обратной пропорции к достоверности.

Например, завтра точно будет от -50° по Цельсию до $+50^{\circ}$. Это достоверные сведения, только вот вряд ли вы в них нуждаетесь. То есть их информативность равна для вас нулю.

Надежность и достоверность систем оценок в фигурном катании

Теперь, когда мы узнали теорию качества измерений, проанализируем системы оценок в фигурном катании, которая описана в предыдущей части, по критериям надежности, информативности и достоверности. Это позволит нам увидеть, как работают требования достоверности и надежности, а также показать алгоритм анализа подобных систем показателей.

Надежность классической 6-балльной системы оценок. Старая система оценок предполагала простое сложение баллов за технику и артистичность в произвольной и обязательной программах, а затем все оценки за обе программы суммировались (или усреднялись, что одно и то же с точки зрения результата).

1. **Правильность отдельных оценок**, то есть приемлемость шкалы (0,0 — 6,0), была не стопроцентной, но к шкале фигуристы, тренеры и судьи приспособились, и как относительный инструмент измерений эта шкала работала. Специалисты отмечали, что примерно в 90 % случаев она позволяла судьям отличать близкие по качеству выступления разных спортсменов или вариации качества выступлений одного и того же спортсмена. Поэтому правильность локальных оценок можно считать равной 90 %¹.
2. **Правильность интегральной оценки** первоначально не вызывала особых сомнений. Действительно, суммируем оценки, тем самым учитываем все стороны мастерства фигуриста. Но мы знаем, что правильность интегральной оценки невелика.

¹ Конечно, я не эксперт в области фигурного катания. Но данное упражнение посвящено не системе оценок в фигурном катании, а показывает алгоритм анализа подобных систем показателей.

Она не дает сбоев в типовом случае, если фигуристы примерно на одном уровне владеют произвольной и обязательной программами. Случай с Беатрис Шубой исключение, но на такие исключения и должна быть рассчитана хорошая система оценок. Поэтому моя оценка правильности интегральной оценки — 50 % (поскольку она работает только в типовых случаях).

3. **Обоснованность отдельных оценок.** Близкие оценки, например, 5,1 и 5,2, отличались не очень четко. За базу один судья брал лучший прокат, оценивая его в 5,8, а другой за то же самое ставил 5,9. Соответственно, и шаг снижения оценки, и конечный результат у них были разными. Градации шкал для всех программ и оценок устанавливались экспертно. Для разных оценок (школы, техники и артистичности в произвольной программе) эксперты и критерии были различны. Это крайне важно учитывать при анализе системы оценок. Тем не менее, система была достаточно хорошо разработана для мастеров фигурного катания. Моя оценка обоснованности локальных оценок — 80 %.
4. **Обоснованность интегральной оценки со временем становилась все сложнее.** Значение (вес, значимость) произвольной программы росло. Да и с самого начала она была важнее для зрителей, а в сегодняшнем мире это самое главное. Поэтому моя оценка обоснованности интегральной оценки — 70 %.
5. **Устойчивость локальных оценок** характеризует степень совпадения результатов измерения при повторных применениях измерительной процедуры. С этим вообще было плохо. Многое зависело от того, какую страну представлял тот или иной спортсмен. Моя оценка устойчивости локальных оценок — 60 %.
6. **Устойчивость интегральной оценки.** Мы уже знаем, что суммирование значений разных показателей, даже если они выражены в одинаковых шкалах, некорректно. Потому моя оценка устойчивости интегральной оценки — 50 %.

Надежность модифицированной 6-балльной системы оценок. Напомню, что изменилось правило подсчета интегральной оценки: стали использовать веса и вес произвольной программы был выше. В результате, по моей оценке, правильность, обоснованность и устойчивость интегральной оценки возросли на 5–10 %. Система явно улучшилась, но до приемлемой системы оценок было еще далеко. Что и доказала Беатрис Шуба.

Мера Шеннона

В теории информации самой знаменитым показателем информативности является мера Шеннона. Она вычисляется как разность энтропий до (H_0) и после измерения (H_1).

Энтропия (H) определяется так:

$$H = \sum_{i=1}^n p_i \log p_i ,$$

где p_i — вероятность события i , которых всего n . Соответственно, информативность (количество информации) равно разнице:

$$I = H_1 - H_0$$

Надежность системы суммы мест. Напомню, что бальную шкалу с большим количеством градаций заменили на порядковую, так как порядковая шкала более надежная, чем бальная. Итог стали подводить по сумме мест с весами, а отдельные оценки судей сохранили 6-бальными для ориентировки судей и спортсменов. В итоге повысилась достоверность и надежность за счет перехода к измерениям в шкалах меньшей информативности. «Эффект Шубы» удалось нивелировать.

Достоверность системы оценок — совпадает ли результат измерений с действительным положением вещей — оценить трудно. Формальных измерителей достоверности нет.

Надежность и достоверность новой системы оценок, основанной на оценке всех элементов программы по отдельности, стала на порядок выше. Во-первых, она в максимально возможной степени защитила фигуристов от произвола судей, то есть повысилась объективность оценок. Во-вторых, — о чудо! — одновременно резко усилилась информативность. Оценки можно сравнивать во времени, их можно раскладывать по элементам. В-третьих, заложен механизм совершенствования. Критерии оценки элементов почти каждый год пересматриваются, дополняются. С точки зрения теории трудно предложить принципиально более

совершенную систему. Фигуристы, тренеры и судьи могут совершенствовать отдельные оценки в рамках принципиально неизменной системы.

Подведем итоги анализа системы оценок в фигурном катании.

1. **Общие закономерности в вычислении качественного интегрального показателя на основе многих частных оценок есть, и они конструктивны.** Их необходимо знать и использовать.
2. **Субъективность системы оценок не может быть преодолена на 100 %, но может быть снижена.**
3. **Надежность, информативность и достоверность системы оценок измерить, определить трудно.** Надо «спускаться» в предметную область и действовать по обстановке.
4. **Надежность системы оценок приходится оценивать по интуиции.** Оказывается, степень разработанности аудита систем оценок близка к нулю. Нет критериев оценки самих критериев надежности и достоверности. Каждый раз придется творить.
5. **Достоверность системы оценок оценить трудно.**

Рекомендации по использованию теории измерений в системе показателей

Прежде всего, надо перенести в практику создания системы показателей деятельности общие требования к качеству — достоверности, надежности и информативности. Пусть и на недостаточно разработанном уровне, но их необходимо учитывать. Все эти показатели комплексные, то есть не сводимые к одному свойству, к одному простому и хорошо измеряемому показателю. Остальные требования к качеству системы показателей деятельности надо взять из конкретных областей, в которых работает фирма.

Надежность — отсутствие отклонений в результатах измерений. В системе показателей деятельности это означает, что два и более эксперта дают одну и ту же оценку, результат измерений не зависит от процедуры. Формальные измерители надежности есть: она измеряется статистическими мерами связи и вариабельности. Однако использовать в системе показателей деятельности это можно только в случае опросов нескольких экспертов, что бывает относительно редко.

Важные общие свойства измерений: как правило, чем больше информативность, тем меньше надежность, а также, чем больше надежность, тем меньше информативность.

Достоверность — результат измерений совпадает с действительным положением вещей. Формальных измерителей достоверности нет. Эксперты должны однозначно понимать, что именно и в какой шкале они оценивают, чтобы избежать ситуаций, когда эксперты оценивают один и тот же показатель, характеризующий объект исходя из различных предпосылок.

Информативность — мера новой информации, заключенная в измерении. Она определяется разрешающей способностью шкалы измерения и нашими исходными представлениями об измеряемой величине.

Если мы уже знаем измеряемую величину, то информативность измерения равна нулю, никакой новой информации мы не получаем. Но со значениями показателей деятельности все ясно: мы потому их и измеряем, что заранее не знаем. Поэтому эта сторона информативности не актуальна. И в системе показателей деятельности информативность сводится к разрешающей способности шкал.

Возможные «противоречия» между показателями измерений. Измерение может быть (а может и не быть, обязательности здесь нет):

- надежным, но недостоверным;
- достоверным, но ненадежным.

Важные общие свойства измерений:

- как правило, чем больше информативность, тем меньше надежность;
- как правило, чем больше надежность, тем меньше информативность.

Итак, в системе показателей деятельности показатели качества измерений и оценок конкретизируются и упрощаются. Основные выводы без строго доказательства:

- надежность сводится к устойчивости;
- правильность сводится к обоснованности;
- информативность сводится к точности шкалы (чем сильнее шкала, тем больше точность);
- многомерность = {количество независимых или частично-зависимых, но дополняющих друг друга показателей для одной цели} .

Очевидно, что в процессе построения в системе показателей деятельности происходит последовательная адаптация шкал и методик измерений под цели и задачи конкретной фирмы. Это облегчает решение задач качества измерений и делает представленные здесь методы необходимыми при проведении оригинальных измерений. Оценки и оценочные шкалы — это база технического инструментария.



Итак, мы видим, что плохое качество измерения показателя дискредитирует всю систему показателей компании. Аудит шкал измерений — задача несложная для квалифицированного консультанта, но очень важно не ошибиться. Поэтому, на мой взгляд, его нужно провести. Вот только наши консультанты такой мелочью не всегда занимаются. На такой работе со шкалами денег не заработаешь, да и некоторая квалификация нужна.

Литература

1. Пфанцагль И. Теория измерений/ Пер. с нем. – М.: Мир, 1976. 165 с.
2. Фишберн П. Измерение относительных ценностей. – в кн.: Статистическое измерение качественных характеристик. – М.: Статистика, 1972, с. 35–94.
3. Орлов А.И. Прикладная статистика. – М.: Издательство «Экзамен», 2004.