

УДК 519.86

КАУЗАЛЬНОСТЬ И ЕЕ ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ

Ефимова А. А.

магистрант

Уральский государственный экономический университет

Россия, г. Екатеринбург

Аннотация

В статье рассмотрены такие понятия как каузальность и корреляция, рассмотрены их различия, обоснован приоритет каузальности как наиболее точного метода определения связности фактов, дополнительно указаны распространённые методы исследования связности факторов. Рассмотрены методы определения связности факторов и прикладное значение каузальности. Приведены примеры возможного использования теории казуальности в таких сферах, как физика, биология, психология, экономика и менеджмент.

Ключевые слова: каузальность, корреляция, каузальные связи, методы оценка связности факторов, применение каузальности.

CAUSALITY AND ITS PRACTICAL APPLICATION

Efimova A. A.

master student

Ural State University of Economics

Russia, Yekateginburg

Annotation

The article discusses such concepts as causality and correlation, considers their differences, substantiates the priority of causality, and also indicates common methods

for studying binding factors. Methods for determining the connectedness of factors and the applied value of causality are considered. Examples are given of the possible use of the theory of casualness in such fields as physics, biology, psychology, economics, and management.

Key words: causality, correlation, causal relationships, methods for assessing the connectedness of factors, the use of causality.

В любой практической и научной деятельности очень часто затрагивается проблема взаимосвязи двух или более факторов, например, спроса и предложения, рабочей недели и производительности труда, курения и продолжительности жизни, кривизна линзы в очках и уровень зрения и тому подобное. И таких примеров можно привести тысячи, если не миллионы.

Такую взаимосвязь в научной литературе часто обозначают термином «корреляция», то есть, статистическая взаимосвязь. В научных работах и статьях такую связь часто считают истинно верной. Так, например, в 2013 году американские ученые Винеи Прасада, Джоэл Йоргенсонб, Джон П.А. Йоанииск и Адам Сифу опубликовали в «Journal of Clinical Epidemiology» исследование, в котором авторы изучили все оригинальные публикации о нерандомизированных исследованиях, опубликованные в 2010 году в таких серьёзных научных журналах как «New England Journal of Medicine», «Lancet», «Journal of the American Medical Association», и «Annals of Internal Medicine». Из 631 публикации 298 (47%) были посвящены обсервационным (не экспериментальным) исследованиям, то есть основанным на коррелировании исследуемых показателей. В 167 (56%) из этих 298 исследований авторы давали лечебные рекомендации, основываясь на результатах своих работ. Только в 24 (14%) из 167 публикаций для проверки рекомендаций было упомянуто о необходимости

провести рандомизированные клинические испытания, в остальных 143 публикациях авторы смело выдавали практические лечебные советы.

То есть, существует большой процент научных исследований, основанных на исследовании корреляции показателей и не подтвержденных экспериментально, такой подход часто может являться неверным и даже опасным как, например, в случае с медицинскими исследованиями, представленными в статье Винея Прасады и его соавторов.

Описание взаимосвязи исследуемых компонентов, пожалуй, будет правильнее рассматривать с помощью такого термина как «каузальность», то есть «причинность».

Итак, для начала следует дать определения корреляции и каузальности.

Корреляция – это статистическая взаимосвязь двух величин, при которой изменения значений одной из этих величин сопутствуют систематическому изменению значений другой величины. [2]

Каузальность – это причинная взаимообусловленность событий, при которой при воздействии одного объекта (причина) происходит соответствующее ожидаемое изменение другого объекта (следствие) [1].

Большинство людей с научным складом ума знают, что корреляция — не значит каузальность. Существует масса примеров, в которых очевидно отсутствует причинное влияние одной величины на другую, хотя присутствует корреляция:

- количество пиратов отрицательно коррелирует с ростом земной средней температуры (мы понимаем, что пираты не влияют на глобальное потепление, но есть какие-то экономические факторы, которые привели к уменьшению количества пиратов и увеличению выбросов углекислого газа в атмосферу);
- количество нобелевских лауреатов связано с потреблением шоколада;

- предпочтение одежды размера XXXL коррелирует с риском сердечного приступа;
- количество церквей в городе может коррелировать с количеством баров (хотя очевидно, что обе эти величины зависят от числа жителей данного населённого пункта) и так далее.

Корреляция может быть ложной и исчезать при добавлении новых данных, а может и вовсе быть иллюзорной в связи с когнитивными искажениями; возможно присутствие артефактов, или ошибка 1 рода, или выборка была слишком маленькой; или использовалось слепое прочёсывание данных и переподгонка (p-hacking/data-mining), а может, повлияла систематическая ошибка отбора или другая ошибка в процессе сбора и обработки данных.

Хорошим примером таких корреляций является проект Тайлера Вигена «Spurious Correlations» [7]. В качестве, конкретного примера можно привести корреляцию таких показателей как потребление сыра на душу населения и количество людей, погибших, запутавшись в собственных простынях (Рис. 1).

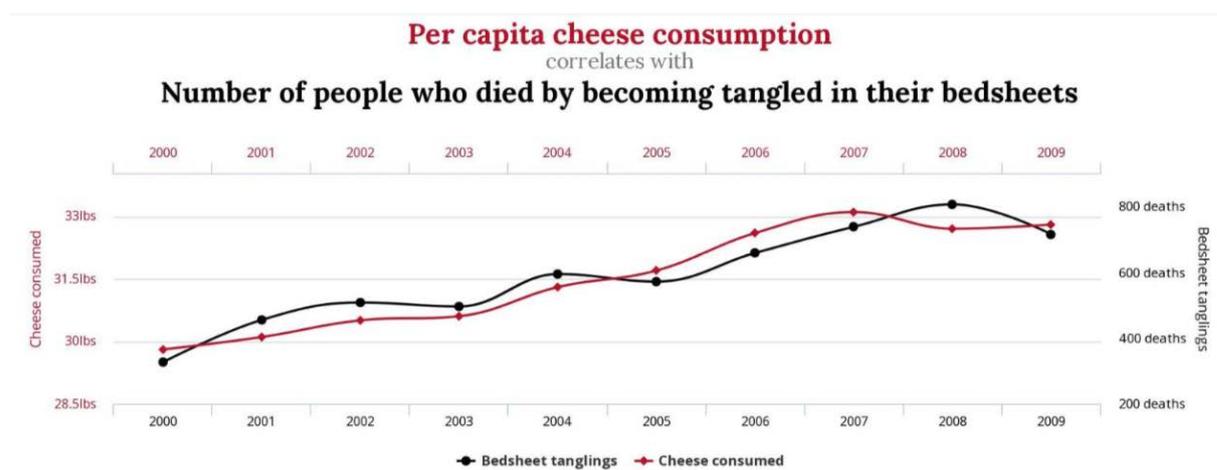


Рис. 1 – Корреляция между потреблением сыра на душу населения и количеством людей, погибших, запутавшись в собственных простынях [7]

Процент корреляции этого фактора в исследовании Тайлера Вигена составил 94,71%, что свидетельствует о сильной связи между факторами, тем не менее, такую связь можно считать абсурдной.

Таким образом, можно заметить, что исследование корреляции факторов может быть абсолютно некорректным, выводы могут быть ложными, это особенно проблематично, когда независимость факторов не очевидна.

Причинами исследования корреляционной связи вместо исследования каузальных связей, вероятнее всего, можно считать:

- во-первых, отсутствие достаточного количества исследований каузальных связей, особенно, в областях технических наук. К сожалению, на данный момент, большинство научных трудов о каузальных связях относятся к областям психологии, философии и метафизики;
- во-вторых, нежелание разбираться в более сложных расчетах, необходимых для вычисления каузальностей связей, по сравнению с вычислением корреляционной связи;
- в-третьих, сложность оценки вероятностных факторов при вычислениях каузальных связей

На данный момент, известными приемами оценки связанности факторов являются следующие методы [4]:

- Кросскорреляционная функция (линейная зависимость). Это наиболее часто используемый алгоритм, как раз-таки определяющий наличие корреляционной связи между исследуемыми показателями.
- Функция когерентности. Обычная функция когерентности определяется как мера причинно-следственной связи между двумя сигналами при наличии других сигналов. Когда функция когерентности равна единице, два сигнала полностью связаны. Если

функция когерентности между двумя входными сигналами одна, эти два входа полностью связаны.

- Функция взаимной информации — это безразмерная величина s (как правило) единицами битов, и ее можно рассматривать как уменьшение неопределенности относительно одной случайной величины, учитывая знание другой. Высокая взаимная информация указывает на значительное снижение неопределенности; низкая взаимная информация указывает на небольшое сокращение; и нулевая взаимная информация между двумя случайными переменными означает, что переменные являются независимыми.
- Синхронизация – подстройка ритмов осциллирующих объектов за счёт слабого взаимодействия между ними.
- Коэффициент фазовой синхронизации. В классическом смысле синхронизация означает регулировку частот периодических самоподдерживающихся генераторов из-за слабого взаимодействия. Этот эффект (также называемый фазовой синхронизацией или захватом частоты) хорошо изучен и находит много практических применений.
- Метод дополнительной системы (измерение обобщенной синхронизация - состояние, при котором и фазы, и амплитуды взаимодействующих систем становятся однозначно связанными)
- Энтропия переноса используется для анализа причинно-следственных связей между переменными подсистемы из данных. Тот факт, что он несимметричен, позволяет определить направление информационного потока. Было показано, что каузальность Грейнджера, которая обычно использует авторегрессию для оценки

взаимодействий, эквивалентна энтропии переноса в случае гауссовых переменных.

- Каузальность по Грейнджеру — это способ исследования каузальности между двумя переменными во временном ряду. Метод является вероятностным описанием каузальности; он использует эмпирические наборы данных, чтобы найти закономерности корреляции. Каузальность тесно связана с идеей причинно-следственной связи, хотя это не совсем то же самое. Переменная X является причинной по отношению к переменной Y , если X является причиной Y или Y является причиной X . Однако в случае каузальности Грейнджера не проверяется истинная причинно-следственную связь.

Из всех приведенных выше методов, только каузальность Грейнджера затрагивает тему измерения каузальности. Также эту тему рассматривали Робин Коллингвуд и Симс.

Каузальность Грейнджера — это процедура «снизу вверх», где предполагается, что процессы генерирования данных в любом временном ряду являются независимыми переменными, затем наборы данных анализируются на предмет их корреляции. Противоположным является метод «сверху вниз», который предполагает, что процессы не являются независимыми; затем наборы данных анализируются, чтобы увидеть, генерируются ли они независимо друг от друга. Нулевая гипотеза для теста является то, что отставали рентгеновское значение не объясняет вариацию y . Другими словами, предполагается, что $x(t)$ не является причиной Грейнджера $y(t)$.

В качестве статистики теста Грейнджера используется статистика F-теста для проверки линейного ограничения на коэффициенты, которая также является статистикой гипотезы о сравнении двух «вложенных» линейных моделей

регрессий: «короткой» и «длинной», «с ограничениями» и «без ограничений» (“short” & ”long”, “restricted” & ”unrestricted”).

В общем виде тест Грейнджера выражается формулой 1:

$$\begin{cases} Y_t = \lambda_0 + \alpha_1 * Y_{t-1} + \beta_1 * X_{t-1} + u_{1t} & R_{YL}^2 \\ X_t = \lambda_1 + \gamma_1 * Y_{t-1} + \delta_1 * X_{t-1} + u_{2t} & R_{XL}^2 \end{cases} \quad (1)$$

Симс в 1972 году предложил другой тест на каузальность по Грэнджеру. Суть теста заключается в построении регрессии одной переменной y на прошлые, текущие и будущие значения другой переменной x и проверке гипотезы об одновременном равенстве нулю коэффициентов при будущих значениях. Если гипотеза принимается, то это означает, что знание будущих значений x не улучшает прогноза y .

Эконометрически данные тесты неэквивалентны, но, по существу, проверяют одно и то же.

Одну из самых интересных теорий каузальности предложил Робин Коллингвуд в своей работе «Просто о метафизике». Коллингвуд замечает, что каузальность имеет три разных значения:

- Каузальность в историческом смысле.
- Каузальность в смысле прикладных наук.
- Каузальность в смысле фундаментальных наук.

В последнем случае определение каузальности сталкивается с серьезными трудностями, причем в настоящее время при обсуждении каузальности чаще употребляют ее в третьем смысле, забывая об оставшихся двух.

Применительно к научной деятельности, теория каузальных связей достаточно слабо развита, поэтому в статье рассмотрены скорее потенциальные

сферы применения каузальности, чем ее практическое применение на данный момент.

Для научного исследования эффективной каузальности причину и следствие лучше всего воспринимать как временные переходные процессы.

Простое наблюдение корреляции почти недостаточно, чтобы установить каузальность. Почти во всех случаях установление каузальности зависит от повторения экспериментов и вероятностных рассуждений. Едва ли когда-либо причинно-следственная связь устанавливается более надежно, чем как более или менее вероятная. Часто наиболее удобно для установления каузальности, если контрастные материальные положения дел полностью сопоставимы и отличаются только одним переменным фактором, возможно, измеренным действительным числом. В противном случае эксперименты обычно трудно или невозможно интерпретировать.

В некоторых науках очень трудно или почти невозможно установить положения вещей, которые проверяют гипотезы каузальности. Ниже приведены примеры таких наук.

Физика. На данный момент актуальным направлением использования теории каузальности в физике является рассмотрение каузальных связей в квантовой механике. Наиболее ранние попытки квантовых причинно-следственных моделей проводились путем определения причинно-следственных структур для квантовых систем, а затем выяснения, какие условно-независимые отношения остались нетронутыми. Однако эти модели не могли выполнить байесовский вывод, поскольку условная независимость больше не являлась предпосылкой для определения общей причины. Мэтью Лейфер и Роберт Спеккенс [6] пытались включить байесовский вывод в квантовую структуру, используя «условные квантовые состояния» вместо условных вероятностей, но этот творческий подход оказался применимым только в ограниченных случаях.

Фабио Коста и Салли Шрапнель [5] отложили проблему условной независимости, чтобы сосредоточиться на причинных вмешательствах. Заменяя «детерминированный» на «унитарный» по принципу Райхенбаха, можно получить новую версию квантовых причинно-следственных моделей. В частности, их квантовая версия принципа Райхенбаха позволила связать условную независимость с квантовыми причинно-следственными связями, подобными описанным в модели Коста и Шрапнеля. Более того, эти условные отношения независимости могут быть использованы для выполнения байесовского вывода. Благодаря таким результатам, можно обнаружить, что квантовая механика имеет причинную интерпретацию, как и классическая механика.

Биология. Принцип биологической относительности заключается в том, что до проведения соответствующих экспериментов, не существует привилегированного уровня каузальности. В мультимасштабных сетях взаимодействия, которые встречаются повсюду в организмах, любые части сети на любом уровне могут влиять на любую другую часть. Принцип основан на математических подходах к пониманию биологических процессов. В то время как дифференциальные уравнения представляют динамику компонентов системы, начальные и граничные условия представляют исторические и контекстуальные факторы, без которых никакие конкретные решения уравнений были бы невозможны. Этот принцип нашел много применений в физиологии и в других областях биологии. Это неудивительно, поскольку математическая точка зрения является необходимой, независимо от того, являются ли компоненты молекулярными (гены, белки и метаболиты), сетями (на всех уровнях), клетками, тканями, органами или какими-либо другими компонентами.

Психология. Поскольку причинное познание настолько вездесуще, психологическое исследование его само по себе является огромной темой, и

буквально сотни людей посвятили его изучению целые карьеры. Причинное познание можно разделить на две грубые категории: причинное обучение и причинное мышление. Первый охватывает процессы, с помощью которых мы узнаем о причинно-следственных связях в мире как на уровне типов, так и на уровне маркеров.

Экономика. Экономика отличается от других социальных наук своей преданностью основной теории, которая в той или иной степени разделяется большинством экономистов. Основную теорию можно рассматривать как формулирование экономических механизмов или структур, мало чем отличающихся от физических механизмов, которые обеспечивают классические иллюстрации причинной структуры. Хотя само понятие экономической структуры, по-видимому, благоприятствует манипулируемости или структурному учету каузальности, с ее фундаментально причинной онтологией. При изучении причинно-следственных связей в экономике можно выделить две проблемы: выводная проблема (как изолируются причины или определяются структуры?) и контрфактивная проблема (как используются знания о причинно-следственной структуре для обоснования ненаблюдаемых результатов?). Статистика и экономика обычно используют уже существующие данные или экспериментальные данные, чтобы вывести каузальность методами регрессии.

Менеджмент. Используемая в управлении и разработке диаграмма Исикавы (Рис. 2) показывает факторы, которые вызывают каузальный эффект. Меньшие стрелки связывают подпричины с основными причинами.

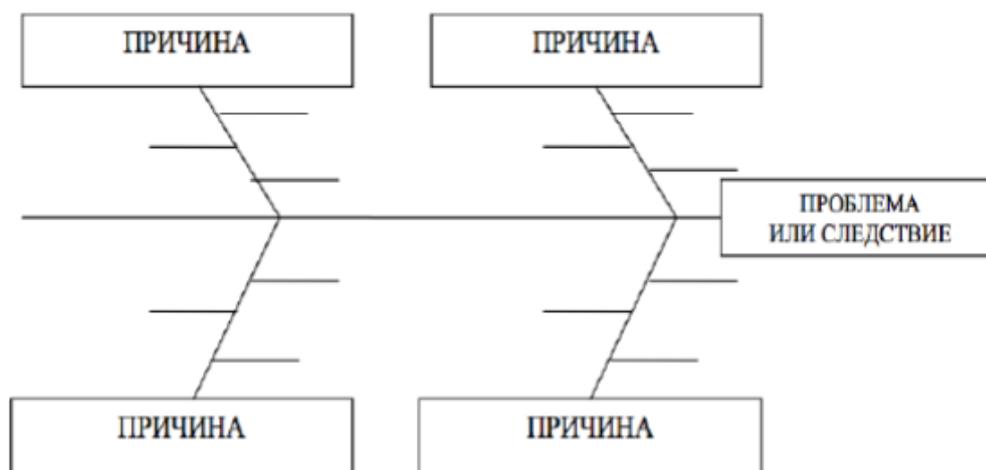


Рис. 2 – Схематическое изображение диаграммы Исикавы [3]

Для контроля качества на производстве в 1960-х годах Каору Исикава разработал диаграмму причин и следствий, известную как диаграмма Исикавы. Диаграмма классифицирует причины на категории. Эти категории затем подразделяются. Метод Исикавы определяет «причины» в мозговых штурмах, проводимых различными группами, вовлеченными в производственный процесс.

Таким образом, можно сделать вывод, что на данный момент каузальность является широко применимой, но слабоизученной областью.

При грамотном применении, теория каузальности связей может найти свое применение в таких сферах как медицина, биология, химия, физика, лингвистика, экономика, юриспруденция и так далее.

К сожалению, рассмотренные в данной статье методы не способны покрыть весь список проблем, возникающий при анализе причинно-следственных (каузальных) связей, слабо охватывает проблему исследования нечетких и вероятностных величин, тяжело анализирует лингвистические переменные.

Сейчас необходимо большое количество исследований в области каузальных связей, чтобы унифицировать методы выявления степени

каузальности связей, применительно к строгим научным дисциплинам, перехода от метафизики и философии к физике и экономике.

Библиографический список:

1. Алекси Ю., Амбрози М., Локайичек М. В. Точность и каузальность в экономике – Киев: Институт общественной трансформации, 2017 – с.4-9
2. Киргизова И.А. Взаимосвязи комплекса факторов окружающей среды и здоровья человека в региональном аспекте/ Киргизова И.А.//ISSN Научный альманах – 2017 – № 4-3(30) [Электронный ресурс] – Режим доступа: <https://ukonf.com/doc/na.2017.04.03.pdf>
3. Логунова О.Е. Применение причинно-следственной диаграммы Исикавы в репутационном менеджменте/Логунова О.Е.//Научные исследования – 2015 – №1 [Электронный ресурс] – Режим доступа: <https://cyberleninka.ru/article/n/primenenie-prichinno-sledstvennoy-diagrammy-isikavy-v-reputatsionnom-menedzhmente/viewer>
4. Митина О.В. Методы исследования каузальных связей// Экспериментальная психология в России: традиции и перспективы – М: Институт психологии РАН, 2019 – с. 139-143
5. Costa F., Shrapnel S. Quantum Causal Modelling//New Journal of Physics – 2016 – [Electronic resource] – URL: <https://iopscience.iop.org/article/10.1088/1367-2630/18/6/063032/meta>
6. Leifer M. S., Spekkens R. W. Towards a Formulation of Quantum Theory as a Causally Neutral Theory of Bayesian Inference//Physical Review A 88 – 2013 – [Electronic resource] – URL: <https://journals.aps.org/pr/abstract/10.1103/PhysRevA.88.052130>
7. Spurious Correlations – [Electronic resource] – URL: <https://tylervigen.com/spurious-correlations>

Оригинальность 89%