

Российская академия наук
Институт экономики

Л.Н. Слуцкий

МОДЕЛИРОВАНИЕ ПРИЧИННЫХ
ЗАВИСИМОСТЕЙ В ЭКОНОМИКЕ

Москва
Институт экономики
2019

Слуцкий Л.Н. Моделирование причинных зависимостей в экономике (научный доклад) – М: Институт экономики РАН, 2019. – 42 с.

С 49

Известно, что причинные связи в социальных науках, включая экономику, носят стохастический характер. В отличие от естественных наук экономические наблюдения практически никогда невозможно воспроизвести в точности при тех же самых условиях, равно как и изолировать эффект действия одного фактора на другой от влияния всех остальных факторов, что создает, на первый взгляд, непреодолимую трудность для установления статистически проверяемых причинных связей между экономическими феноменами. Мы покажем, как эконометристы, начиная с Фишера и Неймана и заканчивая Рубиным, Перлом, Шолкопфом и их последователями, методично преодолевали эту проблему.

В докладе также освещаются вопросы приложения теории статистической причинности к экономическим задачам. Подробно описывается концепция интервенции (оператор «do») в работах Перла и показывается, что данное понятие может эффективно использоваться для вычисления синергетического эффекта, возникающего в архитектуре байесовской сети.

Ключевые слова: байесовская сеть, условная вероятность, инструментальная переменная, интервенция, потенциальный исход, рандомизация, синергетический эффект.

Классификация JEL: C18, C51.

Slutskin L.N. Modeling Causal Dependencies in Economics (report). М: IE RAS, 2019.

It is well known that causal relations in the social sciences including economics have stochastic nature. Unlike natural sciences it is hardly possible to reproduce economic observations under the same conditions or isolate effects of the action of one factor upon the other from the rest. On the surface this fact should be the cause of considerable difficulties on the way of establishing statistically testable causal relations between two economic phenomena. We show how econometricians have step by step overcome this problem beginning with works of Fisher and Neyman and ending with the latest achievements in the field by Rubin, Pearl, Scholkopf and their colleagues.

In the presentation the issues related to applications of statistical causality theory to economics are also considered. Particularly, the concept of intervention (operator “do”) invented by Judea Pearl will be systematically studied. We show how interventions may be effectively used to calculate synergetic effects appearing in the architecture of Bayesian network.

Keywords: Bayesian network, conditional probability, instrumental variable, intervention, potential outcome, randomization, synergetic effect.

JEL Classification: C18, C51.

© Слуцкий Л.Н., 2019

© Институт экономики РАН, 2019

© Валерий В.Е., дизайн, 2007

ОГЛАВЛЕНИЕ

Введение	4
Глава I. Вероятностное обоснование экономических процессов	7
1. Стохастическая экономика	7
2. Инструментальные переменные	9
Глава II. Модель причинности Рубина	13
1. Рандомизация	13
2. Потенциальные исходы	14
3. Средний эффект воздействия	15
Глава III. Байесовские сети	18
1. Марковские свойства сетей	18
2. Оператор «do»	24
3. Определение синергетического эффекта	30
4. Критика байесовских сетей	32
Глава IV. Качественный сравнительный анализ	34
Заключение	37
Литература	40

ВВЕДЕНИЕ¹

По всей видимости, одновременно с появлением на свет *Homo sapiens* у человека появилась потребность в постижении причинности событий и вещей. Во всяком случае, краеугольным камнем одного из самых ранних дошедших до нас трудов античной философии было учение о четырех причинах Аристотеля (*Аристотель*, 1976). Стремление понять причину явлений лежит в основе всех естественных наук. То же относится и к экономической науке. Достаточно сказать, что слово «причина» входит в название основного труда Адама Смита — «Исследование о природе и причинах богатства народов» (*Смит*, 1962).

Безусловно, любой труд, посвященный вопросам причинности, должен включать в себя историю и философскую обоснованность этого понятия. Однако автор, не будучи специалистом ни в истории, ни в философии, поставил перед собой гораздо более скромную задачу — дать представление (и по возможности развить его) о теории и практическом применении основных эконометрических моделей, используемых в настоящее время для описания причинных связей в экономике.

Эконометрическая теория причинности исходит из того, что как сами экономические переменные, так и связи между ними носят случайный характер. В частности, это означает, что действие события-причины *B* на событие-следствие *A* может проявляться при известных обстоятельствах, но может *при тех же самых обстоятельствах* отсутствовать².

-
1. Автор выражает благодарность профессору Александру Либману за внимательное прочтение первоначального варианта доклада и сделанные им ценные замечания и дополнения.
 2. Это может быть объяснено двумя различными способами:
 - причинная связь между *A* и *B* задается вероятностным механизмом, как, например, в квантовой механике;
 - в действительности — *при полном знании* всех сопутствующих обстоятельств — действие *B* на *A* является детерминированным, и представление связи как случайной является лишь приближением к действительности, основанном на отсутствии такого полного знания.При построении эконометрической модели мы обычно исходим из первого подхода, при ее интерпретации — из второго.

Здесь сразу же следует оговориться. Как правило, эконометриста интересует не столько наличие или отсутствие связей между объектами, сколько величина эффекта воздействия одного фактора на другой. Если это воздействие очень мало, то влияние B на A считается несущественным³.

В своей статье Д. Рубин (*Rubin*, 1974) обобщил идеи Дж. Неймана (1923) и Р. Фишера (1925) рандомизации экспериментов на наблюдения, которые невозможно воспроизвести опытным путем. Модель причинности Рубина основана на понятии потенциального исхода A под воздействием причины B , при котором одновременно с наблюдением рассматривается также его мысленный прототип: какой бы был результат A при отсутствии воздействия B ? Рубин реализовал свой подход путем рассмотрения так называемых контрольных факторов, которые должны входить в явном виде в регрессионное уравнение зависимости A от B .

Следующим шагом использования статистических методов для проверки причинных зависимостей в условиях невозпроизводимости наблюдений был подход, в основном реализованный в 1990-х гг. Дж. Перлом (*Pearl*, 2009) и С. Лоритценом совместно с его коллегами (*Cowell et al.*, 2003), который основывается на понятии марковских свойств совокупности переменных (факторов). При этом предполагается, что переменные образуют априорно заданную экономистом архитектуру отношений (сеть). Марковское свойство заключается в том, что непосредственные, т.е. ближайшие по воздействию на A факторы, информационно блокируют все остальные, более удаленные от A причины. Таким образом, сетевой подход является распространением причинных связей с двух факторов A и B на сколь угодно большую группу переменных. Другими словами, чтобы понять механизм воздействия B на A , требуется определить их место в глобальной системе взаимоотношений во всей сети.

Несмотря на значительный прогресс в статистической теории причинности в упомянутых выше исследованиях, основной и в то же время самый простой вопрос, можно ли, обладая необходимым набором статистических наблюдений, решить, какой из двух

3. Наличие такого воздействия определяется посредством различных статистических тестов с использованием p -значений соответствующих статистик.

объектов *A* и *B* является причиной, а какой следствием, вплоть до недавнего времени оставался открытым. Новые подходы к решению этой проблемы наметились в последние десять лет в работах Б. Шолкопфа и его учеников.

Все вышеупомянутые теории будут последовательно рассмотрены в докладе. В первой главе кратко излагается эконометрический подход к экономическому феномену, основанный на теории вероятности и математической статистике. Здесь же будет дано описание метода инструментальных переменных, который содержит в себе все необходимые ингредиенты для понимания изучаемой проблематики. Во второй главе дано развернутое изложение модели причинности Д. Рубина с указанием на связь с ранними работами Р. Фишера и Дж. Неймана. Третья глава посвящена критическому разбору идей Дж. Перла и его учеников, с особым акцентом на концепцию оператора «*do*». Здесь же будет представлен новый подход к определению и вычислению синергетического эффекта в байесовской сети. В четвертой главе описывается методология качественного сравнительного анализа (КСА), предназначенного для выявления причинных связей между категориальными переменными. Хотя КСА не изучает вероятностные свойства факторов и, таким образом, не относится собственно к эконометрике, его методы нашли в последние годы широкое распространение в различных гуманитарных науках. Это также даст нам возможность сравнить вероятностные и невероятностные подходы к теории причинности.

ВЕРОЯТНОСТНОЕ ОБОСНОВАНИЕ ЭКОНОМИЧЕСКИХ ПРОЦЕССОВ

1.1. Стохастическая экономика

В период между тридцатыми и сороковыми годами прошлого века, в основном в работах Р. Фриша, Т. Купманса, Я. Тинбергена, Т. Хаавелмо (*Frisch et al.*, 1948; *Haavelmo*, 1944), окончательно сформировалась эконометрическая наука, трактующая экономику как стохастическую систему с бесконечным (или, по крайней мере, очень большим) числом переменных. Эта система, эволюционируя во времени, находясь в статистической связи со всеми или частью своих предшествующих состояний, представляет собой экономический процесс. Статистическая теория причинности, с одной стороны, существенно опирается на эконометрическую теорию, а с другой — не являясь ее следствием, требует совершенно новых идей и методов для своего обоснования.

Чтобы лучше понять последнее утверждение, предположим, что эконометрист хочет установить причинную зависимость между экономическими переменными A и B . Исходя из вероятностного распределения

$$p(A, B), \quad (1.1)$$

связывающего эти два фактора, можно получить уравнение регрессии A по B :

$$A = f(B) + \varepsilon, \quad (1.2),$$

где ε — ошибка, некоррелированная с B .

Затем исследователь определяет вид функциональной зависимости $f(B)$ и распределение $p(\varepsilon)$ ошибки ε^4 .

Отличие представлений (1.1) и (1.2) и составляет принципиальную разницу между вероятностным и причинным подходами.

4. Нахождение (точнее, оценка) f и $p(\varepsilon)$ по конечному числу наблюдений (A_i, B_i) , $i=1,2,\dots,n$ может оказаться совсем не простой задачей. В прикладных исследованиях f обычно предполагается линейной функцией, а $p(\varepsilon)$ — нормальным распределением.

В то время как распределение $p(A,B)$ задает частоту⁵ совместных наблюдений событий A и B и представляет собой чисто статистическое понятие, при котором асимметрия в записи $p(A,B)$ является не более чем условностью, асимметрия между A и B в формуле (1.2) отражает причинно-следственную зависимость и должна быть задана априорно эконометристом⁶.

Так, на рис. 1 частотное распределение наблюдений представлено графически множеством точек на плоскости. Вопрос же о причинной зависимости является дискуссионным (Маевский, Слуцкий, 2009. С. 16).

Таким образом, статистическая теория причинности опирается на вероятностные методы, изучающие частоту и отношения между наблюдениями, а не наоборот. Причинные связи накладывают дополнительную структуру на вероятностное распределение $p(A, B)$, заключающуюся в указании направления от причины к следствию.

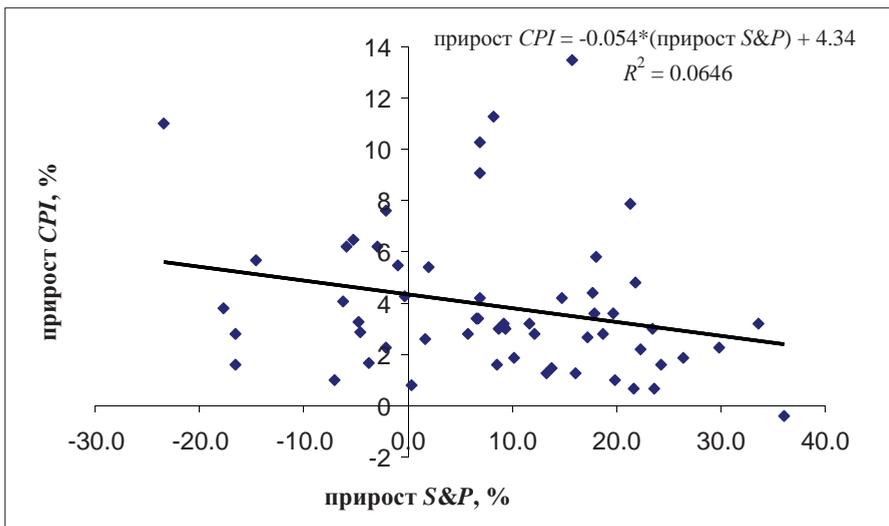


Рис. 1. Частотное распределение двух переменных B (прирост $S\&P$) и A (прирост CPI)

Источник: (Маевский, Слуцкий, 2009. С. 18).

5. Мы здесь исходим из частотного (фриквентистского) подхода к понятию вероятности.

6. Это следует хотя бы из того, что распределение (1.1) может быть также записано в виде: $B = g(A) + \zeta$.

В прикладных исследованиях для определения причинных связей часто строится регрессионное уравнение между объясняемой переменной (Y) и объясняющими факторами X_1, \dots, X_m :

$$Y = k_0 + k_1 X_1 + \dots + k_i X_i + \dots + k_m X_m + \varepsilon. \quad (1.3)$$

Коэффициенты регрессии k_0, k_1, \dots, k_m рассчитываются обычно с помощью метода наименьших квадратов (МНК), который требует знание, по крайней мере, $(m+1)$ наблюдений, включающих в себя как значения объясняемой переменной Y , так и регрессоров X_1, \dots, X_m ⁷. Если коэффициент k_i при переменной X_i равен нулю, то заявляется, что X_i не является причиной (одной из причин) Y . В противном случае считается, что X_i служит причиной для Y с величиной эффекта воздействия равной $k_i \neq 0$. В основе данного метода лежит простой принцип, в соответствии с которым, чтобы узнать, является ли данный фактор X причиной другого фактора Y , надо наблюдать, как меняются значения Y при изменении воздействия X на Y . Коэффициент k_i показывает, насколько в среднем изменилась величина Y при увеличении X_i на единицу (при контроле над всеми остальными регрессорами). Такой подход иногда дает верные результаты, но, как мы увидим уже в следующем пункте, далеко не всегда.

1.2. Инструментальные переменные

Как мы уже отмечали в введении, методология применения инструментальных переменных содержит в себе все необходимые элементы для понимания феномена причинности и сопряженных с этим трудностей.

Впервые метод инструментальных переменных был применен американским экономистом Филиппом Райтом⁸ (*Wright*, 1928) для решения задачи определения импортного тарифа на животные и растительные масла и жиры. Для этого Райту требовалось определить эластичность спроса по цене. Рассмотрим уравнение регрессии:

-
7. Мы не рассматриваем в данной работе проблему определения точности оценок. Также предполагается, что все регрессоры, включая константу, линейно независимы.
 8. Интересный исторический отчет о приоритете открытия метода инструментальных переменных приведен в книге Стока и Уотсона (*Сток, Уотсон*, 2015).

$$\ln q_i = k_0 + k_1 \ln p_i + \varepsilon_i, \quad (1.4)$$

где q_i, p_i – равновесные значения потребленного масла и цены в период i , ε_i – ошибка. Так как p и q в уравнении (1.4) взаимозависимы, то оценка коэффициента k_1 методом наименьших квадратов будет смещенной⁹. Чтобы справиться с этой проблемой, Райт предложил использовать новую переменную, которая бы оказывала влияние на цену p , но была бы некоррелирована с другими переменными, действующими на спрос q , т.е. была бы некоррелирована с ошибкой ε ¹⁰. В качестве такой переменной он взял количество осадков Z , выпавших за соответствующий период. Коэффициент эластичности k_1 оценивается двухшаговым МНК.

Вернемся к задаче установления причинности между двумя переменными A и B . Предположим, что причинная связь возможна только в направлении от B к A ($B \rightarrow A$). Например, B предшествует A по времени, или A не может быть причиной B , основываясь на экономической теории¹¹. Конечно, самым простым решением было бы оценить на равенство нулю коэффициент корреляции между этими переменными, или коэффициент при B при регрессии A по B ¹². Однако здесь следует учитывать возможность наличия *общей причины* (*confounding*) C для обеих переменных A и B .

В качестве примера рассмотрим причинную связь между спросом на мороженое и уровнем ртутного столба в наружном термометре в летнее время. Коэффициент корреляции, возможно, будет весьма значительным, но это, конечно, не означает, что один из них является причиной другого. Общий фактор – наличие жарких летних дней – влияет на обе переменные, тем самым, создавая мнимую корреляцию. Поэтому, многие учебники по эконометри-

9. Другими возможными источниками смещения оценки k_1 будут:
 - 1) систематические ошибки в данных;
 - 2) пропущенные переменные, т.е. те, которые влияют одновременно на p и на ошибку ε .
10. Ошибка ε интерпретируется как агрегированный показатель всех факторов, кроме p , влияющих на q .
11. Также следует упомянуть два важных принципа, сформулированных Т. Купмансом в рамках трудов Комиссии Коулса (*Coortmans*, 1950): 1) факторы, не являющиеся экономическими феноменами, должны рассматриваться как экзогенные; 2) эндогенные переменные могут быть следствиями экзогенных, но не наоборот. Заметим, что ввиду возросшей за последние пятьдесят лет роли, которую играет экологическая проблематика в обществе, первый принцип утратил частично свою значимость.
12. Из математической статистики известно, что оба метода приводят к одному и тому же результату.

ке начинаются с максимы: **корреляция не означает причинность**¹³.

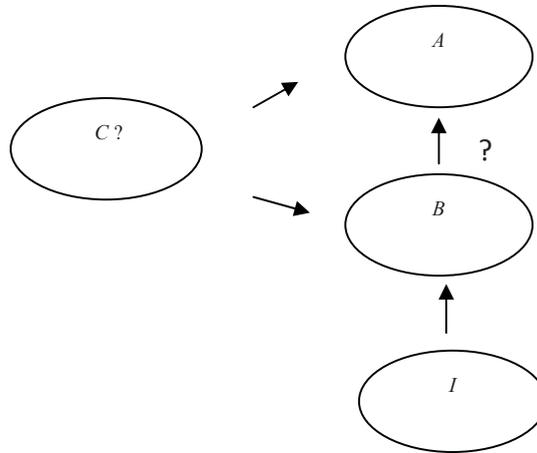


Рис. 2. Инструментальная переменная I для проверки наличия причинной связи по направлению от B к A , если известно, что коэффициент корреляции $\rho(A, B) \neq 0$

Таким образом, если исследователь уверен, что общая причина отсутствует, то коэффициент регрессии при B проверяется на значимость, и на основании этого делается вывод о наличии причинной связи по направлению от B к A ¹⁴. Однако если такой уверенности нет, то он может воспользоваться вышеописанным методом инструментальных переменных. То есть найти переменную (инструмент) I , которая коррелирована с B , но некоррелирована с другими факторами, действующими на A . Другими словами, переменная I связана с A только посредством переменной B (см. рис.

-
13. К сожалению, во многих учебниках по статистике в качестве примера, подтверждающего вышеуказанное правило, приводятся не вероятностные события, а случайные процессы. Например, наличие положительной корреляции между численностью населения страны и ростом цен на длительном промежутке времени. Хотя вывод об отсутствии причинной связи является правильным, такое сравнение не совсем корректно, так как в данном случае речь идет о случайных процессах, для которых корреляция будет не числом, а функцией от времени.
14. Наличие общей причины C еще не означает, что B не воздействует на A . Оба фактора B и C могут быть причинами A .

2). Наличие значимой корреляции между I и A указывает на то, что фактор B является причиной A . В задаче (1.4) Райт с помощью инструментальной переменной Z смог оценить величину эффекта воздействия логарифма цены p на логарифм спрос q при наличии взаимного влияния друг на друга этих двух факторов¹⁵.

Мы можем сформулировать следующие два принципа, на которых основан метод инструментальных переменных:

1) принцип общей причины Райхенбаха (Reichenbach's common cause principle). Если факторы A и B коррелированы, то либо один из них является причиной для другого, либо существует третий фактор C , являющийся общей причиной A и B . При этом оба фактора становятся независимыми, если контролировать C . Последнее условие означает, что значения, принимаемые переменной C , являются тем единственным общим, что связывает A и B ¹⁶;

2) если существует фактор I , который является причиной A , но может воздействовать на A только через B , то B является причиной A .

Как мы видим, задача определения наличия причинной связи от B к A свелась к поиску другой причины, а именно инструмента I , воздействующего на B . К сожалению, такие инструменты часто бывает трудно найти. Поэтому в прикладных исследованиях стараются найти группу регрессоров X_1, \dots, X_m , действующих на A вместе с B , автономно от всех остальных факторов, воздействующих на A , т.е. некоррелированных с ними¹⁷. Коэффициент c при B в уравнении регрессии:

$$A = k_0 + cB + k_1X_1 + \dots + k_iX_i + \dots + k_mX_m + \varepsilon \quad (1.5)$$

представляет собой «истинную» величину эффекта воздействия B на A .

-
15. Другой подход для оценки величины эффекта воздействия одного фактора на другой, при наличии взаимного влияния, основан на рассмотрении системы эконометрических уравнений (Koopmans, 1950).
 16. Мы привели принцип общей причины так, как он был сформулирован Райхенбахом (Reichenbach, 1956). В современной трактовке фактор C считается вектором, состоящим из набора причин X_1, \dots, X_k . Другие возможные причины коррелированности A и B в конечных выборках обсуждаются в (Peters et al., 2017).
 17. В действительности достаточно, чтобы B было некоррелировано с остальными факторами при контроле над X_1, \dots, X_m .

МОДЕЛЬ ПРИЧИННОСТИ РУБИНА

2.1. Рандомизация

Для проверки наличия причинной связи по направлению от фактора B к A можно воспользоваться следующей простой схемой: сравнить свойства A при воздействии на него B и при его отсутствии. В отличие от метода, описанного в первой главе, где воздействие B на A могло меняться непрерывно, здесь B принимает ровно два значения: «включен», «выключен».

Так, например, при тестировании нового лекарства можно сравнить симптомы у пациентов, которым оно было предписано, с больными, которые принимали плацебо. Но здесь следует быть очень осторожным при отборе пациентов для проверки медикамента. Возможно, что симптомы будут различными у мужчин и женщин, зависеть от возраста и т.д. Чтобы нейтрализовать индивидуальные различия и выявить эффективность лекарства, следует выбирать пациентов для исследования случайным образом¹⁸. При таком подходе каждый больной имеет равный шанс стать участником эксперимента.

Описанный нами эксперимент является воплощением идей двух выдающихся статистиков своего времени Дж. Неймана (*Neuman*, 1923) и Р. Фишера (*Fisher*, 1925, 1935)¹⁹. Он получил название *рандомизации*. К сожалению, рандомизацию можно применять лишь в отдельных областях науки, например, в медицинских или биохимических исследованиях.

18. Например, вытаскивая из шляпы после тщательного перемешивания бумажки с фамилиями тех пациентов, которые согласились участвовать в эксперименте. В современных исследованиях для рандомизации используются более продвинутые методы.

19. Следует отметить, что их идеи были во многом предвосхищены в трудах американского философа и статистика Чарльза Пирса (*Peirce*, 1883; *Peirce, Jastrow*, 1885).

2.2. Потенциальные исходы

Рассмотрим следующий простой пример, когда рандомизация неприменима. Предположим, что мы хотим изучить, как курсы переподготовки для сотрудников предпенсионного возраста некоторого предприятия влияют на их дальнейшие возможности продолжения работы по соответствующим специальностям после окончания обучения. Для этого мы исследуем две группы людей: тех, кто закончили обучение на курсах, и тех, кто по тем или иным причинам в программе не участвовали. Далее, предположим, что средняя продолжительность дальнейшей работы по специальности у людей, прошедших обучение, в 1,5 раза выше, чем у тех, кто не обучался. Можем ли мы сделать вывод, что курсы переподготовки оказали существенное положительное влияние на продление времени пребывания в профессии? Хотя это и может показаться парадоксальным, ответ отрицательный. Другие факторы, кроме обучения, могли оказать определенное влияние на результаты эксперимента. Так, например, на курсы записывались лишь те работники, которые считали себя способными составить конкуренцию более молодым специалистам, т.е. наиболее знающие и трудолюбивые. Те же, кто в себя не верил, на курсы не записались, так как считали, что обучение будет только потерей времени. Но возможно представить себе и другую ситуацию. Наиболее знающие и талантливые считали, что курсы им ничего не дадут, что за долгие годы работы они приобрели все необходимые знания и навыки. На курсы же записывались слабые специалисты, которые надеялись с их помощью повысить свою профессиональную пригодность. В таком случае статистика, вероятно, была бы совершенно другой: продолжительность дальнейшей работы по специальности у людей (прошедших обучение) была бы ниже, чем у тех, кто не обучался. И это тоже было бы неверной оценкой полезности и эффективности курсов.

Конечно, запись людей на курсы случайным образом решила бы проблему. Но рандомизированный эксперимент здесь, естественно, неуместен. Было бы неэтично и даже, возможно, незаконно проводить подобные опыты.

Чтобы оценить эффект обучения, можно предложить следующий подход к решению задачи — сравнить каждого участника

программы с ним же самим, при условии, что он на курсах не обучался. Но это, конечно, невозможно! Тем не менее, Д. Рубин (*Rubin*, 1974) предложил рассматривать подобный мыслимый эксперимент, результаты которого он назвал *потенциальными исходами* (*potential outcome*).

Как уже было сказано в предыдущем пункте, рандомизированный эксперимент при определении воздействия фактора *B* на *A* дает хорошие результаты²⁰, потому что он нейтрализует действие всех других факторов, которые могут оказать воздействие на фактор *A*. Таким образом, если контролировать характеристики участников программы, т.е. фиксировать, все факторы, кроме *B*, действующие на *A*, то мы можем считать, что участники определялись случайным образом. Другими словами, предполагается, что сотрудники с одинаковыми характеристиками имели равные шансы участия в эксперименте. В нашем случае этого можно добиться, взяв под контроль такие факторы, как возраст участников, образование, общий стаж и стаж работы по профессии, возможно, и другие значимые. Тогда можно будет сравнить продолжительность дальнейшей работы по специальности у обеих групп с одинаковым набором характеристик. Такое сравнение называется *мэтчингом* (*matching*), а сам эксперимент — *квазиэкспериментом*. В отличие от рандомизированного эксперимента, в случае квазиэксперимента состав участников программы не формируется случайным образом, а зависит либо от желания самих участников, либо от других факторов, не влияющих на результат обучения. Тем не менее для того, чтобы выводы об эффективности обучения можно было распространить на всех работников, каждый из них должен иметь ненулевой шанс участия в программе²¹.

2.3. Средний эффект воздействия²²

В предыдущем пункте мы рассмотрели эксперимент, при котором все участники разбиваются на две группы: эксперимен-

-
20. Качество результатов зависит от количества участников: чем больше участников в каждой группе, тем более объективным будет результат.
 21. Например, если имеется запрет на участие в программе сотрудников с определенным набором характеристик, то выводы об ее эффективности не могут быть распространены на всех работников.
 22. Более полные обзоры различных методов определения среднего эффекта воздействия имеются, например, в статье (*Ениколопов*, 2009) и в монографии (*Wooldridge*, 2010).

тальную, или активного воздействия (участников программы), и контрольную – тех, кто в программе не участвует. Для определения наличия эффекта воздействия участия в программе Рубин (Rubin, 1974) использовал *средний эффект воздействия* (*average treatment effect – ATE*), который определяется по формуле:

$$ATE = \frac{1}{N} \sum_{i=1}^{i=N} [Y_i(1) - Y_i(0)], \quad (2.1)$$

где $Y_i(0)$, $Y_i(1)$ – результаты (время продолжения работы по специальности) для i -го сотрудника, прошедшего обучение, и из контрольной группы соответственно; N – общее количество работников, как участвовавших, так и не участвовавших в эксперименте.

Несмотря на кажущуюся простоту, формула (2.1) имеет одну принципиальную трудность: если i -й сотрудник прошел обучение, то он не может числиться в контрольной группе; и, наоборот, если он находится в контрольной группе, то он, по определению, не участвовал в программе. Таким образом, в каждой разности под знаком суммы в (2.1) один из членов – $Y_i(0)$ или $Y_i(1)$ – фактически отсутствует, т.е. является потенциальным исходом.

Одним из методов нахождения АТЕ является мэтчинг, основанный на сравнении $Y_i(0)$ и $Y_i(1)$ для обеих групп с одинаковым набором характеристик, который был описан в предыдущем пункте. К сожалению, этот метод дает удовлетворительные результаты только для групп с очень большим количеством участников, когда для каждого набора характеристик, представленного в группе активного воздействия, найдутся работники с таким же набором в контрольной группе²³. Более распространенным является метод линейной регрессии с помощью следующего алгоритма:

1) для каждой из групп строятся регрессии по фактическим данным для $Y(0)$ и $Y(1)$ по контрольным переменным $X = (X_1, X_2, \dots, X_k)$:

$$Y_m(0) = a_0 + a_1 X_{1m} + a_2 X_{2m} + \dots + a_k X_{km}, \quad m=1, \dots, N_0; \quad (2.2)$$

$$Y_n(1) = b_0 + b_1 X_{1n} + b_2 X_{2n} + \dots + b_k X_{kn}, \quad n=1, \dots, N_1, \quad (2.3)$$

23. Эту трудность может быть преодолена при определенных условиях (Wooldridge, 2010).

- где $X_{1m}, X_{2m}, \dots, X_{km}$ – набор характеристик для m -го сотрудника из контрольной группы, $X_{1n}, X_{2n}, \dots, X_{kn}$ – набор характеристик для n -го сотрудника, прошедшего обучение; N_0, N_1 – численность контрольной и экспериментальной группы соответственно;²⁴
- 2) для каждого сотрудника i определяется его потенциальный исход, т.е. если он из контрольной группы, то вычисляется $Y_i(1)$ из формулы (2.3) для экспериментальной группы; если же он из экспериментальной группы, то вычисляется $Y_i(0)$ из формулы (2.2) для контрольной группы;
 - 3) находится АТЕ по формуле (2.1);
 - 4) средний эффект воздействия, полученный в п. 3, проверяется с помощью статистического теста на отличие от нуля.

Следует отметить, что подобный алгоритм может быть применен для оценки эффекта воздействия не только для множества всех работников, как в случае с АТЕ, но и для специфических подгрупп, представляющих его подмножества. Например, мы можем вычислить эффекты воздействия отдельно для мужчин и для женщин, а затем их сравнить. Однако надо помнить, что для надежности результатов каждая из подгрупп должна иметь значительное количество участников²⁵.

Модель потенциальных исходов используется в связанных с различными программами исследованиях, не требующих обязательного участия или не носящих универсальный характер. Так, она может быть применена для оценки пилотных проектов. Тем не менее, область приложения модели в экономике остается весьма ограниченной.

24. В действительности, здесь следует производить цензурированную регрессию, так как часть сотрудников на момент сбора информации будут продолжать работать по специальности.

25. Для того чтобы приведенный алгоритм давал достоверные результаты, необходимо чтобы результаты эксперимента определялись исключительно навыками и умениями, полученными участниками программы, и не зависели от побочных постпрограммных влияний, таких как, например, сознательное поощрение со стороны администрации заключения контрактов с сотрудниками, прошедшими обучение, и т.д.

БАЙЕСОВСКИЕ СЕТИ

3.1. Марковские свойства сетей

При сетевом подходе причинная связь по направлению от фактора *B* к *A* изучается в общем контексте отношений, архитектуры сети, между экзогенными и эндогенными переменными, образующими среду, в которой происходит формирование обоих факторов (см. рис. 3)²⁶.

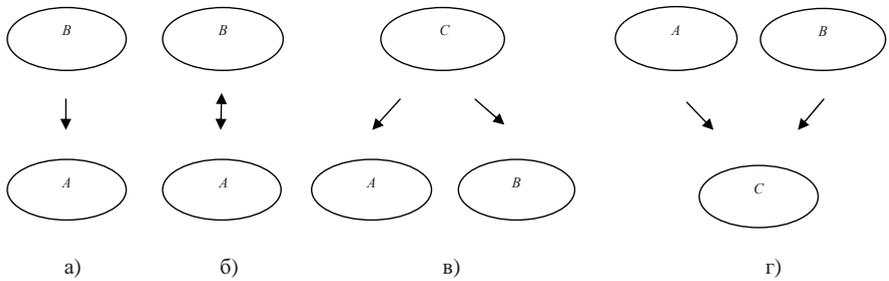


Рис. 3. Различные схемы связей факторов *A* и *B* в сети:

а) B является причиной A; б) A и B воздействуют друг на друга; в) C воздействует на A и B, создавая между A и B ассоциативную связь; г) A и B создают коллизию в C

Вероятностные сети (графы) могут иметь весьма сложный и причудливый характер. Мы будем рассматривать только сети, в которых причинные связи представлены стрелками между вершинами. Отсутствие стрелок означает отсутствие причинной связи²⁷. Наиболее простым примером причинной сети являются *байесовские сети*, которые имеют ряд весьма разумных свойств, значительно упрощающих работу с ними:

26. Если *B* является экзогенной переменной, то рассматривается только формирование фактора *A*.

27. Однако наличие стрелки не означает наличие причинной связи, а только возможное воздействие одного фактора на другой.

- 1) каждая вершина графа представляет собой некий фактор (группу факторов), который является случайной величиной. При этом один и тот же фактор в разные моменты времени может быть представлен различными вершинами;
- 2) не существуют замкнутых путей в графе, т.е. никакой фактор не может быть своей, даже опосредованной, собственной причиной;²⁸
- 3) данные сети обладают марковским свойством.

Чтобы понять последнее свойство введем понятие прямой и опосредованной причины. *Прямая причина* представлена стрелкой от фактора-причины к фактору-следствию. *Опосредованная причина* может быть представлена только следующим друг за другом набором стрелок, состоящим из более одной стрелки. Аналогично определяются *прямое и опосредованное следствие*.

Марковское свойство заключается в том, что прямые причины для фактора Y информационно блокируют все остальные вершины графа, кроме тех, которые являются прямыми или опосредованными следствиями Y . То есть вся информация о факторе Y , содержащаяся в переменных сети, кроме вершин-следствий Y , полностью заключается в его прямых вершинах. Этот факт может быть сформулирован в терминах регрессионных уравнений:²⁹

$$Y = k_0 + k_1 X_1 + \dots + k_m X_m + \varepsilon, \quad (3.1)$$

где X_1, \dots, X_m — полный набор вершин сети, за исключением Y , и не включающий в себя прямые и опосредованные следствия Y ; k_1, \dots, k_m — набор коэффициентов, при соответствующих вершинах; ε — ошибка, только коэффициенты при прямых причинах Y могут иметь ненулевые значения. Если коэффициент при прямой вершине равен нулю, то это означает, что данный фактор не является прямой причиной для Y , и стрелка от него к Y может быть безболезненно удалена из диаграммы³⁰.

-
28. Из свойства 2 следует, что в байесовской сети могут присутствовать только однонаправленные стрелки.
 29. Для простоты изложения, а также ввиду их важности, мы рассматриваем только линейные уравнения связи.
 30. В действительности, коэффициент k_i при факторе X_i может равняться нулю, но он, тем не менее, является одной из причин для Y . Это достигается в состоянии равновесия с действием ряда

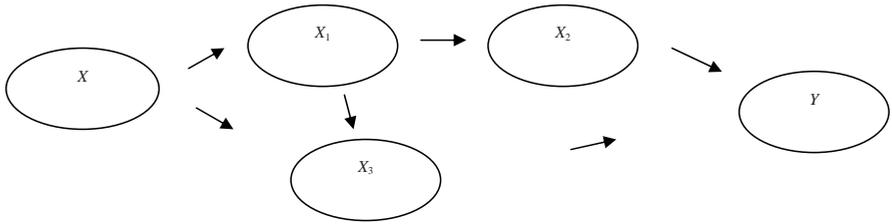


Рис. 4. Пример байесовской сети, состоящей из пяти вершин

Источник: (Слуцкий, 2017. С. 13).

На рис. 4 представлена байесовская сеть, состоящая из пяти вершин: X , X_1 , X_2 , X_3 , Y . Вершина X_1 является прямой причиной для X_2 и опосредованной для Y . Аналогично X_2 — прямое следствие X_1 и опосредованное X . Из марковского свойства следует, что вся информация об Y , содержащаяся в сети, заключена в факторах X_2 и X_3 . Другими словами, если известны значения X_2 и X_3 , то знание об остальных переменных X и X_1 не добавит новой информации о вероятностном распределении Y . Также следует отметить, что есть три пути, ведущих от X к Y : XX_1X_2Y , XX_3Y и XX_1X_3Y и только один путь от X к X_2 — XX_1X_2 .

Примером байесовской сети является модель VAR, где для каждой вершины X_t , представляющей собой вектор случайных переменных, набор прямых причин представлен лагами X_{t-1} , X_{t-2} , ..., X_{t-m} , где m — число лагов, одинаковое для всех t .

Байесовская сеть, может быть представлена аналитически в виде системы *структурных уравнений*, когда каждой вершине, за исключением экзогенных, т.е. не имеющей входящих стрелок, ставится в соответствие уравнение регрессии этой вершины по всем ее прямым причинам. И наоборот, если задана система регрессионных уравнений с взаимно независимыми ошибками³¹, мы можем построить байесовскую сеть, в которой каждая переменная является вершиной графа, а регрессоры для каждого уравне-

других причин, действующих на Y , наряду с X . Такое равновесное состояние не является стабильным, т.е. при малых изменениях в вероятностной структуре сети коэффициент k_i становится отличным от нуля. За деталями мы отсылаем читателя к монографии (Pearl, 2009). В дальнейшем, если это специально не оговорено, мы будем рассматривать только стабильные системы.

31. Экзогенные переменные должны быть независимы как друг от друга, так и от ошибок регрессионных уравнений.

ния будут прямыми причинами для соответствующей объясняемой переменной.

Данное представление байесовской сети можно записать как рекурсивную систему регрессионных уравнений следующим образом. Разобьем вершины графа на соответствующие группы. В первую войдут вершины X_1, \dots, X_{k_1} , не имеющие прямых причин, т.е. экзогенные переменные. Очевидно, что найдется хотя бы одна вершина, обладающая этим свойством. Во вторую группу поместим вершины $X_{k_1+1}, \dots, X_{k_2}$, $k_1 + 1 \leq k_2$, которые в качестве своих причин имеют только вершины из первой группы. Продолжая таким образом, мы получим, что в группу m , $m > 1$ войдут вершины графа, которые в качестве своих причин (прямых или опосредованных) имеют только вершины из предыдущих $(m - 1)$ групп³². Для каждой переменной из группы m можно записать регрессионное уравнение по всем ее прямым вершинам. Таким образом, мы получим рекурсивную систему уравнений:

$$\begin{aligned} X^{(1)} &= \alpha_1 + \varepsilon_1, \\ X^{(2)} &= \alpha_2 + A_{21}X^{(1)} + \varepsilon_2, \\ X^{(3)} &= \alpha_3 + A_{31}X^{(1)} + A_{32}X^{(2)} + \varepsilon_3, \\ &\dots \end{aligned} \quad (3.2)$$

где $X^{(m)}$ – вектор, образованный вершинами группы m , α_i – вектор свободных членов, A_{ij} – матрица коэффициентов, ε_i – вектор случайных ошибок. Число уравнений в системе (3.2) равно числу вершин в самой длинной причинно-следственной цепочке графа. Ошибки $\varepsilon_1, \varepsilon_2, \dots$ являются взаимно независимыми – как между различными группами, так и внутри каждой группы³³. Последнее утверждение эквивалентно марковскому свойству байесовской сети. Мы также заметим, что каждая вершина графа войдет в качестве независимой переменной ровно в одно уравнение системы (3.2).

При заданной системе (3.2) получим байесовскую сеть, если для каждой, не являющейся экзогенной, переменной X_i мы будем рассматривать в качестве прямых вершин все переменные с коэффициентами, отличными от нуля, в уравнении для X_i . Таким

-
32. Вершина группы m характеризуется тем свойством, что самая длинная причинно-следственная цепочка, идущая от экзогенной переменной до этой вершины, имеет ровно $m-1$ стрелку.
33. Условие независимости ошибок внутри группы не является обязательным (Anderson, 2003). В таком случае вершинами сети будут случайные векторы.

образом, можно сделать вывод, что рекурсивная система регрессионных уравнений (3.2) полностью характеризует байесовскую сеть, заданную на графе³⁴.

Значение представления байесовских сетей в аналитической форме в виде рекурсивной системы структурных уравнений состоит, прежде всего, в том, что, переходя от группы более низкого порядка к следующей, мы можем воспроизвести механизм последовательного формирования переменных. На байесовской сети это соответствует продвижению вдоль причинно-следственных цепочек.

В заключение мы приведем группу аксиом, формально обобщающих свойства байесовских сетей, обозначенных в начале главы.

Аксиома 1. *Между двумя вершинами графа X_1 и X_2 может существовать только одна причинная связь, ведущая от причины к следствию.*

Из данной аксиомы следует, что мы не рассматриваем обратные связи между переменными. Это условие представляется на первый взгляд ограничительным. Одним из решений было бы объединить обе (или более) переменные в один вектор. Другой подход, основанный на рассмотрении цепных графов, т.е. таких, в которых также допускаются неориентированные связи, можно найти в монографии (Cowell et al., 2003).

Аксиома 2. *Не существуют замкнутых причинно-следственных цепочек в графе.*

Другими словами, аксиома 2 утверждает, что никакая переменная не может быть своей собственной, пусть даже опосредованной, причиной. Следующая аксиома выражает марковское свойство сети в терминах теории вероятности.

Аксиома 3. *Причинно-следственные связи обладают марковским свойством, т.е. любая переменная X_1 условно независима по ее прямым причинам от любой другой переменной X_2 , за исключением того случая, когда существует цепочка причинно-следственных связей от X_1 к X_2 . В случае, когда множество пря-*

34. При построении рекурсивной системы уравнений мы предполагали, что все вершины сети являются случайными величинами. Распространение результатов на случайные векторы не представляет значительных трудностей.

мых причин является пустым, условная независимость заменяется на обычную (маргинальную).

В то время как первые две аксиомы отражают геометрический характер сети, третья представляет собой ее вероятностную структуру. Еще одна, четвертая, аксиома будет приведена в следующем пункте.

С прикладной точки зрения важность байесовской сети, кроме наглядного представления причинных связей между переменными, заключается в том, что мы можем проследить влияние одной переменной на другую посредством причинно-следственных цепочек (Wright, 1921). Вернемся к байесовской сети, представленной на рис. 4. Чтобы определить влияние фактора X на Y , рассмотрим пути (последовательности) стрелок, ведущих от X к Y : XX_1X_2Y , XX_3Y и XX_1X_3Y . Каждая причинно-следственная пара на диаграмме определяет коэффициент для соответствующей стрелки – величину эффекта воздействия прямой причины на ее следствие. Так, для пути XX_1X_2Y имеются три коэффициента k_{X,X_1} , k_{X_1,X_2} , $k_{X_2,Y}$, а для пути XX_3Y – коэффициенты k_{X,X_3} , $k_{X_3,Y}$. Теперь мы можем вычислить величину эффекта воздействия X на Y вдоль каждого из этих путей, рассмотрев произведение соответствующих коэффициентов:

$$\begin{aligned} k_{X,X_1,X_2,Y} &= k_{X,X_1} \cdot k_{X_1,X_2} \cdot k_{X_2,Y} \\ k_{X,X_3,Y} &= k_{X,X_3} \cdot k_{X_3,Y} \\ k_{X,X_1,X_3,Y} &= k_{X,X_1} \cdot k_{X_1,X_3} \cdot k_{X_3,Y}. \end{aligned} \quad (3.3)$$

Полный эффект воздействия X на Y определяется как сумма величин эффектов воздействия вдоль всех путей, ведущих от X к Y :

$$k_{X,Y} = k_{X,X_1,X_2,Y} + k_{X,X_3,Y} + k_{X,X_1,X_3,Y} \quad (3.4)$$

Коэффициент $k_{X,Y}$ можно интерпретировать как среднее изменение в факторе Y при увеличении X на единицу³⁵ при контроле над всеми остальными переменными сети, т.е. над факторами, не принадлежащими ни одному из рассматриваемых путей³⁶. При

35. При увеличении Y на h единиц коэффициент $k_{X,Y}$ изменяется, соответственно, на $h \cdot k_{X,Y}$ раз.

36. В нашем случае таких факторов нет.

этом марковское свойство является здесь решающим: изменение прямой причины на h единиц вызывает изменение в эффекте на $k \cdot h$ единиц, не влияя при этом на ошибки в регрессионных уравнениях вдоль путей от X к Y .

Таким образом, причинно-следственные цепочки представляют собой действие *передаточного механизма* от фактора X к Y , а коэффициент k_{XY} — его количественное выражение³⁷.

В заключение отметим важное свойство байесовских сетей: вероятностное распределение всего графа можно представить в виде его декомпозиции по условным распределениям на локальных подграфах (Cowell *et al.*, 2003):

$$p(X_1, \dots, X_n) = \prod_{i=1}^n p(X_i | pa(X_i)), \quad (3.5)$$

где для произвольной вершины X_i $p(X_i | pa(X_i))$ означает условное распределение X_i по множеству прямых причин — $pa(X_i)$ ³⁸, а n — число вершин графа. Так, например, чтобы определить вероятностное распределение вершин сети, представленной на рис. 4, достаточно определить вероятностные распределения на вершинах $\{X\}$, $\{X, X_1\}$, $\{X, X_1, X_3\}$, $\{X_1, X_2\}$ и $\{X_2, X_3, Y\}$.

3.2. Оператор «do»

Если посмотреть внимательно на содержание предыдущего пункта, то становится понятным, что байесовская сеть является чисто математическим, точнее, теоретико-вероятностным, понятием. Именно такой подход доминировал в литературе по вероятностным графам до появления монографии Дж. Перла (Pearl, 2009)³⁹. Идеи Перла вдохнули новую жизнь в статистическую теорию причинности. Но это произошло, в первую очередь, не за счет введения новых математических терминов, хотя они тоже там присутствовали, а как результат объединения понятий из различных областей знаний: философии, логики, экономики и компьютерной науки.

37. В прикладных расчетах, кроме полного эффекта воздействия k_{XY} , можно рассматривать частичные эффекты воздействия вдоль произвольных причинно-следственных цепочек, которые встречаются на пути от X к Y . В нашем случае такими частичными цепочками будут XX_1X_2 , X_1X_2Y , X_2Y и т. д.

38. «*pa*» является сокращением от английского слова *parents*.

39. Первое издание вышло в свет в 1999 г.

Во-первых, Перл переосмыслил концепцию модели регрессии⁴⁰. Рассмотрим снова модель:

$$Y = k_0 + k_1 X_1 + \dots + k_i X_i + \dots + k_m X_m + \varepsilon, \quad (3.6)$$

связывающую фактор Y с его причинами X_1, \dots, X_m . Перл утверждает, что знак равенства в уравнении в действительности выступает в качестве оператора присвоения «:=», используемого при написании компьютерных программ⁴¹. Другими словами, уравнение (3.6) является не равенством, связывающим переменную Y с вектором переменных X_1, \dots, X_m , а способом *формирования* фактора Y на основе других факторов X_1, \dots, X_m . Чтобы понять разницу между этими двумя понятиями, рассмотрим следующую процедуру, название которой Перл, по всей видимости, заимствовал из программирования — оператор «do». Действие этого оператора заключается в закреплении или, в терминологии Перла, *интервенции*, за некоторым фактором X_i определенного числового значения x_0 (записывается $do(X_i = x_0)$)⁴². Затем изучается совместное вероятностное распределение остальных переменных сети. Таким образом, X_i из случайной переменной превращается в параметр системы. Традиционное понимание уравнения (3.6) теряет теперь всякий смысл, так как с заменой случайной величины X_i на конкретное число равенство левой и правой частей уравнения становится невозможным. Это легко увидеть, если сравнить соответствующие дисперсии — при присвоении X_i конкретного значения дисперсия правой части изменяется. Если же мы считаем, что фактор Y формируется переменными, стоящими в правой части уравнения, то при интервенции Y уже формируется меньшим количеством факторов (все оставшиеся переменные, за исключением X_i) и, таким образом, отличается от своего предыдущего состояния. Другими словами, вероятностное распределение $p(Y_{\text{до интервенции}}) \neq p(Y_{\text{после интервенции}})$. Уравнение (3.6) преобразуется в уравнение регрессии:

-
40. Перл (Pearl, 2009) утверждает, что идея формирования переменных уже присутствовала неявно в работах Тинбергена, Хаавелмо и др. (Frisch et al., 1948).
41. Интересно отметить, что Перл, являясь по образованию инженером, лишь позже начал работать в области искусственного интеллекта, что и привело его, в конечном счете, к исследованиям по статистической теории причинности.
42. Интервенция, согласно Перлу, может выражаться как в реальном действии, так и быть результатом мысленного эксперимента.

$$Y = k_0 + k_1 X_1 + \dots + k_i x_0 + \dots + k_m X_m + \varepsilon. \quad (3.7)$$

На рис. 5, 6 представлена байесовская сеть до и после интервенции. Обращает на себя внимание тот факт, что стрелка от вершины X_1 к X_3 на втором рисунке отсутствует. Это объясняется тем фактом, что после интервенции значение фактора X_3 остается фиксированным, и фактор X_1 больше не действует на X_3 .

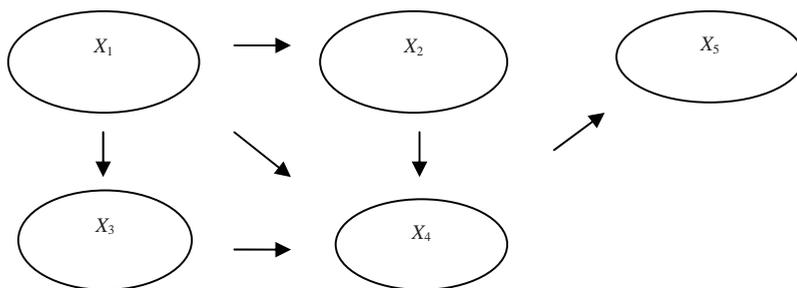


Рис. 5. Байесовская сеть для переменных X_1, \dots, X_5 до действия оператора «do»

Источник: (Слуцкий, 2017. С. 20).

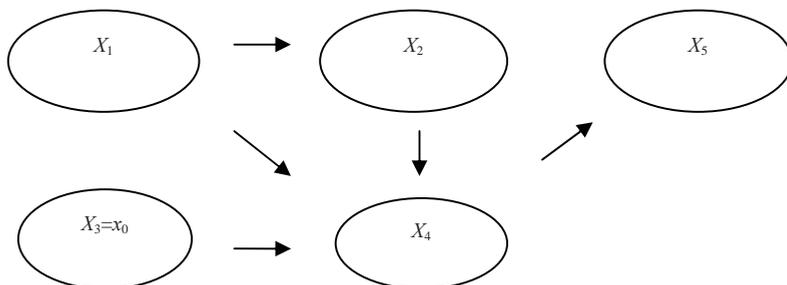


Рис. 6. Байесовская сеть, представленная на рис. 5, после интервенции $do(X_3 = x_0)$

Источник: (Слуцкий, 2017. С. 20).

На первый взгляд, данная процедура ничем не отличается от знакомого в теории вероятности условного распределения переменных системы при $X = x_0$. Однако это не так. Рассмотрим

байесовскую сеть, представленную на рис. 5. Мы хотим изучить совместное распределение переменных X_1, X_2, X_4, X_5 при условии $X_3 = x_0$: $p(X_1, X_2, X_4, X_5 | X_3 = x_0)$. Предположим, что все переменные могут принимать только конечное число значений. Чтобы получить представление о распределении $p(X_1, X_2, X_4, X_5 | X_3 = x_0)$, каждый раз, когда X_3 будет принимать значение x_0 , мы будем регистрировать значения других переменных. Таким образом, после достаточно большого числа наблюдений мы сможем получить довольно полную картину о распределении $p(X_1, X_2, X_4, X_5 | X_3 = x_0)$. Например, если переменные X_1, X_3 связаны сильной положительной зависимостью (коэффициент корреляции $\rho_{X_1, X_3} \approx 1$), то при больших значениях x_0 можно утверждать, что наблюдаемые значения X_1 также будут, в основном, близки к максимальным. При применении же оператора «do», независимо от того, какое значение было присвоено X_3 , маргинальное распределение X_1 не изменится. Как удачно подметил Линдлей (Lindley, 2002. P. 192) в своей рецензии на книгу Перла, разница между условным распределением и оператором «do» — это то же самое, что разница между наблюдением и действием. В то время как при наблюдении исследователь только фиксирует происходящее, при действии он активно вмешивается в ход событий.

Во-вторых, чтобы сделать понятие интервенции операционным, Перл должен был постулировать очень сильные предположения о переменных байесовской сети, а именно — для любой вершины X_i в регрессионном уравнении X_i по своим прямым причинам $pa(X_i)$ все коэффициенты регрессии представляют собой «истинные» значения величин эффектов воздействия соответствующих вершин на X_i . Это предполагает, что все факторы (даже не принадлежащие сети), воздействующие на X_i и отличные от прямых причин, являются независимыми от последних.

В-третьих, причинные связи, образующие байесовскую сеть, могут быть подвергнуты статистическому тестированию. Это достигается путем проверки на значимость коэффициентов уравнения регрессии (3.6), соответствующих непрямым вершинам. До Перла считалось, что наличие причинных зависимостей определяется исключительно из экономических соображений и не поддается математической проверке.

И, наконец, в-четвертых, как нам кажется, наиболее ценным вкладом Перла в статистическую теорию причинности было сближение позиций экономистов и математиков в их понимании механизмов взаимодействия между экономическими феноменами. Кроме того, введение в рассмотрение оператора «do» сделало возможным определять последствия государственного или какого-либо другого вмешательства и показало, что расчеты, основанные на регрессионных уравнениях без учета всей системы взаимных отношений в сети, могут дать неправильные результаты⁴³.

Основываясь на концепции оператора «do», можно дать «математически строгое» определение⁴⁴ причинности (Peters et al., 2017; Слуцкий, 2017. С. 21).

Определение. Вершина X_i является причиной для X_j , если существуют хотя бы два значения $a, b, a \neq b$, так что выполняется:

$$p(X_j | do(X_i = a)) \neq p(X_j | do(X_i = b)). \quad (3.8)$$

Другими словами, фактор X_i является причиной для X_j , если при манипуляции интервенциями, пусть даже виртуально, со стороны X_i меняется вероятностный характер фактора X_j . Таким образом, согласно Перлу, причинность в статистике является онтологическим объектом; в то время как эпистемологичность — лишь следствие. Формула (3.8) также является критерием различия между причинными и ассоциативными связями.

Оператор «do» позволяет измерять причинность. Предположим, мы хотим измерить влияние X_i на X_j . Для этого мы определим сначала маргинальное распределение $p(X_j | do(X_i = x))$ по формуле:

$$p(X_j | do(X_i = x)) = \int p(X_1, \dots, X_n | do(X_i = x)) dX_1 \dots dX_{j-1} dX_{j+1} \dots dX_n, \quad (3.9)$$

а затем рассмотрим разность, которая определяет *полный эффект* (total effect — TE) воздействия X_i на X_j :

$$TE(X_i; X_j) = E[X_j | do(X_i = x + 1)] - E[X_j | do(X_i = x)]. \quad (3.10)$$

43. Тем не менее регрессионные уравнения остаются действенным инструментом для расчетов возможных вариантов и прогнозирования.

44. При этом, согласно выражению известного персонажа из книги Льюиса Кэрролла, следует помнить, что слова имеют лишь тот смысл, который сам автор им придает.

Полный эффект учитывает как прямые, так и косвенные воздействия X_i на X_j . Заметим, что формула (3.10) и обобщение формулы (3.4) на произвольную байесовскую сеть дают одинаковый результат.

Понятие полного эффекта позволяет решить старую проблему⁴⁵ об определении вклада отдельного регрессора в регрессионное уравнение (3.1). Для этого следует сначала, ввиду различия в масштабах, разделить все регрессоры на соответствующие стандартные отклонения, а затем сопоставить между собой полные эффекты. Более полное представление о вкладе регрессора может быть получено путем его декомпозиции на прямое и косвенное воздействия.

Теперь мы можем сформулировать аксиому, связывающую коэффициенты регрессии фактора по своим прямым вершинам с действием оператора «do».

Аксиома 4. Рассмотрим уравнение регрессии фактора Y по своим прямым вершинам X_1, X_2, \dots, X_m , которые независимо от всех других факторов воздействуют на Y :

$$Y = k_0 + k_1 X_1 + k_2 X_2 + \dots + k_m X_m + \varepsilon. \quad (3.11)$$

В таком случае:

$$\begin{aligned} E[Y \mid do(X_1 = x_1 + 1, X_2 = x_2, \dots, X_m = x_m)] - \\ E[Y \mid do(X_1 = x_1, X_2 = x_2, \dots, X_m = x_m)] = k_1. \end{aligned} \quad (3.12)$$

Формула (3.12) аналогична формуле для определения коэффициента регрессии k_1 при факторе X_1 в регрессионном анализе за исключением того, что изменение в среднем Y при изменении X_1 на единицу, когда все остальные факторы принимают те же самые значения, заменяется на условие, чтобы фактор X_1 был фиксирован в течение всего периода наблюдения. Заметим, что равенство (3.12) имеет место только для причинных связей. В случае ассоциативных связей разность в правой части равна нулю, в то время как в регрессионном анализе соответствующая разность — k_1 .

45. Мы отсылаем читателя к статье (Gromping, 2007) и имеющимся там ссылкам по данной проблематике.

3.3. Определение синергетического эффекта

В это пункте мы покажем, как, базируясь на идеи интервенции, можно определить и вычислить синергетический эффект от воздействия факторов X_i, X_j , на объясняемую переменную Y , проявляющийся в передаточном механизме байесовской сети⁴⁶. Сначала заметим, что «выключение» одного из этих факторов, например X_i , можно добиться одним из двух эквивалентных способов: либо положив в операторе «do» $X_i = 0$, либо убрав все входящие и выходящие стрелки из X_i (см. рис. 7).

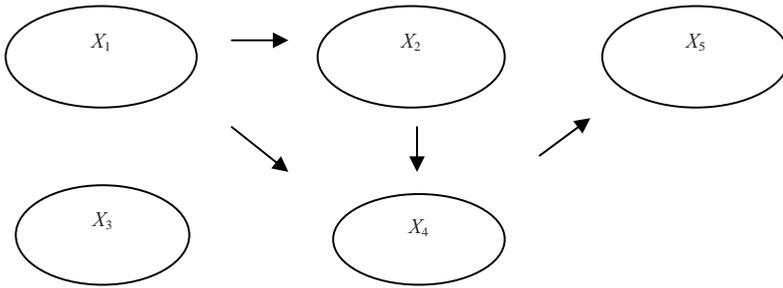


Рис. 7. Байесовская сеть, представленная на рис. 5, при выключенном факторе X_3

В качестве количественной характеристики синергетического эффекта — *synergy effect (SE)* — можно предложить следующую формулу:

$$SE = TE(X_i, X_j; Y) - (TE(X_i, X_j^{\circ}; Y) + TE(X_i^{\circ}, X_j; Y)), \quad (3.13)$$

где $TE(X_i, X_j; Y)$ — полный эффект X_i, X_j на Y ; $TE(X_i, X_j^{\circ}; Y)$ — полный эффект X_i на Y при выключенном X_j ; $TE(X_i^{\circ}, X_j; Y)$ — полный эффект X_j на Y при выключенном X_i . Иначе говоря, мы сравниваем полный совокупный эффект $\{X_i, X_j\}$ на Y с суммой полных эффектов при последовательном выключении каждого из факторов X_i и X_j .

Чтобы сделать формулу (3.13) операционной, мы определим полный эффект $TE(X_i, X_j; Y)$ по формуле:

46. Мы отметим, что синергетический эффект может быть обусловлен другими причинами кроме архитектуры сети, например, совместным распределением вершин $p(X_1, \dots, X_n, \dots, X_n)$.

$$TE(X_1, X_2; Y) = \max\{TE(X_1; Y), TE(X_2; Y)\}. \quad (3.14)$$

В соответствии с формулой (3.14) полный эффект для двух переменных равен максимальному значению среди полных эффектов по каждой переменной.

Пример. Запишем байесовскую сеть, представленную на рис. 8, в виде двух регрессионных уравнений (в целях упрощения расчетов мы предположили, что средние значения для всех факторов равны нулю):

$$X_2 = bX_1 + \varepsilon_1 \quad (3.15)$$

$$Y = c_1X_1 + c_2X_2 + \varepsilon_2. \quad (3.16)$$

В таком случае:

$$TE(X_1, X_2; Y) = c_1 \quad (3.17)$$

$$TE(X_1, X_2; Y) = c_2 \quad (3.18)$$

$$TE(X_1, X_2; Y) = \max\{TE(X_1; Y), TE(X_2; Y)\} = \max\{bc_2 + c_1, c_2\} \quad (3.19)$$

Здесь возможны различные варианты. Например, если $b > 1$, то:

$$TE(X_1, X_2; Y) > TE(X_1, X_2; Y) + TE(X_1, X_2; Y) = c_1 + c_2, \quad (3.20)$$

и синергетический эффект присутствует. При $b = 1$ мы имеем:

$$TE(X_1, X_2; Y) = TE(X_1, X_2; Y) + TE(X_1, X_2; Y), \quad (3.21)$$

и синергетический эффект отсутствует. Наконец, при $b < 1$:

$$TE(X_1, X_2; Y) < TE(X_1, X_2; Y) + TE(X_1, X_2; Y), \quad (3.22)$$

и, таким образом, включение обоих факторов вызывает уменьшение общей суммарной эффективности.

Причина наличия синергетического эффекта при $b > 1$ заключается в том, что посредством передаточного механизма в пути X_1X_2Y воздействие X_1 на Y увеличивается при включении X_2 . Иначе говоря, кроме своего прямого воздействия X_1 *влияет на Y* опосредованно через X_2 .

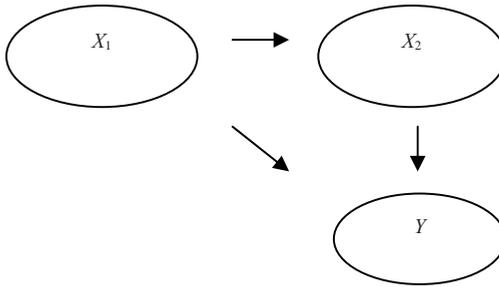


Рис. 8. Вычисление синергетического эффекта от действия двух факторов X_1 и X_2 на Y

3.4. Критика байесовских сетей

Наряду со значительным прогрессом в понимании механизма причинности в работах Перла и его коллег, а также наличием большого количества теоретических статей на эту тему (см. (Pearl, 2009) и ссылки там), байесовские сети до сих пор не нашли заметного применения в экономической науке.⁴⁷ Мы можем дать два следующих объяснения данному факту.

Во-первых, как мы уже отмечали во второй главе, для любой вершины X_i в регрессионном уравнении X_i по своим прямым причинам $pa(X_i)$ все коэффициенты регрессии представляют собой «истинные» значения величин эффектов воздействия соответствующих вершин на X_i . Определение таких значений для всего множества отношений на байесовской сети является, если не безнадежным, то, по крайней мере, очень трудным занятием.

Во-вторых, марковское свойство предполагает, что ошибки всех обозначенных выше регрессий, взаимокоррелированы. Другими словами, ошибка регрессии представляет собой агрегированные автономные эндогенные факторы, формирующие, наравне с экзогенными прямыми причинами, переменную X_i . Такое требование очень далеко от реальности. Как правило, эндогенные факторы для различных переменных системы действуют синхронно или, по крайней мере, тесно коррелированы между собой. Достаточно

47. Тем не менее такие работы имеются. См., например: (Bryant et al., 2006; Zhang et al., 2006).

вспомнить, что известный в эконометрике метод для оценки коэффициентов регрессии – *SUR* (*seemingly unrelated regressions*) – как раз основан на вычислении матрицы корреляций ошибок различных регрессий (Грин, 2016).

В-третьих, как уже было сказано, оператор «do» предполагает реальное или мысленное, т.е. потенциально возможное, вмешательство в процесс формирования переменных процесса посредством придания одному или нескольких из них конкретного значения. Но как в таком случае определить этот оператор для таких экономических показателей, как ВВП или инфляция?

На наш взгляд, можно сохранить то ценное, что содержится в концепции байесовских сетей, отказавшись от ряда исходных предположений. Возможно, при этом придется пожертвовать красотой и целостностью общей теории, но полученные преимущества в прикладных исследованиях их с лихвой окупят.

Например, для вычисления полного эффекта $TE(X_i; Y)$ от интервенции $do(X_i = x_0)$ на переменную Y достаточно рассмотреть только те пути от X_i к Y , в которых влияние X_i на Y существенно (Слуцкий, 2017. С. 25). Также необязательно рассматривать все коэффициенты регрессии для каждой из вершин, встречающейся на путях от X_i к Y , надо принимать во внимание только те, которые входят в формулу (3.4)⁴⁸ для вычисления $TE(X_i; Y)$ ⁴⁹. То же замечание относится и к вычислению синергетического эффекта, рассмотренного в третьем пункте.

48. Точнее, в ее обобщенный аналог.

49. Для определения полного эффекта воздействия фактора X на Y по формуле (3.4) требуется, чтобы переменные, встречающиеся вдоль путей от X к Y , были независимы от ошибок соответствующих регрессий.

КАЧЕСТВЕННЫЙ СРАВНИТЕЛЬНЫЙ АНАЛИЗ

Качественный сравнительный анализ – КСА (*qualitative comparative analysis – QCA*) – был изобретен Чарльзом Рейджином (*Ragin, 1987*) для выявления причинных связей между бинарными переменными, т.е. теми, которые принимают ровно два значения. Этот метод не основан на вероятностных свойствах переменных и в некотором смысле имитирует подходы, применяемые специалистами в гуманитарных науках для анализа различных явлений. Главное отличие КСА от методов, рассмотренных ранее, заключается в том, что переменные-причины не воздействуют на переменную-следствие каждая по отдельности, независимо от остальных, а их эффект может проявляться только при совместном действии.⁵⁰

Вернемся, к примеру, который мы рассмотрели в пункте 2.2. Напомним, что там нас интересовало, как курсы переподготовки для людей предпенсионного возраста могут повлиять на их дальнейшие возможности продолжения работы по специальности после окончания обучения. Сейчас мы зададимся вопросом, какими качествами должны были обладать слушатели курсов, чтобы успешно продолжать работать в течение еще не менее трех лет.

Таблица 1. Характеристики слушателей курсов

№ слушателя	Образование (X1): гуманитарное–0; техническое–1	Пол (X2): женщина–0; мужчина–1	Стаж работы (X3): менее 10 лет–0; 10 и более лет–1	Работа:(Y): менее 3 лет–0; 3 и более года–1
1	1	1	0	1
2	0	1	1	0
3	1	0	0	0
4	1	0	1	1
5	0	1	0	0
6	1	0	1	1
7	1	1	1	1
8	0	0	0	0
9	0	0	1	0
10	1	0	0	1

50. Данное замечание относится только к линейным моделям, рассмотренным в докладе, или их аналогам.

Анализируя табл. 1, можно прийти к выводу, что для этого *достаточно* ($X \rightarrow Y$) два условия:

- 1) слушатель, независимо от пола, должен иметь техническое образование и стаж работы 10 и более лет (слушатели № 4, 6, 7);
- 2) слушатель – мужчина и имеет техническое образование (слушатель № 1, 7).

Таким образом, исследователь может сделать вывод, что слушатели с характеристиками 1 или 2 продолжат работу по специальности в течение последующих трех лет. Когда слушатель – женщина, имеющая техническое образование (№ 3,10), никакого вывода на основании приведенных данных сделать невозможно. Во всех остальных случаях (слушатели № 2,5,8,9) можно сделать вывод, что слушатели с данными характеристиками прекратят работать по специальности менее чем через три года.

Необходимым условием ($Y \rightarrow X$), чтобы слушатель продолжил работу по специальности в течение не менее трех лет, будет наличие технического образования (слушатели № 1,4,6,7,10). Действительно, во всех случаях, когда слушатель продолжил работать в течение этого времени, он имел техническое образование. Однако это условие не является достаточным, так как слушатель № 3, хотя и получил техническое образование, но работал по специальности в течение менее трех лет.

Для определения необходимых и достаточных условий при работе с большими таблицами применяются законы булевой алгебры (*Schneider, Wagemann, 2012*).

Хотя наши рассуждения не основывались на стохастической природе рассматриваемых факторов, теоретико-вероятностные методы могут быть с пользой применены и при анализе табл. 1. Так, слушатели, независимо от пола, имеющие техническое образование и стаж работы 10 и более лет, покрывают 60% случаев продолжения работы в течение трех и более лет; 50% слушателей продолжили работать по специальности после окончания курсов, с одинаковым процентом для мужчин и женщин. Важно также и то, что на основании приведенных вычислений можно будет делать предварительные прогнозы.

С одной стороны, Рейджин придал легитимность простым методам, которыми в своих работах всегда пользовались социологи и экономисты, не склонные обращаться к более изощренным статистическим моделям, а с другой — формализовав и уточнив эти методы, он сделал возможным их применение для очень больших массивов данных. В дальнейшем методология КСА была обобщена на категориальные переменные, т.е. на факторы, принимающие более двух значений, и на нечеткие множества (см.: *Schneider, Wagemann, 2012*). Алгоритмы, используемые в КСА, как правило, применяются при решении задач в области искусственного интеллекта.

В определенном смысле КСА является альтернативой регрессионному анализу. В первую очередь это относится к случаю, когда число наблюдений незначительно превосходит или даже меньше, чем число переменных. В то время как большинство эконометрических методов, используемых для оценок коэффициентов регрессии, основано на усреднении результатов наблюдений, КСА позволяет проследить причинно-следственные цепочки по всем индивидуальным наблюдениям⁵¹. С другой стороны, регрессионный анализ предоставляет возможность сделать выводы и прогнозы, имеющие стохастический характер. Например, проверка статистических гипотез, определение доверительных интервалов заданной точности для различных параметров регрессии и т. д. Также методы регрессионного анализа позволяют работать с непрерывными переменными, т.е. такими, которые принимают целый спектр значений⁵².

-
51. Здесь следует уточнить, что существуют различные методы в регрессионном анализе, также позволяющие выявить влияние индивидуальных наблюдений, например, расчет студентизированных остатков, анализ выбросов и влиятельных наблюдений. Однако их применение носит ограниченный характер.
52. Существуют различные методы, позволяющие переводить непрерывные переменные в категориальные: ниже или выше заданного порогового значения, нахождение наблюдения между двумя соседними квартилями, децилями и т. д.

ЗАКЛЮЧЕНИЕ

Изучение причинных зависимостей в экономике сопряжено с большими трудностями, заключающимися в том, что наряду с рассматриваемыми объектами исследователю приходится принимать во внимание также большое число факторов, действующих на изучаемые переменные, про существование которых либо ничего неизвестно, либо их невозможно или трудно измерить. Единственное, чем он располагает, это результаты наблюдений и, возможно, некоторая экономическая теория. В том случае, когда исследователь уверен, что он знает группу автономных факторов, непосредственно влияющих на изучаемый объект A , то нахождение величин эффектов воздействия сводится к определению коэффициентов уравнения регрессии с зависимой переменной A и факторами, выступающими в качестве регрессоров. В противном случае ему приходится прибегать к одному из методов, изложенных в данном докладе.

В случае, когда один из регрессоров является эндогенной переменной, при систематических ошибках в данных, а также при наличии пропущенных факторов, выходом из положения, возможно, будет нахождение инструментальной переменной. К сожалению, как мы уже отмечали, это может быть очень трудной и даже невыполнимой задачей.

Когда фактор-причина находится в двух взаимоисключающих режимах, модель потенциальных исходов Рубина является подходящим инструментом. Она основана на наблюдениях как различных характеристик объектов, так и эффектов воздействия и представляет собой квазиэксперимент, являющийся обобщением концепции рандомизированного эксперимента. Эта модель с успехом применялась и продолжает применяться в эконометрике, но для довольно специфического ряда задач.

Причинные байесовские сети, базирующиеся на работах Перла и его учеников, позволяют вычислять последствия интервен-

ций, т.е. придания фиксированных значений одной или нескольким переменным сети⁵³. Основываясь на понятии интервенции, можно определять как полный эффект одного фактора на другой в сети, так и вклады отдельных причинно-следственных цепочек. Нами было показано, что оператор «do» может быть эффективно использован для вычисления синергетического эффекта, вызванного архитектурой байесовской сети. На примере простейшей байесовской сети, состоящей из трех вершин, было продемонстрировано, что наличие или отсутствие синергетического эффекта обуславливается как взаимным расположением факторов, так и величинами эффектов их воздействия друг на друга. Последнее обстоятельство может быть использовано в теории эволюционных систем для объяснения образования или упразднения различных институтов. Другими словами, определенная структура добавляется к системе, если она способствует увеличению синергетического эффекта, или удаляется в противном случае. При этом предполагается, что система должна быть достаточной гибкой, чтобы она могла принять подобные изменения.

Метод байесовских сетей не дает ответа на фундаментальный вопрос: при наличии ненулевой корреляции и отсутствии общей причины для переменных A и B можно ли решить, основываясь лишь на наблюдениях и статистических расчетах, какой из факторов является причиной, а какой следствием? Хотя байесовская сеть, как на это неоднократно указывал Перл, допускает частичную проверку, но это относится в основном к сетям, имеющим более двух вершин. Для двух вершин, по мнению Перла, вопрос не может быть решен чисто статистическими методами.

В последние десятилетие наметились, главным образом в работах Шолкопфа и его коллег (*Peters et al.*, 2017), новые подходы к решению данной проблемы. Шолкопф исходит из того, что асимметрия, возникающая в причинно-следственной связи, должна тем или иным образом проявляться также и в статистических свойствах наблюдений. Был произведен буквально мозговой штурм с использованием самых сложных современных математических

53. Перл (*Pearl*, 2009) утверждает, что к оператору «do» не применима «критика» Лукаса, так как она направлена против выводов, основанных на наблюдениях, а не на архитектуре отношений, представленных байесовской сетью.

методов для понимания механизмов действия причинных связей. Достаточно упомянуть теорию алгоритмов, колмогоровскую сложность и машинное обучение, являющееся форпостом теории искусственного интеллекта.

Хотя на данном этапе еще рано судить о прикладных применениях приведенных исследований, в том числе и в экономике, тем не менее можно утверждать, что за последние тридцать лет теория причинности из области философии практически полностью перешла в сферу деятельности математиков и статистиков.

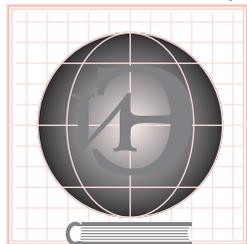
ЛИТЕРАТУРА

- Аристотель* (1976). Сочинения. Т. 1. (Серия «Философское наследие»). М.: Мысль.
- Грин У. Г.* (2016). Эконометрический анализ. Т. 1. М.: Издательский дом «Дело» РАНХиГС.
- Ениколопов Р.* (2009). Оценивание эффекта воздействия // Квантиль. № 6.
- Маевский В. И., Слуцкий Л. Н.* (2009). Инфляция и фондовый рынок: CPI и S&P 500 // Прикладная эконометрика. № 3 (15).
- Слуцкий Л. Н.* (2017). Графические статистические методы для исследования причинных зависимостей. Байесовские сети // Журнал Новой Экономической Ассоциации. № 4 (36). С. 12–30.
- Смит А.* (1962). Исследование о природе и причинах богатства народов. М.: Издательство социально-экономической литературы.
- Сток Д., Уотсон М.* (2015). Введению в эконометрику. М.: Издательский дом «Дело» РАНХиГС.
- Anderson T. W.* (2003). An Introduction to Multivariate Statistical Analysis. Third edition. Wiley-Interscience.
- Bryant H.L., Bessler D.A., Haigh M.S.* (2006). Causality in Futures Markets // Journal of Futures Markets. No. 26 (11).
- Cowell R.G., David A.P., Lauritzen S.L., Spiegelhalter D.J.* (2003). Probabilistic Networks and Expert Systems. Corrected edition. New York: Springer.
- Fisher R.A.* (1925). Statistical Methods for Research Workers. Edinburgh: Oliver and Boyd.
- Fisher R.A.* (1935). The Design of Experiments. Edinburgh: Oliver and Boyd.
- Frisch R., Tinbergen J., Haavelmo T., Koopmans T. C.* (1948). The Autonomy of an Economic Relation. Memorandum Institute of Economics. University of Oslo.

- Gromping U.* (2007). Estimators of Relative Importance in Linear Regression Based on Variance Decomposition // *The American Statistician*. Vol. 61. No. 2.
- Haavelmo T.* (1944). The probability approach in econometrics // *Econometrica*. No. 12 (Supplement). Pp. 1–118.
- Koopmans T. C.* (1950). When is an equation system complete for statistical purposes? In: *Statistical Inference in Dynamic Economic Models*. Cowles Commission for Research in Economics. No. 10. New York: Wiley and Chapman & Hall.
- Lindley D.V.* (2002). Seeing and Doing: The Concept of Causation // *International Statistical Review*. Vol. 70.
- Neyman J.* (1923). On the application of probability theory to agricultural experiments: essay on principles, section 9 // Translated in *Statistical Science*. 1990. No. 5.
- Pearl J.* (1988). *Probabilistic Inference in Intelligent Systems*. Morgan Kaufmann. San Mateo, California.
- Pearl J.* (2009). *Causality: Models, Reasoning, and Inference*. Second edition. Cambridge: Cambridge University Press.
- Peirce C. S.* (1883). *A Theory of Probable Inference*, *Studies in Logic*. Boston: Little, Brown, and Company.
- Peirce C. S., Jastrow J.* (1885). On Small Differences in Sensation // *Memoirs of the National Academy of Sciences*. Vol. 3.
- Peters J., Janzing D., Scholkopf B.* (2017). *Elements of Causal Inference. Foundations and Learning Algorithms*. Cambridge: The MIT Press.
- Ragin C. C.* (1987). *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Oakland, California: University of California Press.
- Reichenbach H.* (1956). *The Direction of Time*. Berkeley: University of Los Angeles Press.
- Rubin D.B.* (1974). Estimating causal effects of treatments in randomized and nonrandomized studies // *Journal of Educational Psychology*. No. 66.
- Rubin D. B.* (2005). Causal inference using potential outcomes: Design, modeling, decisions. 2004 Fisher Lecture // *Journal of the American Statistical Association*. No. 100.

- Schneider C. Q., Wagemann C.* (2012). *Set-Theoretic Methods for the Social Sciences: A Guide to Qualitative Comparative Analysis*. Cambridge: Cambridge University Press.
- Simon H. A.* (1953). Causal ordering and identifiability. In: *Studies in Econometric Method*. Cowles Commission for Research in Economics. No. 14. New York: Wiley and Sons.
- Wooldridge J. M.* (2010). *Econometric Analysis of Cross Section and Panel Data*. Second edition. Cambridge: The MIT Press.
- Wright P.* (1928). *The tariff on Animal and Vegetable Oils*. New York: Macmillan.
- Wright S.* (1921). Correlation and Causation // *Journal of Agricultural Research*. Vol. 20.
- Zhang J., Bessler D.A., Leatham D.* (2006). Does Consumer Debt Cause Economic Recession? Evidence Using Directed Acyclic Graphs // *Applied Economics Letters*. Vol. 13.

Российская академия наук



Институт экономики

Редакционно-издательский отдел:

Тел.: +7 (499) 129 0472

e-mail: print@inecon.ru

Сайт: www.inecon.ru

Научный доклад

Л.Н. Слуцкий

Моделирование причинных зависимостей в экономике

Оригинал-макет *Валериус В.Е.*

Редактор *Ерзнкян М.Д.*

Компьютерная верстка *Борщёва И.В.*

Подписано в печать 11.01.2019. Заказ № 1

Тираж 300 экз. Объем 2,0 уч.-изд. л.

Отпечатано в ИЭ РАН