

2022年度
関西学院大学ロースクール
A日程

一般入試（法学未修者）

論文問題

《10:00～11:30》

○開始の指示があるまで内容を見てはいけません。

【論文問題】

問題文を読んで、〔設問 1〕 および〔設問 2〕 に答えなさい。

〔設問 1〕

なぜデータから因果関係を導くのは難しいのかについて、著者の見解を要約しなさい。(400字程度)

〔設問 2〕

次のような新聞記事があったとする。問題文の著者の見解に従うと、この新聞記事に示されたデータについて、どのように分析することができるか、説明しなさい。(500字程度)

「スマホ使用、少ないと好成绩

中学生の成績はスマートフォンの使用時間が短いほど良くなることが、〇〇市教育委員会の調査で分かった。市内の中学1～2年生3万人を対象に1月、学力テストとアンケートを実施し、関係を分析した。

スマホを持つ生徒は両学年とも9割以上を占めた。1日あたりのスマホの利用時間が30分未満の生徒の数学の平均点は59点、1時間以上2時間未満は57点、3時間以上4時間未満は47点と、長時間になるほど成績が下がった。

国語、英語も同様の傾向。30分未満の生徒と4時間以上の生徒の得点差は、国語11点、数学18点、英語17点に上った。」

問題文

なぜデータから因果関係を導くのはとても難しいのでしょうか？

本章では、この点について3つの具体例を用いて説明します。

1つ目は、企業で働く方の目線から考えたマーケティング戦略の例です。2つ目は、公的機関で働く方の目線から考えた政策形成の現場の例です。また、3つ目の例として教育に携わる機関で働く方の目線から考えた具体例を考えてみたいと思います。

あなたはアイスクリームを売る企業のマーケティング部に所属しています。現在社内では、ウェブサイト上で広告を表示することによって今年夏のアイスクリームの売り上げを伸ばすことができないか、ということが検討されています。あなたは上司から、広告を出すと売り上げがどれだけ伸びるのかデータ分析をしてほしいと頼まれました。

過去のデータを見てみると、次のことがわかりました。

2010年にあなたの会社では、あるアイスクリーム商品についてのウェブ広告を出しました。すると、広告を出さなかった2009年と比較して、2010年の売り上げは40%上がっていました。そのデータの動きをグラフにしたのが図表1-1です。この図では、広告を出した影響で売り上げが伸びたように見えます。そのため、あなたは上司に対し以下のような報告をしました。

「この図を見ていただくとわかるように、広告を出した影響により2010年の売り上げは2009年に比べて40%上がった、ということが分析からわかりました」

さてここで、なぜあなたの結論が間違っている可能性があるのか考えてみてください。どんな可能性が考えられますか？

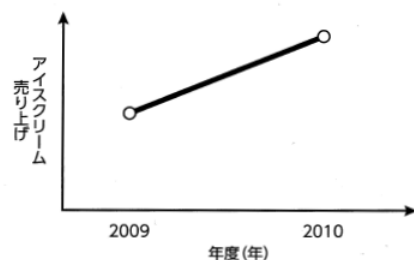
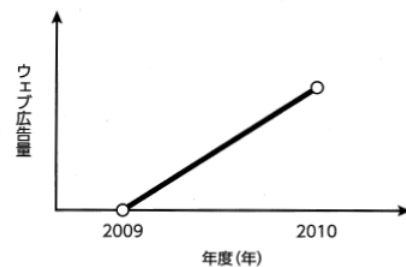
ここでの問題は、

「広告を出した → 広告の影響で売り上げが40%伸びた」

という広告から売り上げへの因果関係（英語では causal relationship、もしくは causality と呼びます）が、あなたのデータ分析結果から導けるかどうかです。

例えば、2010年の夏が2009年の夏よりも猛暑だった場合はどうでしょうか？

図表1-1 広告の影響でアイスクリームの売り上げが伸びた？



実際に日本では2009年は比較的冷夏で、2010年夏は猛暑でした。その場合、40%の売り上げ増というのは、広告の影響ではなく、単に気温が高くなったために消費者がアイスクリームを求めたから、という可能性はないでしょうか？

他にも様々な理由が考えられます。

例えば、日本では2008年の世界金融危機以降、消費が冷え込みましたが、2010年あたりから少しずつ消費が上向きになりました。その場合、40%の売り上げ増は広告の効果ではなく、単に経済が全体的に良くなって消費者がお財布の紐を緩め始めたからだった、という可能性はないでしょうか？

2つ目の例として、政策を実施する政策担当者の抱える課題について考えてみましょう。

あなたは経済産業省の職員で、来年夏の節電対策を考えています。今回のプロジェクトの目的は、電力の価格を上げるとどれだけの節電効果につながるのかについて上司に報告することです。その目的のため、あなたは過去の電力価格と電力消費量のデータを集めました。

データを見てみると次のことがわかりました。

日本のある地域では2012年に電力価格の上昇がありました。仮に、2008年の電力1単位あたりの価格は20円、2012年の価格は25円としましょう。一方、消費量のデータを見ると、2012年の電力消費量は2008年と比較して1時間あたり5kWh下がっていました（注:kWh=キロワットアワーは電力消費量の単位です。日本の平均的なご家庭の使用量は、夏の間は1時間あたり20kWhほどです）。そのため、上司に対し以下のような報告をしました。

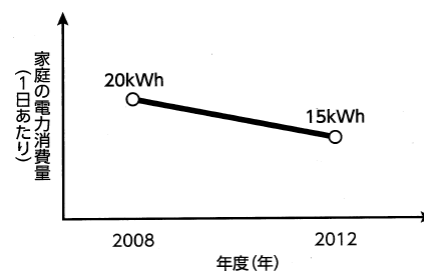
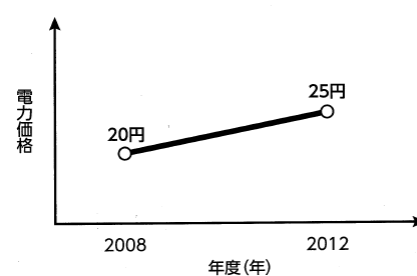
「図表1-2を見ていただくとわかるように、5円の電力価格上昇による影響で、消費量が5kWh下がったことがわかりました。そのため、電力価格を上げれば大きな節電効果が得られると期待できます」

さてここで、なぜあなたの結論が間違っている可能性があるのか考えてみてください。どんな可能性が考えられますか？

ここでは電力の価格が消費量に及ぼした影響を言いたいわけですが、他の要因が色々と考えられないでしょうか。

例えば、2012年は比較的涼しい夏だったため、エアコン利用が減った可能性があります。もしくは、2011年に起きた東日本大震災によっ

図表1-2 電力価格上昇による影響で電力消費量が下がった？



て、消費者の節電意識が高まったのかもしれませんが。そうすると、広告とアイスクリームの例と同様、このデータ分析からでは「電力価格の変化が電力消費量の変化をもたらした」という因果関係を判定できない、という問題が残るのです。

同じようなデータ分析の問題を、教育の例を使って見てみましょう。
先日、以下の新聞記事を目にしました。

「海外留学に力を入れているある大学の調査では、留学を経験した学生が、留学を経験しなかった学生よりも就職率が高いことがわかった。このデータ分析の結果から、留学経験は就職率を向上させるのであると大学は報告している」

留学を経験した学生が、留学を経験しなかった学生よりも就職率が高かったという記事の前半部分は、データが示している事実なのだと思います。しかし、その結果から、

「留学を経験する → 就職率が上がる」

という因果関係を導くことはできるでしょうか？

ここで、留学経験がある学生Aさんと、留学経験がない学生Bさんを考えてみましょう。問題は「留学を経験した」という点以外についても、AさんとBさんは異なる可能性が高いということです。

例えば、Aさんは留学をできるほどの財力が家庭にあった可能性が高いかもしれません。または、留学の奨学金を受けられるほど、もともと成績が良かったのかもしれませんが。さらに、そもそも留学をしたいという強い意志や、好奇心があった可能性もあります。

以上のように、留学をしたという点以外でAさんとBさんに違いがあった場合、2人の就職率に影響を与えたのは留学だったのかもしれないし、それ以外の要因だったのかもしれないのです。

ここまでの3つの例で共通しているのは、「ある要素（X）が結果（Y）に影響を与えた」というX→Yの因果関係が主張されている点です。最初の例では広告（X）がアイスクリームの売り上げ（Y）に影響した因果関係を主張しており、最後の例では留学経験（X）が就職（Y）へ影響した因果関係の主張でした。

ここで「XがYに影響を与えた」という因果関係を示すことが難しい一番の理由は、「Yが変化したのはX以外の他の要因の影響だったのでは？」という問題を排除できないことです。

データ分析者は、XがYに影響した、と主張したわけです。ところが、通常、世の中は実験室のように単純ではありません。Xを発生させたと同時に、色々なこと（XやYではない要素をVと呼ぶことにしましょう）が起こり得ます。

広告の例の場合、広告というXを発生させたと同時に、気温の上昇や経済状況の変化といったVが発生している可能性があります。留学政策の例の場合、留学をしたというXと同時に、留学前からの成績、親の財力、国際的志向といったVがYに影響している可能性があるのです。

その場合、図表1-1や図表1-2にあるように、XとYのデータが同時に動いているように見える場合でも、XがYに直接影響を与えたのではなく、VがXとYの両方に影響を及ぼしただけ、という可能性があるのです。

さらに、場合によっては、実は「YがXに影響を与えたのではないか」という「逆の因果関係（reverse causality）」の可能性を否定できないこともあります。

アイスクリームの例では、

「2010年の初期に猛暑の影響でアイスクリームの売上げが伸びたので、会社としてはその売上金を使ってウェブ広告を始めてみた」

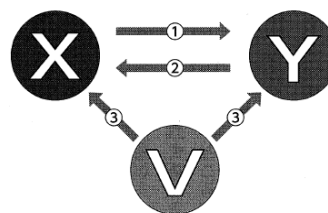
という可能性もあります。この場合、 $X \rightarrow Y$ という因果関係ではなく、 $Y \rightarrow X$ という因果関係が存在しているわけです。

図表1-1や図表1-2のように、2つのデータの動きに関係性があることを、統計学では「相関関係がある」と呼びます。実は、データが手元があれば相関関係を計算することは容易です。例えば、図表1-1や図表1-2のようにグラフを描いて、2つのデータの動きの関連性を調べるのも一つの方法ですし、エクセルなどのソフトウェアを使って相関関係の値を計算することも簡単にできます。

問題は、XとYに相関関係があることがわかって、その結果を用いて因果関係があるとは言えないことです。

相関関係と因果関係が違うことは多くの本で紹介されていますが、初めて聞く方にとっては多少戸惑う点です。わかりやすくするため、図表1-3に「XとYに相関関係がある場合に起こり得る3つの可能性」を示してみました。

図表1-3 データ分析から因果関係を立証することはなぜ難しい?



- ① XがYに影響を与えている可能性
- ② YがXに影響を与えている可能性

③ V が X と Y の両方に影響を与えている可能性

データ分析者にとって非常に厄介なのは、この3つの可能性の全てが図表1-1や図表1-2のデータの動きと整合的であることです。アイスクリームと広告の例に戻れば、図表1-1のデータの動きを見ただけでは、

- ① 広告 (X) が売り上げ (Y) に影響を与えた可能性
- ② 売り上げ (Y) が広告 (X) に影響を与えた可能性
- ③ 他の要因 (V) が広告 (X) と売り上げ (Y) の両方に影響を与えた可能性

という3つの可能性のどれが本当なのか、判定できないのです。

こうやって説明されてみると、データ分析から因果関係を導くことの難しさは直感的に理解できると思います。ところが、ニュースや新聞を見てみると相関関係と因果関係を混同させた怪しい分析結果は世の中に溢れています。さらに問題なのは、怪しい分析結果に基づく単なる相関関係が「あたかも因果関係のように」主張され、気をつけないと読者も頭の中で因果関係だと理解してしまっていることが多いという点です。

以下の例は、実際に著者が見かけたことのある新聞記事の抜粋です。

「ある企業では社長が代わった次の年に株価が上昇した。これは新社長の改革の成果である」

→社長が代わった以外にも株価が上昇した他の要因があり得る。

「政府が数億円かけて実施した補助金政策によって、補助金交付後、各地域で消費が増加した。これは、補助金が地域経済を活性化した証である」

→補助金以外にも消費が増加した他の要因があり得る。

「ある学校では新たなカリキュラムを導入した。すると、生徒の理解度と成績が前年に比べて向上した。よって、新カリキュラムは旧カリキュラムよりも優れていることが示された」

→カリキュラム導入以外にも成績へ影響するような変化があった可能性がある。

「マンションの高層階に住む女性の不妊率が高いことがデータから示された。よって、子供を産みたい女性がマンションの高層階に住むのは危険である」

→マンションの高層階に住む女性と低層階に住む女性では、所得・年齢・職業など

様々な別の要因が違う可能性があり、高層階に住むことが本当の要因なのかは明らかではない。

「電力市場の自由化改革を行った国の電力価格は、行っていない国の価格よりも高い。よって、電力市場自由化改革を行うと電力価格が上がってしまう」

→電力市場の自由化改革を行った国とそれ以外の国では様々な要素が異なるので、自由化改革自体が価格に影響したのかは明らかではない。また、そもそも価格が高い国ほど自由化改革に取り組んだ、という逆の因果関係もあり得る。

このような論調は、一見すると素通りして因果関係と捉えてしまいがちです。しかし、一歩立ち止まってよく考えてみると、「XがYに影響したと結論づけているけれども、他の要因Vも影響している可能性があるのでは？」「もしかしたらYがXへ影響している可能性もあるのでは？」という疑問が出てきます。しかし残念ながら、新聞やテレビで主張されていることの多くは、相関関係を誤って解釈して因果関係のごとく示されているものなのです。

さて、ビジネスの現場や政策決定の過程で、なぜ相関関係だけではなく、因果関係を見極めることが大切なのでしょうか？

この節では、因果関係を見誤るとなぜ問題なのか、そして、因果関係を正確に見極めることがビジネスや政策決定の現場でなぜ大切なのかを見ていきましょう。

先ほどの2つの例を思い出してみてください。

広告とアイスクリームの例では、「広告を打つこと」と「アイスクリームの売り上げ」に相関関係があることは、図表1-1から理解できます。しかし、この分析をもとに「では今年度も数千万円を投じて広告を導入し、売り上げを伸ばそう!」という決断がなされた場合どうなるのでしょうか？

先述したように、もし過去に売り上げが伸びたのが広告の影響ではなく気温や経済活動の変化といった他の影響だった場合、多額の費用を投じて今年度に広告を打っても、売り上げは上がらないこととなります。つまり、数千万円の投資が全くの無駄な出費になってしまうのです。

同様に、留学と就職率の関係を考えてみましょう。

例えば、留学経験と就職率の相関関係をもとに「留学は就職率を上げるので、留学支援政策として補助金を投入しよう」という政策が政府によって打ち出されたとします。しかし、観測された相関関係が、「留学を経験したから」という理由ではなく、他の要因による影響だとしたらどうなるのでしょうか（海外の大学で^{きょうべん}教鞭を執る身としては、留学は素晴らしい経験になることだと思うので、これはあくまで仮の話ですが）。その場合、国民の税金を投じて行われる補助金政策の根拠に誤りがあることに

なります。

ビジネスの現場にしる政策決定の過程にしる、物事を決定する際に鍵となるのは多くの場合「因果関係」であり、相関関係ではないのです。

(中略)

「相関関係は因果関係ではない」という問題を解決する方法の一つとして伝統的に紹介されてきたのは、考えられるだけのVのデータ（他の要因として考えられる要素のデータ）を頑張って集めて、できる限りVの影響を統計分析によって除くという手法です。しかし問題は、どれだけたくさんの種類のVの要素を考慮しても「もしかしたら別の要素も影響したかもしれない」という可能性が無限に出てきてしまうことです。

さらに、Vとして考えられる要素の中にはどうしてもデータとして手に入らない物もあります。アイスクリームの例の場合、気温や経済活動の変数はデータとして集められるかもしれませんが、「消費者の好みが変わったのかも」「アイスを食べる人のうちウェブサイト広告を見る人が増えただけなのかも」など、Vとなり得る追加的要素はいくらでも出てきてしまい、中にはデータとして存在しないものも多くあります。

電力価格の例で言えば、気温などのデータは集められるかもしれませんが、節電意識が変わったのかもしれない、といった要素はデータとして存在しない可能性が高いのです。

同様に、留学と就職率の例で言えば、親の収入や学生の成績などのデータは集められるかもしれませんが、学生がもともと持っていた国際的志向、やる気、潜在的能力などの要素は観測が非常に難しい要素です。

経済学を始めとする社会科学の研究では、できる限りのVのデータを集めてVの影響を排除する統計分析手法が長年開発されてきました。しかし、1980年頃からこういった手法の限界が指摘され始めました。現在では因果関係を求める際には、Vとして考えられるデータをできる限り集めてくることは有用ではあるが、非常に限界があると考えられています。

(以下略)

伊藤 公一朗「データ分析の力 因果関係に迫る思考法」（光文社新書、2017年）より抜粋。出題との関係で必要な補足、省略、変更を施している。

2022 年度入学試験 出題趣旨・解説・講評

【A 日程：論文】

《出題趣旨》

データから因果関係を導くことができるかという、法律家にとっても重要な問題について、著者の見解を正確に理解し、要約した上で、それを具体的な事例に適用する能力を試すものである。

《解答例》

〔設問 1〕

データから因果関係を導くのが難しいのは、要素 X と要素 Y との間で、それぞれのデータの動きに関係性があること、すなわち、両者の間に相関関係があることがわかったとしても、次の 3 つの可能性が考えられるため、X が Y に影響を与えたという因果関係は、必ずしも導かれなからである。すなわち、① X が Y に影響を与えている可能性、② Y が X に影響を与えている可能性、および、③ X および Y とは別の要素 V が X および Y の両方に影響を与えている可能性である。データ分析者にとって非常に厄介なのは、この 3 つの可能性の全てが、相関関係を整合的に説明できることである。このうち、③の可能性を排除するため、考えられるだけの V のデータを集め、統計分析によって V の影響をできる限り除くという手法がある。しかし、V として考えられる要素には、無限の可能性があるほか、性質上データの入手が困難なものもあるため、この手法には限界がある。(393 字)

〔設問 2〕

この新聞記事に示されたデータについては、スマホの使用時間を短くすれば成績が良くなるという因果関係以外に、次のような可能性が考えられる。第 1 に、成績が良い生徒ほど、スマホの使用以外のことに関心があって、スマホを短時間しか使用しないという、逆の因果関係の可能性もある。第 2 に、スマホの使用時間と成績の両方に影響を及ぼしうる様々な要素（著者のいう要素 V）が影響している可能性がある。例えば、塾に長時間通っている生徒は、必然的にスマホの使用時間が短くなるとともに、塾での学習の効果によって成績が良くなっている可能性がある。通塾以外にも、例えば、部活動、生徒会活動、ボランティア活動、読書等の活動に積極的に取り組んでいる生徒は、スマホの使用時間が短くなるとともに、物事に取り組む意欲が高いため、成績も良いという可能性がある。また、スマホを長時間使用しないように厳しくしつける保護者は、教育にも熱心で、それが成績向上の原因になっている可能性もある。以上のことから、スマホの使用時間が短いことと成績が良いこととの間には、相関関

係はあるとしても、前者が後者の原因であるという因果関係があるとは、必ずしもいえない。(498字)

《解説・講評》

〔設問1〕

ポイントは、(1) 相関関係と因果関係の概念について簡潔に説明した上で、相関関係が必ずしも因果関係を意味しないことを指摘する、(2) その理由として、XとYの間に相関関係が認められる場合に、3つの可能性があり、いずれに当たるかを特定することが困難であることを指摘する、(3) XとYの両方に影響を与えるVのデータを集め、その影響を排除する手法の限界を指摘する、という3点である。

上記(2)については、著者自身が3つの可能性を端的に説明しているので、それをそのまま要約に取り入れるのが適切と考えられる。また、著者のいうVについては、XとYの両方に影響を与えるという点が、相関関係を説明する上で重要である。しかし、単にVがYに影響を与えることを指摘するにとどまる答案が多数あった。この点を正確に理解できていないと、後述の〔設問2〕の解答も不正確なものとなる。

上記(3)は、なぜデータから因果関係を導くのが難しいのかを説明する上で、重要な要素であるが、指摘していない答案が多かった。逆に、上記(3)から書き始め、それを中心に据えている答案もあったが、上記(3)は上記(1)および(2)を前提とする議論なので、いきなり上記(3)から書き始めるのは適切ではない。

相関関係が必ずしも因果関係を意味しないというのは、問題文全体を貫くテーマであるにもかかわらず、主たるテーマとは別の論点として相関関係に言及する答案があった。これは、問題文の構造を理解していないものと評価される。

具体例と一般論の区別がついていないと見られる答案もあった。字数に余裕があれば、具体例に言及すること自体は悪くないが、上記(1)～(3)の全てを正確に説明するためには、具体例に多くの字数を割く余裕はないと思われる。

〔設問2〕

問題文の著者の見解に従って分析すると、(1) 成績の良い生徒ほどスマホの使用時間が短いという、逆の因果関係の可能性、(2) スマホの使用時間と成績の両方に影響を与える要素Vが存在する可能性、をそれぞれ指摘して、(3) 新聞記事が主張するような因果関係は必ずしも導かれない、という結論を示すことになる。

上記(2)で、成績に影響を及ぼしうる様々なVを挙げるのみで、それが同時にスマホの使用時間にも影響を及ぼしうることを説明していない答案が多数あった。しかし、新聞記事に示されたデータは、3万人の統計調査によって、スマホの使用時間と成績との間に相関関係があることを示しているから、Vを挙げる際は、スマホの使用時間と成績の両方に影響を及ぼしうることを説明できなければならない。例えば、V

として保護者の財力を挙げる答案があったが、これが成績に影響を及ぼしうるとしても、スマホの使用時間にも影響を及ぼしうることを説明できなければ、不適切である。また、たまたまテストの当日体調を崩した生徒もいるかもしれないとか、スマホを勉強に使っている生徒はむしろ成績が向上するはずであるといった指摘も、統計調査によって示された上記の相関関係を説明するものではないので、不適切である。アンケートに正直に答えず、スマホの使用時間を過少申告している可能性があるという指摘する答案もあったが、これについても、成績の良い生徒ほど過少申告する傾向にあるといえなければ、上記の相関関係を説明できないので、不適切である。

〔設問1〕〔設問2〕の全体を通じた注意点として、思いついた事項からいきなり答案を書き始めるのではなく、どの事項を、どのような順序で、どの程度の字数で書くかを、あらかじめ下書き用紙を使って検討してから、答案を書き始めるべきである。