

政策評価に関する統一研修（中央研修）講演概要

平成28年1月14日開催

講演名：政策評価の質の向上に向けて一データの適正かつ有効な活用の観点から一
講師：早稲田大学政治経済学術院・公共経営大学院 山田 治徳 教授
講義時間：14時50分～16時20分

<本日の内容>

本日の講演は6部構成になっている。

演題は「政策評価の質の向上に向けて一データの適正かつ有効な活用の観点から一」としているが、この質の向上とデータの適正かつ有効な活用という観点からすると、重要なのはやはりスライド番号2番になる。

<政策評価法に見るロジック・モデル>

資料において「1. データの基本もロジック・モデルから」としているように、これは過言かもしれないが、政策評価はロジック・モデルに始まりロジック・モデルに終わる。本日は、ロジック・モデルそのものについて詳細に説明することは避けるが、政策評価、有効なデータの活用という観点から見ても、ロジック・モデルは重要であるということ、まず改めて確認していただきたい。

ロジック・モデルについては政策評価法の第3条で書かれている。「行政機関は、その所掌に係る政策について、適時に、その政策効果（当該政策に基づき実施し、又は実施しようとしている行政上の一連の行為が国民生活及び社会経済に及ぼし、又は及ぼすことが見込まれる影響をいう。以下同じ。）を把握し、また、これを基礎とし」と書かれている。すなわち、政策に基づいて実施し、又は実施しようとしている皆さん方の行政上の一連の行為が住民生活、社会経済に及ぼし、又は及ぼすことが見込まれる影響、こういったロジックの流れをフローチャート上で示したものをロジック・モデルという。

<ロジック・モデルの意義>

ロジック・モデルというのは、もともと何のための仕組みだったかという、政策の論理的形成、政策の意義・概念の明確化及び政策設計における問題点の発見といった政策の

設計図としての役割を持っていた。しかし、ロジック・モデルは単なる設計図としてだけでなく、政策評価においても重要な意義を持っており、政策評価に当たってのポイントの明確化及び政策評価に当たっての指標の設定において重要な役割を果たす。すなわち、皆様が政策を執行する仕事を行う際、これが皆様の意図どおり効果が表われているかどうかを検証するためにも、ロジック・モデルが有用である。

<ロジック・モデルから指標へ—ポイントは？>

政策評価におけるデータの活用について話をするに当たり、ロジック・モデルを用いた指標設定について語るが、ロジック・モデルを用いて指標を設定する際、手順は2つのステップに分かれる。すなわち、意図の明確化と意図の指標化である。

まず、意図の明確化について。意図というのは何かというと、具体的に何をどのような状態にしたいのか、すなわち、対象と成功状態を明らかにしたものをいう。政策評価に当たってはまず意図を明らかにし、その上で意図どおり事が進んでいるかを測定する必要がある。

意図を測定する際に求められるのが意図の指標化である。指標とは意図を具体的に表わしたものであり、定性的なものであれ、定量的なものであれ、エビデンス、データを用いて設定するものである。

<意図の明確化のためにまず何が必要か—1stステップ>

意図の明確化のために、まず何が必要なのかというと、いわゆる使命、ミッションと言われるものの明確化ではないかと思う。それも、単に府省の使命だけではなくて、さらに局の使命、課の使命と明らかにする必要がある。使命、すなわち、それぞれの組織が責任を持って成し遂げなければならない任務、こういった使命を明確にすることは、ロジック・モデルを立てて政策評価を行う上での重要な1stステップになる。

<指標の問題>

これから本日の講演の本題に入る。まず、「指標の問題」について語る。指標に基づく評価というのは、現在の政策評価の大きな特徴である。しかし、実は、成果を指標化するというのは結構難しい。

<成果を指標化するの難しいー「豊かさ指標」の教訓>

「豊かさ指標」の教訓について。かつて内閣府の旧経済企画庁が、この国の豊かさを表わす、いわゆる所得以外の指標の作成を努力されていた。そうした中で同庁は通称「豊かさ指標」、正式名称は「新国民生活指標」と言われるものを、1992年に公表を開始した。これは国民の生活の豊かさを測定するということで作成された指標である。

しかし、この「豊かさ指標」というものに対して疑問が出てきた。総合順位、福井県が5年連続トップになった一方で、埼玉県が6年連続最下位になった。これを受け、当時の埼玉県知事は、企画庁の指標も1つの見方ではあるけれども、各都道府県の豊かさの実態を表わしているとは思えないと意見を表明するなど、様々な意見が出てきた。結果、「豊かさ指標」は99年に公表取りやめになった。成果を指標化するの是非常に難しいという教訓を「豊かさ指標」は与えてくれた。

<一人当たり（単位人口当たり）指標の落とし穴>

「豊かさ指標」は、一人当たり（単位人口当たり）の指標になっている。この一人当たり指標には二つ落とし穴がある。

まず、一人当たり指標は、人口が増加する地域ではこの指標は悪化する、逆に、人口が減少する地域ではこの指標は改善する、こういった特徴がある。しかし、そもそも人の移動というのは暮らしにくい地域から暮らしやすい地域へという形で起こるのであり、一人当たり指標が悪い地域ほど、実は暮らしやすい地域だということになる。

さらに、実は、この1人当たり指標については、もう一つ落とし穴がある。それは、偏ったサンプルからは偏った数字しかでないということである。「豊かさ指標」においては、地方が上位にくることが多い。なぜかというと、地方においては、生活基盤が安定していれば豊かな生活を送りやすい一方で、安定的な生活基盤を確保することが難しい人は都会に流出する。その結果、地方には豊かな生活を送っている人だけが残し、同じ地域における「豊かさ指標」の数値は非常に高くなるという理由がある。

<取組の成果がすぐに指標値の変化として表れるとは限らない>

指標による成果の測定の問題点として述べたいのは、取組の成果がすぐに指標の値の変化として表れるとは限らないということである。特に、これは言うまでもなく、教育とか、福祉とか、そういった分野においては、取組の成果が指標値の成果として表れるに

は、中期的・長期的な時間を要するものが多い。

＜より良い指標のために必要なのは？＞

より良い指標のために必要なものは何か。一言で言うと、これは統計データの精度の向上だと思う。指標を得るために、より多くの統計データをとると、間違いなく統計データの精度は向上する。しかし、これは言うまでもなく、コストの増加に繋がる。このように、統計データの精度向上については精度とコストのジレンマというのが出てくる。

そういった中で、より良い指標のために何が必要なのかということだが、そうすると、これから求められるのは、効率的、計画的なデータの収集と整理である。そのための手段としては、例えば、データの共有化、共同利用が挙げられる。いわゆる原局・原課に負担をかけず、それでいて政策評価というものをよりよいものにするためには、評価担当部署が中心となってこういったことを行う必要があると個人的に考えている。

＜政策評価に求められるもの—客観性と合理性＞

こういった中で、政策評価をより良いものにするために何が必要なのかということで、今日の話のテーマの3つ目になる。資料には「政策評価に求められるもの—客観性と合理性」と書いている。

＜政策評価は誰のため？＝「国民のため」＞

政策評価は誰のためのものかということ、これは「国民のため」である。総務省の作成している『政策評価Q&A（政策評価に関する問答集）』においては、制度の目的について、「国民本位の効率的で質の高い行政を実現すること」、「国民の視点に立ち、成果重視の行政を実現すること」及び「国民に対する行政の説明責任を果たすこと」を掲げており、全て「国民」をキーワードとして説明している。

＜そもそも論で言えば、評価は国民が行うもの＞

非常に乱暴な言いぶりになるかもしれないが、そもそも論で言うと、政策評価というのは、国民が行うものである。PDCAの流れでいうと、企画・立案の部分を行う（Plan）のは議会、国の場合、国会である。国会が決めた政策を執行する（Do）のは行政である。そして、それを評価する（Check）のは、実は主権者たる国民である。国民が、その評価結

果を民意という形でフィードバック（Action）する。

しかし、実際には、このような仕組みにはなっていない。なぜかという、現在の行政は非常に複雑化・専門化しており、国民にはそれを評価するのに必要なノウハウや情報がないからである。

<内部評価を基本とした評価の仕組み>

こうした中で、現在の政策評価制度は内部評価を基本とした仕組みになっている。すなわち、国会が決めた政策を行政が執行し、執行した本人であるところの行政が政策の内部評価、自己点検・評価を行っている。ただし、そもそも論で言うと、評価を行うのは主権者たる国民だということで、行政は政策評価の結果を国民に対して説明する責務を負っている。これを受け、主権者たる国民が、最後、Check、Actionと、民意という形で反映させるといふ流れになっている。

<国民が理解し、納得するためには、何が必要か>

ここで重要なことは、政策評価の結果は国民に対して単に説明すれば良いという話ではなく、その説明について国民が理解し、また、納得することまでが求められることである。

国民が理解し、納得するためには何が必要かという、政策評価法に書かれている。第1条においては「政策評価の客観的かつ厳格な実施を推進」することについて、第3条においても、「客観的かつ厳格な実施を図るため」、「政策効果は、特性に応じた合理的な手法を用いて行わなければならない」ことが定められている。政策評価の説明責任を果たすためには、第1条及び第3条で定められているとおり、客観性及び合理性が鍵になる。

<政策評価における客観性と合理性>

それでは政策評価における客観性と合理性とは何かについて説明する。

<客観性>

まず客観性について、大きく分けると3つに分けられる。まず、その指標というのが目標を適切に反映したものとなっているか。次に、良い結果が表われている指標や事例のみを用いて評価を行っているか。最後に、指標の動きなどに基づいて客観的に評価が行われているかである。このように述べるとどれも当然のこのように思われるが、実際に行お

うとすると、実は難しい。

<これらの評価は客観的と言えますか？>

ここで事例を2つ挙げる。これは、国の政策評価の中で実際に行われた評価の事例である。

事例①は事業化率を指標にとっている。目標値は40%。過去5年間の指標の推移を見ており、過去から順に39.1、38.9、38.3、36.9、32.9と推移している。この評価結果がどのように書かれていたかということ、平成x年度末時点で見ると、事業化率は目標値である40%に及ばないものの、堅調に推移しており、目標はほぼ達成しているものと考えられるという評価がされていた。これは客観的とは言えない。

事例②は充足率を指標としており、目標値は100%である。過去5年間の指標の推移について、80.5、80.6、80.0、81.3、79.6と推移しており、約80%でほとんど変わりなく推移している。これを踏まえた評価結果だが、充足率は、制度発足直後は75%であったが、啓発事業等を実施した結果、充足率は徐々に上昇し、ここ数年は80%前後で推移しており、制度の目的を果たしていると評価できると評価していた。果たしてこれは客観的と言えるだろうか。第三者が見ても納得できるだろうか。

客観性とは何か。一言で言うと、第三者が見て納得できるかということである。国民が納得するためには、この客観性が重要になる。

<合理性>

次に、合理性について述べる。

合理性について、まず問われるのは、正しい手法に基づいてデータの収集が行われているかということである。すなわち、データの信頼性は高いと言えるかということである。データの収集に当たり、正しい手法に基づいて行われているか。例えば、母集団から標本を抽出する場合、無作為抽出、ランダムサンプリングが行われているか。まずこれが合理性の1つ目になる。

次に2つ目、正しい方法論に基づいて評価が行われているか。この場合、方法論ということで、まずは分析枠組の方法論について、次に統計手法の方法論という形で見ていきたい。

<①合理性－分析枠組の問題>

まず、分析枠組という点での問題である。ここではコネティカット州内の交通事故によ

る死者数の推移の分析を例に説明する。

コネティカット州はアメリカのニューヨークに隣接する小さな州であり、州の真ん中を、アメリカの東海岸を縦断する高速道路が走っており、交通事故が問題になっていた。そこで、1955年12月23日に州法が改正され、速度違反者に対する罰則が強化された。しかし、これが隣接する州と比べると非常に厳しい罰則であったため、批判が出てきた。これに対し、これは当時のコネティカット州知事が、資料の図（スライド18番）を用いて説明した。

1955年には州内で年間324人の人が交通事故で亡くなっていたが、罰則を強化した翌年、1956年には、交通事故による死者が284人にまで事故死者が減少した。これは明らかな州法の改正による成果であると説明をした。このようにして、1つの事例をもとに、前と後で比較して評価を行うモデルを、「One case Before-After Design」、「1事例事前・事後比較デザイン」と言い、非常にポピュラーな枠組である。

実は、このコネティカット州の分析には合理性の観点から問題がある。資料のスライド19番は、死亡者の推移を1年間だけでなく5年間にかけての数字をとったものである。これを見ると、1956年の死者数も、1955年からは減少しているものの、それ以前の数字と比べると依然として高いということが分かる。

資料のスライド20番に示した図は、コネティカット州に隣接する州の推移を併記したものである。この図を見ると、コネティカット州の死者数は確かに減少しているが、罰則を強化していない隣接する州においても、実は全て事故死者は減っていることが分かる。そうすると、コネティカット州知事の説明は破綻している

次のスライド21番の図はコネティカット州の交通事故死者数の推移のグラフを51～59年までとったものである。そうすると、罰則を強化した翌年の1956年以降も死者数の減少が続いていることが分かる。

このグラフについて、隣接する諸州と比べたものがスライド22番である。そうすると、コネティカット州においては、継続的に事故死者数の減少が続いている点に特徴があることがよりよく分かる。

「One case Before-After Design」はよく用いられるデザインだが、評価という枠組においては、非常に脆弱なモデルである。分析枠組の合理性という観点で言うと、過去やよそと比較することが、評価においては非常に重要な視点になる。このコネティカット州の事例は、比較の重要性というのを表わす事例である。

<合理性—統計手法の問題>

合理性に関して、次に統計手法の問題について述べる。資料のスライド23番に、実際の新聞記事から抜粋したものを掲載している。1世帯（2人以上）あたりの貯蓄の平均残高が1798万円となったことを記載したものである。ここで問題とするのは、平均という数字が、全体の傾向を表わす適切な指標とは言えないということである。

資料のスライド24番には「貯蓄現在高階級別世帯分布—2014年—（二人以上の世帯）」というグラフを掲載している。横軸は貯蓄残高、縦軸は世帯割合である。これを見ると、先ほどの新聞記事で二人以上の世帯の平均貯蓄額として記載されていた1,798万円は、全体の分布を代表する値とは言えないことが分かる。この分布を見ると、割合が一番高いのは貯蓄額が100万円未満の世帯である。また、中央値という点から見ると1,052万円である。

更に、このグラフは、退職金をもらった高齢者世帯も含まれており、貯蓄残高を押し上げている。スライド25番は、2人以上の世帯のうち、勤労者世帯を見たものである。

平均値は全体を表わす代表的な値とは言えず、分布を表わす代表的な値は中央値、もしくはmode（最頻値）、値の一番高いところになる。こうした基礎的な統計手法が、実は正しく使われていないことがある。

次の事例は「サラリーマンの小遣い 月4万8,800円—民間調べ 今年3,300円増」というものである。これも実際の新聞から抜粋したものである。ここでは、サラリーマンの小遣いが平均月4万8,800円であり、昨年に比べて3,300円増えており、景気回復が会社員の懐を潤し始めていると分析・評価している。これは平均値の動きをもとに評価している。しかし、この分析・評価には問題がある。次のスライドは、この記事の元となった資料である。この表を見ると、83.2%が小遣いの変化無し、8.6%がダウンしており、小遣いがアップしたのはわずか8.2%であることが分かる。

平均というのは、全体の傾向から極端に離れた値（外れ値）の影響を受けるため、全体を代表する値としては好ましくない場合がある。

<代表値>

この全体を代表する値、統計の中で分布を代表する値、これを代表値という。代表値の種類としては、最頻値（mode）、中央値〔中位数〕（median）及び平均値（mean）がある。

代表値の中でも最もポピュラーなのが平均値である。平均の中でも、算術平均、幾何平均といろいろなものがあるが、非常にポピュラーなのは、単純平均、すなわち算術平均で

ある。

<正規分布>

算術平均が最もポピュラーである理由について説明したのが29番のスライドである。正規分布のグラフの特徴は、中央が一番高く、また、両側に向かってだんだん低くなっていき、左右対称の釣鐘型をしていることである。正規分布の場合、この中央の一番高い位置が最頻値 (mode) であり、中央値 (median) であり、さらに単純平均、算術平均である。このため、正規分布の場合は単純平均を用いることで全体の傾向を表わすことができる。

しかし、先ほどの貯蓄の場合や小遣いの場合、分布が正規分布ではなく、算術平均は適切な代表値と言えない。

<合理性—因果関係と相関関係を混同しない>

さらに、合理性という観点で重要なのが、因果関係と相関関係を混同しないことである。相関関係というのは、2つの要素、要素A、要素Bの間に共変関係が見られる関係をいう。ただし、相関がある、共変関係があるという場合、偶然の一致による場合や、因果関係による場合、因果関係の逆転による場合など、様々な場合がある。

因果関係は、相関関係の中でも、要素Aが変化することによって、もう一方の要素Bの変化を引き起こされる関係を言う。因果関係が成立するためには、条件が3つ必要である。まず、要素A、Bが共に変化すること (共変関係) こと。次にAが先に変化することで、時間的にその後でBが変化すること (時間的な先行関係) こと。最後にA以外の要素の変化を固定しても、A、B間に共変関係が存在すること。これらの条件が成り立って初めて因果関係が存在すると言える。

因果関係と相関関係の違いを説明した事例が32番のスライドである。火災現場でよく見かけるちょっと人相の悪い男という要素Aと、火災現場という要素Bがある。この2つの要素を因果関係で捉えると、この男は放火魔になる。しかし、相関関係に過ぎない場合、この男は、消防士かもしれないし、火災現場を取材している記者かもしれない。あるいは、野次馬かもしれない。

33番のスライドは年収と血圧の関係についての事例である。この二つの要素の間には、一見、非常に高い正の相関が見られる。しかし、これは、真の相関ではない。つまり、年齢が上がると年収が上がる。また、年齢が上がると血圧が上がる。結果、年収と血圧の間

に相関が表われている。このような相関は見せかけの相関という。合理性という観点から、相関関係には注意すべき点が多い。

<政策評価の充実と活用に向けて—比較性の向上>

最後に、政策評価の充実と活用に向けて、政策評価をよりよいものにするために何が必要なのかについて申し上げる。評価において重要なのは、やはり比較性を向上することである。比較という場合、2つの視点がある。

まずは経年比較、過去と比較するものである。時系列分析とか縦断面分析ともいう。経年比較をする際には、指標の連続性に留意する必要がある。同じ指標を用いていても、年度により定義が異なっていると連続性は失われ、適切な比較を行うことができない。

さらに横断面分析、国際比較など、よそと比較するものである。横断面分析についても定義の問題があり、多国間で単位が一致せず、適切な比較を行うために指標やデータについて調査・計算を行う必要がある場合がある。しかし、評価を行ういわゆる原局・原課は、時間・費用のコストを費やす余裕がない場合が多い。そのため、評価担当課には、指標やデータを収集・整理し原局・原課に提供をすることが求められると考える。

<政策評価を活かすためには—データを読み解くのは人>

政策評価を生かすためには、データを読み解くのは人であることに留意する必要がある。政策評価という場では、建前としての役所の理屈と本音としての住民感覚との間での議論が生じる。この中で求められるのは、Professionalism=専門家としての立場とAmateurism=住民感覚（「sympathyとempathy」（同情と共感））の両立である。先に触れたとおり、政策評価法の目的においては、国民のため、国民の視点に立って、これが重要であり、専門家としての立場だけではなく、住民感覚が求められる。

しかも、評価は単に弁明、確認のための仕組みではない。評価は改善（進化）のための仕組みである。評価は実はこうした当たり前のことを積み重ねることであるが、これができる組織とできない組織で、間違いなく大きな差が生じる。

本日の講演で、原局・原課に負担をかけずに政策評価をよりよいものにするためには何が必要か、皆様に御理解いただけたものと思う。