HOKUGA 北海学園学術情報リポジトリ
学校法人北海学園
北海学園大学
北海商科大学

| | |
|---|---|
| タイトル | マーケティングにおける予測と数学の関連性について |
| 著者 | 黒田，重雄；Kuroda，Shigeo |
| 引用 | 北海学園大学経営論集，17(1)：59-80 |
| 発行日 | 2019-06-25 |

Hokkai-Gakuen Organization of Knowledge Ubiquitous through Gaining Archives.

《研究ノート》

# マーケティングにおける予測と数学の関連性について

黒　田　重　雄

目　次

はじめに
1．人間は予測で生きている
2．予測のはじまり
3．マーケティングは社会科学になりうるか
4．マーケティングも，予測することを中心テーマとしている
5．ビッグデータの存在が脅威となってきた
6．ポスト・ビッグデータの問題
7．これまで，マーケティングでは，どのような予測法を採用してきたか
8．マーケティングにおける方法論は如何にあるべきか
9．AIでは，どのような予測方法を考えているか
おわりに
注と参考文献

## はじめに

大学教育の在り方に対する要望として，経団連から「文系学生に数学を，理系学生にリベラルアーツ（教養）を」が出されている[1]。

データ時代の人材求める

文系の大学生も数学を学ぶべきだ ──。経団連は若い人材の育成に向け，文系と理系で分かれた大学教育を見直すべきだとする提言をまとめる。近く大学側と対話する場を設け，意見交換をする方針だ。経団連は日本の大企業が加盟し，新卒の採用に大きな影響力を待つ。デジタル分野の人材確保に向け，大学に改革を迫る。

経団連は12月3日に開く正副会長会議で人材の採用や大学の教育改革に関する提言をまとめる。大学との対話は定期的に開き，経団連からは中西宏明会長のほか副会長らが参加する。大学側からは国立・私立大の学長の参加を広く募る。

大学には文系と理系でそれぞれ偏りすぎた教育内容の見直しを迫る。ビジネスの現場ではシェアビジネスやデジタルマーケティングが広がり，統計などの知識が必要と考える経営者は多い。データを扱うために「最低限の数学」を学生が学び続けるよう求める。

理系の学生に対しても「リベラルアーツ（教養）」の充実を求める。グローバルに活動する企業は従業員の国籍が多様になり，他国の文化を理解しながら働く人材が求められる。

文系学生に対する数学の必要性は，昨今のAIのディープ・ラーニング（Deeplearninng：深層学習）の普及からも必要性が叫ばれる素地があることは分かる。

しかしながら，そこまで行くには相当程度のマーケティングというもの理解と数学や統計学を理解していなければならないだろう，というのが本拙論のテーマである。

## 1．人間は予測で生きている

未来は，“絶対”こうなるとは言えないが，人は，“明日”を，“将来”を予測しながら日々行動しなければならない生き物である。

漫然と闇雲に行動していては日常生活を維持できない。人類は，仕事をせずして生きながらえられない仕組みを構築してきたともいえる。つまり，仕事をするためには将来の予測をしっかりしなければならないことを意味する。どうするのか。古来，人々は予測の精度を上げるためいろいろな方法を考えてきた。その最たるものが，天候を予測することであった。

**天気予報について：**

実際上，天気予報の精度は以前に比べてあがっているという[2]。「24 時間以内に 1 ミリ以上の雨があるか」という予報の的中率は，東京で現在 85％となっている。50 年前の的中率が約 70％であったことを考えると精度はあがっていることになるとある。

**地震の予知について：**

これも東日本大震災にときのように，ほとんど当たっていない。予知なく突然やってくるものという印象である。

天気予報，地震予報も現代の技術の粋を集めて検討していると思われるのであるが，こうした自然科学系の予測でもなかなか当たらない。ましてや人間社会を取り扱う社会科学系の将来予測は当たらないのは当たり前といったところである。

この世の中に絶対はない。しかし，このことをどうやって表現するかは簡単ではない。

文豪バルザックは，文字通り『「絶対」の探求』という題名の小説によって，この問題に取り組んだ男の悲惨な一生を描いている[3]。これは，フランス王政復古期の一地方都市を舞台に，科学上の巨大な難問され，万物に共通する物質「絶対」の研究に打ちこむ男パルタザールは探求のはてに「ユーレカ（見つけたぞ）！」と叫んでむなしく息たえる。

「絶対」を解き明かそうとの情熱に憑かれた人間の偉大と悲惨，「絶対」という観念のもたらす恐しい力を旺盛な筆力と緊密な構成で見事に描ききった名作と言われている（同訳書の帯の言葉より）。

未来は，“絶対”こうなるとは言えないが，人は，“明日”を，“将来”を予測しながら行動する生き物である。

一番の基本は，「どのような事業をするか，何を作るか，どんな製品を作るか」である。

マーケティングを科学として認知したいというのには，理由がある。

「予測したい」からである。これからどんな事業をしたらよいか，どのような製品を作ればよいか，を考えたいからである。

人には本能がある。生き続けることと子孫を残すことである。それまで狩猟採集による自給自足の生活に別れを告げ，人類は互いに生きるため，生き残るためのシステムを発明し続けてきた。

これは，今からおよそ 7 千年前（紀元前 5 千年）にいわゆる商人 merchant（後に，ビジネスマン businessman）が生まれたことに端を発している。

商人はどうして生まれたか。紀元前 8 千年のメソポタミヤ地方における農耕生活に発するといわれる。そこではエジプトのナイル（毎年決まって氾濫し農耕が出来た）と違って，チグリス・ユーフラティス（大河）の氾濫が不定期であったため，農耕に支障をきたした。

生活の糧を得なければ生きてはいけない。何をすればよいか。人々が生きるための日常生活品を調達するため遠くへ出かけ物々交換するしかなかった。そのとき他の人に頼まれた物資を運んできたりして「お礼」を受け取っていたが，遠距離を運搬するようになって次第に専門化し，「お礼」がやがて「利益 profit」に転化し，彼らは，“merchant”と呼ばれるようになった存在である。

岩井克人（1992）は，人類の歴史において，

この「遠距離交易」ともう一つの「貨幣の発明」が2大発明と述べている[4]。

確かなのは，その当時の商人にとって，どこに誰がいて何を欲しているかを知ることは至難なことであっただろうということである。わずかばかりの情報を頼りに彷徨い歩いたに相違ない。この状況でも，特に重要だったのは「予測」であり，それなりに懸命に考えて行動に移したに相違ない。また，大きなものをどうやって運ぶかも次の重要問題であったろう。

しかし，彼らは，自分のためと他の人のためこの過酷なことをやり遂げていた。やり遂げなければ皆死ぬしかなかったということもある。この点は今日のビジネスにも当てはまると考えられる。

実際上，これが今日言うところのビジネスの始まりである。ビジネスは，7千年の歴史を持っている。

今日でも，ビジネスにとって重要なのは「予測」である。したがって，もし，ビジネスを学問にするとなれば，「予測」の方法をどうするか，何にするか，が必須の要件であるということになる。

### 井原西鶴と投機

かつて筆者の生家では農家だったので，両親が言っていたことを思い出す。「西の空の夕焼けが美しいので，明日の天気は晴れだ，朝早くから田植えができる」

当然，古でも，例えば，「西の空の夕焼けが美しいので，明日は遠出して狩猟だ」ぐらいは言い合っていたに違いない。生活と予測は切っても切れない糸で結ばれていたはずだからである。江戸時代には先物市場の投機で使われていたことを井原西鶴も「日本永代蔵巻一」（1686）に書いている[5]。

> 波風静かに神通丸
>
> 惣じて北浜の米市は，日本第一の津なればこそ，一刻の間に五万貫目のたてり商も有る事なり。その米は蔵々に山をかさね，夕の嵐，朝の雨，日和を見合せ，雲の立所をかんがへ，夜のうちの思ひ入れにて，売る人有り，買ふ人有り。

天気予報，地震予報も現代の技術の粋を集めて検討していると思われるのであるが，こうした自然科学系の予測でもなかなか当たらない。ましてや人間社会を取り扱う社会科学系の将来予測は当たらないのは当たり前といったところである。

## 2．予測のはじまり

マーケティングは「予測の方法」を求めている。これは，人類が予測しながら生きながらえてきたことと関連している[6]。

人類最初の文明は，メソポタミヤ地方に発生した言われている。このことは，大河の氾濫と関係していて，エジプト文明におけるナイル川とメソポタミヤ文明のチグリス・ユーフラティス川の氾濫の違いに起因している。川が，定期的に氾濫して農業が毎年のように成立したエジプトと川による不定期の氾濫で毎年の農業が不成立であったため「遠距離交易」を活発化させることになったメソポタミヤ地方の違いを生んだ結果であった。

マーケティングを講義する側は，ビジネスマンに対して，天候とビジネス行動の関係ぐらいの話ができる必要性があると考えるのは筆者だけではあるまい。

ひところ，MISとかPOSとか言っていたが，今や，SNSとかスマホとかビッグデータといった言葉が飛び交っている。筆者のような老研究者には新しい言葉についての応接に暇がないどころか，ついていくのもままならないという時代に入っている。

2014年3月の新聞を見ていておや？と思う記事が出ていた。「ビッグデータで創造す

る新時代のマーケティング戦略」というフォーラムの案内である。その文は，

> 世の中のあらゆる情報がディジタル化され，その情報をビッグデータとして集めることがもはや当たり前になってきた今日。集めたデータを「どのように分析し活かしていくか」ということが企業の課題となっています。とりわけ，企業の経営戦略に不可欠なマーケット分野において，ビッグデータをどれだけ効率的に活用できるかということが，企業経営の重要な鍵を握ることは明らかです。そこで本フォーラムでは，ビッグデータを戦略的にビジネスに活かすために必要な，マーケット分野における活用方法に焦点を当てます。ビッグデータの収集，データベース作り，分析・解析，そしてそこからのビジネス戦略策定にあたりどのような手腕が求められるのか？事例を交えてご紹介していきます。

これを見て，ふと，どこかで読んだ記憶があることに気が付いた。しばらくして，それはアメリカにおいて「マーケティング・リサーチ」というものが生まれた記述であることに思いがいたったのである。大不況を経験した後の 1930 年代に生まれた「何をして生きて行くか」から生まれた言葉であった。

マーケティングという言葉は，もっとずっと前に現われていたが，そこでの戦略などは不況期にはほとんど役立たずであった。そこを乗り切る手立てが「リサーチ」であることを認識させたのであった。

上記の文は，まさにそれと一緒の状況を思い起こさせる。80 年前と内容において全く変わっていないのである。

そこで挙げられている調査可能な事柄の「商品，企業組織，市場，人口，富，賃金，価格，１人当たりの消費者収入，生活水準，特定商品の市場，商慣習，購買意欲，潜在市場」等が「ビッグデータ」という言葉で一括されているだけである。

世の中，言葉や状況は変化するが，本質は変わっていないのである。考えてみれば，マーケティング・リサーチと言われるものこそ「自己のビジネスを探すこと」であり，それはすなわち，「マーケティング」のこととなるのである[7]。

マーケティングを学問にするためには，現代のマーケティング・リサーチを研究しなければならないということにもなるのであって，したがって，そこにおける方法論を吟味する必要があるということに繋がっていくのである。

## 3．マーケティングは社会科学になりうるか

大塚久雄（1967）は，ヴェーバーの「社会学」の解釈から[8]，

> その認識の対象は終始生きた，自由な意志をもって行動する人間だ個人というものは，どこまで意識的であるかは別といたしましでも，とにかくたえずさまざまな目的を設定し，その手段を選択して，決断をしつつ行動するのであり，社会現象はそういう生きた人間諸個人の社会的行動の軌跡にほかなりませんから，およそ社会現象のなかには，人間行動における目的―手段の連関，いわゆるテレオロギー，目的論的な関連が奥深く含まれていることは申すまでもありません。したがって社会現象のなかでは，自然現象のばあいとちがって，そうした目的―手段の目的論的な関連というものを辿っていくことができるわけです。
>
> 社会現象のなかにいくらそうした目的論的な関連を追求しても，それだけでは，自然科学が科学であるというような意味において，科学的認識は，したがってすぐれた意味での社会科学というものは，どうしても成りたち

えません。しかし，目的論的関連を含んでなりたっている社会現象を対象としても，われわれは因果性の範疇を使用して十全な意味で科学的認識を成立させることができるんだということを根拠づけているのであります。その解決の仕方を私は「目的論的関連の因果関連への組みかえ」というふうに呼ぶのですが，……，それはどういうことかと申しますと，「社会学とは，社会的行為の〔主観的に思われた〕意味を解明しつつ理解し，それによってその経過と影響を因果的に説明しようとする学問」だということになる。つまり社会科学的な認識のばあいには，そうした自然科学にはみられない，「動機の意味理解」ということがつけ加わってくることになり，そこに科学性を持たせる。

大塚は，なぜ木曜日に人通りが多いかという問題に対して木曜日にはサッカーの試合があるからといったように意味理解が成立することが社会科学を成立させる根拠だと，言った。

その点を具体的に考えてみると，以下のようになる。

事象Ｘと事象Ｙとの間に関係があり（相関係数で確かめる）。

Ｙ＝ｆ（Ｘ）　：　ｆは関数。

Ｘが起これば Ｙが起こる（その逆ではない）という意味理解ならば，

Ｙ＝ａ＋ｂＸ（ａ，ｂ：パラメーター（常数））

という回帰式であらわしてデータにより検証する。データの数が多ければ多いほどよい，のであるが，どれほど多ければよいのかわからない。大事なことは，真の関係を求めようとしないことである。（蓋然性という）

その関係がある程度実証されれば，ビジネスマンにはＸが起これば Ｙが起こる可能性は高いですよ，ぐらいの話をする。

予測法としては，回帰分析が一般的である。経済学では，経済変数の予測に使われる。計量経済学理論として精緻化されている。一般均衡理論を一次の連立方程式体系であらわし，ときに将来の均衡価格（体系）や均衡需要（体系）として求められている。

経営学では，沼上が，「カバー理論」を用いて予測問題に活用することを示唆している。ここでの「カバー理論」は，相関分析と同様の内容を持つものと理解されている。

いずれの学問でも，予測の分析には多変量解析手法が用いられる。つまり，論理実証（経験）主義に基づく帰納法の考え方が採用されている。例えば，経済学では，演繹法が採用されるが，帰納法との併用で用いられることが多い。

社会学においては，富永健一（1999）では，論理実証（経験）主義に基づく帰納法で分析することが示唆されている[9]。

こうして，各種学問では一般的に，統計学における統計的手法が援用される。そこでは，重相関・重回帰分析法（数量化理論第１類），判別分析法（数量化理論第２類），因子分析，主成分分析法（数量化理論第３類），クラスター分析法などを駆使して分析する。

チェーンストアやコンビニエンス・ストアなどでは日々量的・質的データが蓄積され，膨大な量を抱えている。こうしたものに対しては，大量の処理を可能にするコンピュータを用い，例えば，データ・マイニング法などを駆使して分析を行っている。

その中の一つとして，アメリカで消費者の行動で明らかになった有名なものとして以下のような例が紹介されている。

「アメリカの若い男性は，赤ちゃんのオムツを買うとき，一緒に缶ビールを買っていることがデータで確かめられた。そこでドラッグストアなどでは，赤ちゃんのオムツの横に缶ビールを並べて成功した（データ・マイニング法の成功例）」。

## 4．マーケティングも，予測することを中心テーマとしている

マーケティングとは，「どのような事業をするか，何を作るか（どんな製品・サービスを作るか）」である。これは筆者の「マーケティングの定義」である。

**グリーン＝フランク**（P. E. Green & R. E. Frank）（1967）も，「**どこの，誰が，何を，何時，どのようにして」の 5 つが，経営（マーケティング）が営まれてきたと述べている**[10]。

**この**マーケティングを科学として認知したいというのには，理由がある。

「これからどんな事業をしたらよいか，そこではどのような製品を作ればよいか，それをどのようにして売って行けばよいか」について，より科学的に「予測をしたい」からである。

### マーケティングの方法論争はどうなっているのか

マーケティングは「予測の方法」を求めている。マーケティングを講義する側は，ビジネスマンに対して，天候とビジネス行動の関係ぐらいの話ができる必要性があると考えるのは筆者だけであろうか。

人類最初の文明は，メソポタミヤ地方に発生した言われている。このことは，大河の氾濫と関係していて，エジプト文明におけるナイル川とメソポタミヤ文明のチグリス・ユーフラティス川の氾濫の違いに起因している。川が，定期的に氾濫して農業が毎年のように成立したエジプトと川による不定期の氾濫で毎年の農業が不成立であったため「遠距離交易」を活発化させることになったメソポタミヤ地方の違いを生んだ結果であった。

氾濫の起こらないメソポタミヤの人々は，家財道具一式を持ち生活物資との交換を求めて遠くまで彷徨い歩かねばならなかった。

こうした歴史的考察を行うと，約 1 万年前にメソポタミヤ地方で彷徨い歩いた人々は，アメリカにおける大不況期の人々と重なることが示される[11]。大不況期を経験したアメリカでは，“*Marketing Research*”（MR：マーケティング・リサーチ）という研究領域が生まれた。

以降，マーケティング・リサーチは発展の一路といった感がある。20 世紀半ばには日本にもいち早く移入されている。

米国では，AMA（会員は世界中に 3 万人といわれている）から雑誌 “*Journal of Marketing Research*”（*JMR*）が 1964 年に生まれている（同じ AMA から出ている “*Journal of Marketing*”（*JM*）は，1936 年発行されている）。

マーケティングの科学性に関しては，*JMR* の研究が参考となる。*JMR* の動向については，この雑誌の 30 周年記念号として書かれたバス（F. M. Bass）の論文（1993）「マーケティング研究の将来―マーケテイング科学―」に注目したい[12]。その中で，バスは，次のように述べている。

> 1964 年 2 月に発刊されたときの *JMA* 会長の趣旨は，次のようなものあった。
>
> *JMR* の焦点は，マーケティングの研究（research in marketing）における方法論と哲学的，概念的，また技術的諸問題におかれている。こうしたマーケティングにおける科学的方法を広範囲に研究することにより，マーケティング・リサーチ（marketing research）への関心を大いに高めるものとなった[13]。
>
> マーケティング研究の将来の方向性は，過去 30 年間に生み出されてきたもので，マーケティング科学（marketing science）の発展であった。
>
> 「科学」というものは，以下の 3 つの要素，
>
> 1）経験事象の概念化（empirical generalization）
>
> 2）概念化の弁明（generalized explanation）
>
> 3）拡張，修正，最新化の過程（a process of

extension，revision，and updating）を有するものであり，マーケティング研究においても，同様である。

そして，また新しい《現象》の発見によって，マーケティング科学をより一層進展させて行くのである。

こうして*JMR*は，*JM*が広範囲の問題を取り扱っているのに対して，徹底的に科学的な姿勢に基づき，マーケティングの問題にアプローチしてきている。例えば，購買行動過程，消費者の嗜好，ブランド・ロイヤリティ，調査の態度，世論調査，トレンド分析，商品テスト，メディアの選択，広告の測定，コンピュータの応用などである。

## 5．ビッグデータの存在が脅威となってきた

雑誌『Newton ニュートン』（2013 年 12 月号）では，特集「統計の威力」と題する特集を組んでいる[14]。それによると，

ビッグデータとは，現時点では「企業が日々の経営で記録している膨大なデータ」を意味することが多い。たとえば，携帯電話やカーナビゲ―ションシステムに記録される利用者の位置情報や，クレジットカード会社が処理する取引履歴，ウエブサイトに入力される検索ワードなどだ。

……

企業が業務のすみずみにまでコンピュータを取り入れたことで，自動的に蓄積していく「ビッグデータ」出現した。わざわざ時間をかけて統計分析のためにデータを集めなくても，ビッグデータを分析できれば，母集団のデータそのものから，すばやく情報を得られるようになると期待できる。

また，ビッグデータは，以下のような使われ方もある，と例を紹介している。

まだ撮影されておらず，俳優も決まっていないのに，脚本だけで「売れる映画かどうか」が予測できる！（過去の映画から「売れる要素」を見つけ出し，脚本を評価する）

映画会社は，無数の脚本から「売れそうな脚本」を選び，製作費を投じる。莫大な製作費を投じたにも関わらず，売り上げが低く赤字になることや，逆に期待していなかった低予算の映画がヒットすることもある。これまで，興行成績の予測はむずかしかった。

イギリスのエバゴギクス社は，過去の映画の筋を無数の要素に分解し，興行成績と組み合わせて，なんと脚本から興行成績を予測するプログラムをつくりだしたと発表した。ここでは「ニューラルネットワーク」という技術が使われている。ある映画会社の進行中の映画の脚本 9 本の内， 6 本について収益を正確に予測したことが評価された。映画会社が１億ドルの売上を見込む脚本について，エバゴギクス社は 4900 万ドルと予測し，結果は 4000 万ドルに満たなかったという例もある。制作が決定している脚本をどうかえればヒットするかアドバイスできるため，過去の映画脚本をリメイクし，興行成績の予測を発表することで，予測能力を示したいとしている。

何でもビジネスになればよいという世の中。倫理観・道徳観の欠如した日本人の姿がクローズアップする。

そういえば，ホリエモンも，本人は証券取引法違反で逮捕投獄されているが，「まだまだビジネスはある」と言う。

こう考えると，ビジネスの芽はいくらでもありそうに思えてくる。ビッグデータには功罪半ばする要素がありそうだが，われわれは，好むと好まざるに関係なく，ビッグデータに取り囲まれている。

ビッグデータ分析の必要性の増大

世の中，「情報ネットワークの時代」だが，そこには情報量急速拡大して，いわゆるビッグデータの解析が重要性を増してきたと山本強（北大情報工学研究所教授）が解説する[15]。すなわち，情報ネットワークの時代は，

：人に向けた情報配信のネットワーク ── WWW による情報発信，遠隔端末，ファイル共有サービス

：情報収集のネットワーク ── データベースとしての WWW（検索エンジン）ユビキタスネットワークによるオンライン情報集約

インターネット上に人知の理解を遥かに超える巨大データベース空間が存在している。

身近には巨大なデータベースがある。

：ネット情報（WWW コンテンツ，オンライン・データベース）

：POS 情報 ── 大規模店舗における販売実績データ

：自動計測されているデータ ── 気象データ（AMEDAS），交通データ（自動改札データ），通信履歴（WWW アクセスログ，通話履歴）

：GPS ログ情報（カーナビ）

：デジタルカメラ画像，ムービー映像

ここで起こっている問題は，「巨大数値データベース・ブラウジング」である。

二つの問題がある。

（1）時間的制約：データベース・レコード数が巨大

── 中規模スーパーチェーンでも年間 1 億レコード超

── 単純な集計作業でも時間単位の処理時間

（2）多様な可視化要求：構造が単純なるが故に，多様な可視化や分析が重要になる

── フレキシブルな可視化ツールが重要

一方で，「情報科学はどこへ向かうか」として，以下の 3 点を上げる。

・情報の完備性と網羅性の向上

── 巨大ストレージシステムと高速ネットワークが可能にする網羅的データの収集と蓄積

── プロセッサの高速化が可能にするリアルタイム分析と可視化

・持続可能性（Sustainability）

── 情報処理に必要なエネルギーの低下（FLOPS/W の向上）

── 低炭素社会への情報科学の貢献（情報処理による総エネルギー消費の削減）

・システム研究からサービス研究へ

── 物理的実体としてのコンピュータではなく，機能としてのコンピューティングの研究開発

── ハードウエアが見えなくなる“クラウドコンピューティング”の時代

などであった。

クラウドとは，クラウドコンピューティングを略したものであるが，この概念は，データを自分のパソコンや携帯端末などではなく，インターネット上に保存する使い方，サービスのことである（筆者も活用している）。

一方では，消費者は，情報氾濫で悩んでいる。

ビッグデータは，“生ゴミである”と言う人もいる。生ゴミと言う先生も，そうした生ごみの中から自己のビジネス（哲学者）を選択したことにはならないのだろうか。

NHK の『クローズアップ現代』（2012 年 5 月 28 日放送分）で「社会を変える“ビッグデータ”革命」が放映された。

スマートフォン，IC カードなど身近な電子機器から，私たちは膨大な情報を発信している。インターネットで検索した内容，買い物をした商品や価格，駅の改札を通った移動，

さらには病院で受けた検査結果まで，あらゆる情報がデジタル化され記録される時代。生まれるデータの量は，この数年で飛躍的に増え，“ビッグデータ”と呼ばれている。解析不可能だったビッグデータを技術の発達で分析できるようになったことで，生活や社会が劇的に変わりつつある。コンビニでは，購買行動をリアルタイムで捕捉しパターンを発見，利用者が買う商品を事前に予測する。カーナビを使って 100 万台の自動車の位置情報をつかむことで急ブレーキ地点を地図化，“未来の事故現場”を見つけて事前に事故対策をする。アメリカでは医療分野でビッグデータを活用した“先読み”をする医療が加速している。一方で個人の情報が膨大に広がっていくことを懸念する声も。“ビッグデータ”時代の最前線を見ていく。

ここでの“ビッグデータ”（big data）を，ダイヤモンド社編集部では以下のように解説している[16]。

ビッグデータとは何か（ダイヤモンド社編集部による解説）

「ビッグデータ」とは，数多くの情報源から，さまざまな形態で，多くはリアルタイムで収集される膨大なデータ・セットをいう。B2B の場合，ソーシャル・ネットワーク，電子商取引サイト，顧客との通話記録など，多種多様な情報源が考えられる。これは企業の CRM（顧客関係管理）データベースに入っている通常のデータ・セットではない。数十テラバイトからペタバイト級の大容量と複雑さゆえに，データの収集，管理，マイニングには専用のソフトウエアと分析技術が必要である。非構造化データ（ネット上での特定ブランドに関するコメントなど）から営業上の知見を得る，地域の天候パターンを評価してビールの消費量を予測する，競争環境を綿密に理解するなど，その用途は実に幅広い。

（注：テラバイト（terabyte；$2^{40}$）約 1 兆バイト，ペタバイト（petabyte；$2^{50}$）約 1000 兆バイト）

ここで，クラウドとは，クラウドコンピューティングを略したものであるが，この概念は，データを自分のパソコンや携帯端末などではなく，インターネット上に保存する使い方，サービスのことである。

クラウドとは cloud（雲）のことであるが，雲の形はいろいろに変化する。このことから，クラウドは，変化するネットワークにも通じるものである。

## 6．ポスト・ビッグデータの問題

『統計学が最強の学問である』[17] という本がベストセラーになったということもあってか，雑誌『現代思想』（2014 年 6 月号）は，「ポスト・ビッグデータと統計学の時代」と題する特集を行っている。その中で，筆者の関心を呼ぶ評論が掲載されている。

データ科学者の水田正弘は，ビッグデータもスモールデータも本質的には変わらないという[18]。

データを活用する人にとって，スモールデータもビッグデータもすべての全てのデータ（any data）が重要である。統計学は，「従来，扱うことにできなかったデータを解析する」ために，確率論やコンピュータを活用してきた。ポストビッグデータがである any data は，統計学が率先して，情報技術，ビジネス，医療などと協力しつつはってんさせなくてはいけない。データを活用する人にとって，統計学は頼りになる存在である。

社会学者の太郎丸博氏も類似の見解を示している[19]。

そして，数理統計学者の竹内啓氏は，ビッグデータに対する統計学の適用に警鐘を鳴らす[20]。

ビッグデータと統計とはきわめて密接な関

係があると思われるかもしれないが，そこには微妙な問題が含まれている。……。ビッグデータに統計的方法を適用するに当たっては，４つの段階を経なければならない。１．データの吟味，２．モデルの選択，３．手法の選択と適用，４．結果の解釈と判断，である。

マーケティングにおけるビッグデータ

人間社会における，ある問題に対して，ルールが明確に知らされ，データ上の信頼性が間違いないものであれば，予測するために深層学習法は有効であるというのは分かる。

社会科学の場合，上記のことが保証されない場合が普通である。にもかかわらず，現実にはいろいろな意思決定問題に直面する。

経営戦略・ディジタル・マーケティング専攻の牧田幸裕（2018）は，日本経済新聞の『やさしい経済学・デジタル時代のマーケティング戦略』の締の項で，「企業の未来像大胆な構想必要」と書いている[21]。

これまで述べてきたように，デジタルマーケティングは何か特別なことをするのではなく，従来のマーケティングのある部分が進化したものです。これまで行ってきた企業活動をデジタル化や新技術の力で精度を高めたり，利便性を向上させているだけです。

現在，多くの企業がマーケティング分野に限らず，企業全体のデジタル化に取り組んでいます。

しかし，どのようにデジタル化を進めていくべきか決めかねている企業も少なくないようです。その原因は企業の未来像を構想できないことにあります。技術革新のスピードが速すぎて，過去から未来を想像することが難しくなっているからです。

市場環境分析の枠組みに「PEST 分析」（企業を取り巻く政治，経済，社会，技術の 4 つの面から企業への影響を分析する）がありますが，これは過去から現在の変化が業界に与える影響をもとに未来を予測します。過去と現在，そして未来に連続性があれば機能しますが，技術革新のスピードが速すぎると未来の予測が難しくなります。

しかし，現在の技術レベルでは不可能なことも，未来では実現できるのではないかと発想し，未来を構想する。そうすると，実現の時期はずれるかもしれませんが，未来を予測できるようになります。仮説検証型で未来を予測し，企業の未来像を描くのです。

未来は見えないので企業の未来像を描くのは勇気が必要ですが，思い切って未来像を構想する企業はデジタル化を推進できるようになります。多くの企業ができないからこそ，未来像を構想することは企業の競争力を決める重要な要素になり，競合企業から抜け出すチャンスになります。

この連載で解説してきたように，デジタルマーケティングの目的は消費者行動データを分析することで消費者を理解し，それをもとに良い提案をし続けて消費者から高い信頼を獲得することにあります。江戸時代の商人は大福帳に顧客情報を記して顧客理解に努めていました。時代が変わって方法は変化しても，商売の本質は変わりません。

筆者も，現代社会とビッグデータや AI との関係について書いてきている[22]。

マーケティングでも分野によって，ビッグデータというものの受け取り方が違いそうである。

筆者の場合は，企業のグローバル化との関係で「国際（比較）マーケティング」の一分野である「国際市場細分化分析」を行ってきている[23]。

そこでの問題点は，こちらが分析に必要とする変数データの利用が世界各国でまちまちだということである。つまり，ある変数について計上している国としていない国が多々あるということである。変数の定義も一致して

いない。そして，そうした変数データを取る必要のないところもある。

いくら，データが量的・質的に大量（ビッグ）で幅広く細かくなっても，そもそも必要な変数値がないのであれば，分析評価に意味をなさない事態が起こる。

それとインターネット・リサーチが利用できるのは，世界の人々のうちで精々 20％と言われている。後の 80％は関係ないということもある。国際市場細分化分析では悩みの種はつきない。

## 7．これまで，マーケティングでは，どのような予測法を採用してきたか

### 論理実証主義と解釈学

マーケティングでは，これまでどのような方法論が当てはめられてきたかについては，川又啓子（2009）が参照される[24]。

そこでは，規範的，論理実証主義的，論理経験主義，相対主義的などのアプローチをはじめ，解釈主義，反証主義，批判的合理主義などが紹介され，検討されている。

ここで，川又の分類で中心的に取り上げられる「実証主義（論理実証主義，論理経験主義，反証主義）」と「相対主義（パラダイム論や解釈主義（学））」のうち，後者に対しては，「得られた知識の真実が吟味されるような明確な方法基準が存在しないという難点がある」と述べ（このことは「解釈学」側に立つ石井淳蔵（1993）も指摘していることではある），「経営経済学に対する解釈学の本当の意義は，発見的科学領域にある」としている[25]。

論理実証主義や論理経験主義の活用も本来その発見にあるのである。解釈主義の専売特許ではないのである。むしろ後世のマーケティング学者が，マーケティングに適用するに際して厳格に真のモデルの構築を目指したことがあったかもしれない。しかしもし，そこで 100％のモデル（真のモデルともいう）を考えていたとすると間違いになる。応用する側もそう考えてはいけなかったのである。

社会学の方でもこの論理実証主義に対しては，さまざまな立場から批判がなされたという[26]。

> 通常実証主義的とみなされる物理学や化学のような分野ですら，じつは論理実証主義には従っていおらず，真理であることが確かめられていないさまざまな仮定にもとづいて科学的知識が作り上げられていることが明らかになっていった。それどころか，アナロジーやメタファーのようなレトリック科学においては重要な役割を果たすこともある。だから科学は客観的ではなく，イデオロギーや利害関係の産物に過ぎない，と論じるような社会学者まで現れた。すべての知識は相対的で，確かな知識などどこにもない。だから，社会学者は「真理」など追い求めず，「おもしろい」話をすればいいのだ，といった議論まで現れた。

幸い，右で述べたような極端な相対主義に飛びついた社会学者はそれほど多くなかった。科学的な知識がさまざまな仮定にもとづいて作り上げられており，その後の発展によって古い理論や知識が否定されるようなことは，よくあることである。だからといって，真理や客観性やデータによる検証といった原理をすべて捨ててしまう必要はないのである。話は脱線したが，ここでまず確認したいのは，実証主義＝科学ではないということである。多くの（おそらくはすべての）科学の分野はデータのみにもとづいているのではなく，さまざまな仮定や理論的な枠組みにもとづいて作られているので，語の厳密な意味では実証主義的ではないことになる。現代では科学が（そして社会学が）検証されていない何らかの仮定を持ち込むことを全面否定するような議論はきいたことがない。

### 科学も統計，統計も実証主義

確かに科学の多くの分野で統計は用いられ

ているだろう。しかし，すべてではない。実験の再現性が非常に高く測定誤差が無視しうるほど小さければ統計を用いる必要はない。また非常に少数の事例を扱うような分野もあるので，そのような分野では統計が用いられることはなかろう。このように，科学のあらゆる分野が統計学を必要とするわけでない。

事情は社会学でも同じである。

……

さらに，「統計＝実証主義」でもない，実証主義は経験的なデータにのみもとづくことを要求しているだけであって，データを統計的に分析することを常に要求しているわけではない。また，統計学を応用する際には，いくつかの仮定や仮説が必要になる。例えば回帰分析をする際には残差の期待値はゼロで，説明変数とは独立に分布するという仮定が置かれるし，帰無仮説や対立仮説を検討するのは統計学の標準的な手続きである。ここでも極端な実証主義は，統計学とは相容れないのである。

実証主義という言葉をもっとマイルドな意味で使えば（例えばデータのみにもづくのではなく，データを重視するといった程度の意味で使えば），科学は実証主義的であるとか，統計の利用は実証主義的であると言えるかもしれないが，かつては社会学を席巻した実証主義批判の中で批判の的となったのは極端な意味での実証主義であり，端的には論理実証主義であった。

マーケティング現象への帰納法の適用もそういう意味と捉えた方がよいであろう。そこから得られた理論（これも一つの解釈に過ぎないかも知れない）で，現状を理解すると同時に将来を予測する一助にしたいだけなのである。

たとえば，データ・マイニング法活用の典型的な例として出てくるものに以下のようなものがある。

小売りの膨大なデータを調べると，米国における若い夫（男性）は，赤ちゃんのオムツを買いに来たとき，ついでに缶ビールも買っていることが多く観察された。そこで，店では子ども用品売り場のキャッシャーの傍に缶ビールの棚を設置して成功を収めた，という。普段気が付かないことが，膨大なデータ（ビッグデータ）の背後にはあるかもしれないのである。失敗は成功の元という「ことわざ」もあるとおり，失敗は事業につきものということであり，そこからまた新しい事業を展望するものであろう。それの繰り返ししかないということである。

川又は，「むすびにかえて」において，「マーケティングの研究や理論構築の目的は，“科学になる”ことではなく，知的価値の創造であると考える。その目的を達成するのに，もっとも有効な方法論を採用すればよいと考えており，多様な方法論が共存することを認めないものではない」ともいう。

論理実証主義や論理経験主義の活用も本来その発見にあるのである。解釈主義の専売特許ではないのである。ここでは，川又のいう知的価値とは何か，“何のため”の知的価値の創造か，が問われる必要がある。

マーケティングでは「どういう事業をすればよいのか」が基本的に問題であることから，そのための知的価値の掘り起こしであり，そうして収集された情報を活用する“予測の方法”の開発が求められている。

当初は，素朴な手法が用いられたと考えられる。この場合，後世のマーケティング学者が，マーケティングに適用するに際して厳格に真のモデルの構築を目指したことがあったかもしれない。しかしもし，そこで 100％のモデル（真のモデルともいう）を考えていたとすると間違いになる。応用する側もそう考えてはいけなかったのである。

マーケティング現象への帰納法の適用はそういう意味と捉えねばならない。そこから得られた理論（これも一つの解釈に過ぎないかも知れない）で，現状を理解すると同時に将来を予測する一助にしたいだけなのである。

失敗は成功の元という「ことわざ」もあるとおり，失敗は事業につきものということであり，そこからまた新しい事業を展望するものであろう。それの繰り返ししかないということである。

統計学者の田邊國士（2007）[27] は，100％確かなものが考えられないときの予測には，統計科学の優位性を唱えている。佐藤忠彦・樋口知之等も，マーケティング事象に，演繹と帰納の折衷法のあてはめを研究しているというも，そこにあると考えられるのである[28]。

## 8．マーケティングにおける方法論は如何にあるべきか

### (1) オルダーソンの反証主義[29]：

オルダーソンは，「体系より出てくる「命題」（150 本）は実証化によって確かめられ，理論として活用化が図られる」という（彼の命題は，ポパーの「反証可能性」に基づく命題であって〈したがって，帰納主義は想定していない〉と言っている）。

しかしながら，反証主義は，佐和隆光（1993）も言うように，単に科学であるかどうかの問題であって，それ以上何も示唆していないものなのである[30]。

> 要するに反証主義は，科学がしたがうべき「規範」の一つなのである。反証主義は「科学」たるものがあくまでも則るべき方法的立場であって，自然科学の歴史的展開においてそうした方法が有効に機能してこなかったからといって，べつだん反証主義に懐疑的になったり，いわんや「規範」としての反証主義をかなぐり捨てるべきだということにはならない。

その上，反証主義は予測の方法についても何も言っていないのである。筆者としては，反証主義は，佐和も指摘しているように，単に，科学性があるかないかの問題であり，論理実証主義に勝利した科学哲学ではない，という立場である。

経営学者の沼上は，「経営学では，理論はなくても理屈はある」と述べた。そのことから，予測するに際して，「カバー理論」を採用している[31]。

一方で，こうした理屈をつくり上げる当座の論拠としては，既存の経済学，社会学，心理学といったコアとなる学問分野の学問の成果を活用するしかないというのが現状である。これでは，いみじくも森嶋通夫が明らかにしたように，単独の学問とするべく条件を満たさない[32]。

筆者としては，マーケティングを学問に高めたいと考えている。

単独の学問に仕上げるためには，独自の概念を形成し，それと体系化や方法論の問題を一体的にクリヤーしなければならない。これらの問題について，筆者は，これまでも各々についていささか検討を加えてきている。

人が自己のビジネスを探索する（すなわち，マーケティング）ために，まずやることは「予測」である。ある時代，ある地域において自分は何をして，何を生活の糧として生きて行くのかを考えるに当たって，関連ある情報を収集し，整理し，組合せたりしながら意思決定を下していくことになろう。

こうした一連の考え方をどのように現わしていくかは方法論の問題となる。筆者は，帰納法で行くしかないと考える。論理実証主義ないし論理経験主義の立場である。

**黒田の結論：**論理実証主義（帰納と演繹の折衷説）：

実証化で確かめようとしているのは，その時点の関係を説明し，近い将来の状況を予測し，行動の意思決定の方向を検討する情報の収集の一環で，予測力を増そうとするものである。決して，真理を追究するものではない。

### (2) 統計科学の採用

「統計科学」については，北川源四郎の解説が参考となる[33]。

**【情報化に伴う研究対象の拡大と科学的方法論の変化】**

19 世紀までの科学研究は，基本的には機械論的世界観に基づき，デカルト・ニュートンパラダイムのもとで発展してきた。演繹的な理論科学においては，科学の言語としての数学が重要な役割を果たした。ところが，19 世紀中盤のダーウィンの進化論は，実世界そのものが進化し変化する現象であることを意味していた。このような変化に触発されて 1891 年 K. ピアソンは実世界のあらゆるものが科学研究の対象たりうることを宣言して，科学の文法を提唱した記述統計学およびその後の数理統計学はこの科学の文法を実現する方法として，発展してきたものと考えることができる。

この実験科学の方法の確立によって，生物学だけでなく，経済学，心理学など，確率的現象を伴う実世界，すなわち生命や社会が科学的研究の対象となることとなった。

20 世紀後半になると，コンピュータの発展によって計算能力が飛躍的に発達した。これによって，従来の理論科学の解析的方法では限界があった非線形現象，複雑系，多自由度システムの解明に数値計算やモンテカルロ計算が適用され，計算科学が急激に発展した。

しかしながら，21 世紀の現在，IT の飛躍的発展はもうひとつの発展の契機ともなりつつある。爆発的に膨張しつつあるサイバー世界の利用技術である。既に述べたように，IT 技術の発展は，あらゆる分野で，大量かつ大規模なデータを生み出し，巨大なサイバー世界を構築しつつある。

今後の科学・技術の発展は大規模データの活用抜きには語れない。大量データの活用技術に支えられた科学的方法論をここでは「第 4 の科学」と呼ぶことにする【図 1 】。

統計科学におけるデータ科学あるいはデータサイエンスの発展形ともいえる。言うまでもなく，第 1 と第 2 は，理論科学と実験科学である。これらは演繹的（原理主導型）および帰納的（データ主導型）方法とも呼ばれ，車の両輪のように 20 世紀の科学研究を推進する原動力となった。

しかし，20 世紀後半，解析的方法に基づく理論科学の限界を補う方法として計算科学が確立し，非線形現象や複雑系あるいはシステムの予測や現象解析において，輝かしい成功を収めた。

従来の理論科学と実験科学の方法論が研究者の知識と経験に依拠するものとすると，計算科学と第 4 の科学は Cyber 世界（IT）が可能とする新しい演繹的および帰納的方法論である。計算科学が発達し確立した現在，この第 4 の科学を戦略的に推進することが，情報時代の科学研究をバランスよく進展させるために不可欠である。NSF が 2008 年度の新し

【図 1 】「第 4 の科学」の位置づけ

| | | | |
|---|---|---|---|
| 第 1 の科学の方法 | 演繹的 | （理論科学） | （計算科学） |
| 第 2 の科学の方法 | 帰納的 | （実験科学） | （第 4 の科学） |

い重点課題として“Cyber-enabled Discovery and Innovation”を掲げていることは注目に値する（NSF，2008）。

地球環境シミュレーション等の分野では理論モデルとデータの情報を統合するデータ同化の方法が盛んになっているが，一般的に言えばこれは原理主導型の方法とデータ主導型の方法の統合技術といえる。

今後の知識社会の発展のためには，従来むしろ意識的に別個に考えられてきた二つの方法論の統合も重要な課題である。しかし，“統計科学”の立場から技術的にみれば遂次フィルタリングの方法と見なすことができる。

統計科学の研究者にとっては極めて自然なことが，今後の科学技術のキーテクノロジーともなりうるのである。

記述統計学に代表される推測統計学以前の統計学は，対象の観測に基づくものであったが，推測統計学では少数ながら周到に設計した実験データにもとづいて科学的推論を実現しようとしたものであった。

しかしながら，近年の情報化によって再び，必ずしも厳密に設計されていない大量・大規模データからの知識獲得が重要になっている。大量・大規模情報の適切な利用技術の開発なくしては，データ量の著しい違いはあるが一種の記述統計学に先祖返りした感もある。

これに関連して，堀田情報・システム研究機構長は生物学の変遷についで，興味あることを述べている。堀田氏によれば，19 世紀までは博物学であった生物学は，20 世紀初頭に科学的研究の手法を取り入れ実験科学となったが，今やヒトゲノム解読のようにすべての情報を読み取れるようになっている。これはある意味で現代版の博物学に戻りつつあるともいえる。

統計学者の細谷雄三も同様の見解をあらわしている[34]。

今日経験分析の範型とみなされているのは，経験データの生成過程を確率モデル類で表現して，そのモデル類の上で推定・検定などの統計的推測を行うことである。しかし社会現象では，データがよく制御された対照ランダム実験から生成されたものではないような事例が，むしろ一般的である。社会現象に適合する確率モデルについて意見の一致が得られにくい状況に対して，いかに経験分析を適用して，現象の全体像を記述したらよいであろうか？

……

科学用語とは対照的に，日常言語は基本的に非専門家間のコミュニケーションのために自然発生的に構成される人々は，各人さまざまなグループに同時に属して，それぞれの一員として複雑な関わりのなかで相互に（必ずしも）整合しない多様な意思決定をしながら生活している。日常言語は社会の多様性に直面して専門化せずに話ができるように作られている。厳密性は犠牲にしているが，適用性や頑健性がある。またその多義性ゆえに，現実の使用に耐える言語になっている。日常言語を用いて統計的証拠をいかに非専門家に説得すればよいかという視点から，高度化・狭隘化し過ぎてしまった統計的諸方法を日常性と接点をもつ言語に戻すことも必要であると思われる。

ザイゼルと分野は異なるが経験的・行動的ファイナンス分野の研究者にロバート・J. シラーがいる。彼は著書で，株価や不動産価格などの資産価格について，人々がなぜ不合理な期待を形成するのかをかれ独自のサーベイ調査にもとづいて解明し，その中でかれは最近の不動産バブル崩壊を予想していたシラーの経験分析はザイセルと驚くほど類似しており，高度な数学や精密な計量経済学的方法は使わない。独自のサーベイ調査にもとづく単純な数値結果を大変有効に使う。使われる統計数字が生きていて，説得力がある。かれら

に共通する接近法は「緩やかな」経験分析と特徴づけることができよう。

かれらの関心は，適用する方法の整合性や統一性にはなく，分析対象の全体像にある。

こうしたことは，筆者としては，文化勲章受章者の篠原三代平教授が，かつてエコノメトリックス（計量経済学）をもじった言葉で，経済変数間の個別の関係を積み上げる考え方で理論を構成する方式を採用していたことから，「メノコメトリックス」学者と呼ばれていたこともあるが，それに通ずるものがあると考えている。

マーケティング・リサーチの考え方や方法もそうした考え方の一環なのである。

## 9．AI では，どのような予測方法を考えているか

### (1) 人工知能（AI）の方法論

近年は，人工知能（AI）における「ディープラーニング（深層学習法）」がクローズアップしている。

将棋や囲碁では，ディープラーニング（*Deep Learning*：「深層学習法」）を駆使して人工知能がプロの強者を次々に撃破していることが報道されている。これが評判になってか否か，やがては（2045 年），人工知能が人間を上回る「シンギュラリティ（*singularity*：技術的特異点）」がやってくるという人もいる[35]。

松尾は，人工知能と人間の関係についても書いている。

それによると，「深層学習法」とは，分けることであり，主成分分析（数量化理論第Ⅲ類）の繰り返しを行って，結果の頑健性を求めるものである，という。

筆者も，一種の「深層学習法」を実行しているとは気づかずに，ある実際のマーケティング現象を理解すべく，判別分析や，分類法の成分分析・主成分分析（数量化理論第Ⅰ類・第Ⅱ類）を用いてきた。国内市場細分化や国際市場細分化を求める際である。

そこでは，未だ満足のいく結果を得られていないと考えている。なぜならば，使用するデータの信頼性や比較のためのデータの網羅性（欠落性）にかかわる問題が常に付きまとっているからである。

この点は，筆者も最近の論文で明らかにしてきている。

ビッグデータが生み出される時代なのだから，これを分析することが重要といっても，その個々のデータの妥当性や信頼性が問われれことが多々存在することに留意する必要がある。

いかに高度な技術を駆使しても，たった一つのデータの不確実性や欠落が分析結果にゆがみが出ることは当然のこととなる（これでは，いつまで立っても「頑健性 robustness」は生まれない）。

この点は，たとえば，数学者オスカー・モルゲンシュタイン（Oskar Morgenstern）（フォン・ノイマンとの共著“*Theory of Games and Economic Behavior*”（1944 年）で知られる）が，著書“*On the accuracy of economic observations*”（1950 年）で明らかにしていることと軌を一にしている。2 元連立方程式において，係数のほんのわずかな違いが，解答における符号が正負逆転することなどを説明した。

### (2) A. ゴブニックの AI における方法についての見解

—ボトムアップ方式とトップダウン方式—

A. ゴブニック（2018）によると，問題解決に，ボトムアップ方式とトップダウン方式がある，という[36]。

特に，AI を復活させた 2 つのアプローチ例として，「ディープラーニング（深層学習）」と「ベイズ法」を挙げる。

たとえば，一つの典型的な意思決定問題を考えてみる。

〈バス会社が，ある新興住宅地に新しい路線を開拓するかしないかを判断する〉

これは，バス路線を開拓「する」か「しない」か，の二者択一の意思決定問題である。これをどういう方法で解決を図るか，について大きく二つ方式が考えられる。

(a)「深層学習法」で考える

これまでのすべての路線のデータを収集する。それをいくつかの特徴に基づき分類する。その上で，新しい路線がどの分類に相当するかを評価する。もし，分類 $i$ に相当するとなれば，採算性からみて，十分路線開拓の意味がある，との解釈となるであろうし，そうでなければ路線開拓を止めるという決断を下すことになる。

(b) ベイズ推定法で考える

人間社会の場合，何度も同じことを検証できない。せいぜい１回か，２回である。そんなとき「ベイズ推定法」は効果を発揮すると考えられている。

一つの予想を立てる。それに対して，一回リサーチを行う。その結果を利用して，初めの予想を修正するという考えである。

上記の問題の解決を図るために，

(a) では，筆者も分析してきている[37]。

(b) では，上記の問題の解決を図るには，どう考えていくのか。

バス路線開拓に当たっては，採算性が第一である。これについても様々な状況が考えられる。外出の少ない高齢者が多いとか学童の少ないとかといった年齢構成に偏りがあったり，近くに買い物に便利なスーパーが出来ていたりで，バスを利用する必要の少ない人々が多い，とか，最初のうちは物珍しさで多くの人が乗ってくれるだろうが，そのうちに，徒歩や自転車に切り替えるかもしれない，とかあるだろうし，問題を解決するべく要素は複雑に絡み合うと考えられる。

しかしここでは簡単に考えて，まず，路線バスが走ったときの沿線住民の乗車率のみを検討することから解決を図ると考える（採算性を考慮した乗車率であり，つまり，これだけ乗ってくれなければ採算は取れないというレベルのもの）。

そこで，路線住民に対して，開通した場合，「乗る」か「乗らない」かのリサーチを１回行う。その結果を利用して，初めに設定した乗車率を変更するというやり方である。

こうした問題設定では，ベイズ確率を利用して解決を図る典型的な例となる。

### (3) 分け方の考え方と方法

松尾豊（2016）は，「分け方」の方法についても書いている[38]。

> 新聞記事をカテゴリに分けることを考える。まずはコンピュータに訓練用のデータを読み込ませて，記事に出てくる単語をもとに，何らかの空間をつくる。たとえば，記事に出てくる単語から最も頻出するものを 100 個選んで，それで 100 次元の空間をつくると，１つの記事は，この空間上の１つの点として表すことができる。この空間では，同じ単語が出てくる記事は近くに，出てこない記事は遠くになるようにマッピングされる。
>
> 新聞記事には，「政治」「科学」「文化」というカテゴリがつけられているとしよう。
>
> ひと通りマッピングが終わったら，次に，新しいテストデータを読み込ませて，どのカテゴリに分類されるかを見る。下の図の真ん中の■がテストデータだったとして，これが３つのカテゴリのうちのどこに分けられるか。図のように線引きされていれば，テストデータは「政治」に分類されるだろう。この線を

どのように引くかによって，分け方が変わる。つまり，「分ける」ということは，分けるための「線を引く」ことと同じなのである。（筆者注：線形判別分析法）

最終的に，「国」「政府」「予算」「行政」「与党」などの単語が出てきたら「政治」，「宇宙」「物理」「生命」「細胞」「コンピュータ」などの単語が出てきたら「科学」，「音楽」「美術」「絵」「彫刻」「アニメ」などの単語が出てきたら「文化」といった具合に，コンピュータが学習してくれたら OK だ。線の引き方にはいろいろな方法があり，それぞれ異なる仮説に基づいている。ここでは代表的な分類のしかたを 5 つ紹介する。

①最近傍法，②ナイーブベイズ法，③決定木，④サポートベクターマシン，⑤ニューラルネットワーク

## お わ り に

### (1) ビッグデータと官庁統計の問題 ── 厚生労働省の毎月勤労統計調査における不適切な方法による調査について

新聞報道によると，2019 年 1 月，厚生労働省が都道府県を通じて労働者の給与や労働時間の変化を調べている「毎月勤労統計」とよばれる調査が，長年にわたって決められたやり方と違う方法で行われ，その結果，データが誤っていたことが発覚した。

毎月勤労統計は，失業中の人などに給付される雇用保険などの給付金の算定にも使われていた。データが誤っていたことによりのべ約 2 千万人の給付金が少なくなってしまい，少なくなった総額は数百億円に上るという。

今回の毎勤統計の不適切な官庁による統計問題に関して，筆者も思い当たることがある。筆者は，大学で近代経済学や数理統計学をかじって来ていたので（学部卒業論文名は，「自己相関の一考察」），大学院に入って，経営統計学の分野を研究したいと，伊大知良太郎先生にその旨申し出たところ，統計の理論はよく知らないので，これをまず初めに読んで，1 か月後に，それについて発表しなさい，と言われ『家計調査年報』を与えられた。このときの筆者の感想は，こんなことをやっていたら，同期に入った連中から研究で 1 歩も 2 歩も遅れてしまうであった。

しかし，これを読んで驚いた。全国には 5 千万世帯があって，そのうち 8 千世帯しか調べていないものであった。それも標本理論というものに則って選ばれた世帯の，所得，消費支出，貯蓄など多様な家計項目をアンケート調査項目について，調査員が一軒一軒回って面接調査方式をとっていた。しかも，8 千世帯の半数は，半年ごとに入れ替えるという抽出法であった。

この統計は，世界に冠たる統計であるということも，その時知った。日本の家計データは，『家計調査年報』を参照して分析されていることの意味をはじめて教わった気がしている。その後の筆者の分析は，『家計調査年報』の数値を大いに活用するとともに，そこから得た知識を活用していることは言うまでもない。

今だに伊大知先生のご指導には感謝してもしきれないご恩を感じている。

とにかく，官庁統計というものは，このように作成されていると考えていたが，今回の厚労省の不正が発覚したことで，ところによっては，そうでもないことが分かった。

日本統計学会からも，いち早く「声明」が発せられている[39]。

---

平成 31 年 1 月 28 日

厚生労働省毎月勤労統計調査における
不適切な方法による調査に関する声明

日本統計学会

今回，厚生労働省毎月勤労統計調査において不適切な調査が行われていたことが発覚した。信頼性の高い公的統計の提供のために政府において原因が究明され，その結果が公表されるとともに，再発防止策が講じられることを強く希望する。原因究明と再発防止策の検討の過程において，求めがあれば，専門的な見地から本学会が協力することを表明する。

---

これらのことがらは，データの量に関係なく生ずる問題であるということである。統計学を適用することはそう簡単なことではないことが理解できる。

一方，官庁統計に，こうした事態が起こったことから，竹内は，「データ学のすすめ」を提案している[40]。

最近「統計不正問題」がやかましく論じられているが，私はこの際［データ学］というものを提唱したい。このようにいうと「データ・サイエンス」（データの科学）がもてはやされている現在，改めて何をいうのかと怪しまれるかもしれない。

しかし「データ・サイエンス」或いは「データ・アナリシス」はほとんどの場合，与えられたデータから出発して，それをどのように扱うかの方法を論ずるものとなっており，データそのもの，そのあり方，品質等を論ずることはあまりない。しかし具体的な統計分析を料理することにたとえるならば，よい料理を作るには，優れた調理法とともによい食材をえらぶことが大切であるように有効な統計分析を行うには良いデータを選んで適切な解析を行わなければならない。利用できるデータが与えられている場合には，その性質をよく吟味して，それに適した解析法をえらばなければならない。その場合は，有り合わせの材料から食事を作るようなものである。いずれにしても良い料理人は食材をよく吟味して調理法を選ぶように優れた実際統計家は統計素材をよく調べて解析法を決めるものである。しかしデータを吟味する過程はしばしば「経験と勘」によるものとされ，定式化されて説明されることはほとんどない。しかしそれを何らかの形で一般化し定式化して示すことはできないだろうか。それがデータ学である。

### (2) 文系学生に数学を教えるということ

マーケティングを科学として認知したいというのには，理由がある。

「これからどんな事業をしたらよいか，そこではどのような製品を作ればよいか，それをどのようにして売って行けばよいか」について，より科学的に「予測をしたい」からである。

本拙論の冒頭の現今の社会が要請する大学卒業生が，文系・理系の枠を取り払った教育が施させれている必要があるということは分かる。しかし，ことは簡単ではない。

結局，文系・理系を問わず，大学生には「数学」と「リベラルアーツ（教養）」の教育が求められるということであろう。

しかしながら，マーケティング教育にふさわしい「数学」とはどういうものなのか，が問われる。

まず言えることは，マーケティング分野でいえば，そもそも科目の性質上，数学・統計学が苦手な教師が多いのである。実際のビジネス・マーケティングに精通していて，数学・統計学の方は若干かじった程度の人が担当している場合が多いのである。

一方，経営学の方に数学を専門とする人が採用された場合には，ややこしい数式展開を念頭に置かれるので，実務にどう応用するかがおろそかになる場合がでてくる。結果として，実務と専門がなかなか結び付かないと

いった現実問題がでてきている。

現状では，マーケティング担当者の大半は，必要に応じた数学・統計学活用で済ませている（筆者もこれと同列にあった，詳しくは専門書を見なさいというアドバイス付きで）。

また一方，リベラルアーツ（教養）の教育は，本来，大学で基本的に教えられるべきものである。ただ，リベラルアーツ（教養）とは何か，をしっかり共有しなければならないだろうが。

いずれにしても，上記のことが満たされるためには，マーケティング教育にとって，どういう人が望ましいのかといった教師の資格から検討することにならざるを得ないと筆者は考えている。

### (3) 筆者のマーケティング研究の始まりは数理統計理論の当てはめからであった[41]

筆者は，マーケティング現象に対する数理統計理論の当てはめから研究を開始した。若かりしころ，数理統計学を少し勉強していたこともあって，大学における講義科目「マーケティング」を担当することになって，マーケティング現象に対する統計学理論のあてはめを研究対象としていた。「マーケティングにおける購買者行動論」という形で何編か発表している。

とりわけ，「購買者行動の銘柄推移問題」に確率過程論のマルコフ過程（マルコフ連鎖）のあてはめの研究について発表したときは，フロアの一人から「君のマーケティングは統計学だね」と言われている。

つまり，マーケティング関連の購買者行動に関する数学モデルとして，

（統計的確率的モデル・数学的解析法）

マルコフ連鎖モデル

エーレンバーグ（A.S.C. Ehrenberg）の反復購買モデル

線形学習モデル

（シミュレーション・モデル・数値解析法）

モンテカルロ法に基づくモデル

組み替え法

ブートストラップ（bootstrap）法

などを取り扱っていたことに由来すると思われる。

しかし，これらの研究過程で，エーレンバーグ（A. S. C. Ehrenberg）の著書“*Repeat Buying*”（反復購買）を読んで，家計簿にあらわされた膨大なデータから購買者の反復購買についての精密な理論［負の 2 項分布モデル（*Negative binomial Distribution*（NBD））］を見出していくという，そうした地道に研究を進める彼の研究姿勢に，筆者は非常な感銘を受けた記憶がある。

その後，マーケティングの方法論について研究，「マーケティング学では如何なる分析方法を採用すべきか」の研究を進めている。

## 注と参考文献：


（1）「文系学生も数学を ── 経団連，大学に改革提言へ」『日本経済新聞』，2018 年 12 月 1 日，土曜版（1 面）。

（2）ニュートン編集部（2006）「天気予報：天気予報はどれほど当たる？言い伝えは本当？」『ニュートン』，教育社，2006 年 6 月号，pp.108-109。

（3）バルザック（1834）『「絶対」の探求』，（水野亮訳），2009 年発行，岩波文庫。

（4）岩井克人（1992）『ベニスの商人の資本論』，筑摩書房。

（5）井原西鶴（1686）「波風静かに神通丸」『日本永代蔵』，巻一（三）（堀切　実訳（2009）『新版・日本永代蔵―現代語訳付き―』，角川文庫，pp.18-23。）

（6）黒田重雄（2009）「商学とマーケティングの講義ノート（1）」『経営論集』（北海学園大学），第 6 巻第 4 号（2009 年 3 月），pp.163-184。

（7）「マーケティング・リサーチ」は，「マーケティング」そのものである。

（8）大塚久雄（1967）『社会科学の方法』，岩波新書，pp.60-64。

（9）富永健一（1999）『社会学講義―人と社会の学

—』，中公新書，p.4。
（10）Green, P.E, and R.E. Frank（1967）, *A Manager's Guide to Marketing Research; Survey of Recent Developments,* John Wiley & Sons, Inc.（グリーン＝フランク著（土岐　坤訳）（1969）『マーケティング・リサーチはどこまで進んだか』，ダイヤモンド社。）
（11）黒田重雄（2009）「商学とマーケティングの講義ノート（2）」『北海学園大学経営学部経営論集』第7巻第1号（2009年6月），pp.123-142。
（12）Bass, F.M.（1993）, "The Future of Research in Marketing: Marketing Science," *Journal of Marketing Research,* Vol.30, February, pp.1-6.
（黒田重雄他著（2001）『現代マーケティングの基礎』，千倉書房，pp.204-205。）
（13）質問表形式の実査は，1824年（今からおよそ180年前）に新聞において用いられていた。マーケティング・リサーチとしての始まりは，1910年頃であり，C.C. パーリン（Charles C. Parlin）の農機具製造業者の経営活動調査研究が嚆矢であるという説が一般的である。
（14）「（情勢判断・意思決定の数学）統計の威力」『Newtonニュートン』，2013年12月号，株式会社ニュートンプレス，pp.22-53。
（15）山本　強（2014）「今そこにあるビッグデータ── デジタルデータで見る自然現象と社会現象──」，日本マーケティング協会北海道支部4月期セミナー，2014年4月22日，於：電通恒産札幌ビル。
（16）ダイヤモンド社編集部（2012）「ビッグデータで営業の精度を高める『DIAMONDハーバード・ビジネス・レビュー』，第37巻第12号，2012年12月，pp.92-104。
（17）西内　啓（2013）『統計学が最強の学問である』，ダイヤモンド社。
（18）水田正弘（2014）「ビッグデータブームを考える」『現代思想』，2014年6月号，青土社，pp.69-79。
（19）太郎丸　博（2014）「統計・実証主義・社会学的想像力」『現代思想』，2014年6月号，青土社，pp.110-121。
（20）竹内　啓（2014）「ビッグデータと統計学」『現代思想』，2014年6月号，青土社，pp.28-37。
（21）牧田幸裕（2018）「企業の未来像大胆な構想必要」（やさしい経済学・デジタル時代のマーケティング戦略（7））『日本経済新聞』，2018年7月16日付（朝刊）。
（22）筆者による人工知能（AI）関連論文
（論文）「人工知能（AI）の技術的発展が経営やマーケティングへどう影響を及ぼすかについての覚書」『開発論集』（北海学園大学開発研究所紀要），第100号（2017年9月），pp.161-195。
（論文）「ビッグデータ，人工知能（AI），そしてマーケティング学—人工知能の技術的発達とマーケティングへの影響に関する一考察—」『北海学園大学経営学部・経営論集』，第15巻第4号（2018年3月），pp.147-170。
（論文）「これまでのマーケティング方法論と近年のAIにおける二つのアプローチ法との関係についての一考察—特に，ディープラーニングとベイズ推定法の考え方を中心に—」『北海学園大学経営学部・経営論集』，第16巻第3号（2018年12月），pp.43-55。
（23）筆者による国際市場細分化分析関連論文等
（論文）「比較マーケティングと国際市場細分化」『経済学研究（北海道大学）』，第45巻・第2号（1995年6月），pp.94-108。
（著書）『比較マーケティング』，1996年10月1日発行，千倉書房。
（論文）「比較マーケティングの研究方向に関する一考察」『経済学研究』（北海道大学），第47巻・第2号（1997年9月），pp.84-90。
（論文）「比較マーケティングの研究方向に関する一考察」『流通研究』（日本商業学会誌），第1巻・第1号（創刊号）（1998年3月），pp.19-32。
（論文）「比較マーケティング研究とグローバル・マーケティング」『経営論集』（北海学園大学），第1巻・第1号（創刊号）（2003年6月），pp.69-89。
（論文）「地域の国際マーケティングに関する一考察—北海道における貿易活性化の必要性をめぐって—」『経営論集』（北海学園大学），第2巻・第3号（通巻第7号）（2004年12月），pp.55-73。
（論文）「国際市場細分化を中心とする実証化—1991年時点と2000年時点の比較分析—」『経営論集』（北海学園大学），第2巻・第4号（2005年3月），pp.141-159。
（研究報告書）「比較マーケティング研究における一展開—国際市場細分化を中心とする実証化—」，平成15～16年度科学研究費補助金（基盤研究（C）（2））研究成果報告書（2005年3月），平成17年3月。
（論文）「比較マーケティングにおける国際市場細分化分析のビジュアル化」『経営論集』（北海学園大学），第3巻第1号（2005年6月），pp.1-38。
（24）川又啓子（2009）「方法論争の展開」（嶋口充輝監修『マーケティング科学の方法論』，白桃書房，

第 1 章所収，pp.3-30）。
（25）石井淳蔵（1993）『マーケティングの神話』日本経済新聞社。
（26）社会学からの批判
（27）田邊國士（2007）「ポスト近代科学としての統計科学」『数学セミナー』，第 46 巻 11 号（通巻 554 号），2007 年 11 月号，pp.44-49。
（28）（a）佐藤忠彦・樋口知之（2008）「動的個人モデルに消費者来店行動の解析」『日本統計学会誌』，第 38 巻，シリーズ J 第 1 号，pp.1-19。
（b）佐藤忠彦・樋口知之（2008）「返答：マーケティングもデータ同化へ」『日本統計学会誌』，第 38 巻，シリーズ J 第 1 号，pp.31-38。
（29）Wroe Alderson（1965），*Dynamic Marketing Behavior*, Richard D. Irwin, Inc.（オルダーソン著（田村正紀・堀田一善・小島健司・池尾恭一）（1981）『動態的マーケティング行動―マーケティングの機能主義理論―』，千倉書房）。
（30）佐和隆光（1993）「夢と禁欲」（浅田彰・黒田末寿・佐和隆光・長野敬・山口昌哉『科学的方法とは何か』，中公新書（1986 年初版））。
（31）沼上　幹（2000）『行為の経済学』，白桃書房。
（32）森嶋通夫（2010）「付記・社会科学の暗黒分野」『なぜ日本は没落するのか』，岩波現代文庫，pp. 197-206。
（33）北川源四郎（2008）「知識社会における統計科学の役割」『日本統計学会誌』，第 37 巻・第 2 号，pp.201-211。
（34）細谷雄三（2010）「巻頭随筆・社会諸現象の緩やかな経験分析」『日本統計学会会報』，No.143，pp.1-3。
（35）シンギュラリティ：2005 年にアメリカの未来学者レイ・カーツワイルが出した説。
（36）A. ゴブニック（2018）「（特集・AI の新潮流）子供の脳に学ぶ AI」『日経サイエンス』，2018 年 2 月号，pp.28-41。
（原文：Gopnik, Alison.（2017），“Making AI more Human: Artificial intelligence; Child psychology; Machine learning; Thought & thinking; Cognitive robotics”，*Scientific American*, Jun2017, Vol. 316 Issue 6, pp.60-65.）
（37）筆者による「国際市場細分化分析」
（38）松尾　豊（2016）『人工知能は人間を超えるか―ディープラーニングの先にあるもの―』，株式会社 KADOKAWA。
（39）赤坂昌文（日本統計学会会長）（2019）『日本統計学会会報』，No.179（2019 年 4 月 30 日），pp.3-4。
（40）竹内　啓（2019）「データ学のすすめ」『日本統計学会会報』，No.179（2019 年 4 月 30 日），pp.1-2。
（41）黒田重雄（1980）『消費者行動と商業環境』，北海道大学図書刊行会，pp.266-278。