

報告

数学教育学会後援筑波大学・アジア太平洋経済協力国際会議(3)  
 椿 広計講演「統計的予測と不確実性のマネジメント」

松寄昭雄、小篠拓央\*

概要：本稿は、筑波大学 APEC 国際会議でなされた椿広計による講演をまとめたものである。講演では、未来予測研究の一貫としてビジネス・サイエンスにみられる事例から、(1) 労働収入と労働力に関する実験上の法則を導出し、(2) 物件購入額に対する景気の影響を前提とした場合における、購入に関する意思決定の樹形図が紹介された。  
 検索語：データ中心科学、科学の文法、科学的意思決定

Abstract: This is the report of the lecture by Hiroe Tsubaki on the title ‘Statistical Prediction and Uncertainty Management’ which explained the grammar of science and the uncertainty management and scientific decision process with decision tree analysis and value of additional information. This lecture was done for the conference theme ‘Mathematics Education for Future Prediction’.

Keywords: data-centric science, grammar of science, scientific decision

1. はじめに

本稿は、「未来を築くための数学教育」をテーマとして実施された表記会議における椿 広計による基調講演録である。同講演では、未来シナリオ作成に至る基幹的方法として、ビジネス分野での統計利用法が概説された。

2 統計数理研究所のミッション

統計数理研究所 (Institute of Statistical Mathematics : ) は、1944 年に設立され、2004 年に情報・システム研究機構 (Research Organization of Information and Systems : ) として統合された 4 研究所の 1 つである。ROIS のデータ中心科学のリサーチコモンズプロジェクトは、これら統合された研究所が、新たな研究分野を開くことや、統合された学際的研究の実践への貢献を目的としている (図 1)。

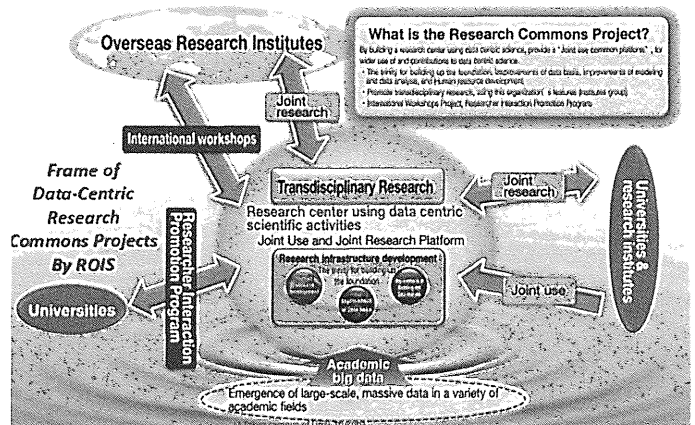
統計数理研究所のミッションは、統計とデータ・サイエンスの理論と

応用に関する研究を遂行することであり、数多くの研究者がデータ中心科学に従事しており、情報・システム研究機構の人文社会科学と地球環境科学システム分野の学際的研究プロジェクトを支援している。

3 科学の文法

統計と社会の間のインターフェースとしての科学の文法について、留意すべき 2 つの典型的な誤解がある。1 つ目は、実際のビジネスが科学の域を超えるという誤解である。なぜ、ビジネス・サイエンスがあるかという、それはビジネスが科学にとっての探究材料であるからで

図 1 ROIS データ中心科学リサーチコモンズプロジェクトの枠組み



\* (Akio Matsuzaki) (Hiro Ozasa)  
 埼玉大学

ある。2 つ目は、統計がある種の応用数学であるという誤解である。もちろん、ある意味でそれは誤解ではなくとも、なぜ、統計的な探究方法があるかという点、統計が科学の文法であり、数学が科学の言葉であるからだ。この意味で、統計の役割は、解釈の際に意味を構成することにある。

Einstein (1940) は、科学的活動とは、我々の感覚経験の混沌とした違いを、論理的に規則正しい思考のシステムに対応づける試みであると定義した。Galton (1883) は、統計的科学的とは、事実のなかに結びつけられた幅広いグループに関する情報を、議論に適した、端的で簡潔な表現に要約する方法であると定義した。『科学の文法：The Grammar of Science』(Pearson, 1892) では、次の通り言及されている：1 つは、すべての科学の総体が、科学の方法において構成されており、探究材料における構成ではないことである。もう 1 つは、科学の分野には際限がなく、その探究材料は無限であり、自然現象のすべてのグループ、社会生活のすべての相、そして過去もしくは現在の発展のすべての段階が、科学にとっての探究材料になっていることである。

『科学の文法』では、人間が自然に対してある法則を与える科学の方法、つまり、予測もしくは解釈の方法として、科学的活動のプロセスを次のようにまとめられた：(1) 観察された現象の原因と結果を分析することによって、ある仮説上の法則を計画し、(2) 実験上の法則を得るために、そのモデルや仮説上の法則を関係する事実へと適正化し、そして、(3) 事実の分類の必要性を明らかにするために、得られた法則の実践を確認する。その後、このプロセスは、PDC (A) サイクルとしてまとめあげられ、統計学者らが予測や解釈に用いる基礎的なマネジメント・サイクルとしてその用途が広まった。

では、科学の法則が社会における予測に用いられる例として、Hooke の法則がビジネスにおいて見られるかという点、残念ながらそうとは言えない。ここで、労働収入と労働力の関係を例として、PDC (A) サイクルを通じたビジネス・サイエンスの構築を確認する (図 2)。

まず、ステップ (1) として、労働力を原因、労働収入を結果として分析し、モデルをつくり、記述手法を考える。ばらつきの記述にはヒストグラムや標準偏差、関連性の記述には散布図や相関がある。例えば、労働収入に対数を用いると、ヒストグラムが正規分布する (図 3)。

また、この事象にバネの伸びで知られる Hooke の法則を仮定すると、比例関係 ( $Stress = Const \times Strain$ ) に対応する関係式として、一次

関数の比例定数が 0 の場合 ( $y = ax + b, a = 0$ ) や指数関数における冪乗が 1 の場合 ( $y = ax^b, b = 1$ ) が考えられる。そこで、軸となる入力変数 (労働力) と出力変数 (労働収入) のそれぞれに対して、対数を用いて変換すると、散布図に比例関係をみることができ (図 4)。

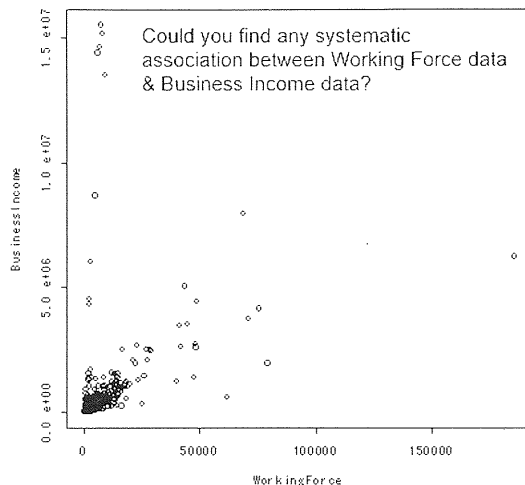


図 2 労働力 (横軸) と労働収入 (縦軸) の散布図

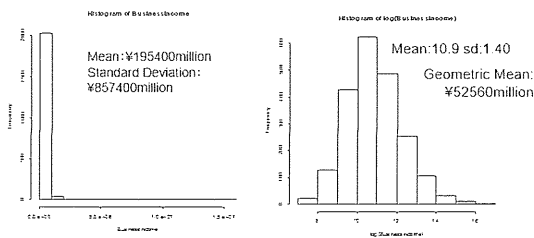


図 3 労働収入のヒストグラム (左) と労働収入の対数のヒストグラム (右)

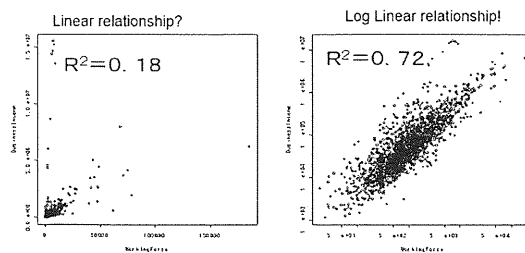


図 4 「労働力」と「労働収入の対数」の散布図 (左) と「労働力の対数」と「労働収入の対数」の散布図 (右)

次に、ステップ (2) として、実験から法則を得る。このステップは、統計におけるテクニカルな部分であり、分析方法にもとづいて予測することはもちろん重要であるが、現象を解釈することも同じく重要である。ここで、回帰分析を行うとき、観測値とモデルの予測値との差を残差とみる。この残差を用いてモデルをフィットさせる (図 5)。

そして、ステップ (3) として、フィットさせたモデルを確認する。確認の目的は、データを

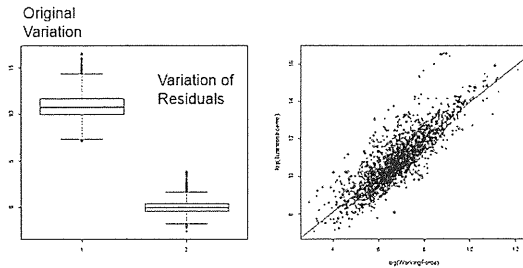
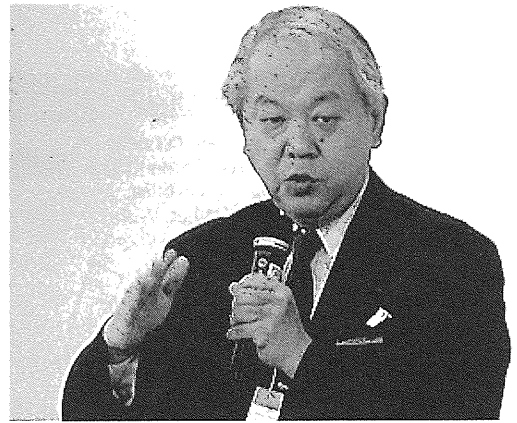


図 5 元の変数を用いた場合と残差を用いた場合のばらつきを示す箱ひげ図 (左) と適正化されたモデル (右)

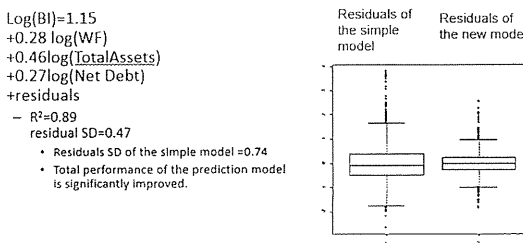


図 6 変数を追加したモデル (左) と追加前後の場合のばらつきを示す箱ひげ図 (右)

分類する必要があるか否かを明らかにすることである。残差分析結果から観測値が予測値を大きく上回る (下回る) データに注目する。例えば、業種について分類するとき、それらの実際のビジネスに応じた変数 (例えば、総資産: Total Assets, 純負債: Net Debt) を追加する (図 6)。

この科学的な活動プロセスを通じて、ビジネスにおける事実から実験上の法則を見出すことができた。PDCA サイクルで言えば、Action の段階の役割は、実験上の法則を確認した後の改善である (A→P)。また、PDC (A) サイクルよりは、PPDAC サイクルの方が教育の分野にとって馴染みがある。PPDAC サイクルには、要求 (もしくは期待) と現実とのギャップを確認する日々のマネジメント・サイクルもある (図 7)。

#### 4 不確実性のマネジメントと科学的意思決定プロセス

『科学的経営学入門: Introduction to Management Science』(Taylor, 2009) では、未来に対する意思決定方法としてのマネジメント・サイエンスが注目されている。そこでは、どのように数学を現実に対して用いるかが注目されている。意思決定の構成要素には、意思もしくは行為 (D) と自然状態 (Y) がある。Y には、未来において起こりうる実際の出来事や、通常起こる出来事の変数と確率論によって定式化しえる変化量が含まれる。ここでは、不確実性を前提としたシナリオにおいて意思決定を行う。このような科学的意思決定はどのように行われるのだろうか。

ここでは、「アパート」「オフィス」「倉庫」それぞれの支払項目を D、「好景気」「不景気」を Y、「利益 (損失)」を D と Y の関数として、確率による意思決定について考える。いくつかの経済予報にもとづいて、単純に、好景気が 6 割 (不景気が 4 割) で起こるとする。そこで、それぞれの景気の場合における支払金額を項目ごとに考えると、不確実を前提とした景気状態に対して、それぞれの支払いに対する利益 (損失) の期待値が計算できる (図 8)。

それぞれの支払項目に対して期待値を求める方法より、樹形図を描くことができる (図 9)。この樹形図は、不確実な未来に対する体系的な選択肢、もしくは、シナリオとも言える。現実のリスクを伴うシナリオは、理論上描かれるものより、さらに複雑になるかもしれない。ここでの確率は、通常、不確実性において予想されたものである。ここで、意思決定における統計的予測の役割として重要なことは、「景気」「不景気」を記述しうるモデルの構築ではなく、「景気」「不景気」にもとづいて、不確実性とともな確率について述べた点である。

(■ : 意思の節, ● : 確率の節)

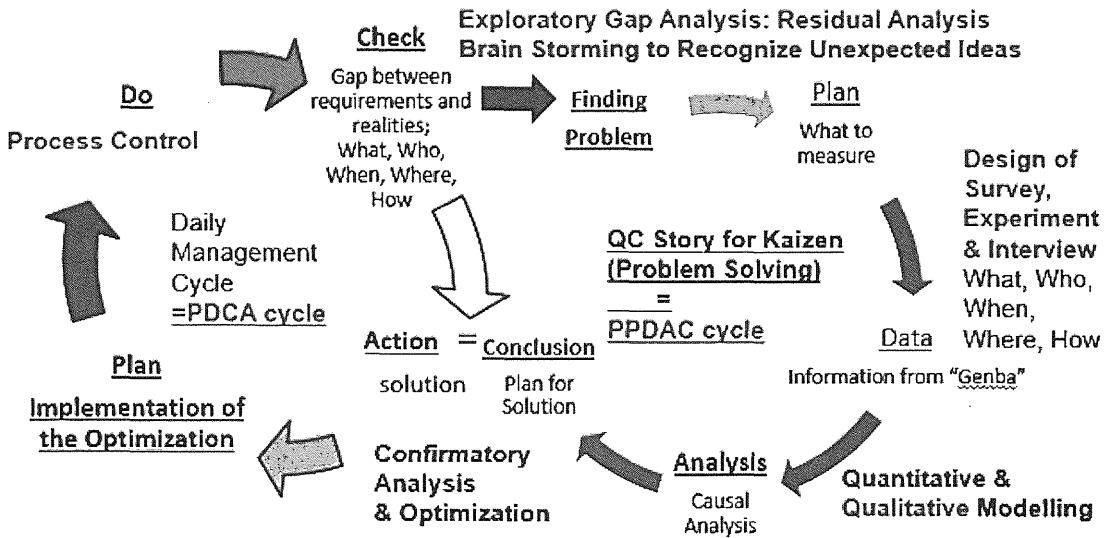


図7 PPDAC をもとにしたマネジメント・サイクル (Tsubaki, 2014)

Decision (Purchase)	State of Nature	
	GOOD ECONOMIC CONDITIONS .60	POOR ECONOMIC CONDITIONS .40
Apartment building	\$ 50,000	\$ 30,000
Office building	100,000	-40,000
Warehouse	30,000	10,000

The expected value (EV) for each decision is computed as follows:

$$EV(\text{apartment}) = \$50,000(.60) + 30,000(.40) = \$42,000$$

$$EV(\text{office}) = \$100,000(.60) - 40,000(.40) = \$44,000$$

$$EV(\text{warehouse}) = \$30,000(.60) + 10,000(.40) = \$22,000$$

図8 自然状態に対する確率による支払表と期待値

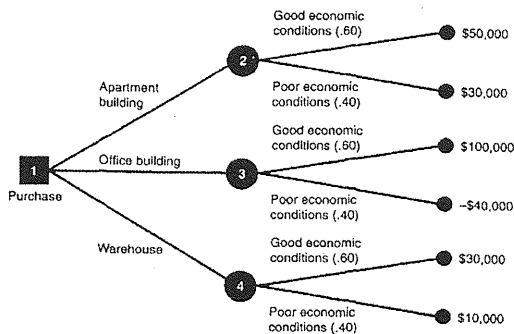


図9 支払いに関する意思樹形図

註：本講演は英語でなされた。本稿はその要旨であり、文責は松寄昭雄、小篠拓央にある。講演は、科学研究費基盤研究 A「グローバル社会における未来構築型数学教育実現への国際先導研究」(研究課題番号：26245082)、研究代表者磯田正美の一貫として実施された。講演の詳細を示す ppt ファイルは以下にある。

<http://www.criced.tsukuba.ac.jp/math/apec/apec2015/>