

## ガイドラインに関する統計学的研究

研究分担者 鎌谷直之 東京女子医科大学附属膠原病リウマチ痛風センター 客員教授  
分科会長・研究分担者 山中 寿 東京女子医科大学附属膠原病リウマチ痛風センター 教授

研究要旨：日本と英米の様々な統計学に関する教科書の比較、および多くの方面の日本人、外国人との対話により日本人と欧米人、特に英米人との不確実性と多様性に関する認識の枠組みの違いが明らかになった。それは3つに集約され、それぞれ言語に反映されている。1. 集合と要素の区別の問題、2. 任意の要素と特定の要素(変数と値)の区別の問題、3. 質的对象と量的対象の区別の問題、である。これらの点を知り教育を進めれば、日本においても疫学や統計学の重要性が理解され、ガイドラインも有効に用いられる事が示唆される。

### A. 研究目的

ガイドラインの作成、解釈などには統計学的考察が欠かせないが、日本社会は統計学的考察が弱いと言われる。その理由は不確実性と多様性を認識する枠組みの不足によると考えられる。本研究の目的はその理由を明らかにし、解決法を提案することである。上記の問題はガイドラインの作成、医師による解釈と説明、患者や家族の解釈など様々な場面に関係する。

### B. 研究方法

英米、更に日本の確率論、統計学、生物学の解説書、教科書などを精査する。欧米、更に、日本の確率論、統計学、疫学、言語学、医学、精神医学、心理学、数学などさまざまな分野の専門家との対話により問題を抽出し、仮説を構築し、それを検証する。

(倫理面への配慮)

厚生労働省の「疫学研究に関する倫理指針」など、様々な倫理指針、ガイドライン等を遵守する。

### C. 研究結果

日本と英米の様々な統計学に関する教科書の比較により、欧米の教科書ではコルモゴロフの公理的確率論を前面に出して説明を行っているが、日本の教

科書の手法は様々であることがわかった。例えば、コルモゴロフによる確率論では標本空間 (sample space) が一つの試行の全ての結果の集合として定義されるが、日本の教科書では事象を分割して、それ以上に分割できなくなった物を根源事象と言う、と説明している。

また、日本人と欧米人（特に英米人）の間で、認識する対象物の捉え方に違いがある事が示唆された。それを3つに分類すると、1. 集合と要素の区別の問題、2. 任意の要素と特定の要素(変数と値)の区別の問題、3. 質的对象と量的対象の区別の問題、である。その3つのいずれの場合も、日本社会では区別が明瞭ではなく、英米人では、特に知的な人々の間では明確である。

それは、言語に反映されている。上記の1の問題は、複数と単数の違い、2の問題は、不定冠詞と定冠詞のちがひ、3の問題は可算名詞と不可算名詞の違いである。

例を取ると、「敵が来る」という情報を伝える場合、英語では「An enemy comes.」か「Enemies come.」のどちらかを選択する必要がある。また、英語には集合名詞 (collective noun) が存在し、集合を認識対象とする事が可能である。また、本 (book) の場合、英語では「A copy of the book」という言い方

をしばしば用いるが、この場合、copy は要素、book は集合、あるいは集合のラベルと考える事ができる。Book は単に「モノ」というより、内容である「情報」であり、copy の集合である。集合と要素の区別は、不確実性と多様性を理解する上で不可欠の概念である。

また、a copy of the book という場合と、the copy of the book では意味が異なる。A copy は集合である copies の「任意」の一つであるが、the copy は「特定の」一つである。これは「変数と値」の概念と同等である。変数と値の区別は不確実性と多様性を理解する上で不可欠の概念である。

また、日本語では、「水」も「りんご」もモノであり本質的な違いは無いが、英語では water と apple には、可算か不可算かの大きな違いがある。これは、変数が連続か離散かの違いに関係しており、不確実性と多様性を理解する上で不可欠の概念である。

コルモゴロフの確率論では、一つの試行、一つの結果、全結果の集合である標本空間、標本空間の部分集合である出来事、出来事の集合である 集合体、集合体から  $[0,1]$  への写像である確率関数、標本空間、集合体、確率関数を結合した確率空間が定義される。これらの構造を理解する上で、集合と要素、変数と値、質と量、を区別する事は肝要である。

日本語に集合と要素、変数と値、質と量を区別し認識する概念が乏しい事を英米の統計学者、遺伝学者に説明したところ、それは重要な発見であると言う意見であった。それらの概念なしに確率論、統計学、遺伝統計学の理解は困難ではないかと言う意見が多く聞かれた。

以上の点を、実例を示して説明する方法により、不確実性と多様性を理解できるようになり、疫学研究の実施、結果の解釈、疫学研究の理解、ガイドライン作成、ガイドラインを用いた説明、説明の理解がより深まると考えられる。

#### D. 考察

以上述べた 3 つの点を、項目を挙げ、言語の比較

を行いながら実例を挙げて説明する事により、中等教育、高等教育、社会教育における不確実性と多様性の理解が深まると期待される。

日本社会はこれまで、存在が確実である「モノ」を認識対象としてきた。モノではない「情報」について、認識が深いとは言えない。そのため、先進国では中心になっている情報産業や健康産業の発達が妨げられており、医療において効率的な発展が妨げられている。

モノを対象とした製造業と違い、人間を対象とした情報産業や医療では不確実性と多様性が顕著である。従って、予測には確率の概念が必要である。しかし、確率は「集合の関数」であり、集合の概念なしに理解する事は困難である。また、個人の未来予測のためには集合の中の任意の要素である変数の概念なしに、確率的予測の概念を理解する事はできない。更には、疾患などの質的对象物と、重症度などの量的対象物を区別する必要がある。

これらの概念は、実はガイドラインを作成し、理解する上で不可欠の概念である。

#### E. 結論

日本社会が統計学的データを理解するためには、1. 集合と要素の区別の問題、2. 任意の要素と特定の要素(変数と値)の区別の問題、3. 質的对象と量的対象の区別の問題、を理解する必要がある。

#### F. 健康危険情報

無し。

#### G. 研究発表

##### 1. 論文発表

・鎌谷直之 確率を真に理解するには BIO Clinica 28, 13, 2013

##### 2. 学会発表

・鎌谷直之、「膨大なゲノム情報の利活用に必要な遺伝と多様性の統合」第 58 回日本人類遺伝学

会、2013年11月20-23日、仙台

H. 知的財産権の出願・登録

なし