

書評

計量情報学－図書館／言語研究への応用

影浦峠著 著. 丸善. 2000年. 182p. 2,600円(本体)

駿河台大学文化情報学部 岸田 和明

本書で意味するところの「計量情報学」とは、一般には、計量書誌学と呼ばれている図書館情報学の下位領域が発展したものと見なすことができる。この計量書誌学が探究する基本的な問題としては、いわゆる「集中と分散の現象」や「引用分析を用いた学問領域の構造の析出」を挙げることができる。このうち前者に関しては、単語の使用頻度や研究者の発表論文数が具体的な問題として取り上げられ、さまざまな角度から研究が進められてきた。その結果、ジップの法則やロトカの法則と呼ばれる一種の規則性が普遍的に観察されることが知られており、これらの法則のさまざまな性質等が明らかにされている。

本書はこのような計量書誌学的な法則について知るための格好の学術書である。特に、統計的な理論面からきちんと押さえたいという人には最適であろう。さらに、本書の大きな特徴は、計量情報学をこのような計量書誌学と「計量言語学」とを幅広く含むものとして捉え、計量言語学における成果を大幅に取り入れている点にある。この点、本書を計量言語学の問題を知るための本として使うこともできる。著者の影浦氏は、現在、国立情報学研究所で助教授として勤務されており、図書館情報学と深く関わりながらも、これまで、専門用語をはじめとする言語学的な問題に関する研究成果を数多く出されきた(本学会でも何度か研究報告をされている)。このあたりのことが本書にも色濃く反映されているようである。

具体的には、本書では、計量情報学的なデータから算出される統計量の「標本量依存性」の問題に数多くのページが割かれている。これは、計量情報学の基本的な問題であり、計量情報学が扱うデータが、標本平均の分布に

対して正規分布を仮定できるような通常の統計学的データではないことに起因する問題である。誤解を恐れずに簡単に説明すれば、標本のサイズを大きくしていくと系統的に統計量が変化していくという性質こそが標本量依存性であり、この結果、そのような現象を実証的に分析するには、通常の統計学にはない、いくつかの工夫あるいは解釈が必要になる。本書では、まずこの問題の導入に始まり、その具体例、その解釈をめぐる議論、そのような性質を持つデータに対する分析方法などが丁寧に解説されていく。そして最後には、その分析方法の1つであるZipf族LNRE(Large Number of Rare Events)モデルの説明で締めくくられている。

のことからわかるように、本書は計量情報学の諸問題を幅広く概観したものではなく、最近の話題の1つに焦点を当て、それを基本から高度なレベルの研究成果まで、根気よく解説した学術書である。そのため、数式も数多いが、その説明・展開等は非常に親切である。類書の中には、数式の説明を簡便に済ませようとして、途中の展開や添え字を省略したために、かえってわかりにくくなっているものが見受けられるが、本書にはそのようなことはない。記号の説明も十分で、何の記号かが明瞭でないためのフラストレーションもおこらない。

このようなレベルの計量情報学(あるいは計量言語学)の単行書を日本語で読めるというのは幸せなことである。この点、著者と出版社とに敬意を表したい。