

金融データサイエンス・概論 —オルタナティブデータと機械学習によるアルファの追求—

中 妻 照 雄

目 次

1. データサイエンスの時代
2. データ駆動によるアルファの追求
3. 情報効率性の隙間
4. オルタナティブデータの利活用
5. 「アート」としてのデータサイエンス

本稿では、まず近年におけるデータサイエンスの隆盛を支える技術革新について振り返るとともに、金融市場におけるアルファの追求と情報効率性の観点から金融の実務において機械学習をはじめとするデータサイエンスの手法を応用する意義を論じる。続いて、最近注目を集めているオルタナティブデータの金融における活用事例を紹介するとともに、データサイエンスの手法を実践する際の課題と注意点、そして今後の方向性について述べる。

1. データサイエンスの時代

過去10年余りの間に、データサイエンスに対する人々の関心は飛躍的に高まった。まずDavenport and Patil [2012] が*Harvard Business Review*で、“Data scientist: The sexiest job of the 21st century”（データサイエンティストは21世紀で最もセクシーな職業）という論文を発表し、「データサイエンティスト」および「データサイエンス」という言葉が一躍世界的に有名となった。さらに*The Economist*の2017年5月6日号に“The world’s most valuable resource is no longer oil, but data”（世界で最も価値ある資源は石油

ではなくデータである）という特集記事が掲載され、ビッグデータの重要性が広く認知されることとなった。

このような傾向は今も続いている。特定のトピックに対する人々の関心度を示す手近な指標として、Googleトレンドで「Data Scientist」および「Data Science」というトピックの検索動向を調べてみると、2012年からの10年余りで、全世界でのこれらのトピックに対する関心度は20倍以上になっていることが分かる。

近年における「データサイエンスの時代」の到来とその発展を支えている技術的要因としては、データの加速度的蓄積、コンピューターの高性能



中妻 照雄（なかつま てるお）

慶應義塾大学経済学部教授。1998年ラトガーズ大学経済学博士課程修了（Ph.D.）。一橋大学経済学研究所専任講師を経て、現職。主な著書に、『Pythonによるベイズ統計学入門』（朝倉書店、2019年）がある。