

# ビッグデータと景気分析

跡見学園女子大学 マネジメント学部教授 山澤 成康

夏目漱石の小説には株の話がよく出てくる。どれも威勢のいい話である。「株」という単語の使用回数は、『三四郎』と『野分』は2回、『それから』は4回である。

『吾輩は猫である』は最も多く、6回出てくる。同小説が連載されたのは1905年から1906年で、当時東京には3つの鉄道会社があった。鉄道とはいっても、路面電車のことだ。そのうちの一つの株を持っていると非常に儲かるという話が載っている。猫の飼い主の苦沙弥(くしゃみ)先生の友達である迷亭先生は、「半株だって千年も持つてうちゃ倉が三つぐらい建つ」と大げさに言っている。また、株に興味がない苦沙弥先生のことを「株といえど大根の兄弟分ぐらいに考えているんだから」と馬鹿にしている。

他には、「保険」という単語は、株より多く15回出てくる。ほら吹き迷亭先生が考えた存在しない謎の料理「トチメンボー」は38回も出てきている。

小説のほか、ホームページに使われている言葉、新聞記事、SNS(ソーシャル・ネットワーキング・サービス)の言葉など、大量の文章情報(ビッグデータ)が簡単に入手できるようになった。ソフトウェアの発達で分析も容易になった。文章から単語を取り出して言葉の数を数えるだけでなく、それを元に応用的な分析もできる。

経済統計では、内閣府の景気ウォッチャー調査が言葉の宝庫である。景気ウォッチャー調査は、景気を敏感に感じとることができるタクシーの運転手やスーパーのレジ担当者などに景況を聞いて数値化したものだ。同時に「景気判断理由集」として、判断した理由を文章の形で集めている。たとえば、2015年11月調査では、小売店店員の「品質が良く、やや高価格の商品や地方のお取り寄せ商品など、価格にあまり左右されない商品の動きが良い。」という声や、タクシー運転手の「11月に入り一気に量が減

り、客の動きが読めない。深夜は更に悪くなっており、最悪の月である。12月に期待したい。」といった声載っている。

2015年1月から11月までを集計すると約2万3000個の文章になる。最もよく使われる言葉のベスト5は、「前年」「客」「売上」「良い」「販売」となる。言葉の数としては、「良い(2018個)」の方が「悪い(1032個)」より多く、「増加(1731個)」の方が「減少(1716個)」より多いので、全体的な景気は良かったと考えられる。

次に、一つの文章の中で同時に使われる言葉は何かを探り出してみる。共起ネットワーク分析というものだ。

「景気」「回復」「良い」の3語は一つの文で同時に使われることが多く、景気回復を表している。一方で「厳しい」「状況」「続く」の組み合わせや「変化」「ない」という組み合わせも多く、一部の業種、地域では回復が進んでいないことがわかる。また、「単価」「上昇」「来客」「減少」が同時に使われることも多く、2015年の特徴を表している。「人材」「不足」「確保」「できない」という組み合わせは、現在の人手不足の状況を表している。「求人」「増加」といった雇用の好転を表す繋がりもある。言葉の分析から分かることも多い。

最も身近な言葉の分析法は、検索サイトの検索件数だろう。ある言葉がどれだけ使われたかを時系列で把握できる。『吾輩は猫である』が書かれてから110年目の2015年は猫がヒットした年だった。猫の動画が話題になり、猫と遊ぶ猫カフェが話題になった。検索サイトでの検索件数をみると、2014年までは「犬」と「猫」の検索件数にはそれほど差がなかったが、2015年になってから猫の検索件数は急上昇し、犬の検索件数の2倍以上になっている。