

今泉忠：「ビールの売上げは予測できるか：みんなの意見を反映するには」

オムニバスによる問題解決学総論の B クラスの第 5 回目として今泉教授が登壇。今回は、「ビールの売上げは予測できるか」といったテーマに、未来を予測したいという問題（課題）があった場合にどういったアプローチが取り得るかについて、専門の統計学の立場から話が始まる。以下、講義のポイントをまとめてみよう。

まずは、以下のような質問からスタート。

- 1) 5 万円のデジタルカメラを、A：「20%ポイント還元で買う」、B：「4 万 1500 円で買う」。どちらを選びますか、
- 2) 遺産の一部をもらえることになった。ただし、配分方として、A：「100 万円確実にもらえる」、B：「150 万円を 75%の確率でもらえるか、0 円となる」。どちらを選びますか、
- 3) 陶器店に入ったならば、入り口にあった壺を壊してしまった。お店から壺代金 100 万円弁償を迫られた。A：「100 万円を払う」、B：「裁判では費用と合わせて計 150 万円を支払う確率 75%で、勝てば支払いはなし。どちらを選びますか。

これらの質問への回答を通じて、意思決定が意外と合理的では無いことを体感させる。必ずしも人の意見は一貫性があるとは限らない。実際、2) と 3) の回答で一貫性がない回答が多数観られた。

これを元に、「人の意見（意思決定）は時と場合によって変わるものだ」という話へ。では、なぜ人の意見は使い分けられるのだろうか。この要因として「外部からいろいろなデータが入り、それを情報として扱うからだ」という。たとえば、遺産の例では収入という情報が入り、壺の例では出費という情報が追加される。だからこそ、2) と 3) で回答の一貫性が損なわれる意思決定も普通に行われてしまう。そして「判断データが増えると、視点や立場が変化しやすい」とも指摘された。

これらの話を受け、「みんなの意見を「予測」に使えないか」という話に進む。まずはビールの売上げのデータから売上傾向をグラフ化（折れ線グラフ）してみる。すると、8 月や 12 月に売れるという季節性があることがわかる。統計学的には、こういった季節性を加味して予測するという方法がある。でも、ここでは統計学の専門家とは異なるアプローチができないかを考えてみる。

例えば、kizasi.jp (<http://www.kizasi.jp>) というブログから話題を知り、兆しを見つけ

るというサイトで、「ビール」という言葉がどれくらいブログで語られるかを知ることができる。いわば、クイズミリオネアでいうところの「オーディエンス」という方法に近い。実際のデータを集めてグラフにしてみると、ブログでの「ビール」という言葉の出現数と、ビールの売上との関係が強いことがわかる。ただし、例外は、12月。12月はブログでの出現数に対して、ビールの売上げが非常に多い。こういう「外れ値」を理解するには、単にオーディエンスに頼るだけでは上手くいかない。この例で、12月だけ外れている、つまり「オーディエンスの意見だけでは、当たらない部分がある」という問題を発見できたということになるだろう。

問題を見つけられれば、さらに「なぜ12月だけ外れるのか」を考えるヒントを考えることができるだろう。では、データの専門家ならこのデータを用いて、さらにどうこの問題にアプローチするだろうか。

kizashi.jpのデータには、ビールという言葉だけではなく、ビールと一緒に出てくる単語はどんな言葉があるかについても知ることができる。だとすれば、ビールと一緒に出てくる言葉を「12月とそれ以外の月で共通する言葉」と「12月に顕著に多く出てくる言葉」とに分類することで、何かがわかるだろう。実際、12月に多く出てくる言葉は、“忘年会”、“クリスマス”といった言葉が見つかる。これらをヒントに、どうして12月のみ話題に出ている以上にビールが売れるのかのヒントが見えてくるはずだ。

<まとめ>

これらの話を通じて強調したいことは以下の2点だとのこと。

一つは、問題解決には全体を考える。つまり問題を発見するには、木を見るより森を見るということ。そのためには、さまざまな情報を得るということも有効だとのこと。クイズ\$ミリオネアの4つのライフライン的な手法も問題解決のツールとして紹介された。

もう一つは、データを活用して比較してものを考えること。せっかく「森」視点で得られたデータも、比較しなければ見えないことがある。先ほどの12月のビールの例がその例だ。

いずれにしろ、「今の社会を膨大に蓄積されているデータをうまく活用する問題解決の仕方というものに興味を持って欲しい」というまとめがなされた。