

パーソナルアシスタント町田通信コラム

<第112回> 統計学的な考え方

知人のことを「おかしな人」と思っていたら、意外に「しっかりした人」だった。こんな経験はありませんか？他人のことや、本当のこと、そして自分のことも、よくは分からないものです。そんな時に役に立つのが統計学的な考え方です。心理統計学の初歩を使って統計学的な考え方に触れてみたいと思います。

よく使われる「普通ではない」とか「一般常識がない」などの表現の「普通」や「一般」とは何のことでしょう？統計学では「対象の全員」を「母集団」と表現し、「普通」や「一般」は母集団の特徴の指標として「平均値」や「中央値」、「最頻値」などと表現されます。例えば、従業員が11人の会社を例に考えてみると、従業員の年齢の、

「平均値」は従業員の年齢の合計を従業員の数で割った値

「中央値」は従業員の年齢を小さい順に並べ6番目（中央）の人の年齢

「最頻値」は従業員の年齢で一番多いもの

となります。

それでは「普通」との表現に一番近い指標はどれでしょう？

<平均値>従業員11人中10人が25歳で、1人だけ80歳の会社の従業員の平均年齢は30歳です。「平均値」は母集団の中に特別にはずれた値が混じると、それに大きく引っ張られます。

<中央値>従業員の年齢が20歳~23歳5人、25歳1人(6番目)、75歳~80歳5人、だった場合「中央値」は25歳です。もし、6番目の人が60歳だったとすると「中央値」は60歳です。

<最頻値>従業員の年齢が20歳だけ2人いて、その他は全員違う場合「最頻値」は20歳です。

上記の例を考えると3つの指標と「普通」はどれも合わないように感じます。合う場合があり、合わない場合もあることが実際ではないでしょうか。つまり、統計学的な考え方をすると「普通」や「一般」とは概念であり、現実に存在するものではないこととなります。さらに、「普通ではない」との表現を平均値や最頻値からズレてい

ることとして使うことにも矛盾があります。母集団を均して作った指標の値と個人の値はズレがあって当然だからです。統計学的な考え方の良さは人を型に分けて評価するとき生ずる不整合や矛盾に気付けるようになることだと私は思います。つまり「分かったつもりの矛盾」に気付けることです。

近年、統計学は様々な業界で使われています。企業のマーケティングにはもちろん、人工知能開発にも地球温暖化予測にも使われています。「統計学が最強の学問である」西内啓著には「統計学はどんな権威もロジックも吹き飛ばして正解（最善の道）を導き出してくれる」と書かれています。興味の湧いた方は読んでみては如何でしょう？
(TK)