

擬似相関の利用価値

「身長が高いほど体重も増加する」といった“相関関係”を扱う際には、「相関関係はあっても因果関係は無い」という“擬似相関”の可能性についても注意を払わなければならないと言われている。擬似相関を排除したい最大の理由は、観測された相関関係が擬似相関であった場合、大きな誤解を招いてしまう危険性があるからだ。

一例を示そう。イギリスで行われたある調査の結果、「コウノトリの巣が減少した町では、出生率が低下する」という相関関係が見られたという。では、この調査結果を元に“子供の減少を食い止めるにはコウノトリの巣を増やせばよい”と考えてもいいのだろうか。言うまでも無く、答えはノーである。コウノトリの棲息域は地方に多く、出生率は都市部で低くなる傾向にある。すなわち、調査結果は“都市化の進行具合”が“コウノトリの巣の数”に置き換わっただけで、本質的には「都市化が進んだ町では、出生率が低下する」という相関関係を示しているに過ぎないのである。

上記は少々極端な例であるが、擬似相関は巷に実に多く溢れている。「成績向上のため毎朝ごはんを食べましょう」というキャンペーンを耳にする事もあるが、これもその1例として挙げられよう。

“朝食を摂るかどうか”と“成績の良し悪し”には確かに相関関係が認められ、「脳にエネルギーが与えられる事で、授業に集中できるようになる」といった解説もまことしやかに行われている。当

然こうした効果は期待できるのだろうが、一方朝食を摂るかどうかは生活習慣によって左右されるため、“朝食を必ず食べる”という行動は“きっちりとした生活習慣が身に付いている”事の表れであるとも考えられる。そして、この“生活習慣”という要素は、睡眠時間や自宅学習の時間など、集中力や成績に直結する要素とも密接に関係している可能性が高い。成績向上に大きく影響するのは“朝食”なのか“生活習慣”なのか。この点について、先のキャンペーンは多くを語っていない。

以上、2つの例を用いて擬似相関の弊害を示したが、この問題の本質は相関関係を因果関係そのものだと誤解している点にある。裏を返せば「擬似相関はノイズであり、利用価値が無い」という意味ではないのである。先のコウノトリの例で言えば、ベビー用品のDMを送るにあたり、他の情報が何も無ければコウノトリの巣の数が多い地域に重点的に送る事でヒット率は高まるだろうし、朝食の例では、入試面接官が「毎朝ご飯を食べるかどうか」という質問の回答を参考にすることは、優秀な生徒を効率的に集めるという観点では有効であろう。“擬似相関に気をつける”とは、つまるところ「相関関係と因果関係は別物であるという事を忘れるな」という意味であり、両者を誤解無く用いる限りにおいては、擬似相関も有効利用する事ができる。何事も“使いよう”である。

(須貝悠也)