

研究員 の眼

平均にもいろいろある

求めたいのはどの平均？

保険研究部 上席研究員 篠原 拓也

(03)3512-1823 tshino@nli-research.co.jp

自然科学でも社会科学でも、データをあれこれと分析して、そこから仮説や命題を裏付けたり否定したりするような結果を導く。こうしたデータの分析のうち、頻繁に行われるのが、「平均をとる」という作業である。通常、平均は、データを抽出した元の集団の傾向を表すものと考えられている。このため、確率論や統計学では、平均に関する定理が多く、その考察が欠かせないものとなる。

平均は、小学校の算数で割り算をマスターした後に、高学年くらいから身につけるものだ。各データの値を足し算した結果を、データの個数で割り算して、平均が計算される。たとえば、ある学校のクラスで生徒の平均身長を求めるには、各生徒の身長を合計を生徒の人数で割ればよい。これは、「算術平均」と呼ばれる。もっとも単純で、わかりやすい平均である。

しかし、算術平均が役に立たない場合もある。たとえば、ある高級青果物店では、リンゴを1個400円、ミカンを1個100円で売っているとす。ある日、この店でリンゴが200個、ミカンが600個売れた。リンゴとミカンをあわせた平均単価はいくらだろうか。この場合、2つの単価の算術平均である250円（ $= (400+100) \div 2$ ）には意味がない。平均単価を求めるには、合計の売上高を、合計の売上個数で割り算する必要がある。売上高は、リンゴ8万円、ミカン6万円で、合計14万円。これを合計の売上個数800個で割り算した、175円が平均単価となる。これは、「加重平均」と呼ばれる。

また、成長率や伸び率など、物事の変化を表す率の平均には、別のものが必要になる。預金の利率で考えてみよう。預金の利率は、複利で表すと扱いやすい。これは、最初に預けた元本と前年度までの利息の合計に対して当年度の利息が付く、という利殖の仕組みを反映した数値だ。ある銀行の外貨預金で、利率が1年目は30%、2年目は2%と、大きく変化したとしよう。このとき、2年間の平均利率は、いくらだろうか。算術平均の16%ではない。元本として100万円を預けた場合で考えてみよう。1年後には、元本に30%の利子がついて130万円となった。2年後には、この130万円に2%の利子がついて、132万6,000円に増加した。つまり、2年間で1.326倍に増えたわけだ。1年間でみると、1.326の平方根をとって、1.152倍ほどに増えることを意味する。つまり、平均利率は約15.2%となる。これは、「幾何平均」と呼ばれる。

つぎに、速さや濃度や圧力といった、何かを何かで割り算して得られる比について、平均を求めることを考えてみよう。ここでは、小学生向けの時間・距離・速さの問題が有名だ。家から学校まで1.5kmの距離を、徒歩で往復するとしよう。行きは時速5km、帰りは時速3kmで歩いたとすると、往復で平均の速さはどれくらいだっただろうか。ここであわてて算術平均をとり、時速4kmと答えてはいけない。まず、往復にかかった時間を考えてみる。1.5kmを、時速5km、3kmで割り算する。行きは0.3時間、帰りは0.5時間かかったことになる。つまり、往復で3kmの距離を、0.8時間かけて歩いたことになるから、3kmを0.8時間で割り算して、平均の速さは、時速3.75kmとなる。これは、「調和平均」と呼ばれる。(実は、距離が1.5kmでなくて、どんな長さであっても、調和平均は変わらない。調和平均を割り算で求めるときに、往復の距離3km(割られる数)と、往復の時間0.8時間(割る数)は、ともに片道の距離1.5kmによって決まる。このため、割り算をすると、互いに打ち消しあうのだ。)

さらに、あるデータをとったときに、各データの算術平均からの乖離がどれくらいあるかという、データのブレについて平均をとることを考えてみよう。ブレの平均をとるときには、算術平均ではうまくいかない。たとえば、身長が1.6、1.7、1.8mのAさん、Bさん、Cさんの3人について、ブレの平均を考えてみよう。この3人の平均身長は1.7mなので、ブレは、Aさんはマイナス0.1m、Bさんはプラスマイナス0m、Cさんはプラス0.1mとなる。このブレの算術平均をとると0mとなる。

しかし、これでは3人の身長がいずれも1.7mでブレがない場合と同じ結果になってしまう。そこで、ブレを二乗したものの算術平均をとり、その平方根をとる。この例では、ブレの二乗の算術平均は0.00666… ($= ((\text{マイナス } 0.1\text{m})^2 + 0\text{m}^2 + 0.1\text{m}^2) \div 3$)。その平方根は、約0.082mとなる。これは、「二乗平均平方根」と呼ばれる。統計学などで、よく目にするものだ。

最後に、ある会社の職員の平均年収を考えてみる。実は、平均年収というときの平均には、いろいろな前提が置かれているので注意が必要だ。たとえば、組合員平均。労働組合に加入している職員の平均年収だ。組合に加入していない役員や管理職は対象外となる。そのため、組合員平均の平均年収は、実態よりも小さな金額になる。また、平均をとる対象は正社員だけか、それとも派遣社員・パートスタッフ・アルバイト職員も含むのかによっても、平均年収は大きく違ってくる。派遣社員等の職員占率が高い会社で平均年収をみる際には、対象職種に目をこらす必要がある。さらに1人、2人の高給役員のために、平均年収が歪むという問題も出てくる。中小企業の、ある会社を考えてみよう。

この会社はよくあるワンマン経営で、社長の年収は1億円と突出して高い。一方、50人いる従業員の年収は、一般的な給与所得者の水準だ。役職や勤務年数などによって異なり、10人の管理職は平均700万円、40人の非管理職は平均500万円となっている(この平均は、算術平均)。このとき、社長も含めた51人の算術平均をとると、平均年収は約725万円となる。社長の年収が、平均年収を引き上げている。だが、ほとんどの職員の年収は、725万円未満である。この会社の職員の年収を表す数値としては意味がないかもしれない。そこで代わりに、51人を年収が高いほうから順に並べたときに、ちょうど真ん中の26番目となる職員の年収をみしてみる。このほうが、この会社の平均的な職員の年収として意味があるだろう。この値は、「中央値」と呼ばれる。中央値は、平均順位のデータの値を表す。

以上のように、ひとくちに平均をとるといっても、いろいろな種類がある。何かのデータを収集して平均をとる場合、どの平均を使うべきか、よく考える必要があると思われるが、いかがだろうか。