

研究員 の眼

実数で見えるか、比率で見えるか どうすれば、より深い数量比較ができるか？

保険研究部 主任研究員 篠原 拓也
(03)3512-1823 tshino@nli-research.co.jp

医療や介護などの社会保障政策を検討する際は、複数の地域の実態を比較して、それぞれの特徴を把握するということが行われる。その場合、市町村の比較を、実数で行うこともあるが、そうすると、人口や面積など、各市町村の規模の影響を受けてしまう。これでは、適切な比較とは言えない。

そこで、実数でそのまま比較するのではなく、単位人口や単位面積あたりの比率に直して、比較しようということになる。比率を使えば、各市町村の規模の影響を受けないため、適切に比較できるのでは、と考える訳だ。

例えば、糖尿病の予防策の優先度を判断するために、A市と、B町の住民の、糖尿病の状況を比較することになったとしよう。疾病関係の調査を行ったところ、次のデータが得られたとする。

	A市	B町
(1) 糖尿病患者数	10,000人	110人
(2) 人口	500,000人	5,000人
糖尿病割合 (1)/(2)	2%	2.2%

A市は、地方の中核都市で、人口は500,000人。一方、B町は人口5,000人の典型的な規模の町だ。糖尿病患者の実数は、A市の方が圧倒的に多い。しかし、これは、人口が多いので当然と言える。そこで、患者数を人口で割り算して、糖尿病割合という比率で見てみよう。すると、B町の方が、この比率が高い。つまり、B町の方が、糖尿病になりやすい、ということがわかる。

そこで、糖尿病の予防策は、B町から優先的に行うことになる。しかし、ここで、ふと疑問が湧いてくる。患者の実数では、何十倍も多いA市よりも、B町の方を優先することになるが、本当にこれでいいのだろうか。つまり、実数よりも、比率を重視すべきなのだろうか、という疑問である。

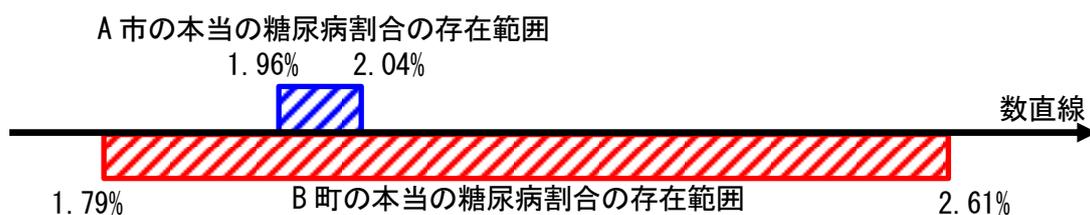
実は、同様のことは、スポーツで選手のパフォーマンスを比較する際にも見られる。選手を比率だけで比較していくと、いろいろと問題が生じる。代表的なのは、野球の打率ランキングだろう。各選手の打率を単純に比較すると、1打数1安打の選手は、打率10割となる。この選手をそのまま1位としてしまつては、もっと多くの打席に立って多くの安打を放った(凡退もした)他の選手と、適切に比較したことにならない。そこで、プロ野球等では、あらかじめ試合数の3.1倍などと、規定打席数を定めておいて、規定打席数を満たした選手だけを、打率ランキングの対象としている。

他のスポーツでも、個人成績をランキング形式で比較する際には、同様の数量基準を置いていることがある。例えば、バスケットボールのフリースロー成功率では、フリースローの成功数が一定数以上あることが、ランキングに入る要件となる。また、バレーボールのアタック決定率では、アタックの打数が、一定数以上あることが要件となる。

このように見ていくと、実数か、比率か、どちらか一方だけを見ても、適切な比較をしたことにはならないことがわかってくる。それでは、どうしたらよいだろうか。

数理統計学では、観測されたデータの信頼度を、問題にする。即ち、得られたデータには、たまたま生じたブレが含まれているだろう、と考えるのである。そして、そのブレは、母集団の大きさによって変わる。母集団が大きいほど、ブレは小さくなる。

そこで、先ほどの、A市とB町の糖尿病の比較に、話を戻そう。データから得られた糖尿病割合のブレ幅を考えてみる。このブレ幅は、数理統計学では、信頼区間と呼ばれる。例えば、本当の糖尿病割合が、95%の確率で存在する範囲として、信頼区間を考えてみよう。この信頼区間を、数直線上に図示すると、次のようになる。



こうしてみると、糖尿病割合は、B町の方が高い可能性が大、となる。しかし、A市の信頼区間が、B町の信頼区間の中に納まっているため、絶対にそうだとは言いきれない。A市の方が糖尿病割合が高い、という可能性も、それなりにあることがわかる。即ち、A市とB町のいずれか一方の糖尿病割合が、他方よりも必ず高いとは言えない。そのことを踏まえて、糖尿病予防策を検討すべきであろう。

このように、定量的な比較をする際に、それぞれの信頼区間を描いてみることで、実数か、比率かという二者択一ではなく、より深い比較が可能になると思われるが、いかがだろうか。