

第8回

クロス集計を使ってみよう

仮想に基づいて作ったクロス集計の例です。仙台市で100人にプロ野球3チームを挙げて好きな球団を選んでもらいました。集計したところ、楽天が43人、日本ハムが35人、ソフトバンクが22人でした。さすがに仙台では楽天ファンが多いようです。ここから、「地元の球団を応援する傾向がある」と考えたくくなります。果たしてそう考えてよいのでしょうか？

そこで、札幌と福岡の市民それぞれ100人に同じように聞いて集計し表にしてみました。右下表がクロス集計表です。なるほど、地元のチームの人気がいちばんになっています。クロス集計表からは、行(表側)と列(表頭)の因子(項目)の間に関係があるかどうかを探ることができます。

集計表があれば、グラフの作成や分析がエクセルのような表計算ソフトを使って簡単にできます。都市名を

日本公衆衛生学会理事・評議員
(一財)宮城県成人病予防協会学術・研究開発室長
小島 光洋

横軸に置いて、それぞれの都市の上に球団を並べた棒グラフを作ると視覚的な説明が可能になります。行と列の項目間に関係の有無はクロス集計が得意とするところですが、 χ^2 (カイ2乗)検定という方法を使うと、この例には100回に1回も起きないという数の偏りがあることが分かりました。どうやら、「住んでいる都市と好きな球団の間には関係がある」と考えてよいようですが、この例は仮想データです。念のため。

表●仮想のクロス集計

	日本ハム	楽 天	ソフトバンク	小 計
札 幌	42	33	25	100
仙 台	35	43	22	100
福 岡	27	31	42	100
小 計	104	107	89	300