

## 7. チェビシエフの定理

これまで、標準偏差は、ばらつきの度合いを示すデータの要約値であること、そして、この値が小であればあるほど、データのばらつきは小で平均への集中度は良く、この値が大であればあるほど、データのばらつきは大で平均への集中度は悪い、と判定できることなどについて説明してきました。

しかし、標準偏差の持つ意味は、これだけのものではなくて、実はもっと重要な意味を持つものであることが明らかにされています。

これを明らかにしたのは、**チェビシエフ**（ソ連1821～1894）で、その内容は、標準偏差を単位として、データの平均への集中度を測ることができる、というものです。この内容は、次の定理の形で述べられています。

(チェビシエフの定理)

どんなデータにおいても、平均値  $\bar{x}$  からのずれが、標準偏差  $\sigma$  の  $k$  倍より小のものは、全体の  $1 - \frac{1}{k^2}$  以上ある。(ただし  $k > 1$ )

この内容について説明しますと、例えば、

- $k = 1.5$  のときは、 $1 - \frac{1}{1.5^2} \doteq 0.56$  ですから、平均からのずれが、標準偏差の1.5倍以内のもの、すなわち、区間  $(\bar{x} - 1.5\sigma, \bar{x} + 1.5\sigma)$  内にあるものは、全体の約56%以上ある。
- $k = 2$  のときは、 $1 - \frac{1}{2^2} = \frac{3}{4} (0.75)$  ですから、平均からのずれが、標準偏差の2倍以内のもの、すなわち、区間  $(\bar{x} - 2\sigma, \bar{x} + 2\sigma)$  内にあるものは、全体の  $\frac{3}{4}$  以上 (75%以上) ある。
- $k = 3$  のときは、 $1 - \frac{1}{3^2} = \frac{8}{9} (\doteq 0.9)$  ですから、平均からのずれが、標準偏差の3倍以内のもの、すなわち、区間  $(\bar{x} - 3\sigma, \bar{x} + 3\sigma)$  内にあるものは、全体の  $\frac{8}{9}$  以上 (ほぼ90%以上) ある。

というもので、どんなデータにおいても成立するところに大きな意義があります。