

用語

尖度・歪度 (Kurtosis, Skewness)

尖度および歪度は、確率分布の非ガウス性すなわちガウス分布からのずれ具合を表すために用いられる尺度である。尖度は分布の裾の広がりを表す指標で、ガウス分布では3である。尖度が3より大きいときはガウス分布に比べて裾が重くなり、尖度が3より小さいときはガウス分布に比べて裾が軽くなる。それに対し、歪度は分布の非対称性を表す指標で、左右対称なガウス分布では0である。歪度が0より大きいときは分布が右に歪み、0より小さい場合には分布が左に歪むようになる。

いま、 x_i を各データの値、 m をデータの平均値、 N をデータの個数、 σ をデータの標準偏差とすると、尖度 K と歪度 S は以下の式で表される。

$$K = \frac{M_4}{\sigma^4}, \quad S = \frac{M_3}{\sigma^3}$$

$$M_n = \frac{1}{N} \sum_{i=1}^N (x_i - m)^n \quad (n = 3, 4)$$

$$m = \frac{1}{N} \sum_{i=1}^N x_i$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - m)^2}$$

図1に、平均が0、標準偏差が1であり、(i) 尖度が3、歪度が0の確率分布(ガウス分布)、(ii) 尖度が10、歪度が0の確率分布、(iii)

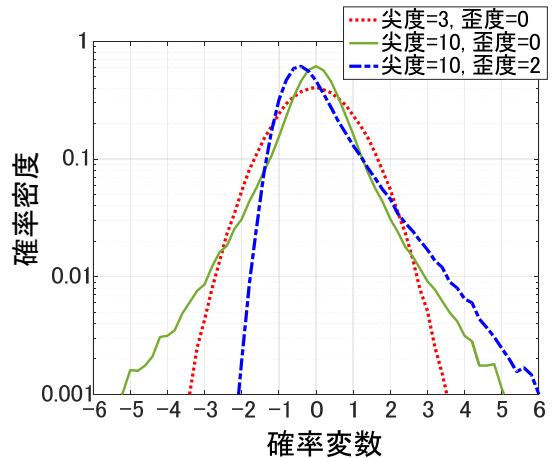


図1 尖度・歪度と確率分布の関係

尖度が10、歪度が2の確率分布の一例を示す。

(i) 尖度が3、歪度が0の確率分布(点線)をみると、確率変数の絶対値が3を超える値はほとんど発生しない。それに対し、(ii) 尖度が10、歪度が0の確率分布(実線)では分布の裾が重くなり、確率変数の絶対値が3を超えるような中心値から外れた値の発生頻度が高くなる。また、(iii) 尖度が10、歪度が2の確率分布(一点鎖線)では右側に歪んだ左右非対称な分布となる。確率分布の右側では中心値から外れた値の発生頻度が高くなるのに対し、確率分布の左側では確率変数の絶対値が3を超えるような値は発生せず、中心値に近い値が発生していることがわかる。

尖度・歪度と確率分布には上述したような関係があるため、尖度と歪度をみることで、分布がガウス分布と比べてどのようにずれているのか把握することができる。

(地独) 大阪府立産業技術総合研究所 細山亮