

Lesson 9. 頻度と分布

§ D. 正規分布

正規分布
Normal Distribution

- 最もよく使われる連続確率分布
- 釣り鐘形の曲線
- - から + までの値を取る
- 平均 mean = 中央値 median
= 最頻値 mode
- 曲線より下の面積は1に等しい

正規分布 (続き)

- 確率変数Xは以下の確率密度関数に当てはまる時に正規分布をとる。

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

- π は円周率
- 手計算では困難なので式を知らなくても良い。

正規分布のパラメータ

- 母集団の分布の平均 $= \mu = \frac{\sum_{i=1}^n x_i}{N}$

- 母集団の分布の標準偏差

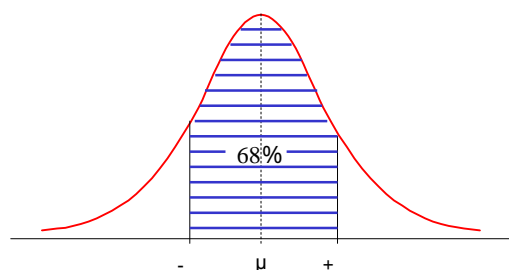
$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{N}}$$

- 母集団の分布の分散 $= \sigma^2$

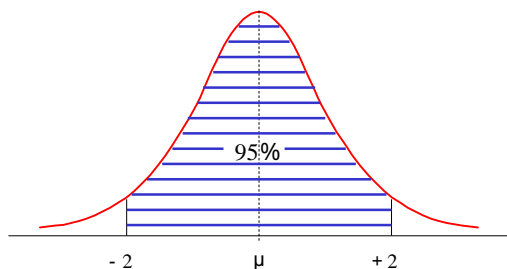
正規分布の性質

- 曲線以下の面積の約 68% は $\mu \pm 1$ の間にはいる。
- 曲線以下の面積の約 95% は $\mu \pm 2$ の間にはいる。
- 曲線以下の面積の約 99% は $\mu \pm 3$ の間にはいる。
- 平均や標準偏差の値にかかわらず、曲線以下の全面積は確率1である。

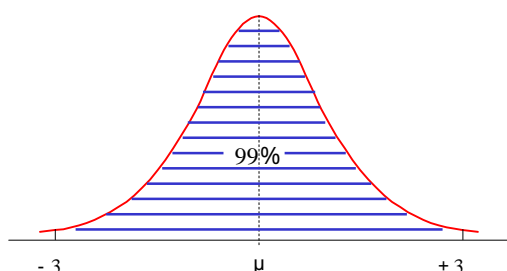
正規分布の面積



正規分布の面積 (続き)



正規分布の面積 (続き)



標準正規分布

Standard Normal Distribution

- 確率変数の正規分布のうち、左右対称で平均 = 0、分散 = 1 となるものを標準正規分布という。
- つまり、 $N(0, 1)$ の正規分布である。
ただし、 $N(\mu, \sigma^2)$
- 統計の基本となる正規分布である。

標準化 Standardization

- 一般の正規分布は、その平均と標準偏差より標準正規分布に変換できる。これを**標準化**という。
- 確率変数 X は以下の式で標準正規分布をとる確率変数 Z に変換される。

$$Z = \frac{X - \mu}{\sigma}$$

パラメトリック

- 統計確率が正規分布に従う検定手法を**パラメトリック検定**という。(正規分布をパラメーターとしているという意味)
- 正規分布に従わないものを**ノンパラメトリック検定**という。
