

データの分布

道具としての統計学

奥田 千恵子

福井 克法

目次

- ▶ 二項分布
- ▶ 正規分布
- ▶ 二項分布と正規分布の関係性
- ▶ 正規分布のまとめ

二項分布とは

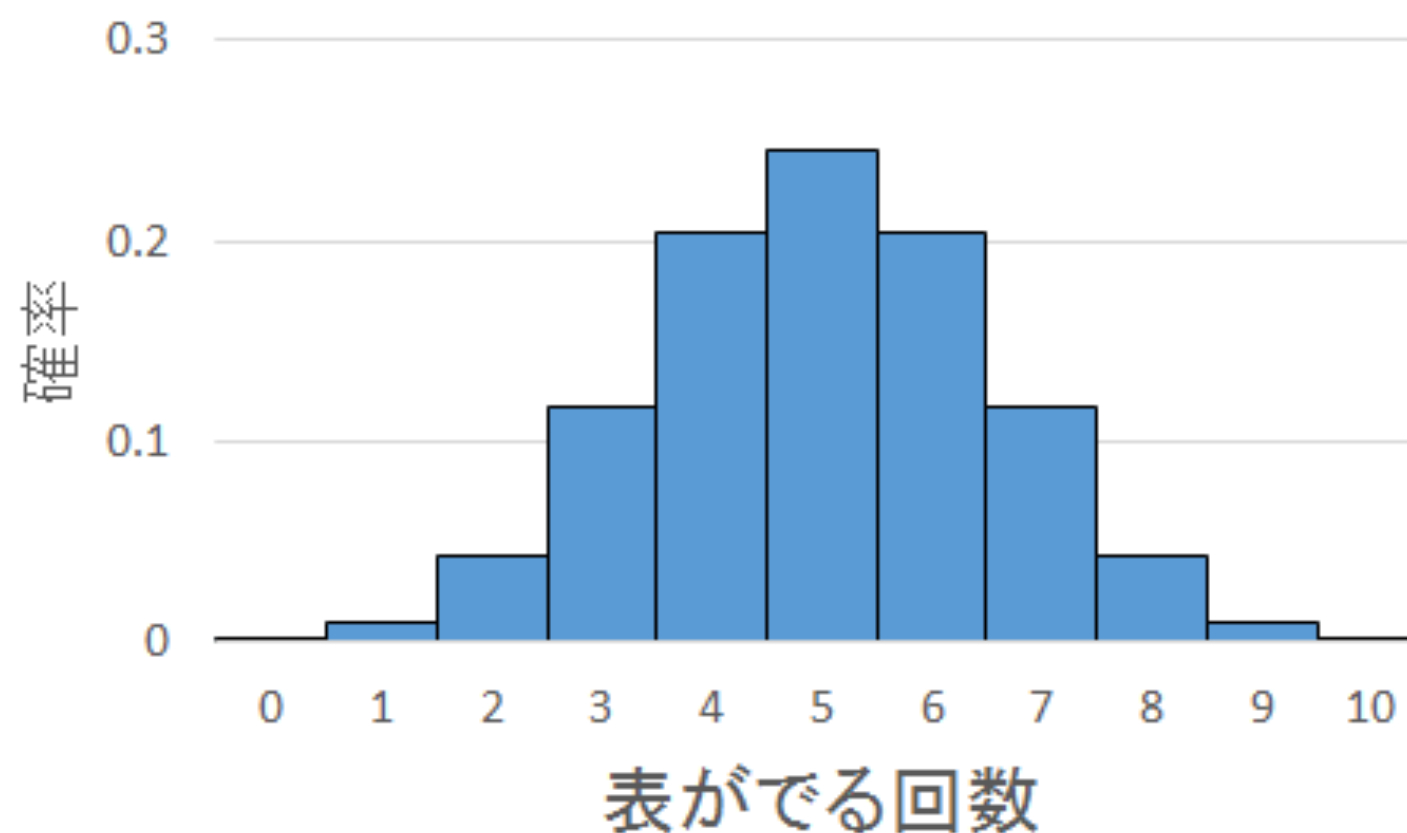
結果が成功か失敗のいずれかである n 回の独立な試行を行ったときの成功数で表される離散確率分布である。

- ・ 可能性が二つしかない場合の確率分布

例：（コイン投げ）

二項分布のグラフ

- ▶ 例：10回投げて何回表が出るか？
- ▶ 5回が一番確率が高くなるようなグラフとなる



二項分布の確率計算の例

- ▶ 10個のコインを投げた時、
表が出る確率を求める方法

1.x個を1組として

取り出す組み合わせの数を計算

$${}_{10}C_x = \frac{10!}{x!(10-x)!}$$

2.表が出る確率(p) 裏(1-p)

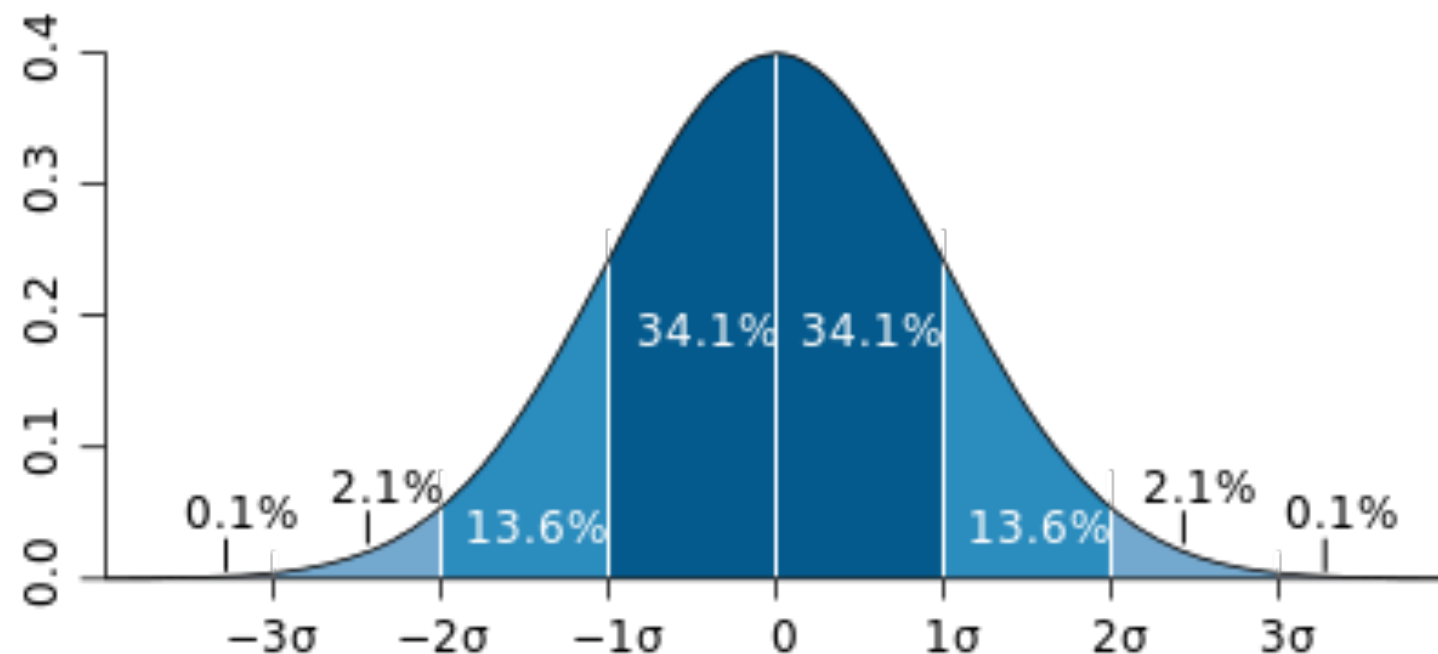
表が出る確率P(x)

$$\begin{aligned} P(X) &= {}_{10}C_x \times p^x (1-p)^{10-x} \\ &= {}_{10}C_x \times \left(\frac{1}{2}\right)^{10} \end{aligned}$$

x	${}_{10}C_x$	P(x)
0	1	0.000977
1	10	0.009766
2	45	0.043945
3	120	0.117188
4	210	0.205078
5	252	0.246094
6	210	0.205078
7	120	0.117188
8	45	0.043845
9	10	0.009766
10	1	0.000977

正規分布とは

- ▶ 正規分布 (Normal distribution)
もっとも一般的な分布
- ▶ 左右対称・釣り鐘型の性質をもつ分布
- ▶ 平均値付近の付近に集積するようなデータの分布を表した
連続的な変数の確率分布



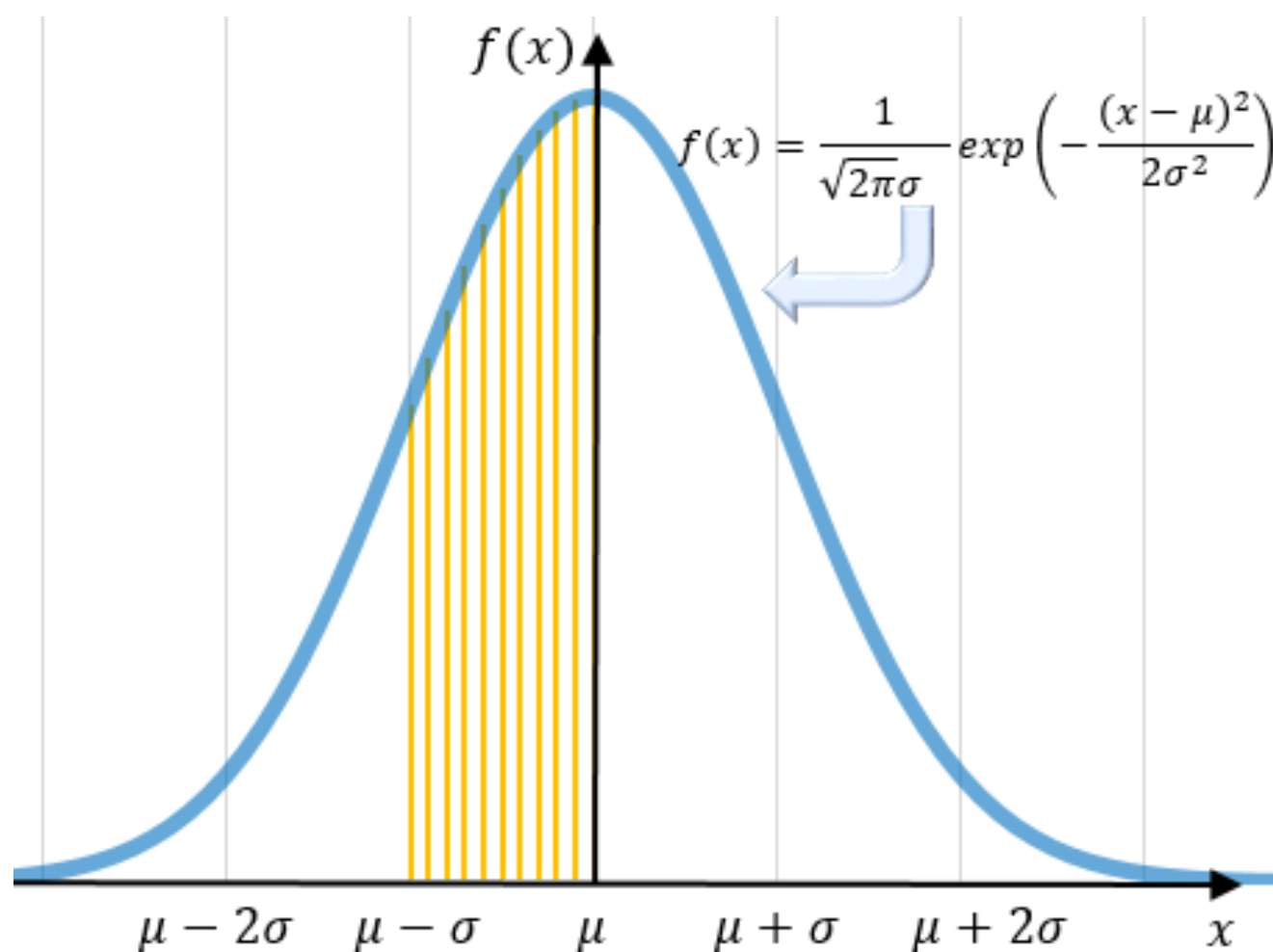
正規分布のパラメータ

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

x : 変数 σ^2 : 分散 σ : 標準偏差 (分散の正の平方根)

正規分布のグラフと確率

• 正規分布
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right)$$



「 $\mu - \sigma$ 以上 μ 以下になる確率」 =

『オレンジで塗りつぶされた部分の面積』

正規分布の利用例 1

▶ 例題)

日本の成人男性の平均身長 $\mu=171\text{cm}$ 、標準偏差 $\sigma=6\text{cm}$ とする。ランダムに一人選んだ場合、身長が165cm以上171cm以下である確率は何%か？

→ 正規分布に従う場合

確率は

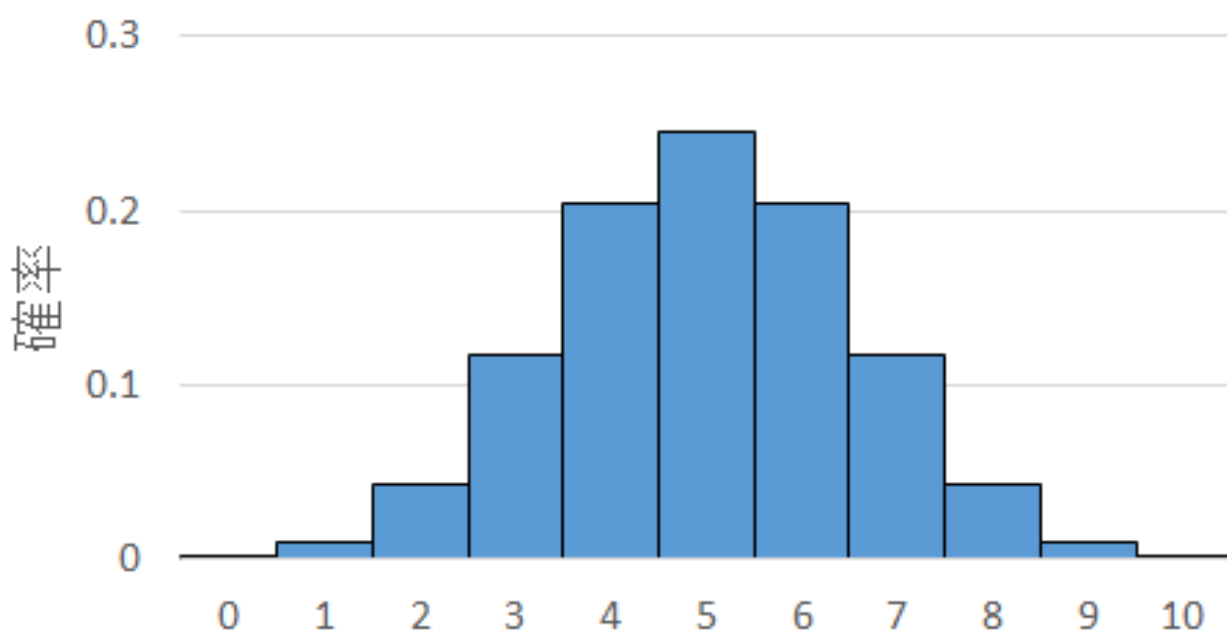
「 $\mu - \sigma$ 以上 μ 以下になる確率」 → グラフの対象部分の面積

面積 約0.3413 → 34%

正規分布と二項分布の関係性

▶ (例) コイン投げ

n 個のコインのうち表の個数、 x は0から n 個までの $n+1$ 通りの値をとる



10枚投げた場合

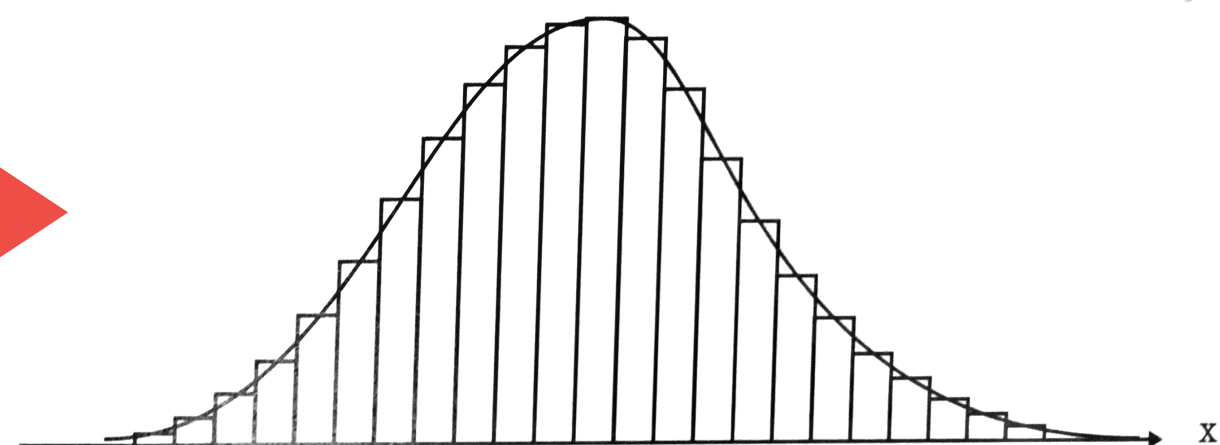
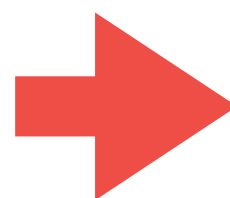


図 5-3

投げるコインの数 n を増やしていく

- ・ 正規分布に近づいていく

→ 二項分布の極限の形は正規分布と言われている

正規分布のまとめ

- ▶ 世の中の社会現象や自然現象の中には、その確率変数が正規分布に従うとみなせるものが多く存在する
- ▶ 平均と標準偏差が分かれば、
その現象がどのくらいの確率で起こるかを計算できる場合が多い