情報理論03

藤田一寿

ver. 20210205

確率変数

- 各値に確率が与えられている変数
  - 離散型と連続型がある。
  - サイコロの場合は出る目が確率変数となり、離散型である。

確率分布

Xが離散値をとる場合、確率変数Xのそれぞれの値に対する確率

$$P(X = x_k) = f(x_k) \qquad (k = 1, 2, \ldots)$$

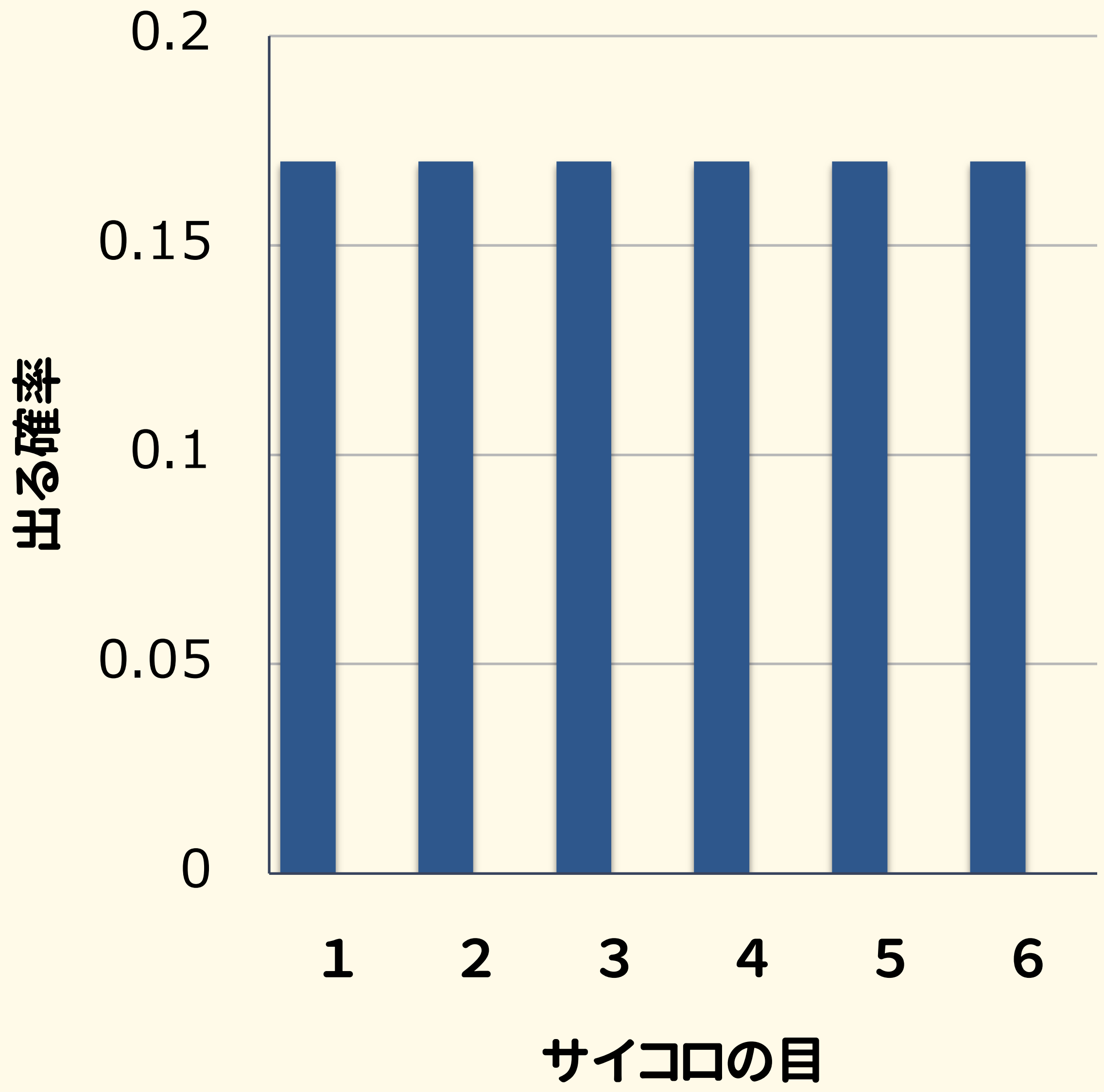
をXの確率分布という。

$f(x_k) \geq 0$ かつ $\displaystyle\sum_{k=1}^{\infty} f(x_k) = 1$ を満たす。

x_kが離散値なのでfは離散確率分布である。

離散確率分布の例

- サイコロの目の出る確率の分布
- 離散一様分布



確率密度関数

Xが連続の場合、Xがaからbの値をとる確率 は

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

と表される。

xが連続値な場合、fは確率密度関数と呼ばれる。

f(x)は $f(x) \geq 0$ かつ $\int_{-\infty}^{\infty} f(x)dx = 1$ の条件をもつ。

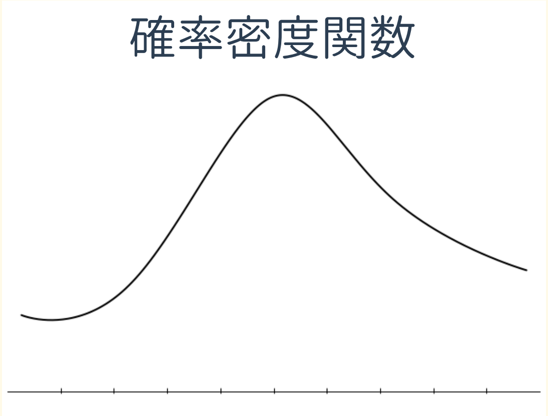
確率密度関数の注意

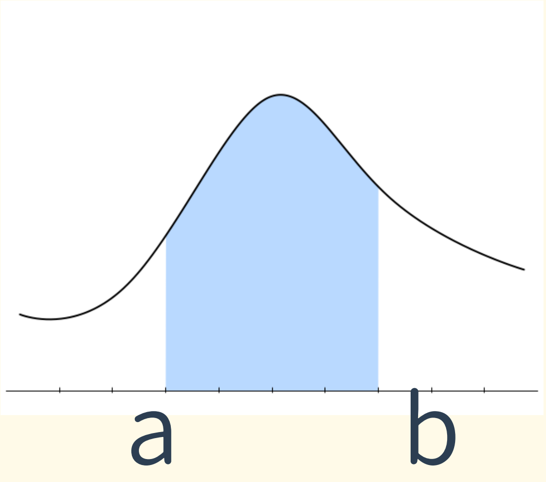
$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

なので、Xがaとなる確率は

$$P(X = a) = \int_a^a f(x)dx = 0$$

となる。確率密度関数fは密度である。そのため、積分しないと確率にならない。これが、離散確率分布との大きな違いである。



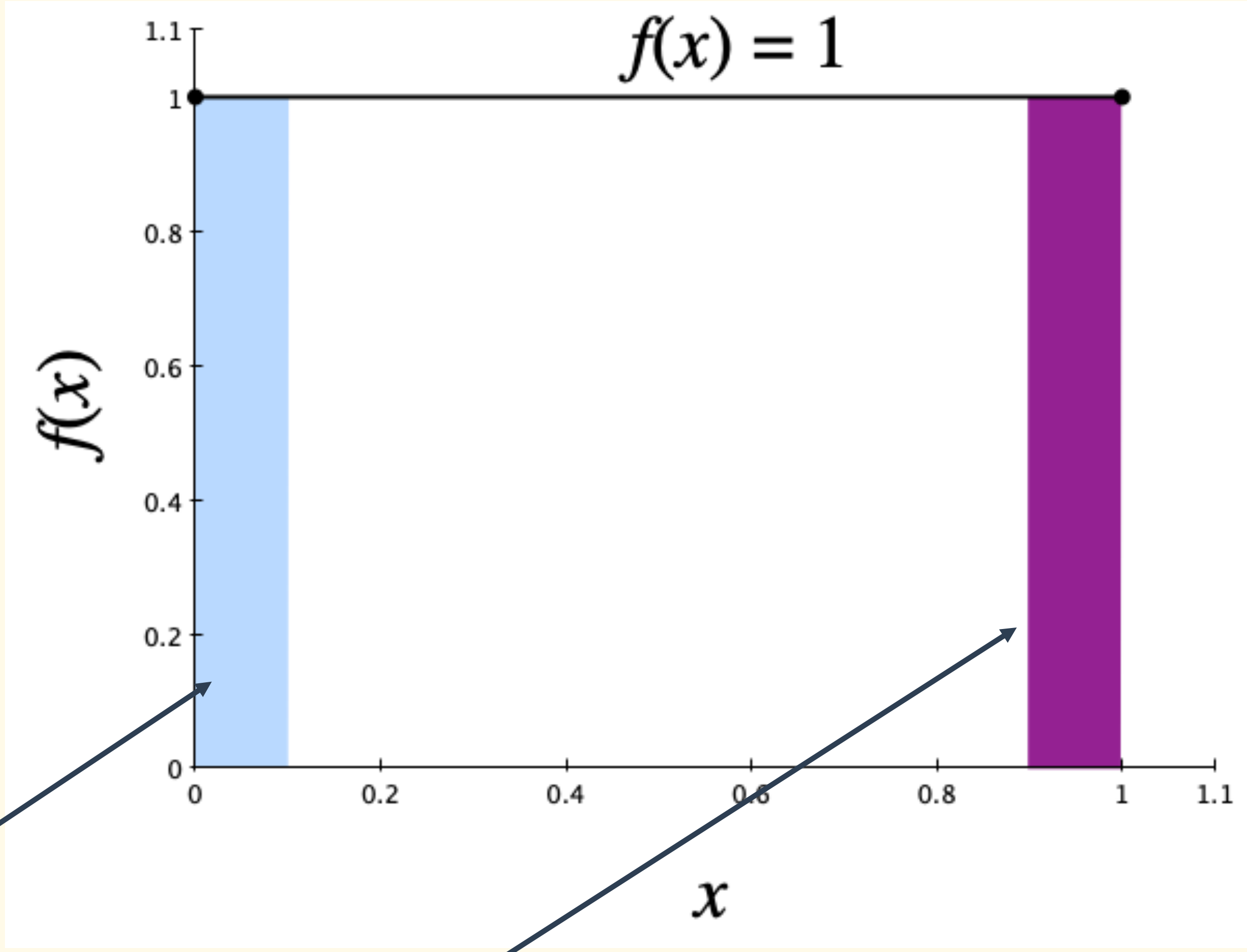


連続一様分布

xが0から1までの値を取る場合の、連続一様分布

$$f(x) = 1$$

$$P(a \leq x \leq b) = \int_a^b dx$$



$$P(0 \leq x \leq 0.1) = 0.1$$

$$P(0.9 \leq x \leq 1) = 1 - 0.9 = 0.1$$

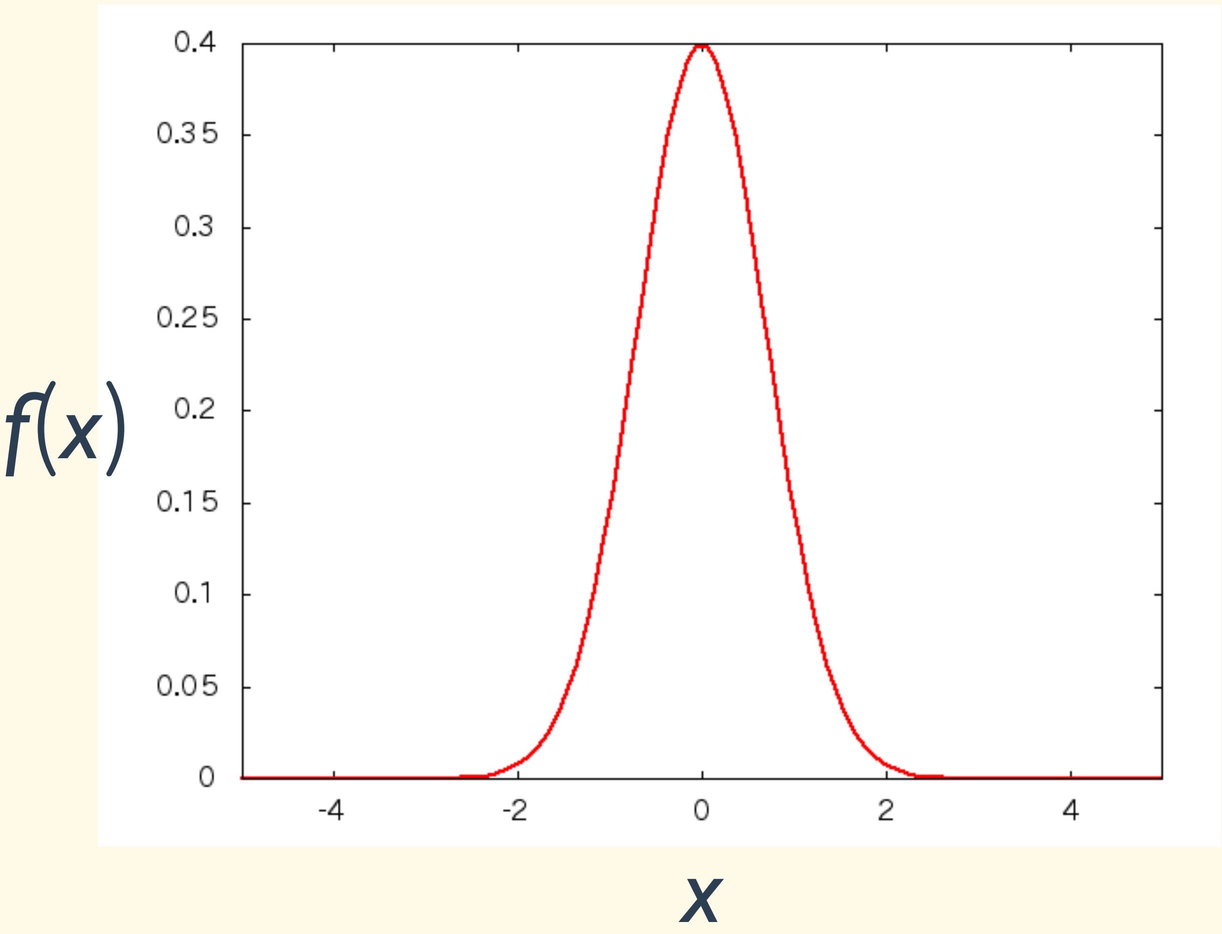
代表的な確率分布

- 離散
  - 二項分布
  - ベルヌーイ分布
  - ポアソン分布
- 連続
  - ガウス分布
  - 指数分布
  - ガンマ分布

ガウス分布(正規分布)

- 最も重要な確率分布

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



σは標準偏差,μはxの期待値となる.

期待値

確率変数g(X)の期待値は、確率変数g(X)に対し、その値が取る確率の重み付き和で表される。

$$E(g(X)) = \sum_x g(x)f(x)$$ 離散

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx$$ 連続

期待値と統計量

g(x)	g(x)の期待値と関係する統計量
x	平均
$(x-\mu_x)^2$	分散
$(x-\mu_x)^3$	歪度
$(x-\mu_x)^4$	尖度

期待値と標本平均(平均)の違い

- xの期待値と標本平均(平均)は異なる。
- xが離散値の場合、無限回試行したときの標本平均がg(x) = xの期待値と考えて良い。発生回数をnとした場合n/NのN無限大の極限が確率となるため。
- 著者がxの期待値のつもりで平均を使っているのか、標本平均のつもりで平均を使っているかをよく注意して本や論文を読む必要がある。

分散

- 確率分布のバラつき具合を表す。

$$V(X) = \sum_{x} (x-\mu)^2 f(x)$$ 離散

$$V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$$ 連続

$$\begin{aligned} V(X) &= E((X-\mu)^2) \\ &= E(X^2) - E(X)^2 \end{aligned}$$

問題

- $V(X) = E(X^2) - E(X)^2$ を導出せよ。

大数の法則

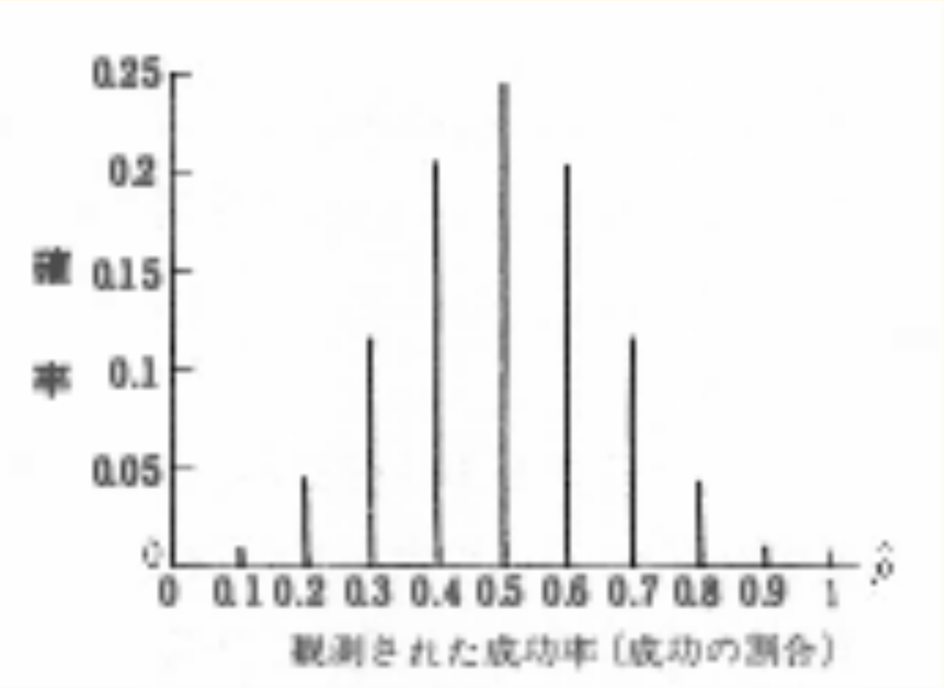
- 試行回数を増やせば増やすほど真の確率に近づく。

この講義では詳しい内容は省略する。

コイン投げの例

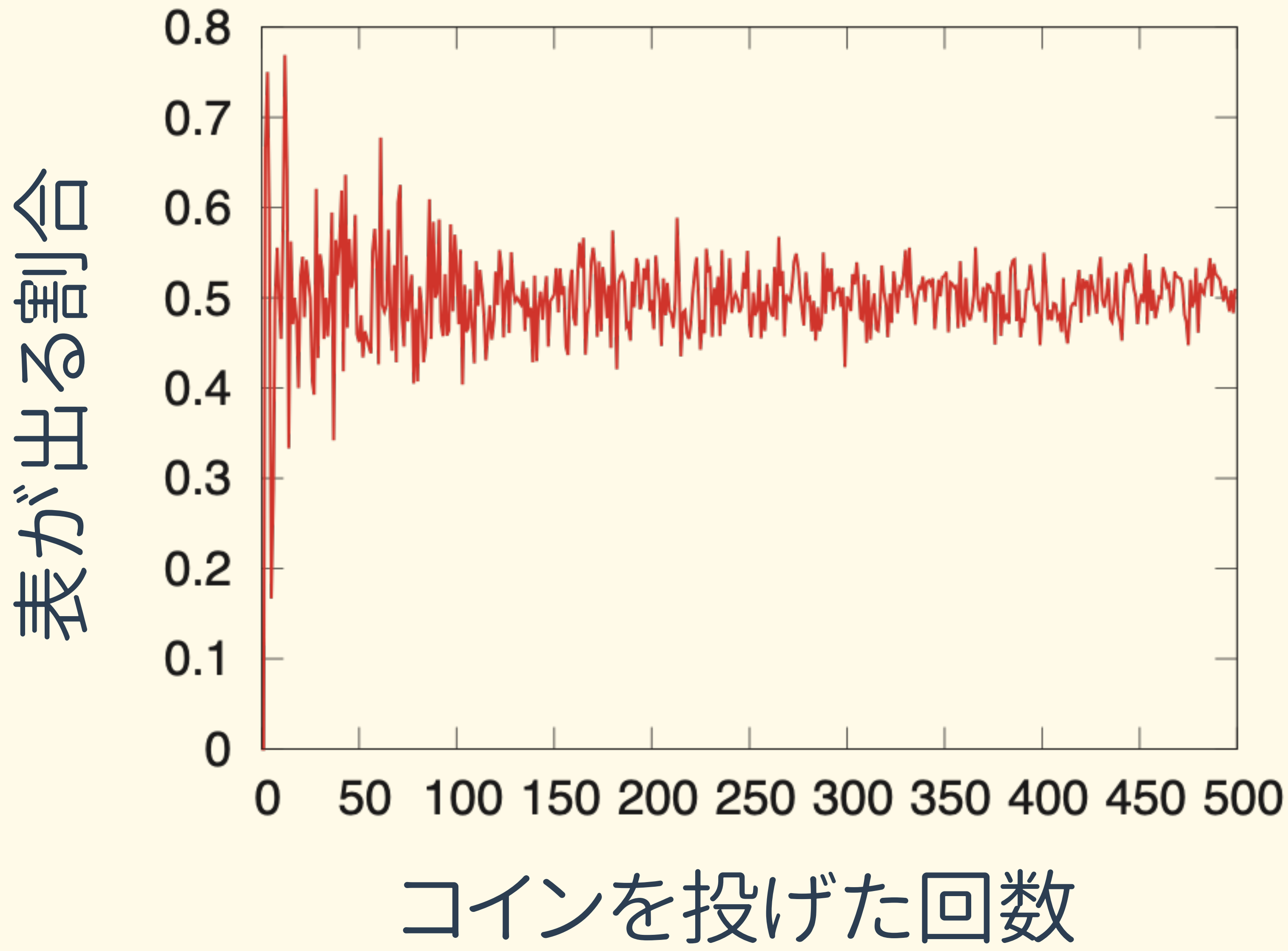
- 理想的なコインを投げることを考える。
- コイン投げで表が出る確率は0.5と考えられる。
- しかし、実際に投げた場合表が出る割合が0.5となるとは限らない。

10回コインを投げた時の表が出る割合の分布



(松原ら,1991)

回数を増やすとどうなるのか



コインを投げる回数を増やすと0.5に近づく。

独立性

$P(X|Y) = P(X)$ の場合、XとYは独立である。このとき、YによらずXが決まる。

XとYが独立な場合、

$$P(X,Y) = P(X \mid Y)P(Y) = P(X)P(Y)$$

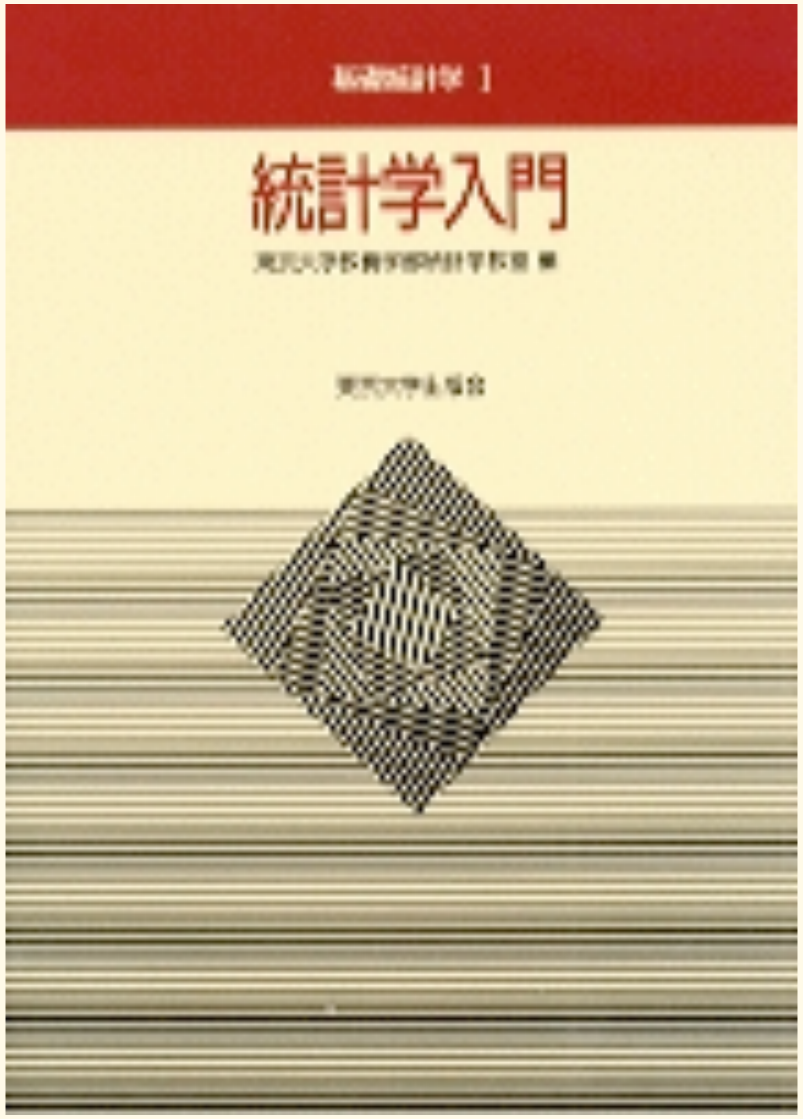
が成り立つ。

問題

- XとYが独立な場合、次の式が成り立つことを示せ。

$$E(X+Y) = E(X) + E(Y)$$
$$E(XY) = E(X)E(Y)$$
$$V(X+Y) = V(X) + V(Y)$$

参考文献



C. M. Bishop
パターン認識と機械学習

松原仁ら
統計学入門