

確率と統計(0)

「確率変数と確率分布(第5章)」

- 担当教員: 杉山 将 (計算工学専攻)
- 居室: W8E-505
- 電子メール: sugi@cs.titech.ac.jp
- 授業のウェブサイト:
<http://sugiyama-www.cs.titech.ac.jp/~sugi/>

確率変数と確率分布

- **確率変数**(random variable): とる値に対して確率が与えられている変数
- **実現値**: 確率変数が実際にとる値
- **確率分布**(probability distribution): 確率変数の実現値と確率の関係を関数として表現したもの
- 確率変数は大文字で, 実現値は小文字で表わすことが多い

離散型の確率変数と確率関数 32

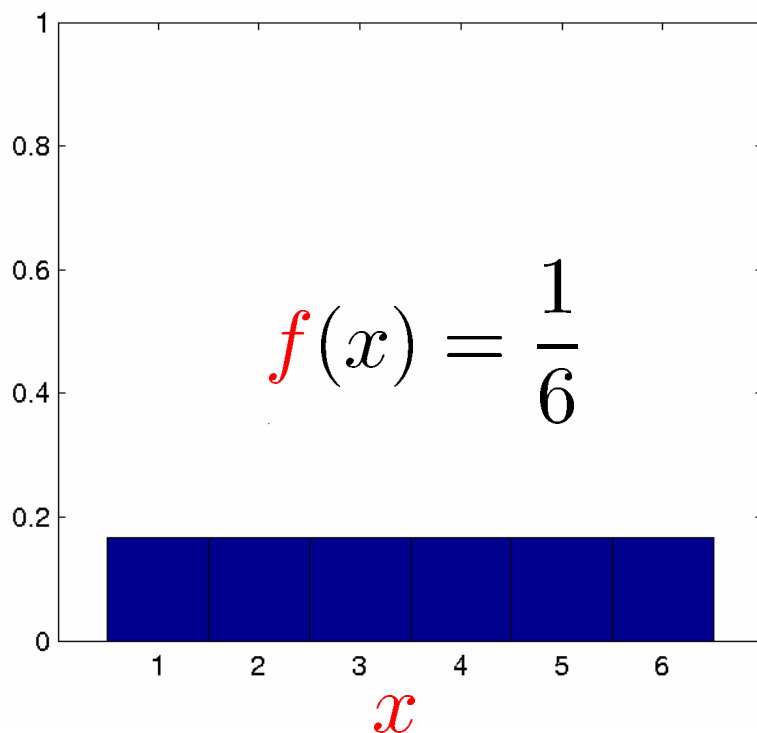
- 離散型(discrete type)確率変数: 可算集合の中の値をとる確率変数
- 離散型の確率変数の確率分布: 確率変数がそれぞれの値をとる確率

$$P(X = x) = f(x)$$

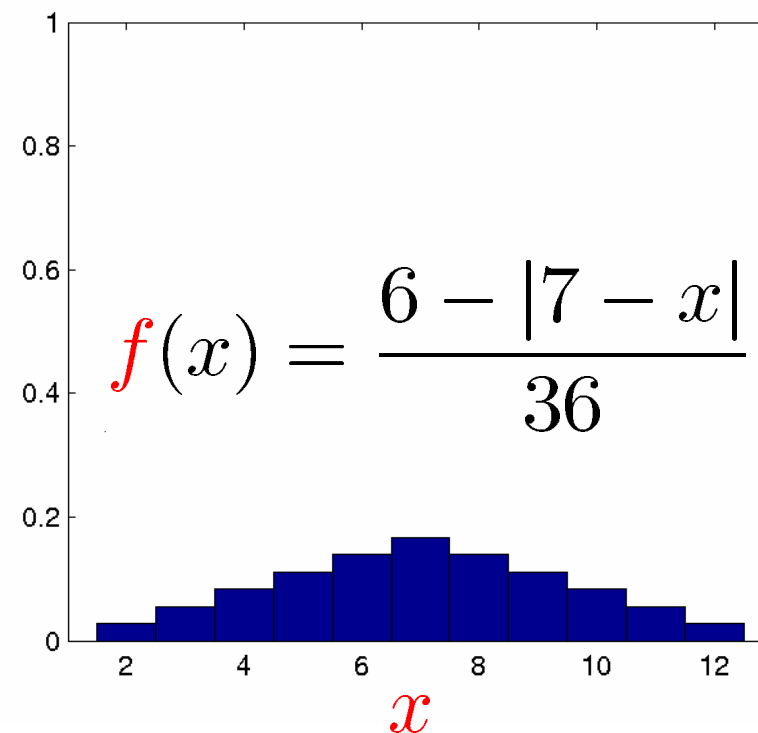
$f(x)$: 確率関数(probability function)

$$f(x) \geq 0, \quad \sum_{x \in \Omega} f(x) = 1$$

離散型の確率分布の例



さいころの出る目



二つのさいころの
出る目の和

主な離散型の確率分布

- 一様分布
- 超幾何分布
- 二項分布
- ポアソン分布
- 負の二項分布
- 幾何分布

連続型の確率変数と確率密度関数³⁵

- 連続型(continuous type)確率変数: 連続値をとる確率変数
- 連続型の確率変数の確率分布: 確率変数が a 以上 b 以下の値をとる確率

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

$f(x)$: 確率密度関数(probability density function)

$$f(x) \geq 0, \int f(x) = 1$$

- 連続型の確率変数がある値 a をとる確率はゼロ

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

累積分布関数

- 連続型の確率変数が x 以下の値をとる確率

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(u) du$$

$F(x)$: 累積分布関数(cumulative distribution function)

$$F'(x) = f(x)$$

- 広義単調増加:

$$x_1 < x_2 \implies F(x_1) \leq F(x_2)$$

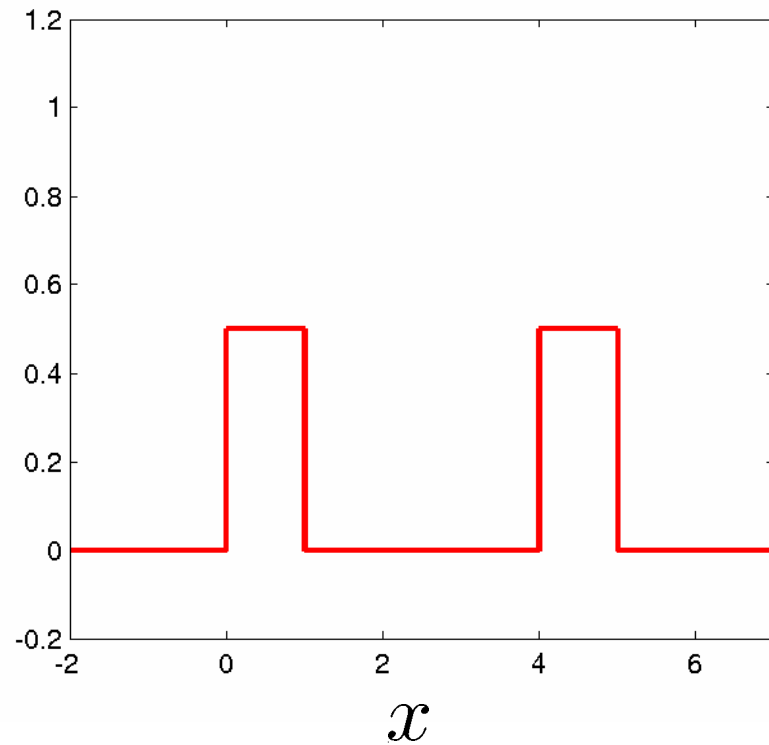
- 範囲: $x \rightarrow -\infty \implies F(x) \rightarrow 0$

$$x \rightarrow \infty \implies F(x) \rightarrow 1$$

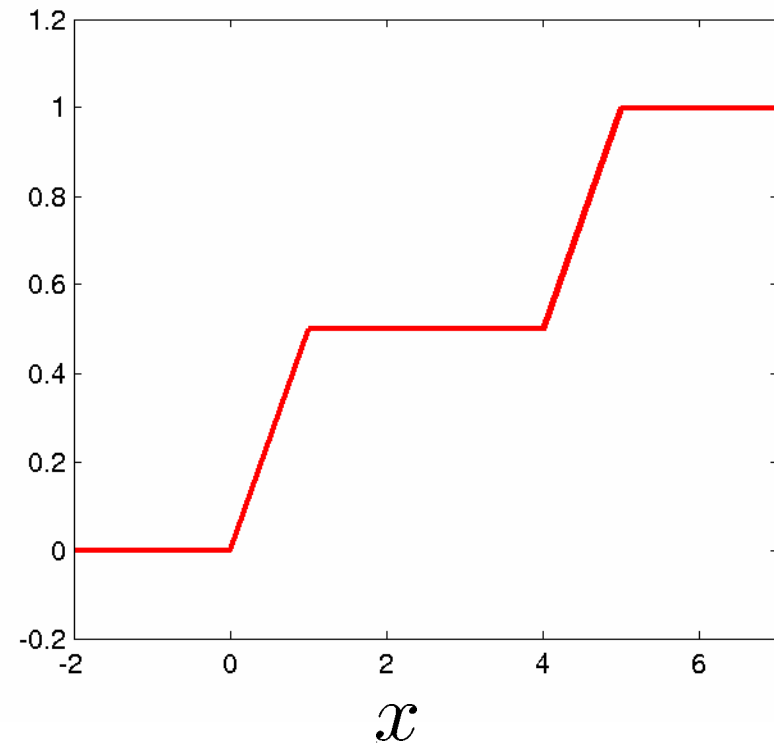
- 右連続:

$$\epsilon \rightarrow +0 \implies F(x + \epsilon) \rightarrow F(x) \text{ for any } x$$

確率密度関数と累積分布関数の例³⁷



確率密度関数



累積分布関数

主な連続型の確率分布

- 一様分布
- 正規分布
- 指数分布
- カイ二乗分布
- ガンマ分布
- ベータ分布
- コーシー分布
- t 分布

確率変数の性質を表わす指標 39

- **期待値(expectation):** 確率変数の値の平均 (正確には確率による重み付きの平均)
- 確率変数 X の期待値を $E(X)$ で表す

- 離散型:

$$E(X) = \sum_x x f(x)$$

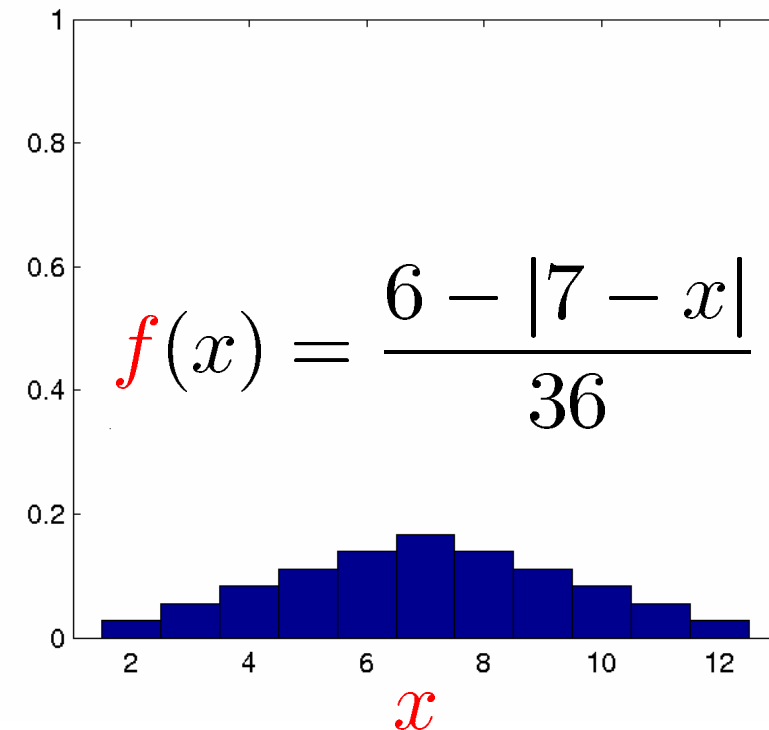
- 連続型:

$$E(X) = \int x f(x) dx$$

- 以後, 説明を簡単にするため, 連続型の確率変数を主に扱うことにする. 離散型を考える場合は, 積分を和に変更すればよい.

例

- 二つのさいころの目の和の期待値を求めよ



$$E(X) = \sum_{x=2}^{12} x f(x) = 7$$

二つのさいころの
出る目の和

期待値作用素の表記について ⁴¹

- $E(\cdot)$ は確率変数 X に関する期待値を表す . 即ち

$$E(\cdot) = \int \cdot f(x) dx$$

- 正確には , $E(\cdot)$ を $E_X(\cdot)$ と表記すべきであるが , 簡単のため省略している .

期待値演算の性質

- 定数は期待値をとっても値は変わらない

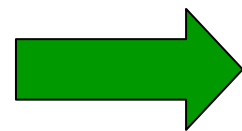
$$E(c) = c$$

- 定数を足した期待値は, 期待値に定数を足したものと等しい

$$E(X + c) = E(X) + c$$

- 定数倍の期待値は, 期待値の定数倍と等しい

$$E(cX) = cE(X)$$



期待値演算は線形

証明

$$\blacksquare E(c) = \int cf(x)dx = c \int f(x)dx = c$$

$$\begin{aligned}\blacksquare E(X + c) &= \int (x + c)f(x)dx \\ &= \int xf(x)dx + c \int f(x)dx \\ &= E(X) + c\end{aligned}$$

$$\begin{aligned}\blacksquare E(cX) &= \int cx f(x)dx \\ &= c \int xf(x)dx = cE(X)\end{aligned}$$

その他のよく用いる指標

■ 中央値(median):

$$P(X \leq x) = \frac{1}{2} \text{ となる } x$$

■ 最頻値(mode):

$$f(x) \text{ を最大にする } x$$

宿題

■ $[a, b]$ 上に定義された確率密度関数 $f(x)$ を考える

1. 次の二乗誤差 $J_1(y)$ を最小にする y を y_1 で表す:

$$y_1 = \operatorname{argmin}_y J_1(y) \quad J_1(y) = \int_a^b (x - y)^2 f(x) dx$$

このとき, y_1 は X の期待値 (つまり, $y_1 = E[X]$) であることを示せ.

2. 次の絶対誤差 $J_2(y)$ を最小にする y を y_2 で表す:

$$y_2 = \operatorname{argmin}_y J_2(y) \quad J_2(y) = \int_a^b |x - y| f(x) dx$$

このとき, y_2 は X の中央値 (つまり, $F(y_2) = 1/2$) であることを示せ