

指定された分布に従 う乱数

モデリングとシミュレーション

2017年度

1

一様乱数から指定された分布に従う乱数へ

- シミュレーションにおいて、乱数の分布が指定される場合がある
 - 指数分布に従う乱数
 - でたらめなに事象が発生する
 - 正規分布に従う乱数
 - ある平均値のまわりのでたらめな揺らぎ
- 変換法
- 棄却法

確率分布

■ 実数に対する分布

■ 実数は連続であることに注意

■ 「ある値になる確率」はあり得ない

■ 確率変数 $X \in [a, b)$ に対する確率分布

■ $a \leq X < x$ という「範囲」に対して確率を定義

$$F(x) = P(a \leq X < x)$$

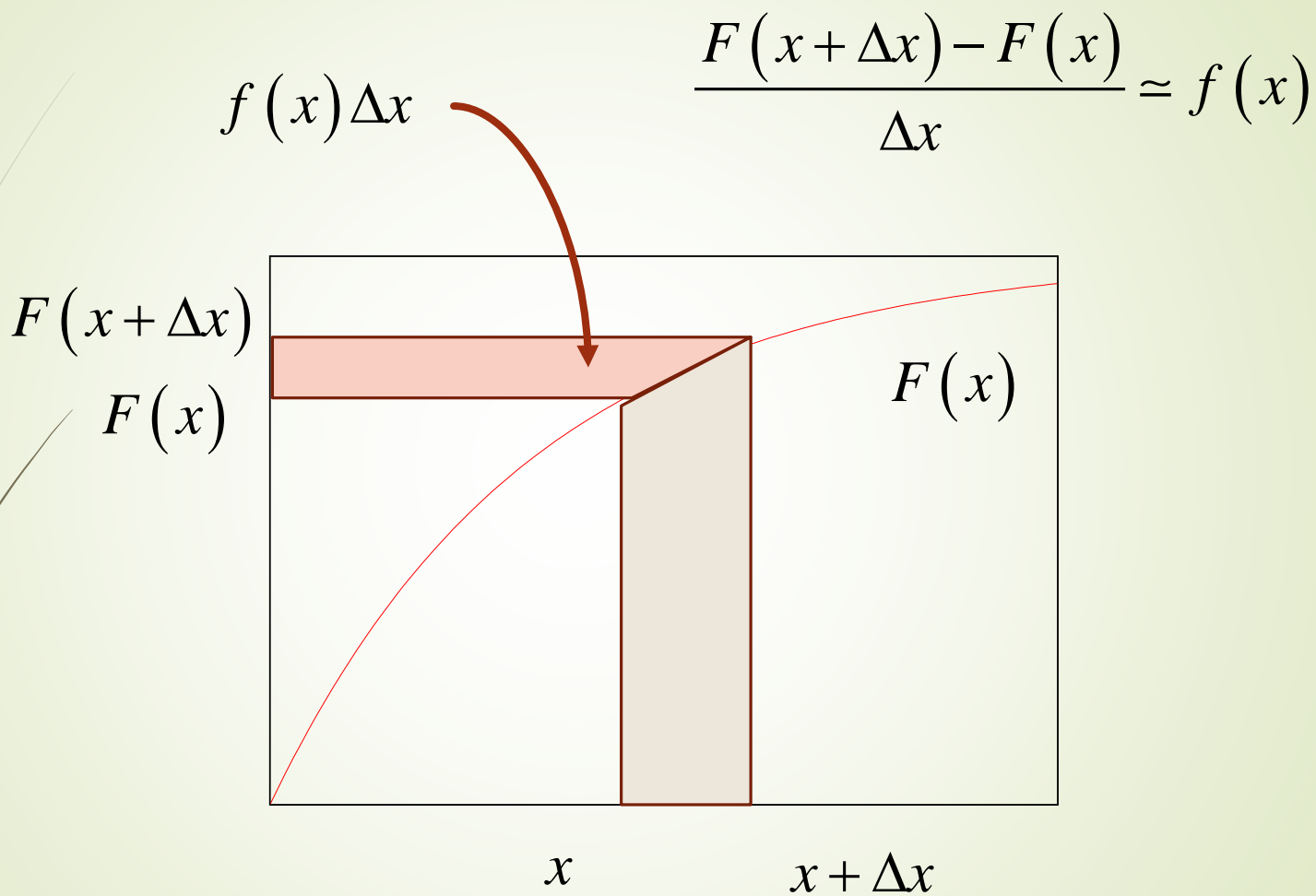
確率密度

- ➡ 微小区間($X \in [x, x + \Delta x)$)に入る確率

$$f(x) \Delta x = \frac{d}{dx} F(x) \Delta x$$

$$F(x) = \int_a^x f(y) dy$$

- ➡ $f(x)$ を確率密度と呼ぶ
 - ➡ 確率そのものでないことに注意



変換法

➡ $r \in [0,1)$ の一様乱数を生成

➡ x へ変換 $x = F^{-1}(r)$

➡ $[x, x + \Delta x)$ に入る確率

➡ $[F(x), F(x + \Delta x))$ の長さ

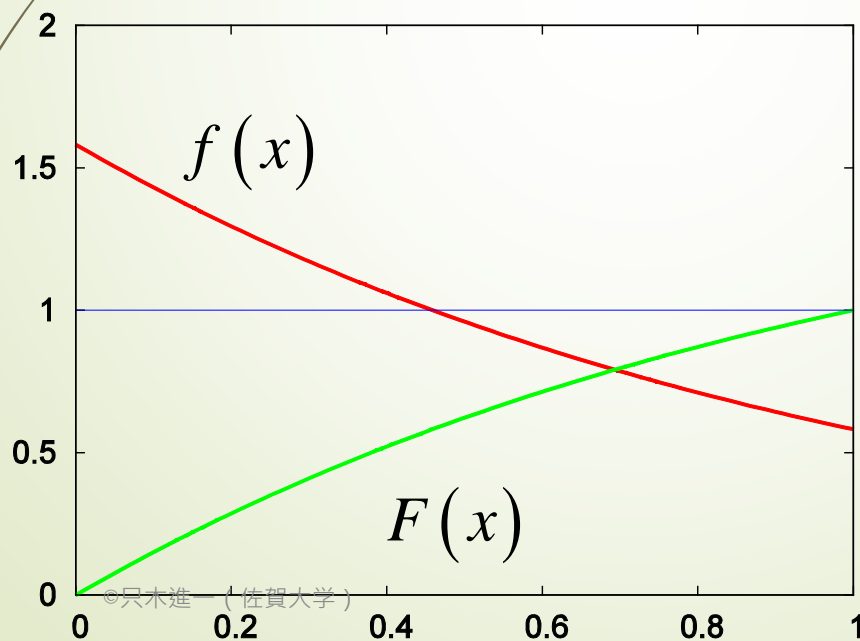
$$F(x + \Delta x) - F(x) = f(x)\Delta x + O(\Delta x^2)$$

➡ つまり、 $f(x)$ に従う乱数

例：指数分布

$$f(x) = Ae^{-x}, \quad 0 \leq x < 1, \quad A = \frac{e}{e-1}$$

$$F(x) = \int_0^x f(y) dy = A(1 - e^{-x})$$



$$F^{-1}(r) = -\ln(1 - r/A)$$

変換法の困難さ

- ▶ 変換法が使えるためには
 - ▶ 確率分布 $F(x)$ の表式を得られる
 - ▶ 確率密度の不定積分の表式が得られる
 - ▶ 確率分布の逆関数 $F^{-1}(x)$ の表式を得られる
- ▶ **これらは、かなり特殊な場合**

正規分布：特殊な例

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

- ▶ 不定積分は誤差関数と呼ばれ、表式は知られていない
 - ▶ 数表があるのみ
 - ▶ 単純な変換法は使えない
- ▶ 標準的な確率分布
 - ▶ よく使われる

➡ 二次元正規分布を考える

$$f(x, y) dx dy = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) dx dy$$

➡ 極座標へ変換 $\begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases}$

$$f(r) r dr d\theta = \frac{1}{2\pi\sigma^2} r \exp\left(-\frac{r^2}{2\sigma^2}\right) dr d\theta$$

➡ θ 方向には一様であること

➡ θ 方向に積分 $f(r) r dr = \frac{1}{\sigma^2} r \exp\left(-\frac{r^2}{2\sigma^2}\right) dr$

➡ 動径方向の確率分布は積分可能

$$\rho = \int_0^r f(r') r' dr' = \int_0^r \frac{1}{\sigma^2} r' \exp\left(-\frac{r'^2}{2\sigma^2}\right) dr' = 1 - \exp\left(-\frac{r^2}{2\sigma^2}\right)$$

$$r = \sqrt{-2\sigma^2 \ln(1 - \rho)}$$

■ 二つの乱数 (u, v) , $0 \leq u, v, < 1$

$$r = \sqrt{-2\sigma^2 \ln(1-u)}, \quad \theta = 2\pi v$$

$$x = r \cos \theta$$

棄却法 (rejection method)

- ➡ 変換法で生成できない分布に対応
- ➡ 効率は悪いが、応用範囲が広い

1. $f(z)$ の変域 : $z \in [a, b)$ 、値域 : $0 \leq f(z) < m$
2. 乱数生成 $(x, y) \in [0, 1)$
3. $z = (b - a)x + a$
4. $y < f(z)/m$ ならば z を採用
5. それ以外ならば棄却
6. 次の乱数を生成 : 2へ戻る

ヒストグラムと規格化

- ➡ ヒストグラムは区間毎の頻度、または相対頻度
- ➡ 確率密度関数と比較するには
 - ➡ 確率密度関数は規格化されている：積分すると1になる
- ➡ ヒストグラムを規格化するには？

■ ヒストグラムの規格化

■ bin の幅 w

■ bin の頻度 h_i

■ $s = \sum h_i$

■ $\frac{h_i}{s}$: 相対頻度

■ $g_i = s^{-1}w^{-1}h_i$

$$\sum wg_i = \frac{1}{sw} \sum wh_i = 1$$

