確率と統計(0)
「確率不等式と計算機による乱数の発生」

- 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

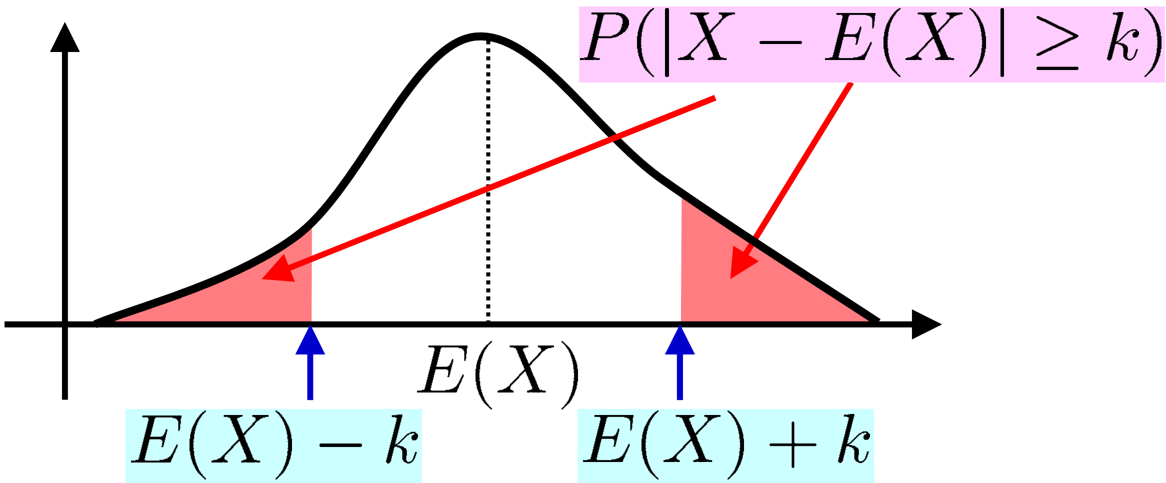
- 授業のウェブサイト：
  http://sugiyama-www.cs.titech.ac.jp/~sugi/

チェビシェフの不等式

- **チェビシェフの不等式**(Chebyshev's inequality)

$$P(|X - E(X)| \geq k) \leq \frac{V(X)}{k^2}$$

- **確率分布の具体的な形は分からないが期待値と分散が分かるとき，チェビシェフの不等式によって確率の上限が計算できる**
- **チェビシェフの不等式はいかなる確率変数に対しても成立する！**



チェビシェフの不等式の使用例

- ある試験の点数の平均が６０点，分散が３０であった．６０±１０点の範囲に含まれる人は全体の何パーセント以上か？

- 宿題！

チェビシェフの不等式 (証明)

$$P(|X - E(X)| \geq k) \leq \frac{V(X)}{k^2}$$

$$\begin{aligned}
V(X) &= \int_{-\infty}^{\infty} (X - E(X))^2 f(X) dX \\
&\geq \int_I (X - E(X))^2 f(X) dX \\
&\geq k^2 \int_I f(X) dX \\
&= k^2 \; P(|X - E(X)| \geq k)
\end{aligned}$$

$$I = \{X : |X - E(X)| \geq k\}$$

その他の便利な不等式

- **マルコフの不等式**(Markov's inequality)：

$$P(X \geq a) \leq \frac{1}{a}E[X] \quad \text{for any } a > 0$$

- **ジェンセンの不等式**(Jensen's inequality)：

$$E[h(X)] \geq h(E[X])$$

$h(x)$ ：**凸関数**

- **ヘルダーの不等式**(Hölder's inequality)：

$$E[|XY|] \leq \left(E[|X|^p]\right)^{1/p}\left(E[|Y|^q]\right)^{1/q}$$

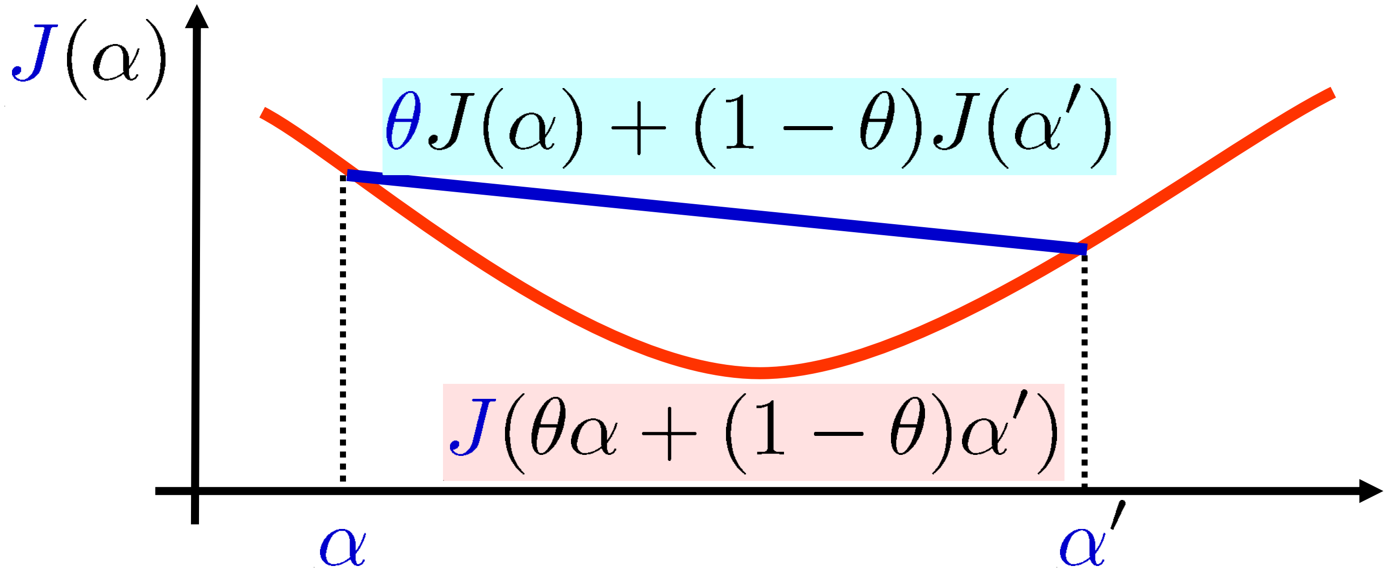
$$\text{for any } p, q > 0 \text{ such that } 1/p + 1/q = 1$$

特に $p = q = 2$ **の場合を シュワルツの不等式**(Schwarz's inequality)**とよぶ**

凸関数

- 任意の α, α' と任意の $\theta \in (0,1)$ に対して以下の式が成り立つとき，$J(\alpha)$ は**凸関数** (convex function)であるという

$$J(\theta\alpha + (1-\theta)\alpha') \leq \theta J(\alpha) + (1-\theta)J(\alpha')$$



計算機による乱数の生成

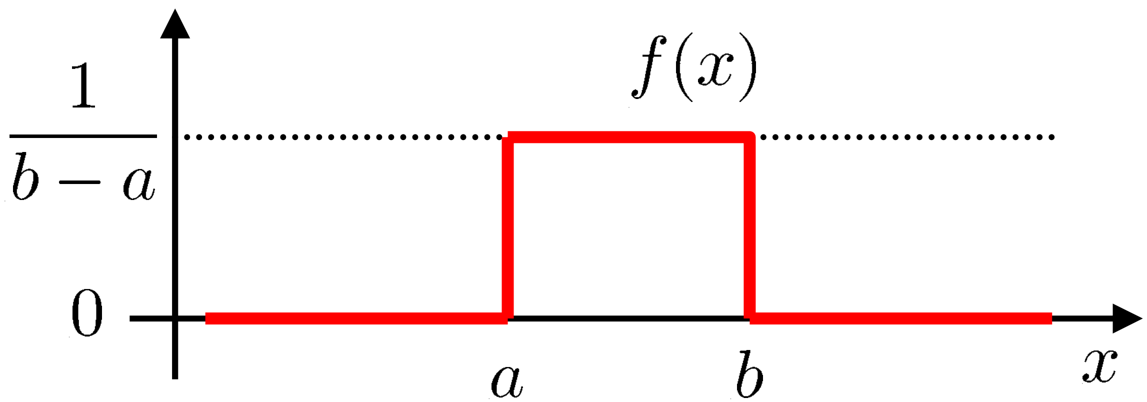
- 乱数を計算機内で生成するのは非常に難しい！
- 乱数を生成するための専用のハードウェアもある．
  例：電子素子の熱雑音などの物理現象を利用
- 一般には，擬似乱数を用いることが多い．
  例：C言語のrand関数は一様擬似乱数を生成する
- 一様擬似乱数や正規擬似乱数を生成する関数は，大抵の計算機言語で用意されている．
- それ以外の任意の分布に従う乱数も作りたい！

一様分布

- **連続一様分布**(uniform distribution of continuous type)：**確率密度関数が一様**

$$f(x) = \begin{cases} \frac{1}{b-a} & (\text{ for } a \leq x \leq b) \\ 0 & (otherwise) \end{cases}$$

- (a, b) **上の一様分布を** $U(a, b)$ **と表す.**



乱数の生成法1

- **逆関数法**(inverse transform sampling)：ある確率密度関数 $f(x)$ に従う確率変数を生成する方法

1. $u \sim U(0,1)$ を発生させる．
2. $v = F^{-1}(u)$ は $f(x)$ に従う．

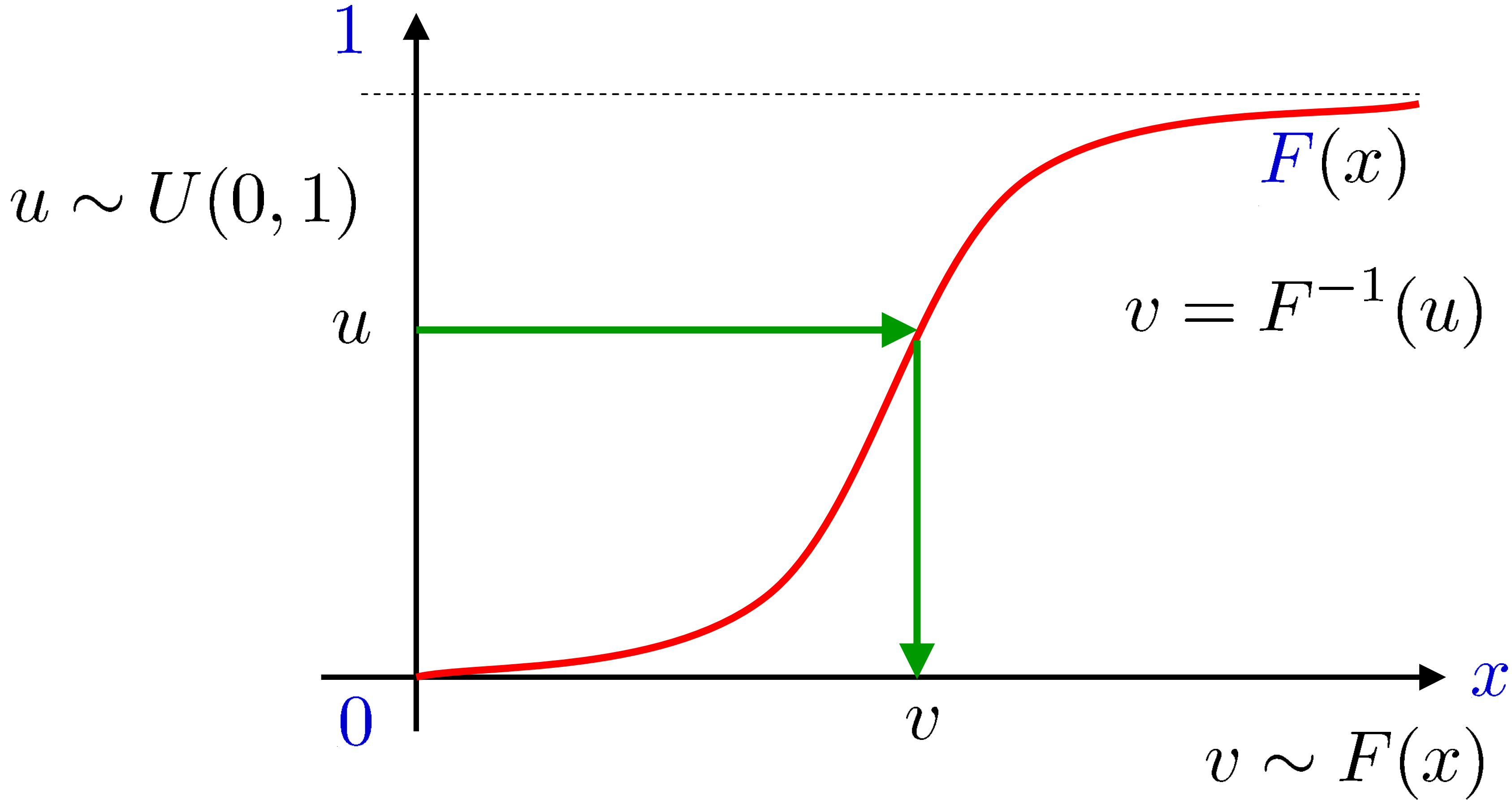
$F^{-1}(u)$ ：$F(x)$ の**逆関数**

$F(x)$ ：$f(x)$ の**累積分布関数**

$$u = F(x) \qquad x = F^{-1}(u)$$

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

逆関数法 (続き)



逆関数法 (証明)

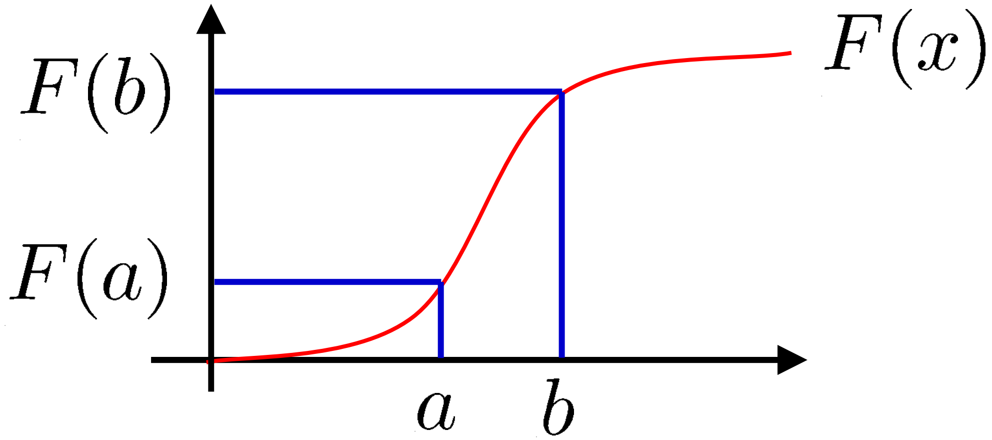
- $P(v \le c) = F(c), \forall c$ を示す.

$$\begin{aligned} P(v \le c) &= P(F^{-1}(u) \le c) \\ &= P(u \le F(c)) \\ &= P(0 \le u \le F(c)) \\ &= F(c) \end{aligned}$$

$v = F^{-1}(u)$

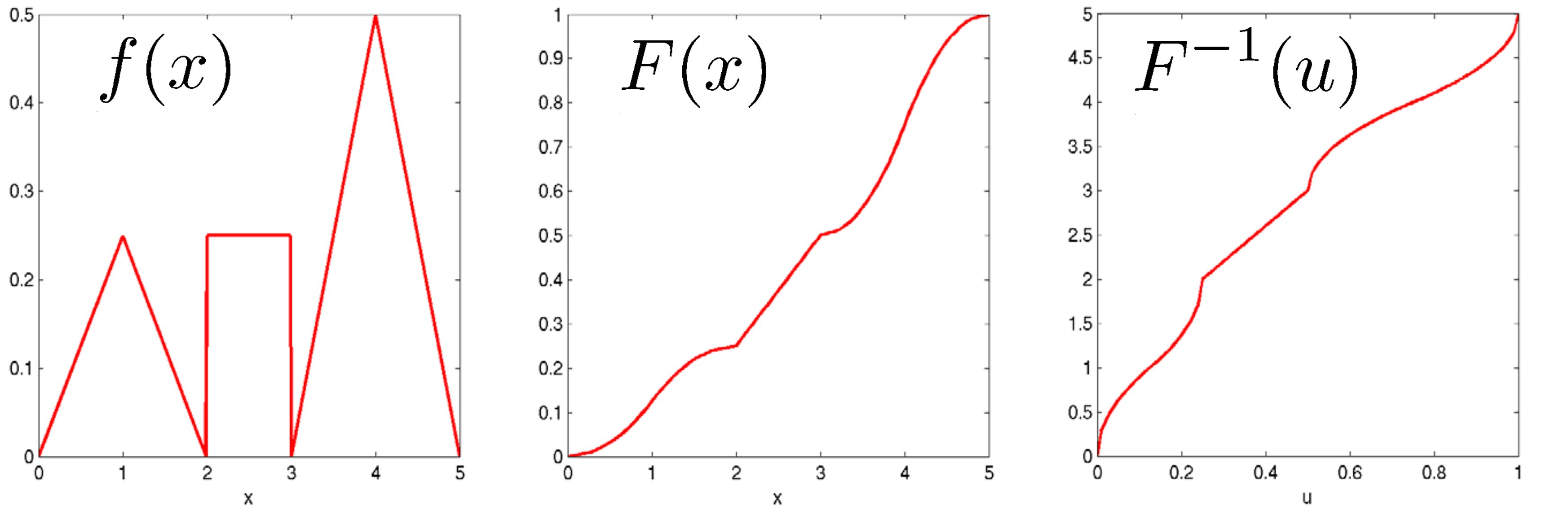
$a \le b \Rightarrow F(a) \le F(b)$

$u \sim U(0,1)$



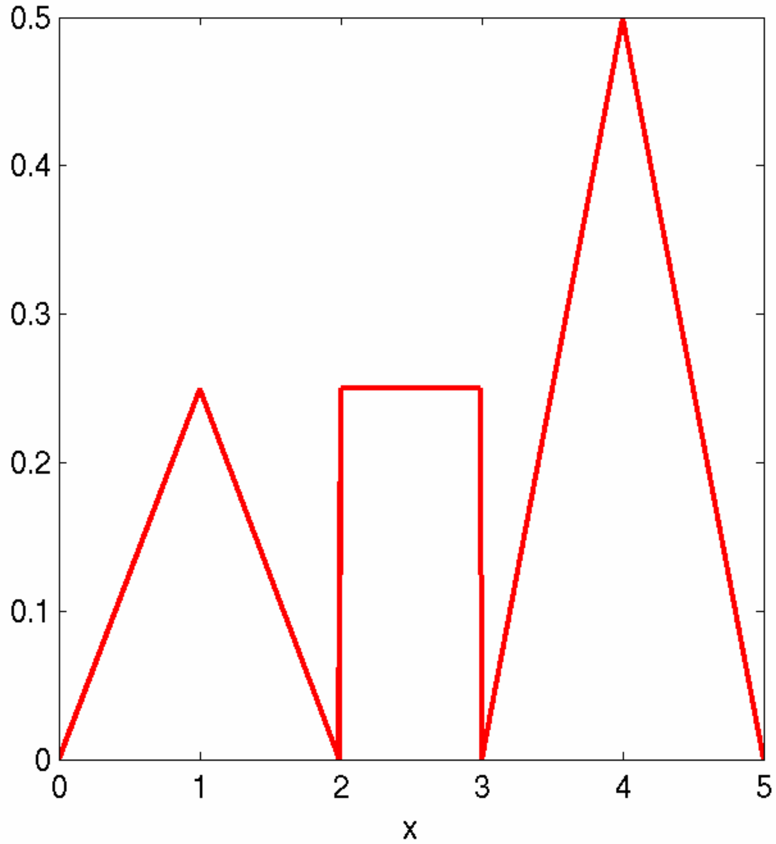
逆関数法による乱数生成の例

$$x = F^{-1}(u)$$

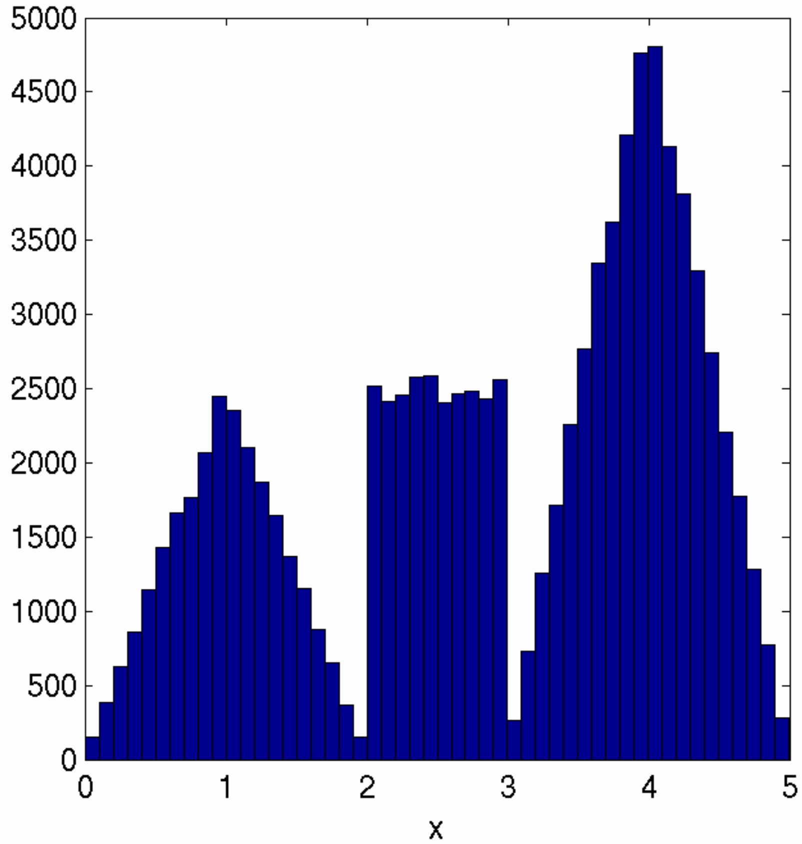


確率密度関数　　累積分布関数　　累積分布関数の逆

逆関数法による乱数生成の例 (続き)



確率密度関数



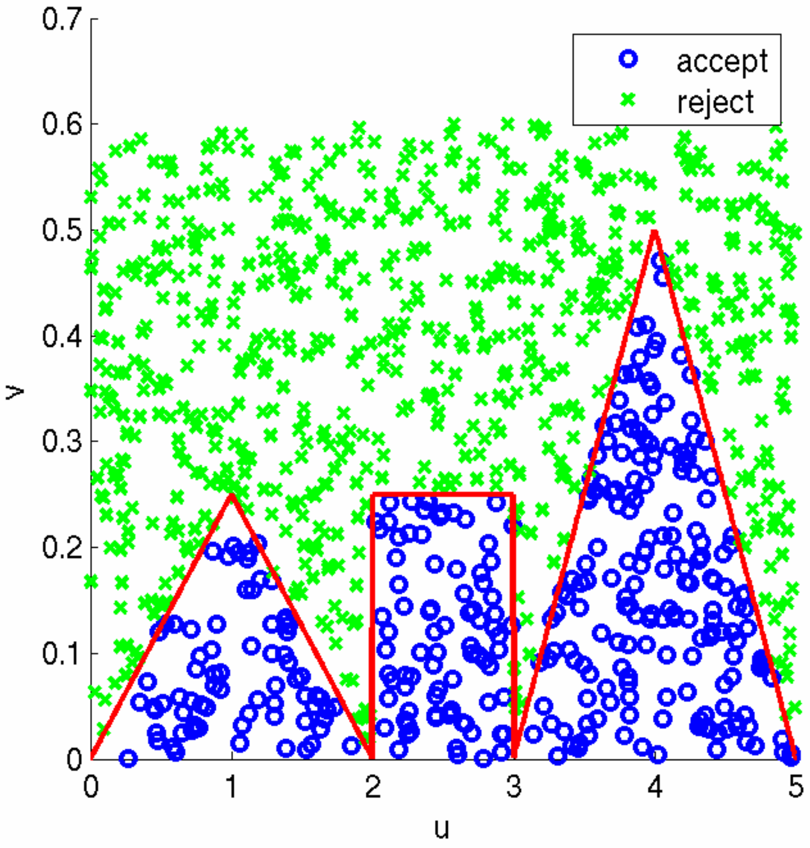
生成した乱数の
ヒストグラム

乱数の生成法2

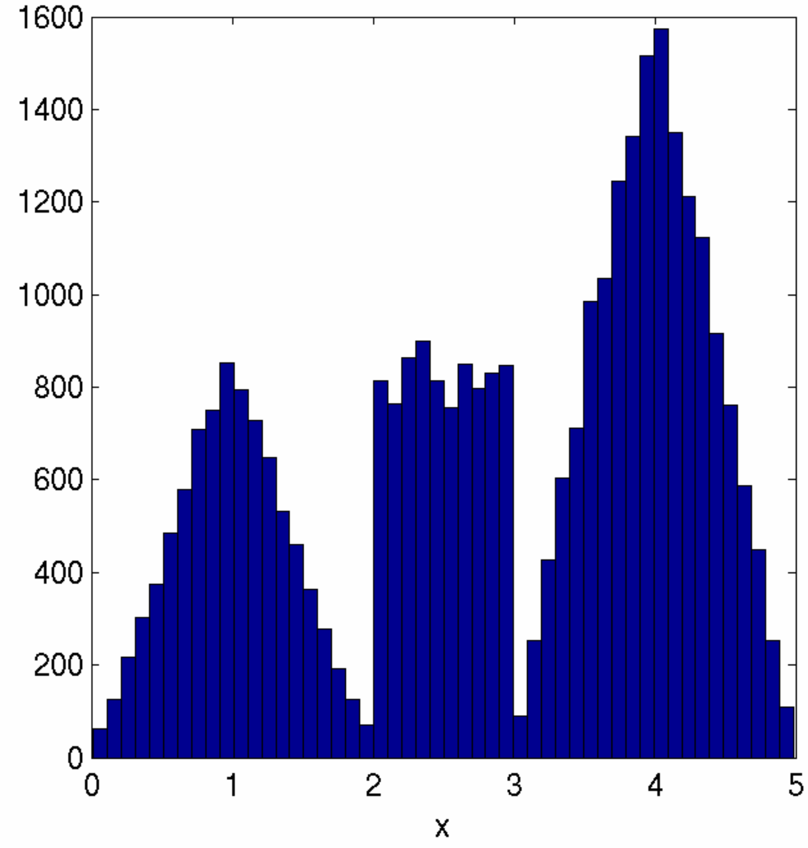
- 棄却法(rejection method)：有限区間 $[a,b]$ 上の確率密度関数 $f(x)$ に従う乱数を生成する

1. $u \sim U(a,b)$ を発生させる．
2. $v \sim U(0, \max_x f(x))$ を発生させる．
3. $v \leq f(u)$ ならば u を採択(accept)し，そうでなければ棄却(reject)し，1．からやり直す．

棄却法による乱数生成の例



確率密度関数



生成した乱数の
ヒストグラム

逆関数法と棄却法の問題点

- **逆関数法**:
  - 逆関数がきれいな形で求まらないことがある.
- **棄却法**:
  - 定義域が有限の乱数しか発生させることができない.
  - 棄却域が大きい場合, たくさんの乱数を発生させるのに時間がかかる.

宿題

1. ある試験の点数の平均が６０点，分散が３０であった．６０±１０点の範囲に含まれる人は全体の何パーセント以上か？
2. 授業で扱っているトピック，授業の進め方，講義内容の難易度，講義資料，宿題などについて意見を述べ，この授業を評価せよ．
   - 分かりにくかったところがある場合は，スライドの何ページのどの部分がどのように分かりにくかったのか具体的に記述すること．
   - 悪いところはどうすれば改善するかも考えよ．
   - 良いところ，このまま続けるべきところも述べよ．

- ２．も採点対象なので，真剣に書くこと．

連絡事項

- 来週の授業は，**情報工学科計算機室**で行う．
- 内容は，今週の授業で習った乱数の発生法に関する**計算機実習**．
- **10時40分**までに計算機室に集合すること．
- 情報工学科計算機室の**アカウントを持っていない学生**はあらかじめ相談に来ること．