

確率と統計(○)

「確率不等式と擬似乱数」

- 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/>

講義計画(シラバス)

164

- 確率と統計の基礎
- 確率変数, 確率分布
- 積率, 積率母関数
- 離散型の確率分布の例
- 連続型の確率分布の例
- 確率不等式, 擬似乱数
- 多次元の確率分布
- 大数の法則, 中心極限定理
- 統計的推定, 仮説検定

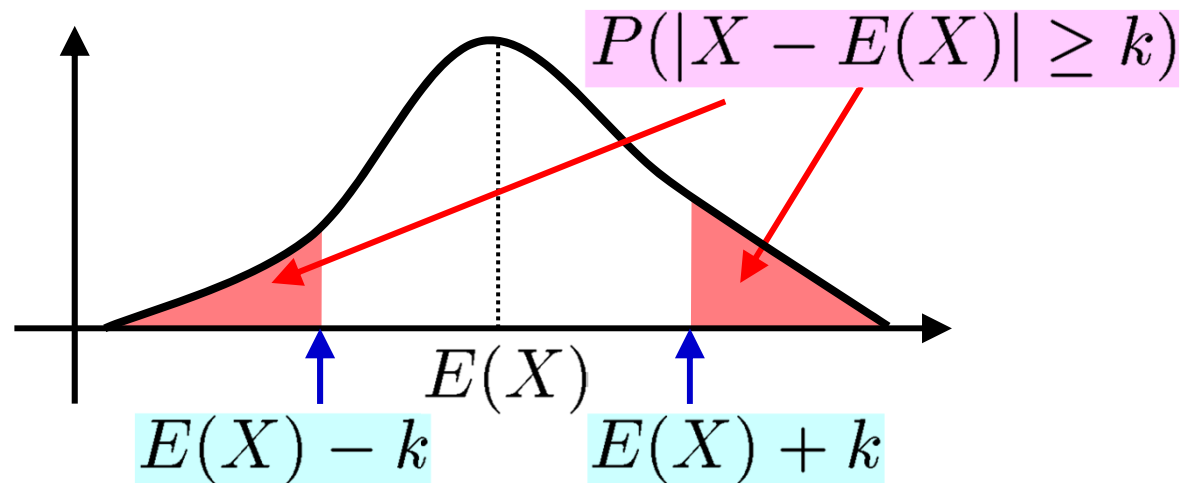
チェビシェフの不等式

165

■ チェビシェフの不等式(Chebyshev's inequality)

$$P(|X - E(X)| \geq k) \leq \frac{V(X)}{k^2} \quad k > 0$$

- 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(証明) 166

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

$$V(X) = \int_{-\infty}^{\infty} (X - E(X))^2 f(X) dX$$

$$\geq \int_I (X - E(X))^2 f(X) dX$$

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

$$\geq k^2 \int_I f(X) dX$$

$$= k^2 P(|X - E(X)| \geq k)$$

チェビシェフの不等式の使用例 167

演習

1. ある試験の点数の平均が60点, 分散が20であった. 65点以上または55点以下の人は全体の何パーセント以下か?
2. ある試験の点数の平均が60点, 分散が30であった. 点数が50点より高く70点より低い人は全体の何パーセント以上か?

その他の便利な不等式

168

- 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):

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

for any $p, q > 0$ such that $1/p + 1/q = 1$

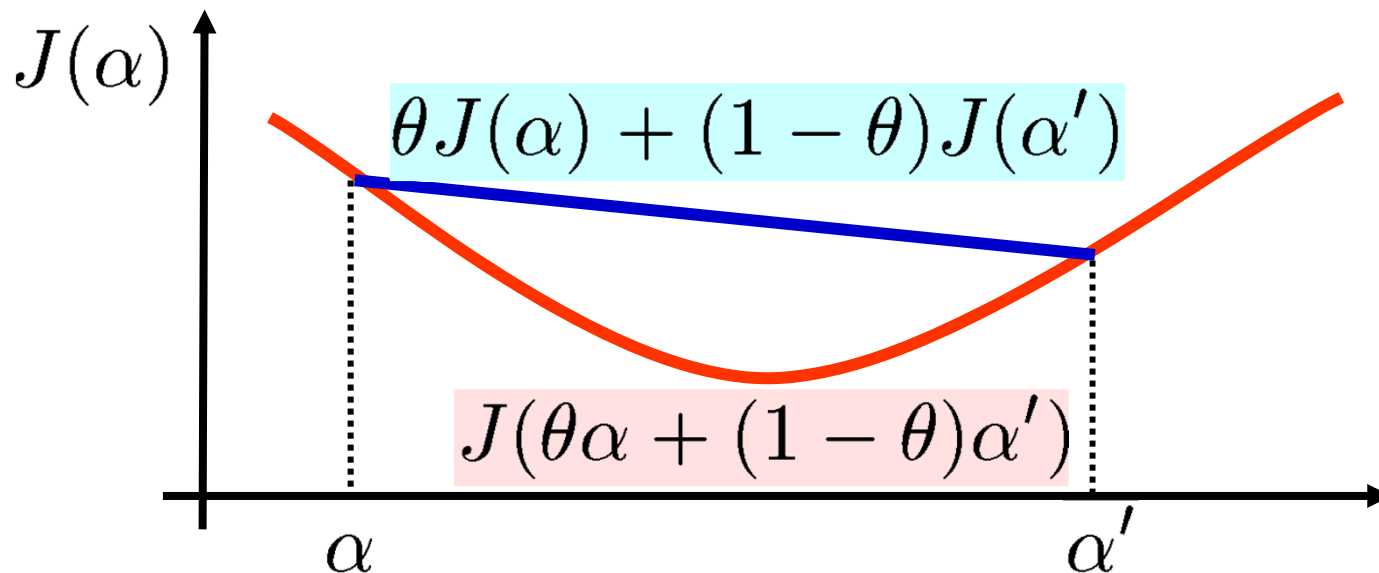
特に $p = q = 2$ の場合をシュワルツの不等式
(Schwarz's inequality)とよぶ

凸関数

169

- 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

170

- 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！

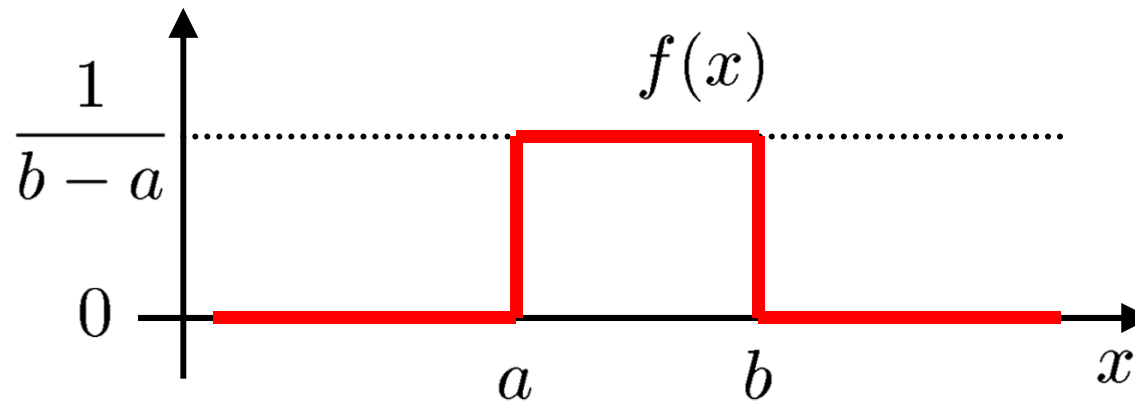
一様分布

171

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

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

- (a, b) 上の一様分布を $U(a, b)$ と表す.



■ 逆関数法(inverse transform sampling):

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

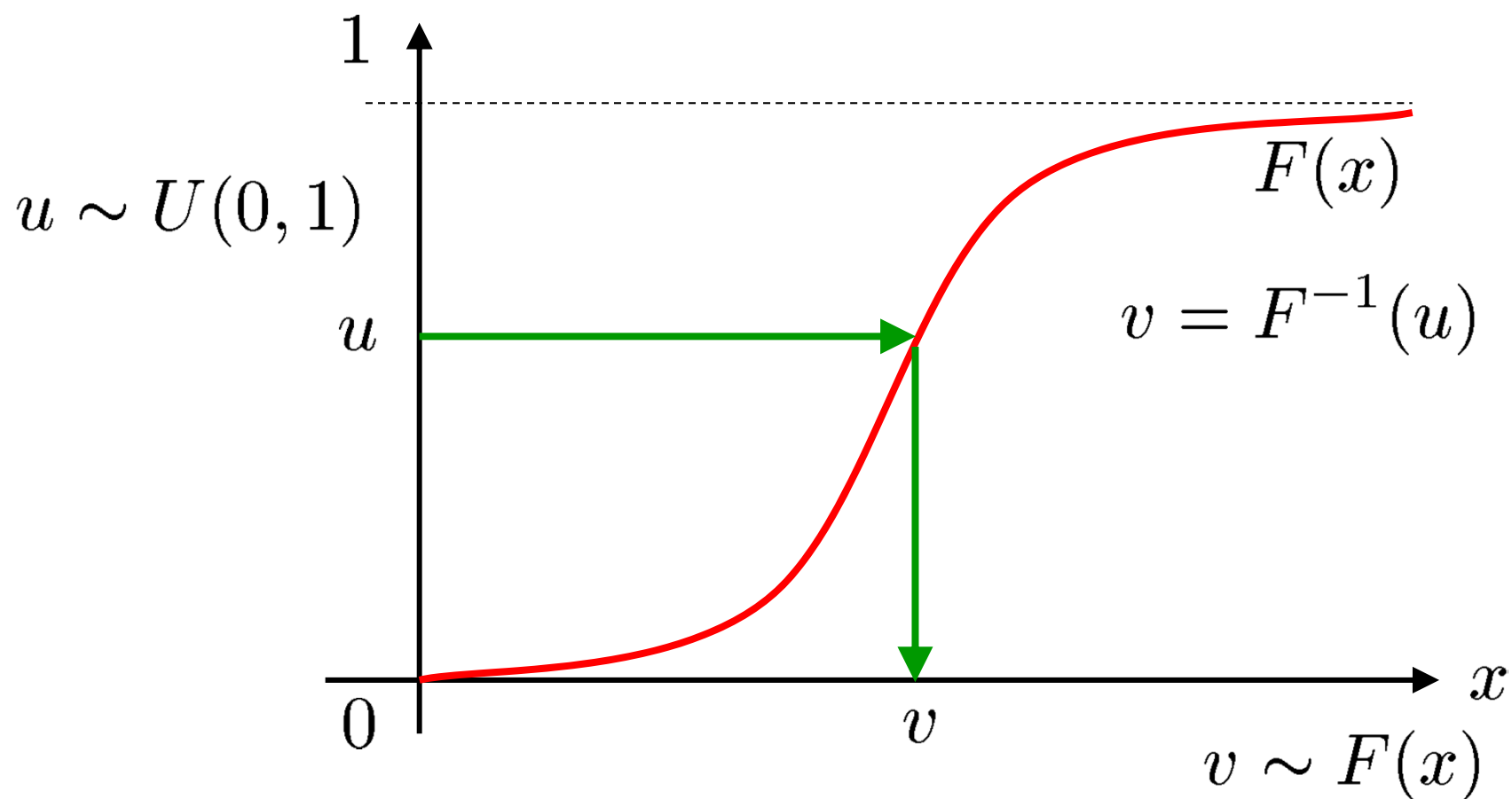
$$u = F(v) \quad v = F^{-1}(u)$$

$F(x) : f(x)$ の累積分布関数

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

逆関数法(続き)

173



逆関数法(証明)

174

■ $\forall c \ P(v \leq c) = F(c)$ を示す.

$$P(v \leq c) = P(F^{-1}(u) \leq c) \quad v = F^{-1}(u)$$

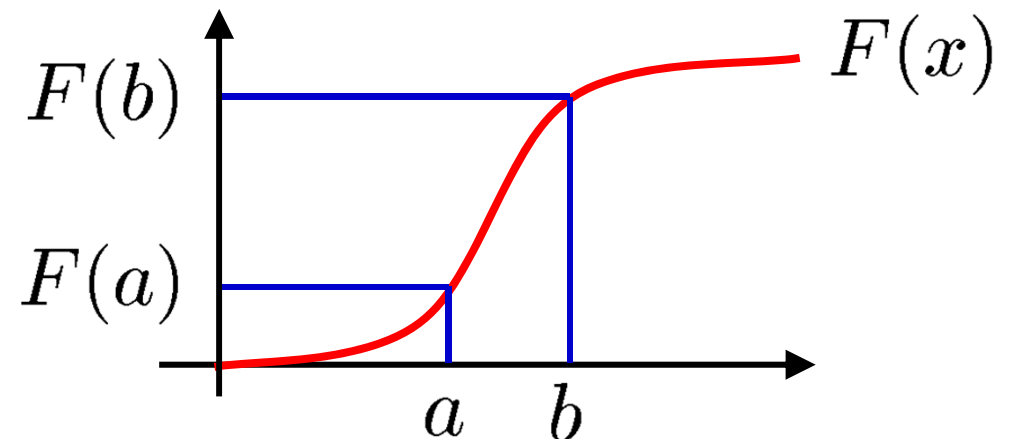
$$= P(u \leq F(c)) \quad a \leq b \Rightarrow F(a) \leq F(b)$$

$$= \int_{-\infty}^{F(c)} g(u) du \quad g(u) = \begin{cases} 1 & (0 \leq u \leq 1) \\ 0 & (\text{otherwise}) \end{cases}$$

(u の確率密度関数)

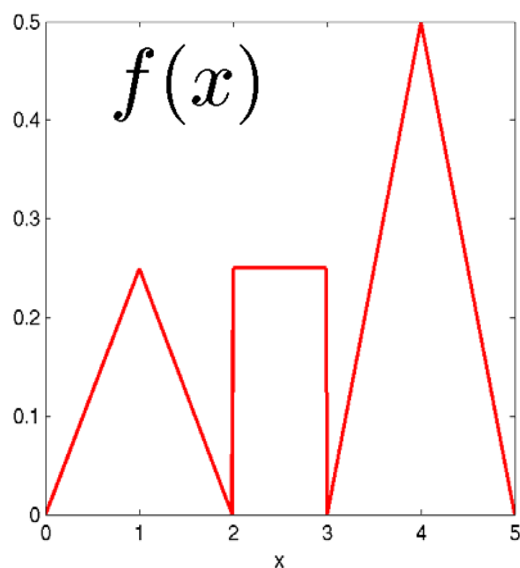
$$= \int_0^{F(c)} du \quad u \sim U(0, 1)$$

$$= F(c)$$

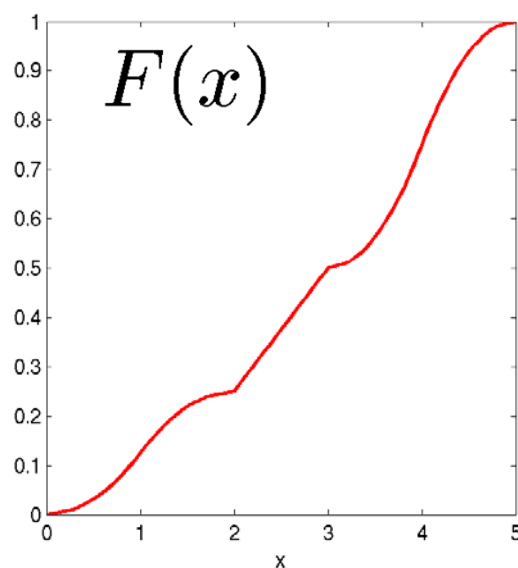


逆関数法による乱数生成の例 175

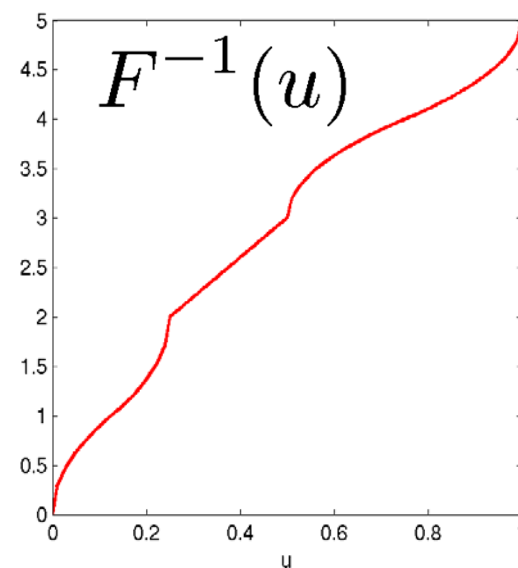
$$x = F^{-1}(u)$$



確率密度関数

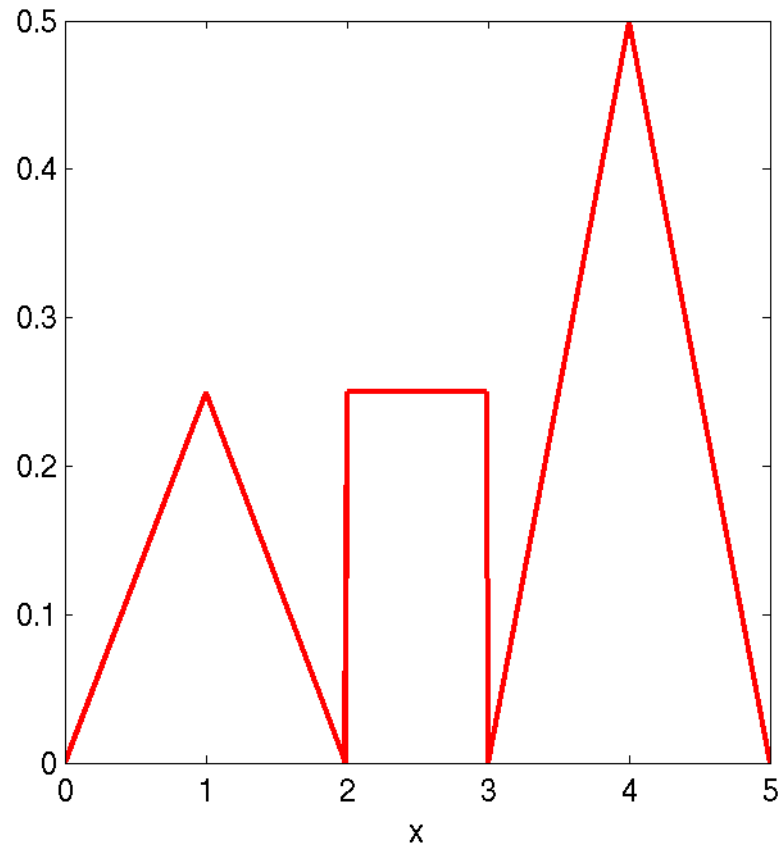


累積分布関数

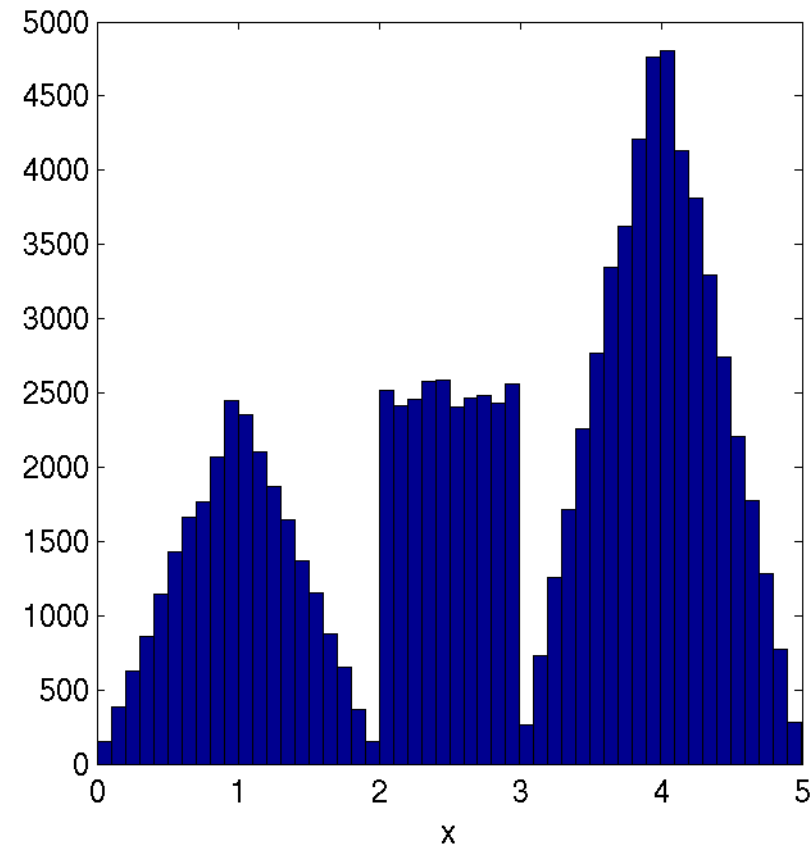


累積分布関数の逆

逆関数法による乱数生成の例(続き)¹⁷⁶



確率密度関数



生成した乱数の
ヒストグラム

■ 棄却法(rejection sampling):

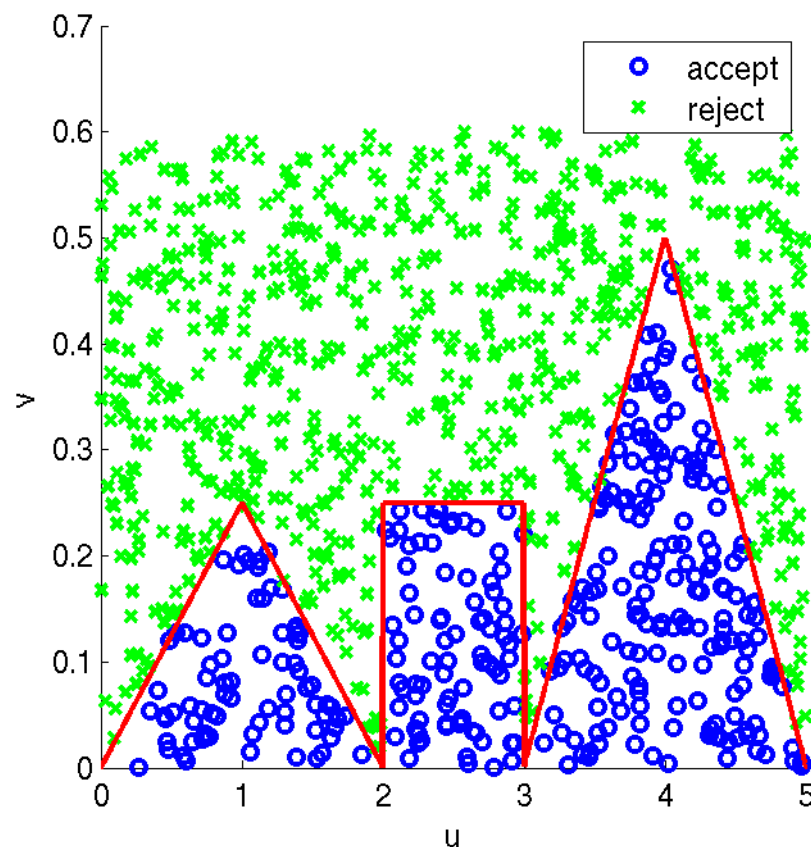
1. $u \sim U(a, b)$ を発生させる.

$(a, b) : f(x)$ の定義域

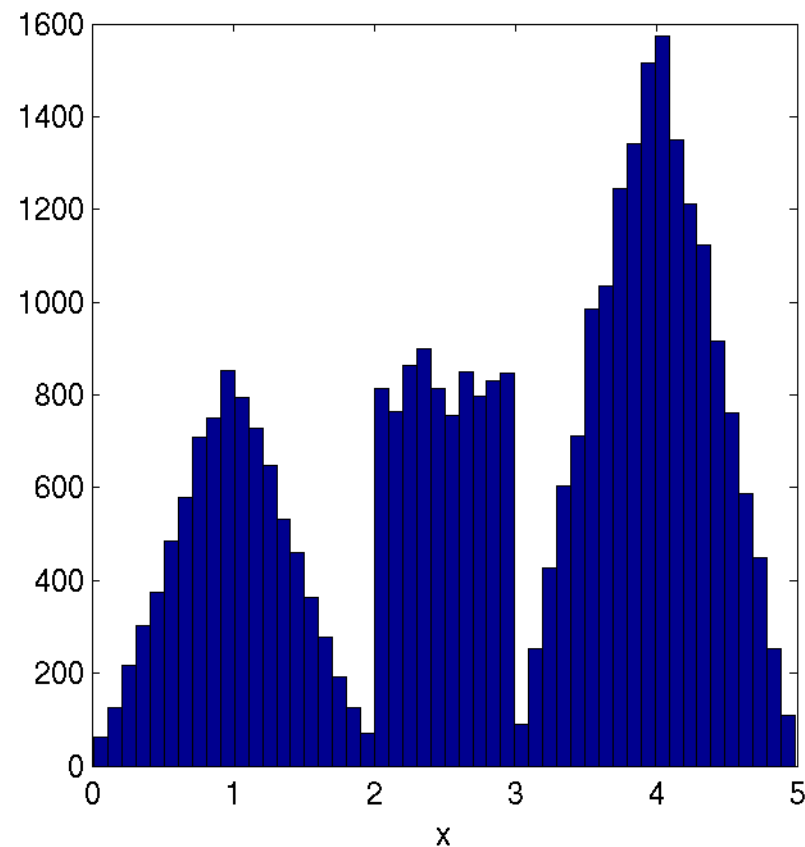
2. $v \sim U(0, \max_x f(x))$ を発生させる.
3. $v \leq f(u)$ ならば u を採択(accept)し,
そうでなければ棄却(reject)する.
4. 1. にもどる.

棄却法による乱数生成の例

178



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 179

■ 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:

- 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.

- 確率不等式
 - チェビシェフの不等式
- 計算機による乱数の生成
 - 逆関数法
 - 棄却法

- 歪んだ6面体のさいころ(出る目は1, 2, 3, 4, 5, 6)がある. この変なさいころの出る目は

- 期待値が2. 2
- 分散が1

であるという. チェビシェフの不等式を用いて以下の問いに答えよ.

1. 6が出る確率は最高でいくらか?
2. 1, 2, 3のいずれかが出る確率は最低いくらか?
3. 2が出る確率は最低でいくらか?

連絡事項

182

- 次回の講義は5月16日(水) ←金曜日授業
- 5月16日(水)の授業は、情報工学科計算機室で行う
<http://www.csc.titech.ac.jp/>
- 内容:乱数の発生法に関する計算機実習
- 資料を事前にOCWに公開するので、目を通しておくこと.
- 10時45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、授業終了後に相談に来ること