Universal hash families and
the leftover hash lemma,
and applications to cryptography
and computing

D.R.Stinson

2011年12月20日
黒澤研究室
M1 小笠原琢磨

目次

1. Extractorとは
2. 準備
3. (k-ε)-Extractor

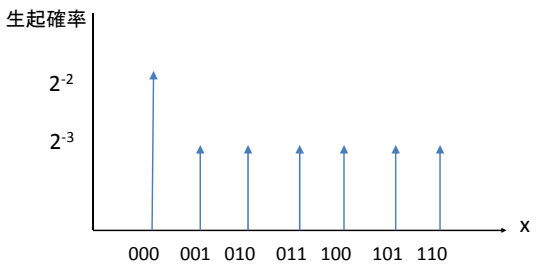
Extractorとは

・Extractor（**Ext**）は、ある確率分布から選んだ要素(**x**)と乱数(**seed**)を入力すると、一様に近い分布から要素(**y**)を選び出力するハッシュ関数

(x,seed) → Ext → y

・使用例：生体認証

集合Xの確率分布(=p)

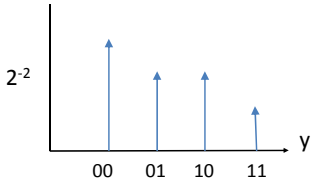


良いハッシュの場合

x←確率分布pに従って選ぶ
f:ハッシュ関数

x → f → **y**

集合Yの確率分布



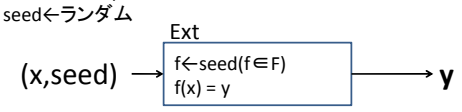
Extractor

Ext

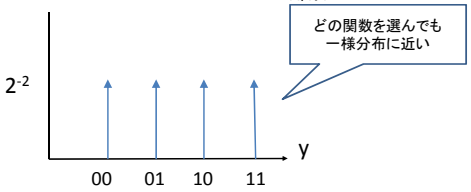
(x,seed) → f←seed(f ∈F)
f(x) = **y** → **y**

Extractor

x←確率分布pに従って選ぶ
seed←ランダム



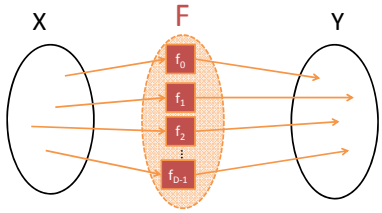
fを選んだときの集合Yの確率分布: $\chi_f(y)$



準備

ハッシュ族

- |X|=N,|Y|=Mとする。
- それぞれの関数fにおいて、f : X → Yとなるようなfがd個ある集合Fを、(D;N,M)ハッシュ族という。



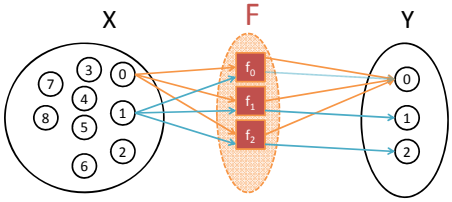
δ-Uハッシュ族

- δ:ハッシュ族の衝突確率
- $x_1,x_2 \in X$ $(x_1 \neq x_2)$
- $f(x_1)=f(x_2)$ $f \in F$
  と衝突する関数fの数を数える。
  この衝突数が高々δD個ある集合Fを、
  δ-U(δ-universal)(D;N,M)ハッシュ族という。

δ-Uハッシュ族　例

$\frac{1}{3}$-U(3;9,3)ハッシュ族　の場合　(δD=1)

F＼X	0	1	2	3	4	5	6	7	8
f_0	0	0	0	1	1	1	2	2	2
f_1	0	1	2	1	2	0	2	0	1
f_2	0	2	1	1	0	2	2	1	0



衝突確率

- (X , p)を有限確率空間とする。
- 確率分布p上の衝突確率を

$$\Delta_p = \sum_{x \in X} (p(x))^2$$

と定義する。
・入力Xの確率分布を考えるときに使用

Reniy entropy

- 確率空間(X,p)のReniy entropy : $h_{Ren}(p)$を

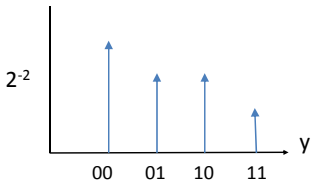
$$h_{Ren}(p) = -\log_2 \Delta_p$$

と定義する。

確率分布の定義

$$q_f(y) = \sum_{x \in f^{-1}(y)} p(x)$$

$$\chi_y(f) = q_f(y)$$



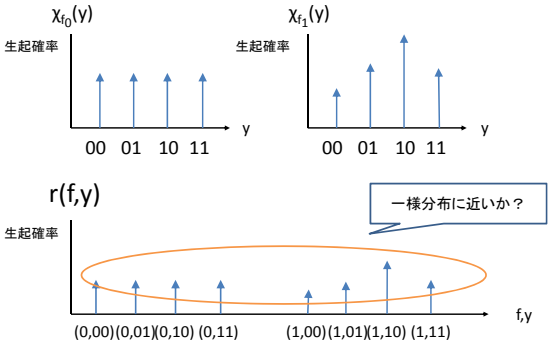
確率分布の定義

f←ランダム(f∈F)
x←確率分布p(x∈X)
と選んだときのfとf(x)の対の集合をF×Yとする。
このF×Y上の確率分布をrとし、

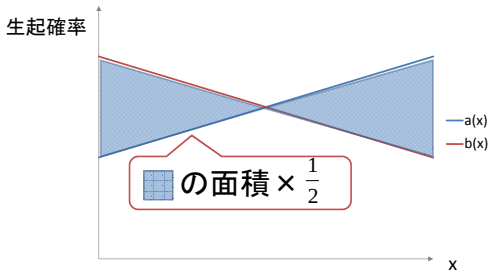
$$r(f,y) = \frac{\chi_y(f)}{D}$$

と定義する。

確率分布の定義



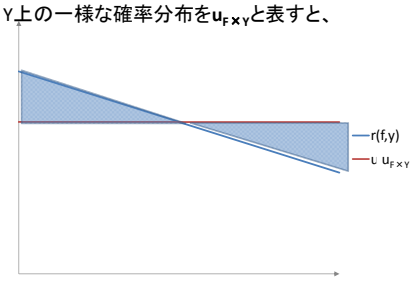
Statistical Distance



$$d(a,b) = \frac{1}{2}\sum_{x \in X} |a(x) - b(x)|$$

Statistical Distance

F×Y上の一様な確率分布を$u_{F\times Y}$と表すと、



$$d(r,u_{F\times X}) = \frac{1}{2}\sum_{f \in F}\sum_{y \in Y}\left|r(f,y) - \frac{1}{DM}\right|$$

(k,ε)-extractor

(k,ε)-extractor

kを、$h_{Ren}(p) \geqq k$を満たす数とする。

もし、 $h_{Ren}(p) = -\log_2 \Delta_p$

$$d(r, u_{F\times Y}) < \varepsilon$$

を満たすならば、δ-U(D;N,M)ハッシュ族Fは(k,ε)-extractorとなる。

ExtractorのStatistical Distance

確率空間$(F\times Y, u_{F\times Y})$の確率変数をY(f,y)と表す。

Yを $Y(f,y) = \left| r(f,y) - \frac{1}{DM} \right|$

と定義する。

Yの期待値をE(Y)とする。

ExtractorのStatistical Distance

$$d(r, u_{F\times Y}) = \frac{1}{2} \sum_{f\in F} \sum_{y\in Y} \left| r(f,y) - \frac{1}{DM} \right|$$

$$= \frac{DM}{2} \times \frac{Y(f,y)}{DM}$$

$$= \frac{DM}{2} \times E(Y)$$

ExtractorのStatistical Distance

$$E(Y^2) = \frac{1}{DM} \sum_{f\in F} \sum_{y\in Y} \left(r(f,y) - \frac{1}{DM} \right)^2$$

$$= \frac{1}{DM} \sum_{f\in F} \sum_{y\in Y} \left(r(f,y)^2 - \frac{2}{DM} r(f,y) + \frac{1}{(DM)^2} \right)$$

$$= \frac{1}{DM} \left(\Delta_r - \frac{1}{DM} \right)$$

$$= \frac{\Delta_r DM - 1}{(DM)^2}$$

$$\sum_{f\in F} \sum_{y\in Y} (r(f,y))^2 = \Delta_r$$

$$\sum_{f\in F} \sum_{y\in Y} r(f,y) = 1$$

$$\sum_{f\in F} \sum_{y\in Y} \frac{1}{(DM)^2} = \frac{DM}{(DM)^2} = \frac{1}{DM}$$

ExtractorのStatistical Distance

$$E(Y^2) = \frac{\Delta_r DM - 1}{(DM)^2}$$

$$E(Y) \leq \sqrt{E(Y^2)}$$

$$= \frac{\sqrt{\Delta_r DM - 1}}{DM}$$

$$d(r, u_{F\times Y}) = \frac{DM}{2} \times E(Y) \leq \frac{\sqrt{\Delta_r DM - 1}}{2}$$

イェンゼンの不等式

$$E(f(Y)) \leq f(E(Y))$$

f(x)= $-x^2$とすると、

$$(E(Y))^2 \leq E(Y^2)$$

$$E(Y) \leq \sqrt{E(Y^2)}$$

衝突確率Δ_r

$$\Delta_r = \sum_{f\in F}\sum_{y\in Y}(r(f,y))^2 = \sum_{f\in F}\sum_{y\in Y}\frac{(\chi_y(f))^2}{D^2}$$

$$\Delta_p = \sum_{x\in X}(p(x))^2$$

$$r(f,y) = \frac{\chi_y(f)}{D}$$

$\sum_{f\in F}\sum_{y\in Y}(\chi_y(f))^2$ を求める。

衝突確率Δ_r

$$\sum_{y\in Y}\sum_{f\in F}(\chi_y(f))^2 = \sum_{y\in Y}\sum_{f\in F}\left(\sum_{x\in f^{-1}(y)}p(x)\right)^2$$

$$\chi_y(f) = \sum_{x\in f^{-1}(y)}p(x)$$

衝突確率Δ_r

$$\sum_{y\in Y}\sum_{f\in F}\left(\sum_{x\in f^{-1}(y)}p(x)\right)^2$$

$$= \sum_{y\in Y}\sum_{f\in F}\left(\sum_{x\in f^{-1}(y)}p(x)\right)\left(\sum_{x\in f^{-1}(y)}p(x)\right)$$

$$= \sum_{y\in Y}\sum_{f\in F}\sum_{x\in f^{-1}(y)}(p(x))^2 + \sum_{y\in Y}\sum_{f\in F}\sum_{x_1\in f^{-1}(y)}\sum_{x_2\in f^{-1}(y),x_2\neq x_1}p(x_1)p(x_2)$$

同じxでyが等しくなる確率

異なる2つのxでyが等しくなる確率

衝突確率Δ_r

$$\sum_{f\in F}\sum_{y\in Y}\sum_{x\in f^{-1}(y)}(p(x))^2$$

$$= \sum_{f\in F}\sum_{x\in X}(p(x))^2$$

$\sum_{x\in X}(p(x))^2 = \Delta_p$

ハッシュ関数の数=D

$$= D\Delta_p$$

衝突確率Δ_r

$$\sum_{f\in F}\sum_{y\in Y}\sum_{x_1\in f^{-1}(y)}\sum_{x_2\in f^{-1}(y),x_2\neq x_1}p(x_1)p(x_2)$$

$$= \sum_{f\in F}\sum_{x_1\in X}\sum_{x_2\in X,x_2\neq x_1}p(x_1)p(x_2)$$

Δ_pの余事象（$1-\Delta_p$）

δ-Uハッシュ族は$f(x_1)=f(x_2)$のような組み合わせが高々δD個ある

$$\leq \delta D(1-\Delta_p)$$

衝突確率Δ_r

$\sum_{y\in Y}\sum_{f\in F}\sum_{x\in f^{-1}(y)}(p(x))^2 = D\Delta_p$

$\sum_{y\in Y}\sum_{f\in F}\sum_{x_1\in f^{-1}(y)}\sum_{x_2\in f^{-1}(y),x_2\neq x_1}p(x_1)p(x_2) \leq \delta D(1-\Delta_p)$

$$\sum_{y\in Y}\sum_{f\in F}\sum_{x\in f^{-1}(y)}(p(x))^2 + \sum_{y\in Y}\sum_{f\in F}\sum_{x_1\in f^{-1}(y)}\sum_{x_2\in f^{-1}(y),x_2\neq x_1}p(x_1)p(x_2)$$

$$\leq \delta D(1-\Delta_p) + D\Delta_p$$

$$= D(\delta + (1-\delta)\Delta_p)$$

衝突確率Δ_r

$$\Delta_r = \sum_{f \in F} \sum_{y \in Y} \frac{(\chi_y(f))^2}{D^2}$$ より、

$$\therefore \Delta_r \le \frac{(\delta + (1-\delta)\Delta_p)}{D}$$

Extractorのstatistical distance

$$d(r, u_{F \times Y}) \le \frac{\sqrt{\Delta_r DM - 1}}{2}$$

$$\Delta_r \le \frac{(\delta + (1-\delta)\Delta_p)}{D}$$

$$d(r, u_{F \times Y}) \le \frac{\sqrt{M(\delta + (1-\delta)\Delta_p) - 1}}{2}$$

Extractorのパラメータ

$$h_{\mathrm{Ren}}(p) = -\log_2 \Delta_p \ge k$$

より、

$$\Delta_p = 2^{-k}$$

とすると、

$$d(r, u_{F \times Y}) \le \frac{\sqrt{M(\delta + (1-\delta)2^{-k}) - 1}}{2}$$

$$< \frac{\sqrt{M(\delta + 2^{-k}) - 1}}{2}$$

$$\le \varepsilon$$

$$d(r, u_{F \times Y}) \le \frac{\sqrt{M(\delta + (1-\delta)\Delta_p) - 1}}{2}$$

Extractorのパラメータ

$$\frac{\sqrt{M(\delta + 2^{-k}) - 1}}{2} \le \varepsilon$$

$$\sqrt{M(\delta + 2^{-k}) - 1} \le 2\varepsilon$$

結論

もし、

$$\sqrt{M(\delta + 2^{-k}) - 1} \le 2\varepsilon$$

ならば、δ-U(D;N,M)ハッシュ族は
(k,ε)-extractorとなる。

ご清聴ありがとうございました。