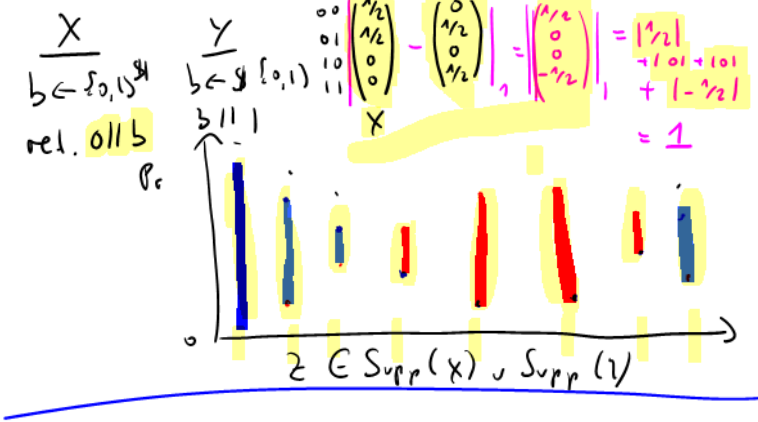
Recap statistical distance

$|X-Y|_1$

$$SD(X,Y) := \frac{1}{2} \cdot \sum_{z \in Supp(X) \cup Supp(Y)} |Pr[X=z] - Pr[Y=z]|$$

New perspective: $Supp(X) \subseteq \{0,1\}^n \supseteq Supp(Y)$



$(a-b)^2 = a^2 - 2ab + b^2$

① SD as $\frac{1}{2}|.|_1$

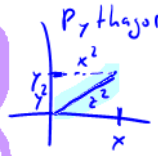
② $\|.\|_1 \leq \sqrt{dim} \cdot \|.\|_2$ (2^m (universe s.t.))

③ CP

④ $SD \leq \sqrt{dim} \cdot \sqrt{CP}$

Bound SD from uniform distribution

→ • $\|v\|_1 \leq \sqrt{2^n} \cdot \|v\|_2$ $\quad v \in \mathbb{R}^{2^n}$



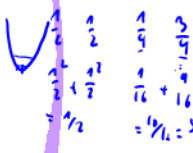
• Collision probability

$$CP(X) := Pr_{x \in \#X} [X = x'] \quad \sqrt{\sum v_i^2}$$

$$x' \in \#X \quad \sum_x (Pr[X=x])^2$$

Example: $CP(U_n) = 2^{-n}$

$= 2^n \cdot (2^{-n})^2 = 2^{-n}$



Lemma: X, $Supp(X) \subseteq \{0,1\}^n$

$$SD(X, U_n) \leq \frac{1}{2} \cdot \sqrt{2^n} \sqrt{CP(X) - 2^{-n}}$$

Proof: $SD(X,U_n) = \frac{1}{2}|X-U_n|_1 \leq \frac{1}{2}\sqrt{2^n}\,\|X-U_n\|_2$

$(\|X-U_n\|_2)^2$

$\overset{def}{=} \sum_{x \in \{0,1\}^n} (Pr[X=x] - \overbrace{Pr[U_n=x]}^{2^{-n}})^2$

$= \sum_{x \in \{0,1\}^n} (Pr[X=x])^2 - 2 \cdot 2^{-n} Pr[X=x] + 2^{-2n}$

$(Pr[X=x])^2$ → $\sum_{x \in \{0,1\}^n} = CP(X)$

$= CP(X) - 2 \cdot 2^{-n} + 2^{-2n} \cdot 2^n$

$= CP(X) - 2^{-n}$

$\{0,1\}^n \to |X-Y|_1$ vectors with 2^n entries

$$SD(X,Y) := \frac{1}{2} \cdot \sum_{z \in Supp(X) \cup Supp(Y)} |Pr[X=z] - Pr[Y=z]|$$

Lemma: X, $Supp(X) \subseteq \{0,1\}^n$

$$SD(X, U_n) \leq \frac{1}{2} \cdot \sqrt{2^n} \sqrt{CP(X) - 2^{-n}}$$

extractors:

Thomas Viedick

☺

⑤ 1-universal hash

⑥ 2-universal hash

⑦ leftover hash-lemma

Leftover Hash Lemma

Let $h : \{0,1\}^n \times R_m \to \{0,1\}^m$ be a 2-universal hash-function.

X, $Supp(X) \subseteq \{0,1\}^n$, $H_\infty(X) \geq m$

Then, (U_m uniform over $\{0,1\}^m$, $+2\log_2 1/\varepsilon$)

$$SD((h(X,R_m),R_m),(U_m,R_m)) \leq \varepsilon$$

$\forall x \neq x', \forall z,z'$

$$Pr_{r \leftarrow \$ R_m}[h(x,r)=z \wedge h(x',r')=z'] = 2^{-2m}$$

$\{0,1\}^n \to |X-Y|_1$ vectors with 2^n entries

$$SD(X,Y) := \frac{1}{2} \cdot \sum_{z \in Supp(X) \cup Supp(Y)} |Pr[X=z] - Pr[Y=z]|$$

Lemma: X, $Supp(X) \subseteq \{0,1\}^n$

$$SD(X,U_n) \leq \frac{1}{2} \cdot \sqrt{2^n} \sqrt{CP(X) - 2^{-n}}$$

Proof: $SD((h(X,R_m),R_m),(U_m,R_m))$ (known)

$$\leq \frac{1}{2} \cdot \sqrt{2^{2m}} \cdot \sqrt{CP((h(X,R_m),R_m)) - 2^{-2m}}$$

$$CP((h(X,R_m),R_m))$$

$$= Pr_{\substack{x \leftarrow \$ X \\ x' \leftarrow \$ X' \\ r \leftarrow \$ R_m \\ r' \leftarrow \$ R_m}}[h(x,r)=h(x',r') \wedge r=r'] = Pr[h(x,r)=h(x',r') \mid r=r'] \cdot Pr[r=r']$$

uniform distribution over $\{0,1\}^n$, so $CP(R_m) = 2^{-m}$

$$= Pr[h(x,r)=h(x',r)] \cdot Pr[r=r']$$

$1 - CP(X)$

$$= 2^{-m} \left(Pr[h(x,r)=h(x',r) \mid x \neq x'] \cdot Pr[x \neq x'] + Pr[h(x,r)=h(x',r) \mid x=x'] \cdot Pr[x=x'] \right) / CP(X)$$

$$= 2^{-m} \left(2^{-2m} \cdot (1 - CP(X)) + 1 \cdot CP(X) \right)$$

$$= 2^{-m} \left(2^{-2m} + CP(X) \cdot (1 - 2^{-2m}) \right)$$

$\leq 2^{-H_\infty(X)}$

$$\leq 2^{-m} \left(2^{-2m} + 2^{-H_\infty(X)} \right)$$

→ Exercise Sheet 5

$$SD \leq \frac{1}{2} \cdot \sqrt{2^{2m} \cdot \left[2^{-3m} + 2^{-(m+H_\infty(X))} - 2^{-2m} \right]}$$

$$= \frac{1}{2} \cdot \sqrt{2^{-m} + 2^{m-H_\infty(X)} - 1} \quad +\leq 0$$

$$\leq 2^{\frac{m-H_\infty(X)}{2}} \leq \varepsilon \iff H_\infty(X) \geq m + 2\log_2 1/\varepsilon$$

Q & A session:

$x \neq x'$ $\forall z,z' \in \{0,1\}^m$

$$Pr_r[h(x,r)=z \wedge h(x',r')=z'] = 2^{-2m}$$

$$\downarrow \quad Pr_r[z,z'] = 2^{-2m}$$

$z \| z' \in \{0,1\}^{2m}$

$$Pr[z] = \sum_{z'} Pr[z \| z']$$

$$2^m = \sum_{z'} 2^{-2m}$$

$$= 2^m \cdot 2^{-2m} = 2^{-m}$$

Equivalent terms
- 2-wise uniform
- pairwise independent
- 2-uniform

→ Lecture notes

$$Pr[X=x \wedge X'=x']$$

$x \leftarrow \$ X$

$x' \leftarrow \$ X'$

$$= Pr[X=x] \cdot Pr[X'=x']$$

$$Pr[X=X']$$

$$= \sum_x Pr[X=x \wedge X'=x]$$

$$= \sum_x Pr[X=x] \cdot Pr[X'=x]$$

$$= \sum_x (Pr[X=x])^2$$

$\forall x \; \forall z \quad Pr_{r \leftarrow \$ \{0,1\}^m}[h(x,r)=z] = 2^{-m}$

Example: $h(x,r) := r$

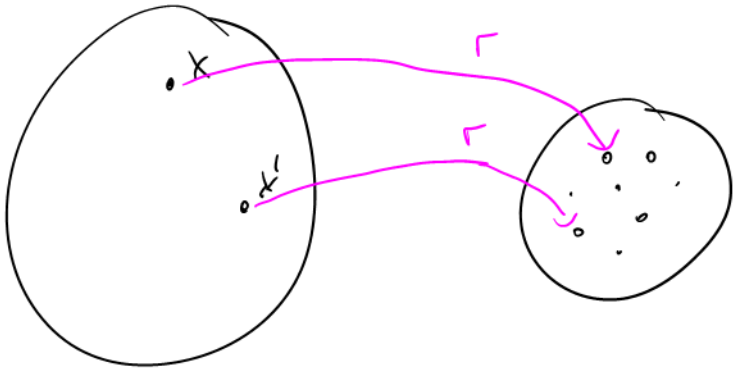
$h(x,r) := x \oplus r$

1-universal

statistical $X \neq U^n$

$h(x,r) = \langle x, r \rangle$

$= \bigoplus x_i \wedge r_i$



- x, x' fixed
- Pick random r

"$\equiv$" random z and random z'

2-wise uniform hash-function

$SD(X,Y) \geq CD(X,Y)$

$X \overset{\text{stat.}}{\approx} Y \Rightarrow X \overset{\text{comp}}{\approx} Y$

$\not\Leftarrow$

Example:

PRG g

X

$x \leftarrow \$ \{0,1\}^n$

$y \leftarrow g(x)$

ret. y

2^n

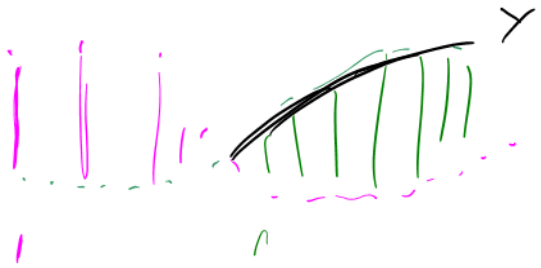
for each $t \in \{0,1\}^{2n}$

Y

$y \leftarrow \$ \{0,1\}^{2n}$

ret. y

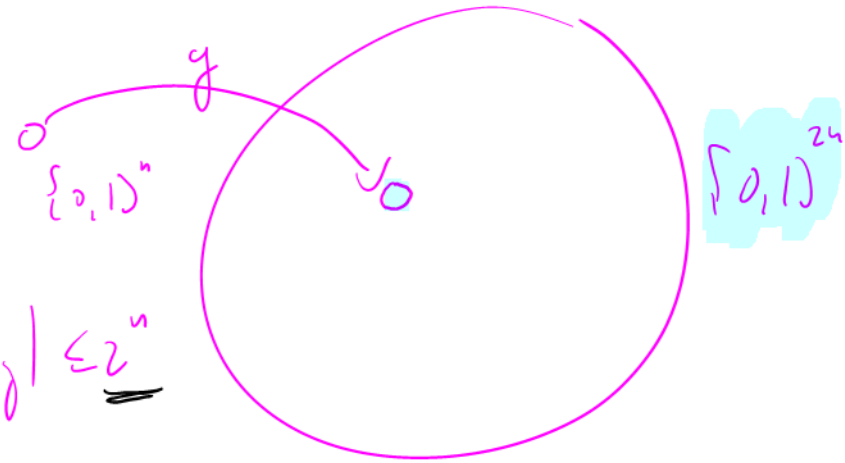
$=$



$SD(X,Y) \geq 2^{-2n} \cdot (2^{2n} - 2^n)$

$= 1 - 2^{-n}$

2.



$|\text{img}| \leq 2^n$

$2^{2n} - 2^n \cdot \Pr[g(x) = t] = 0$