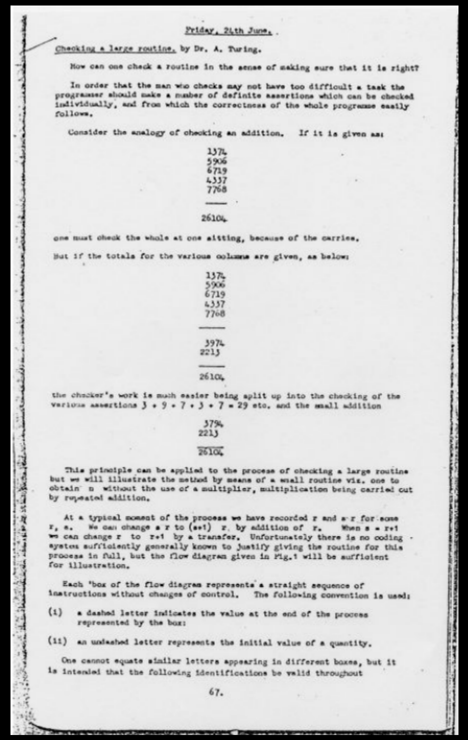
automated verification + differential privacy

aws albarghouthi university of wisconsin–madison

calvin smith

justin hsu



Friday, 24th June.

Checking a large routine, by Dr. A. Turing.

How can one check a routine in the sense of making sure that it is right?

In order that the man who checks may not have too difficult a task the programmer should make a number of definite assertions which can be checked individually, and from which the correctness of the whole programme easily follows.

Consider the analogy of checking an addition. If it is given as:

1374
5906
6719
4337
7768

26104

one must check the whole at one sitting, because of the carries.

But if the totals for the various columns are given, as below:

1374
5906
6719
4337
7768

3974
2213

26104

the checker's work is much easier being split up into the checking of the various assertions 3 + 9 + 7 + 3 + 7 = 29 etc. and the small addition

3794
2213

26104

This principle can be applied to the process of checking a large routine but we will illustrate the method by means of a small routine viz. one to obtain n without the use of a multiplier, multiplication being carried out by repeated addition.

At a typical moment of the process we have recorded r and s r for some r, s. We can change s r to (s+1) r by addition of r. When s = r+1 we can change r to r+1 by a transfer. Unfortunately there is no coding system sufficiently generally known to justify giving the routine for this process in full, but the flow diagram given in Fig.1 will be sufficient for illustration.

Each 'box of the flow diagram represents a straight sequence of instructions without changes of control. The following convention is used:

(i) a dashed letter indicates the value at the end of the process represented by the box:

(ii) an undashed letter represents the initial value of a quantity.

One cannot equate similar letters appearing in different boxes, but it is intended that the following identifications be valid throughout

67.

program
logics

abstract interp
model checking

industrial
tools

1949 1960 1970 1980 1990 2000

```
assert(x != null)
```


algorithm
property
ϵ-dp
verifier
proof


algorithm
property
accuracy
verifier
proof

Proof of Report Noisy Max

Proof. Fix $D = D' \cup \{a\}$. Let c, respectively c', denote the vector of counts when the database is D, respectively D'. We use two properties:

1. *Monotonicity of Counts.* For all $j \in [m]$, $c_j \geq c'_j$; and
2. *Lipschitz Property.* For all $j \in [m]$, $1 + c'_j \geq c_j$.

Fix any $i \in [m]$. We will bound from above and below the ratio of the probabilities that i is selected with D and with D'.

Fix r_{-i}, a draw from $[\mathrm{Lap}(1/\varepsilon)]^{m-1}$ used for all the noisy counts except the ith count. We will argue for each r_{-i} independently. We

use the notation $\Pr[i|\xi]$ to mean the probability that the output of the Report Noisy Max algorithm is i, conditioned on ξ.

We first argue that $\Pr[i|D, r_{-i}] \leq e^{\varepsilon} \Pr[i|D', r_{-i}]$. Define

$$r^* = \min_{r_i} : c_i + r_i > c_j + r_j \ \forall j \neq i.$$

Note that, having fixed r_{-i}, i will be the output (the argmax noisy count) when the database is D if and only if $r_i \geq r^*$.

We have, for all $1 \leq j \neq i \leq m$:

$$\begin{aligned} c_i + r^* &> c_j + r_j \\ \Rightarrow (1 + c'_i) + r^* \geq c_i + r^* &> c_j + r_j \geq c'_j + r_j \\ \Rightarrow c'_i + (r^* + 1) &> c'_j + r_j. \end{aligned}$$

Thus, if $r_i \geq r^* + 1$, then the ith count will be the maximum when the database is D' and the noise vector is (r_i, r_{-i}). The probabilities below are over the choice of $r_i \sim \mathrm{Lap}(1/\varepsilon)$.

$$\begin{aligned} \Pr[r_i \geq 1 + r^*] &\geq e^{-\varepsilon} \Pr[r_i \geq r^*] = e^{-\varepsilon} \Pr[i|D, r_{-i}] \\ \Rightarrow \Pr[i|D', r_{-i}] \geq \Pr[r_i \geq 1 + r^*] &\geq e^{-\varepsilon} \Pr[r_i \geq r^*] = e^{-\varepsilon} \Pr[i|D, r_{-i}], \end{aligned}$$

which, after multiplying through by e^{ε}, yields what we wanted to show: $\Pr[i|D, r_{-i}] \leq e^{\varepsilon} \Pr[i|D', r_{-i}]$.

We now argue that $\Pr[i|D', r_{-i}] \leq e^{\varepsilon} \Pr[i|D, r_{-i}]$. Define

$$r^* = \min_{r_i} : c'_i + r_i > c'_j + r_j \ \forall j \neq i.$$

Note that, having fixed r_{-i}, i will be the output (argmax noisy count) when the database is D' if and only if $r_i \geq r^*$.

We have, for all $1 \leq j \neq i \leq m$:

$$\begin{aligned} c'_i + r^* &> c'_j + r_j \\ \Rightarrow 1 + c'_i + r^* &> 1 + c'_j + r_j \\ \Rightarrow c'_i + (r^* + 1) &> (1 + c'_j) + r_j \\ \Rightarrow c_i + (r^* + 1) \geq c'_i + (r^* + 1) &> (1 + c'_j) + r_j \geq c_j + r_j. \end{aligned}$$

Thus, if $r_i \geq r^* + 1$, then i will be the output (the argmax noisy count) on database D with randomness (r_i, r_{-i}). We therefore have, with probabilities taken over choice of r_i:

$$\Pr[i|D, r_{-i}] \geq \Pr[r_i \geq r^* + 1] \geq e^{-\varepsilon} \Pr[r_i \geq r^*] = e^{-\varepsilon} \Pr[i|D', r_{-i}],$$

Proof of Exponential Mech

Proof. For clarity, we assume the range $\mathcal{R}$ of the exponential mechanism is finite, but this is not necessary. As in all differential privacy proofs, we consider the ratio of the probability that an instantiation

of the exponential mechanism outputs some element $r \in \mathcal{R}$ on two neighboring databases $x \in \mathbb{N}^{|\mathcal{X}|}$ and $y \in \mathbb{N}^{|\mathcal{X}|}$ (i.e., $\|x - y\|_1 \leq 1$).

$$\begin{aligned} \frac{\Pr[\mathcal{M}_E(x, u, \mathcal{R}) = r]}{\Pr[\mathcal{M}_E(y, u, \mathcal{R}) = r]} &= \frac{\left(\frac{\exp(\frac{\varepsilon u(x,r)}{2\Delta u})}{\sum_{r' \in \mathcal{R}} \exp(\frac{\varepsilon u(x,r')}{2\Delta u})}\right)}{\left(\frac{\exp(\frac{\varepsilon u(y,r)}{2\Delta u})}{\sum_{r' \in \mathcal{R}} \exp(\frac{\varepsilon u(y,r')}{2\Delta u})}\right)} \\ &= \left(\frac{\exp(\frac{\varepsilon u(x,r)}{2\Delta u})}{\exp(\frac{\varepsilon u(y,r)}{2\Delta u})}\right) \cdot \left(\frac{\sum_{r' \in \mathcal{R}} \exp(\frac{\varepsilon u(y,r')}{2\Delta u})}{\sum_{r' \in \mathcal{R}} \exp(\frac{\varepsilon u(x,r')}{2\Delta u})}\right) \\ &= \exp\left(\frac{\varepsilon(u(x, r') - u(y, r'))}{2\Delta u}\right) \\ &\quad \cdot \left(\frac{\sum_{r' \in \mathcal{R}} \exp(\frac{\varepsilon u(y,r')}{2\Delta u})}{\sum_{r' \in \mathcal{R}} \exp(\frac{\varepsilon u(x,r')}{2\Delta u})}\right) \\ &\leq \exp\left(\frac{\varepsilon}{2}\right) \cdot \exp\left(\frac{\varepsilon}{2}\right) \cdot \left(\frac{\sum_{r' \in \mathcal{R}} \exp(\frac{\varepsilon u(x,r')}{2\Delta u})}{\sum_{r' \in \mathcal{R}} \exp(\frac{\varepsilon u(x,r')}{2\Delta u})}\right) \\ &= \exp(\varepsilon). \end{aligned}$$

Similarly, $\frac{\Pr[\mathcal{M}_E(y,u)=r]}{\Pr[\mathcal{M}_E(x,u)=r]} \geq \exp(-\varepsilon)$ by symmetry. □

[dwork/roth book]


IO examples
property
ϵ-*dp*
synthesizer
algorithm
+
proof

```
function IDC
  (iter : Nat[i]) (eps : num[e])
  (db :[2 * i * e] db_type) (qs : query bag)
  (PA : (query bag) -> approx_db
      -> db_type -o[e] Circle query)
  (DUA : approx_db -> query -> num -> approx_db)
  (eval_q : query -> db_type -o[1] num)
  : Circle approx_db {
  case iter of
    0      => return init_approx
  | n + 1 =>
      sample approx = (IDC n eps db qs PA DUA);
      sample q = PA qs approx db;
      sample actual = add_noise eps (eval_q q db);
      return (DUA approx q actual)
  }
```

Figure 11. Iterative Database Function in *DFuzz*

[Gupta, Roth, Ullman, TCC 2012]

short-term vision aid algorithm designers

long-term vision put theorists out of work

1 automatic proofs of accuracy [POPL19]

2 automatic proofs of differential privacy [POPL18]

theme

get rid of probability! long live logic!

```
{x > 0}
 y = x + 1
{y > 0}
```

$$\longleftrightarrow \quad x > 0 \land y = x + 1 \implies y > 0$$

solve with an **SAT/SMT** solver

```
{0 < p < 1}

 x ~ flip(p)

{x = true} @ 1-p
```

challenge how do we check this with first-order logic?

```
{0 < p < 1}

 x ~ flip(p)

{x = true} @ 1-p
```

idea axiomatization

```
{0 < p < 1}

w = 0
assume(x = true)
w = w + 1-p

{x = true & w <= 1-p}
```

$$w = 0 \wedge x \wedge w' = w + 1 - p \Longrightarrow x \wedge w' \leq 1 - p$$

challenge many different axiomatizations

```
{0 < p < 1}

 w = 0
 assume(x = true)
 w = w + 1-p

{x = true & w <= 1-p}
```

```
{0 < p < 1}

 w = 0
 assume(x = false)
 w = w + p

{x = true & w <= 1-p}
```

challenge many different axiomatizations

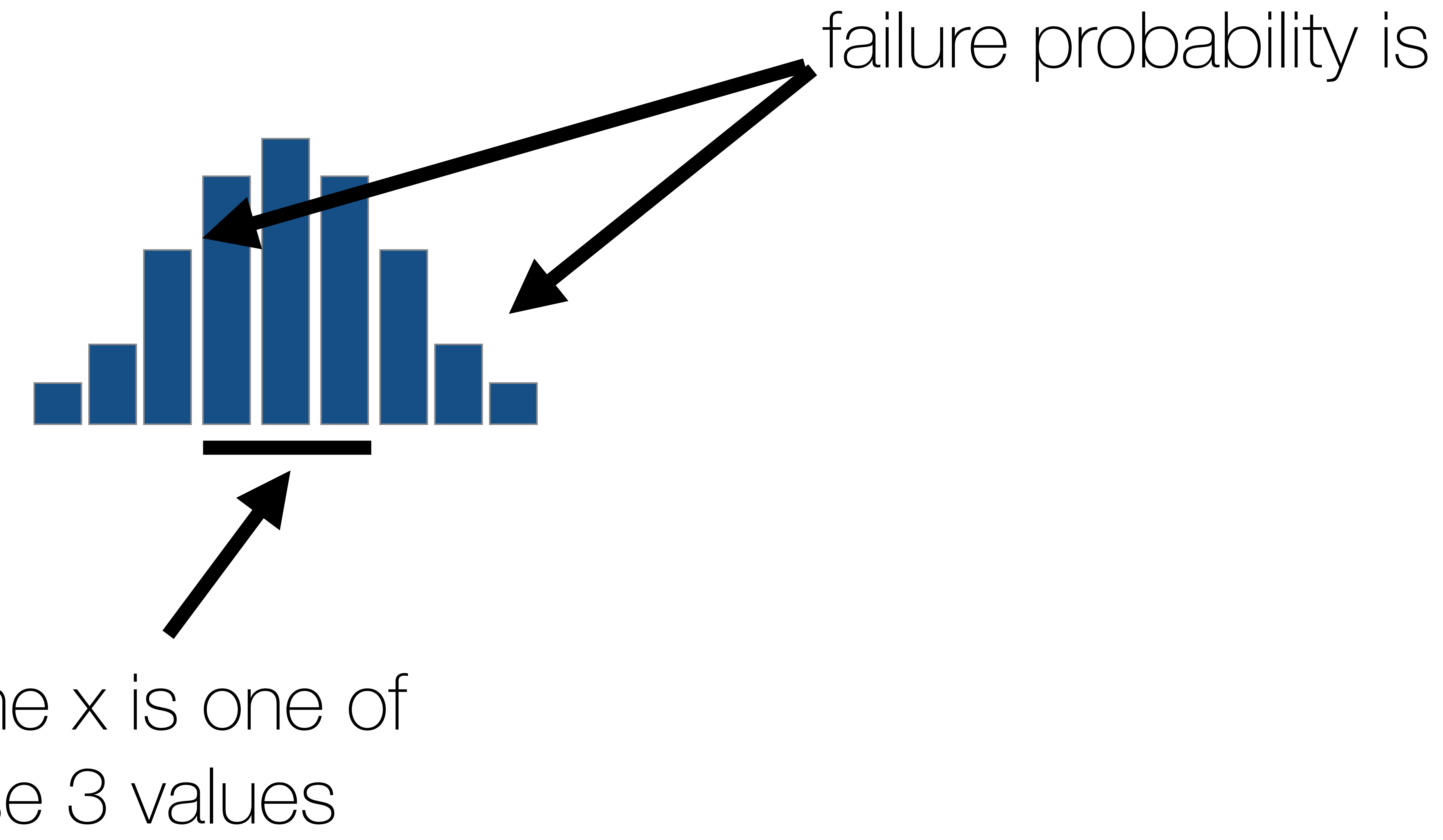
```
{0 < p < 1}

 w = 0
 assume(true)
 w = w + 0

{x = true & w <= 1-p}
```

challenge many different axiomatizations

`x ~ dist`



```
{0 < p < 1}

 x ~ flip(p)

{x = true} @ 1-p
```

idea synthesize axiomatization

```
{0 < p < 1}

 w = 0
 assume(phi(x))
 w = w + pr(not phi(x))

{x = true & w <= 1-p}
```

$$\exists \varphi \,.\, \forall w, w', x \,.$$

$$w = 0 \wedge \varphi(x) \wedge w' = w + pr(\varphi(x)) \implies x \wedge w' \leq 1 - p$$

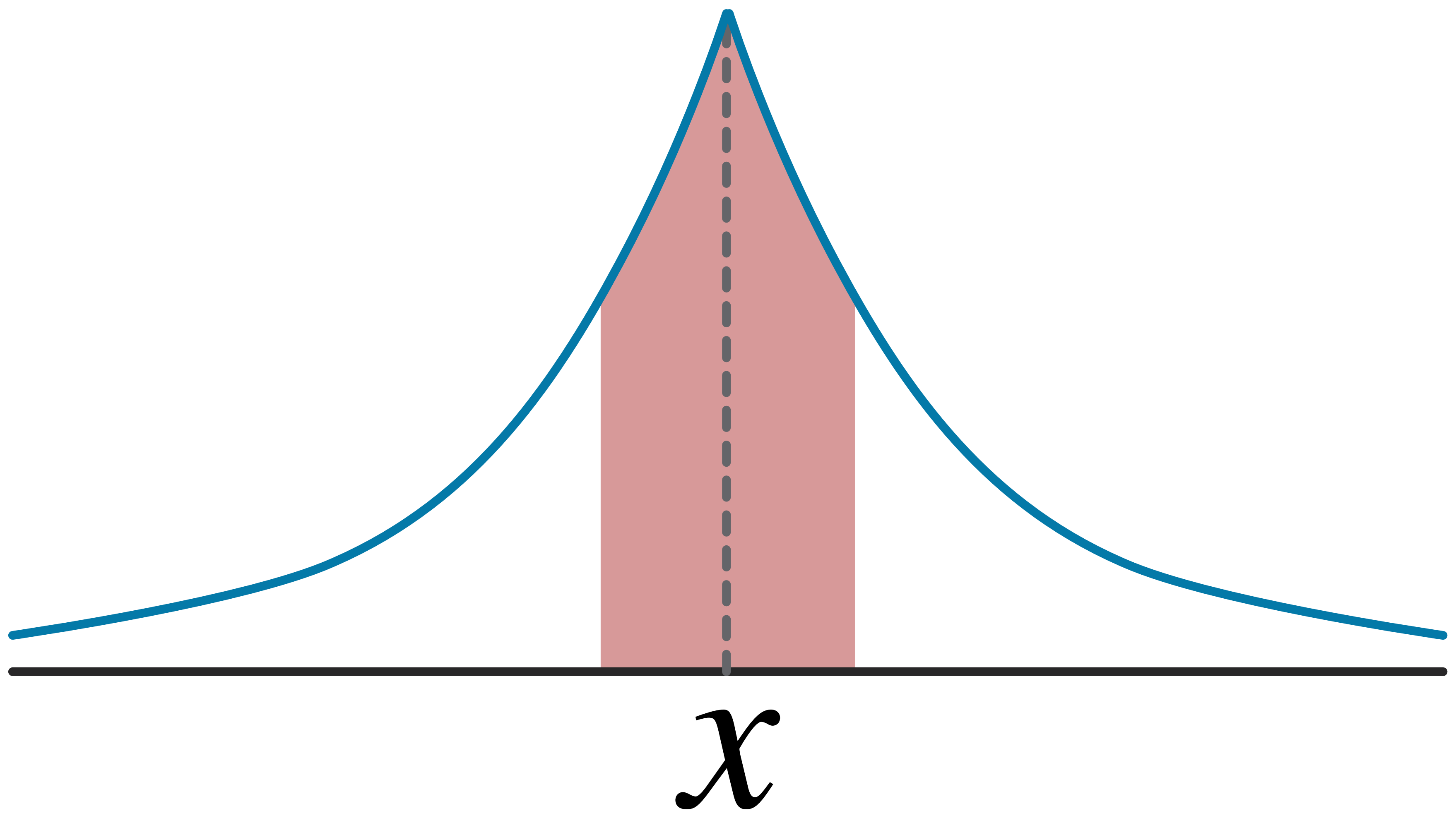
```
y ~ Lap(x,s)
```

axiom family

$$|x - y| \leq s \cdot log\left(\frac{1}{f(V_I)}\right)$$

with failure probability

$$f(V_I) \in (0,1]$$



```
{0 < p < 1}

def rnm(q):
  i, best, r = 0

  while i < len(q)
   d ~ Lap(q[i], 2/ε)

   if d > best || i = 0
    r = i
    best = d

   i = i + 1

  return r
{∀j. q[r] >= q[j] - 4/ε log (len(q)/p)} @ p
```

$$|q[i] - d| \leq \frac{2}{\epsilon} \cdot log\left(\frac{len(q)}{p}\right)$$

with failure probability $\frac{p}{len(q)}$

1 automatic proofs of accuracy [POPL19]

2 automatic proofs of differential privacy [POPL18]

theme

get rid of probability! long live logic!

$$p : D \to \Delta(\mathbb{Z})$$

$$\forall d, d', a, \epsilon \,.\, \mathsf{adj}(d, d') \Rightarrow$$

$$\mathbb{P}[p(d) = a] \leq e^{\epsilon} \cdot \mathbb{P}[p(d') = a]$$

problems

proving differential privacy is hard and error-prone [lyu et al. 16]

existing automated techniques only work for simple algorithms

goal

automatically prove differential privacy of advanced algorithms

key ideas

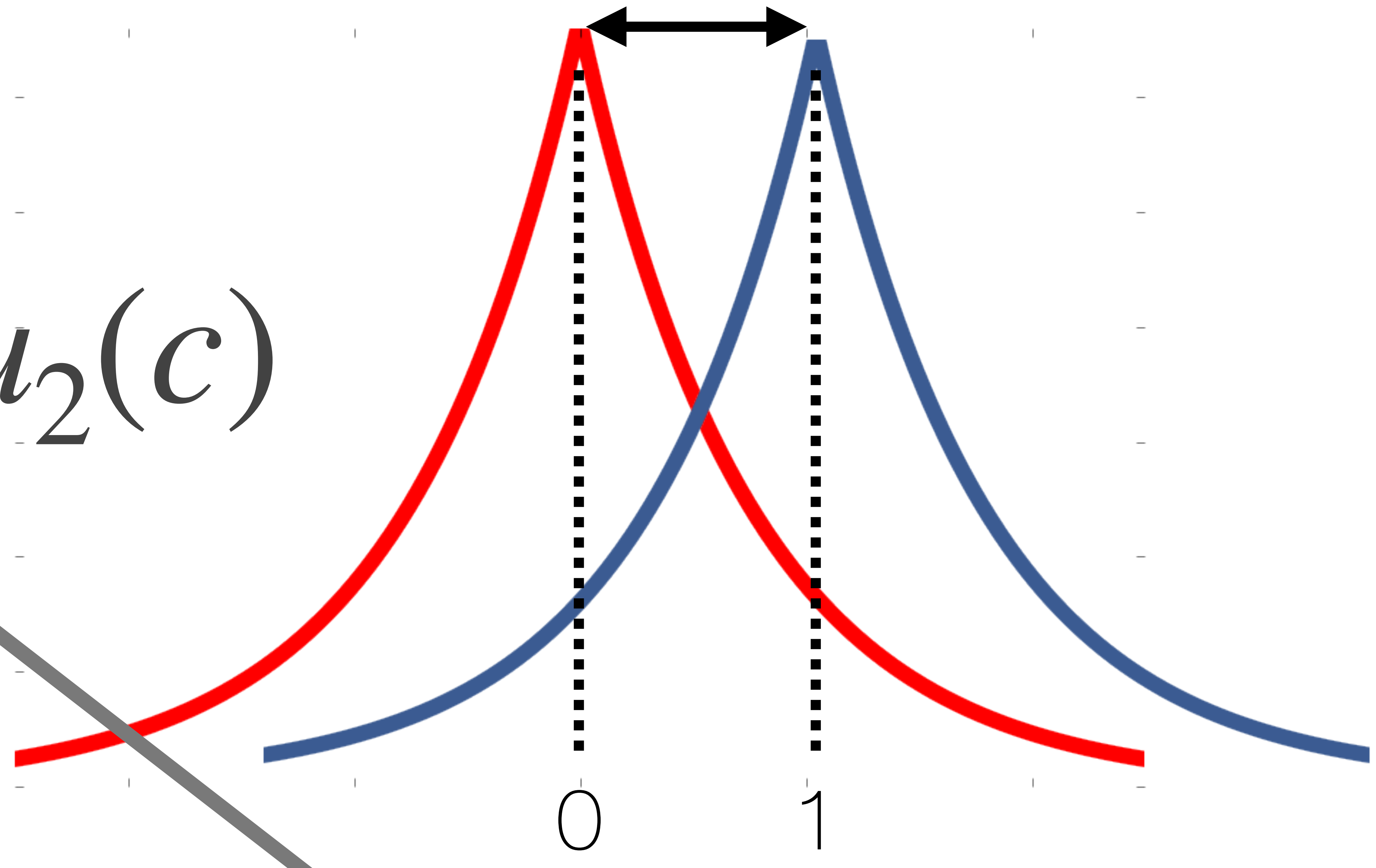
view differential privacy **coupling proofs as games**

solve a program **synthesis/verification problem**

$$\exists q \, . \, \forall x \, . \, \varphi(q, x)$$

variable approximate couplings

scale of distributions is ***1/y***

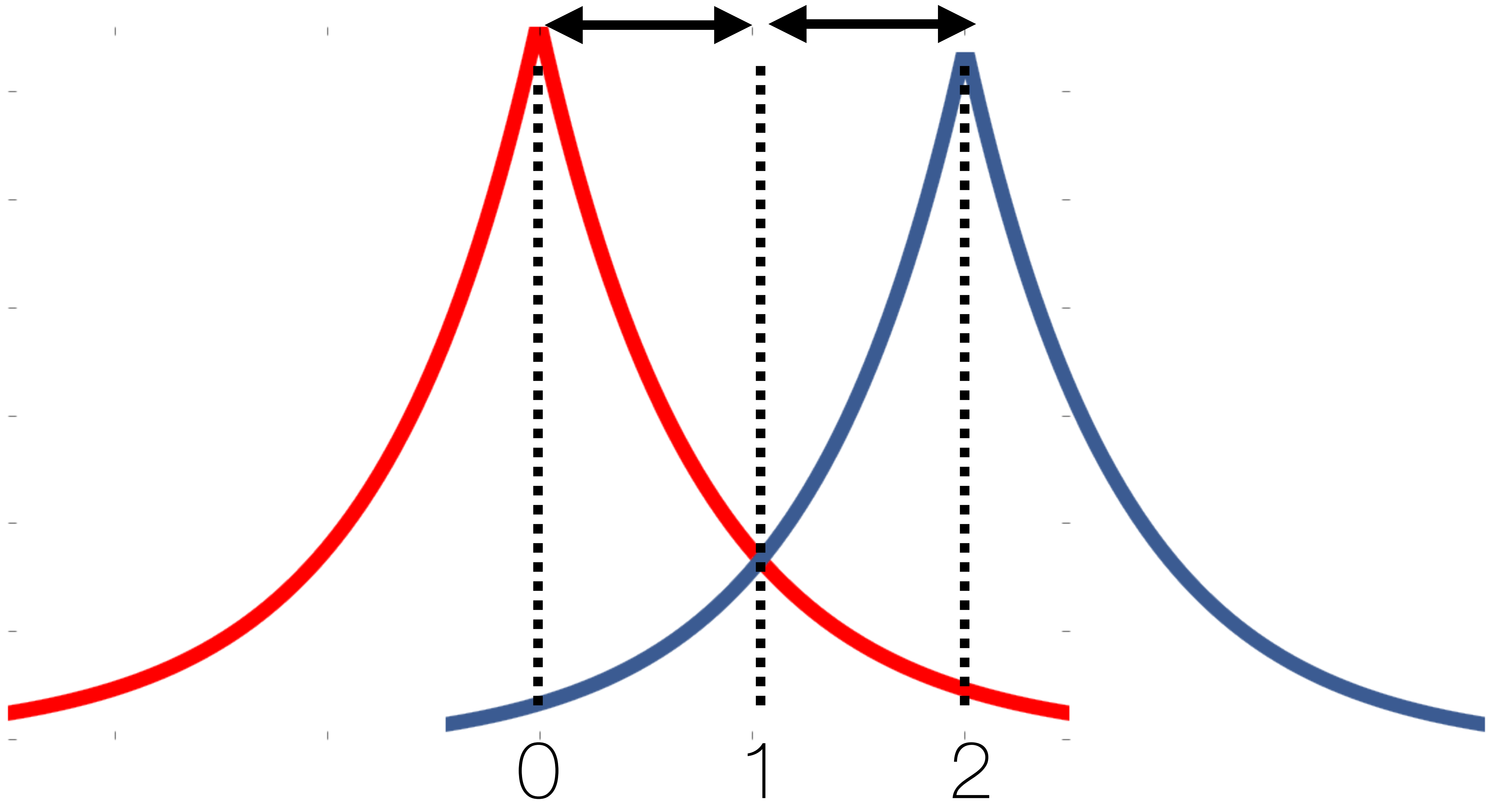


$$\mu_1(c) \leq e^y \cdot \mu_2(c)$$

$$\{(c, c, \boxed{y}) \mid c \in \mathbb{Z}\}$$

variable approximate couplings

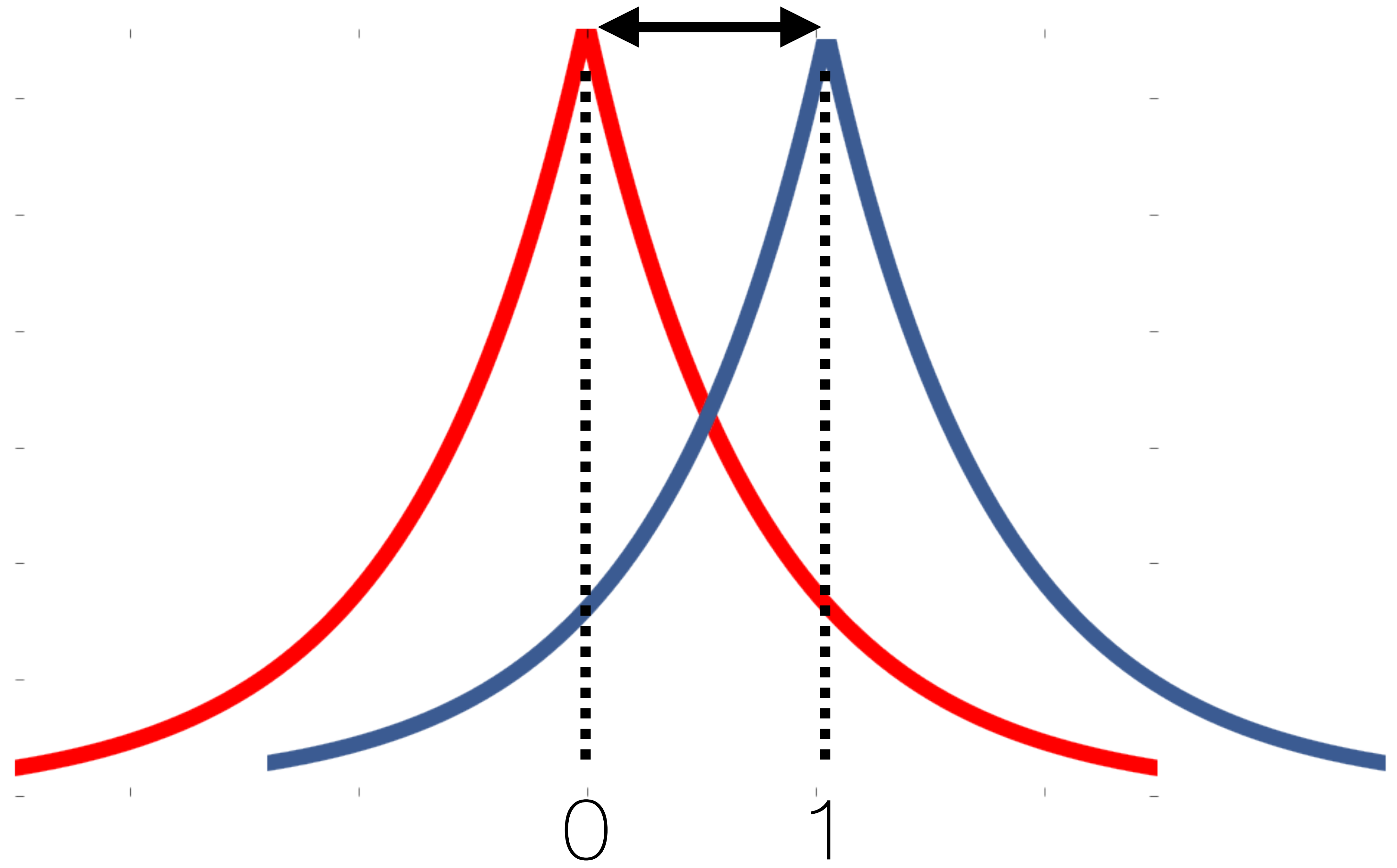
scale of distributions is ***1/y***



$$\{(c, c, 2y) \mid c \in \mathbb{Z}\}$$

variable approximate couplings

scale of distributions is ***1/y***



$$\{(c, c+1, 0) \mid c \in \mathbb{Z}\}$$

proof rule

p is DP if $\forall d, d', \epsilon \,.\, \exists \mathscr{C} \,.$

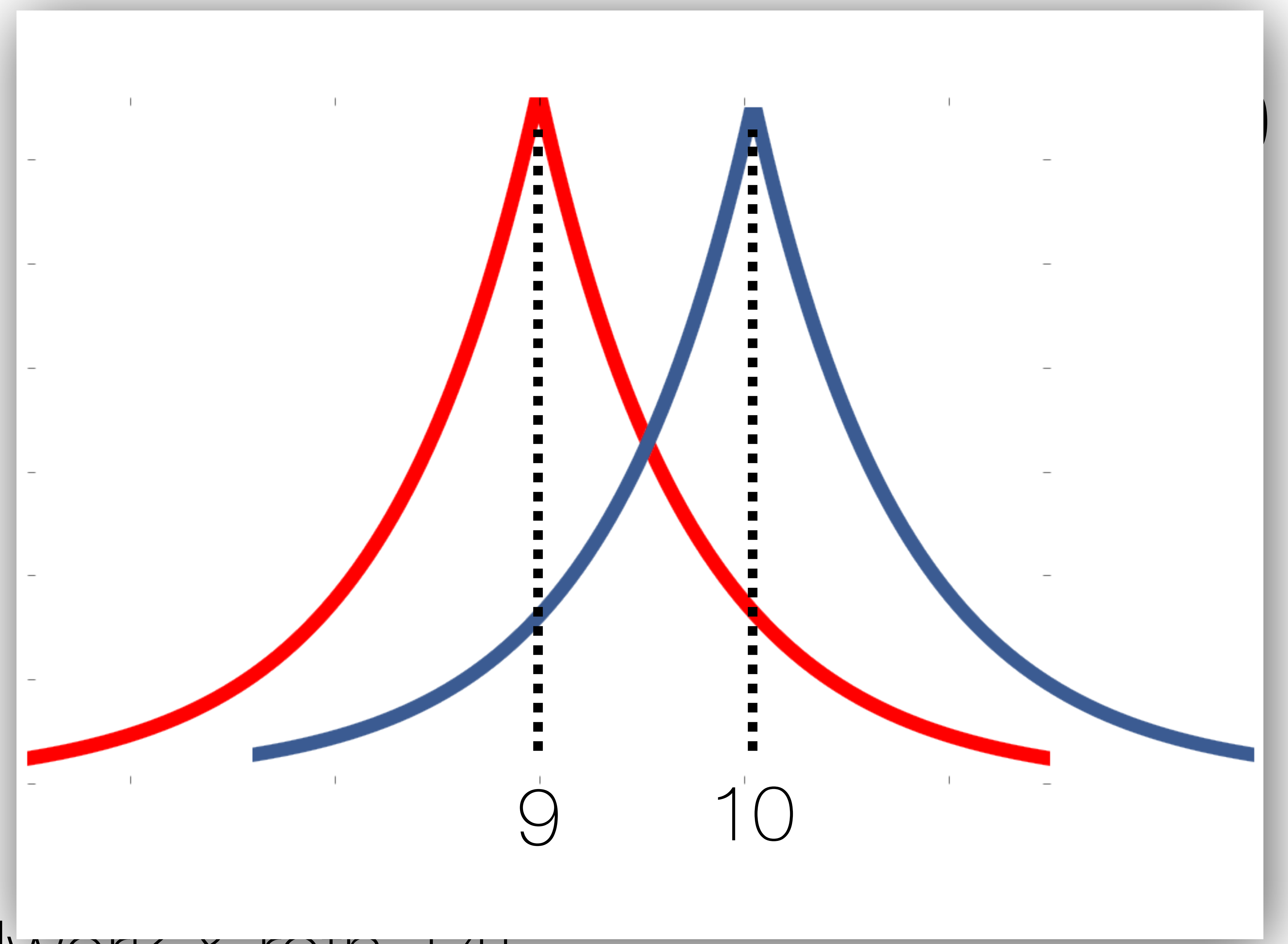
$\mathscr{C}$ couples $p(d), p(d')$

$\mathscr{C} = \{(c, c, y) \mid y \leq \epsilon\}$

let’s play!

```
def rnm(q):
  i, best, r = 0

  while i < len(q)
   d ~ Lap(q[i], 2/ε)
```



[dwork & roth 14]

```
q1 = [9, 0]          q2 = [10, 1]
```

cost = 0

```
r1 = 0               r2 = 0


r1 = 0               r2 = 0
d1 = c               d2 = c
```

cost = ε/2

non-deterministically pick from

$$\{(c, c, \epsilon/2) \mid c \in \mathbb{Z}\}$$

cost = ε

{r1 = r2 && cost <= ε}

our game strategy

in every iteration, couple samples using

$$\{(c, c, \epsilon/2) \mid c \in \mathbb{Z}\}$$

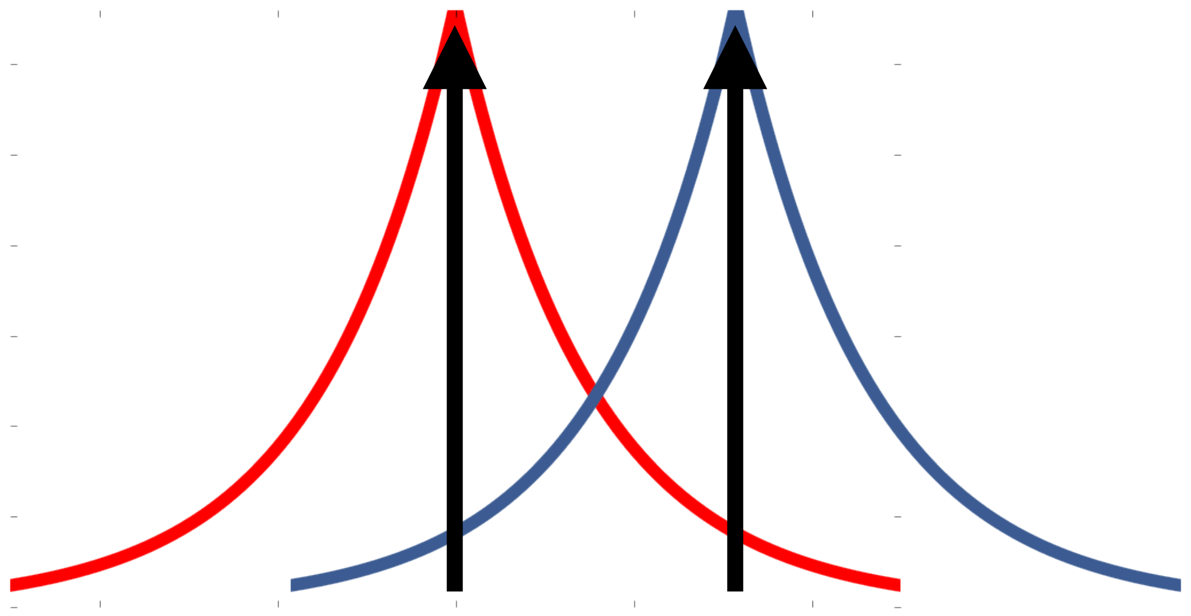
$$\frac{n \cdot \epsilon}{2}$$ differential privacy

a winning strategy

use this coupling in 1 iteration only

$$\{(c, c+1, \epsilon) \mid c \in \mathbb{Z}\}$$

in all other iterations pay **zero cost**



winning strategies are programs

if *condition*
  use *coupling C1*
else
  use *coupling C2*

evaluation

PartialSum	Compute the noisy sum of a list of queries.
PrefixSum	Compute the noisy sum for every prefix of a list of queries.
SmartSum	Advanced version of PrefixSum that chunks the list [Chan et al. 2011; Dwork et al. 2010].
ReportNoisyMax	Find the element with the highest quality score [Dwork and Roth 2014].
ExpMech	Variant of ReportNoisyMax using the exponential distribution [Dwork and Roth 2014; McSherry and Talwar 2007].
AboveThreshold	Find the index of the first query above threshold [Dwork and Roth 2014].
AboveThresholdN	Find the indices of the first N queries with answer above threshold [Dwork and Roth 2014; Lyu et al. 2017].
NumericSparse	Return the index and answer of the first query above threshold [Dwork and Roth 2014].
NumericSparseN	Return the indices and answers of the first N queries above threshold [Dwork and Roth 2014; Lyu et al. 2017].

1 automatic proofs of accuracy [POPL19]

2 automatic proofs of differential privacy [POPL18]

theme

get rid of probability! long live logic!