Verifying Properties of Distributions

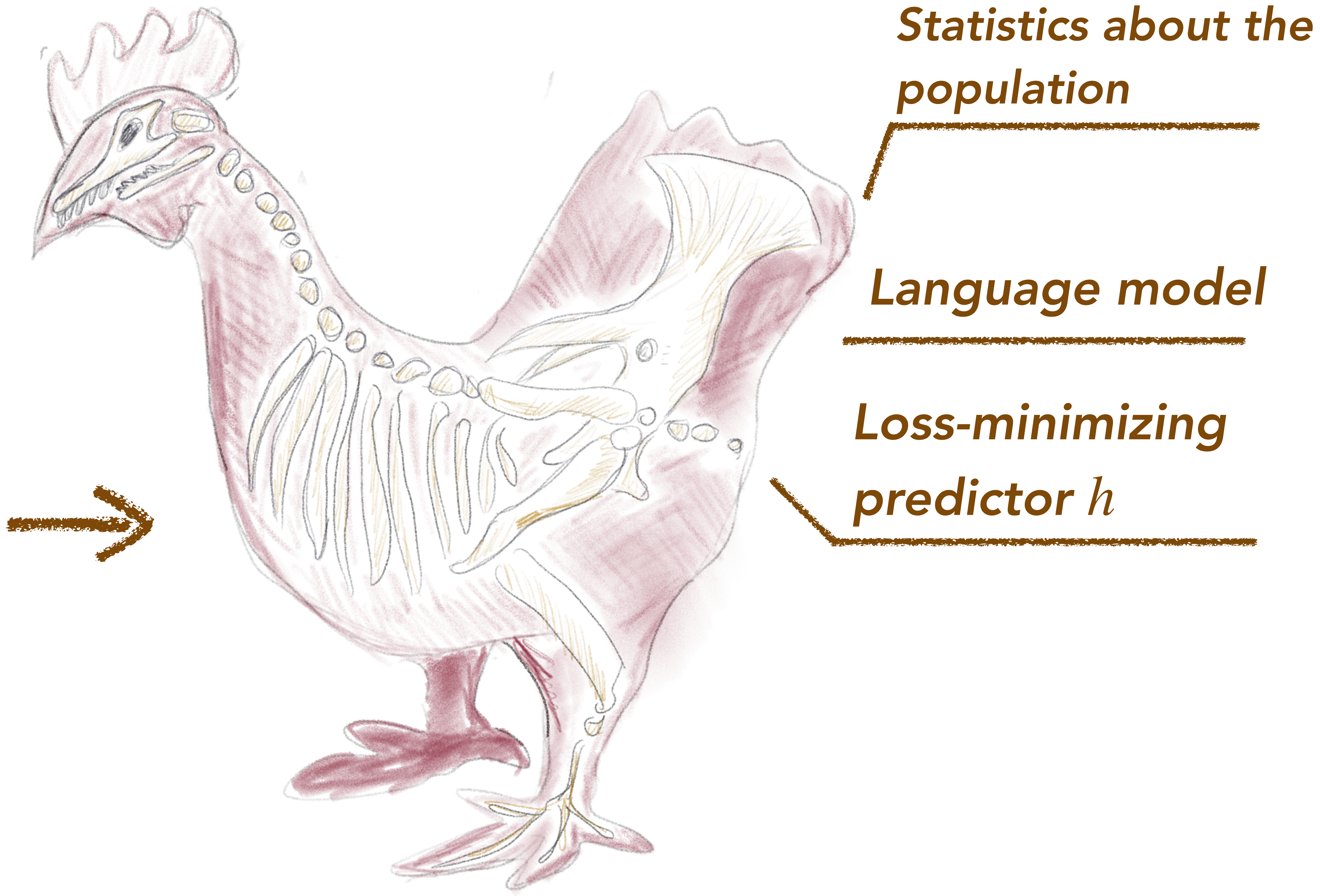
Based on works with Guy Rothblum

MIT, September 2025

Motivation: Data Science Pipeline

Gather valuable dataset

Analyze using sophisticated or extensive algorithm

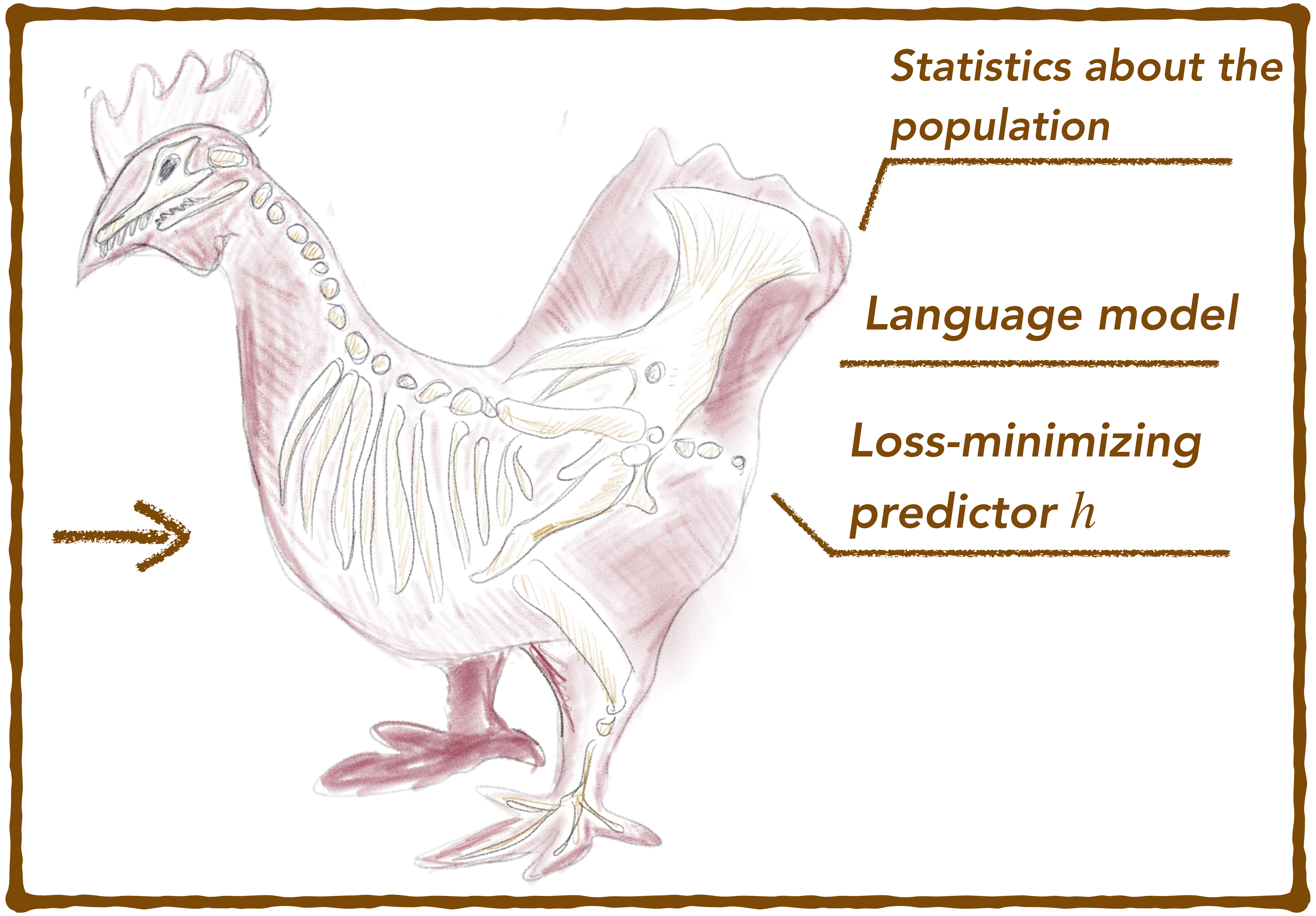


Arrive at useful conclusions

Motivation: Data Science Pipeline

Gather valuable dataset

Analyze using sophisticated or extensive algorithm



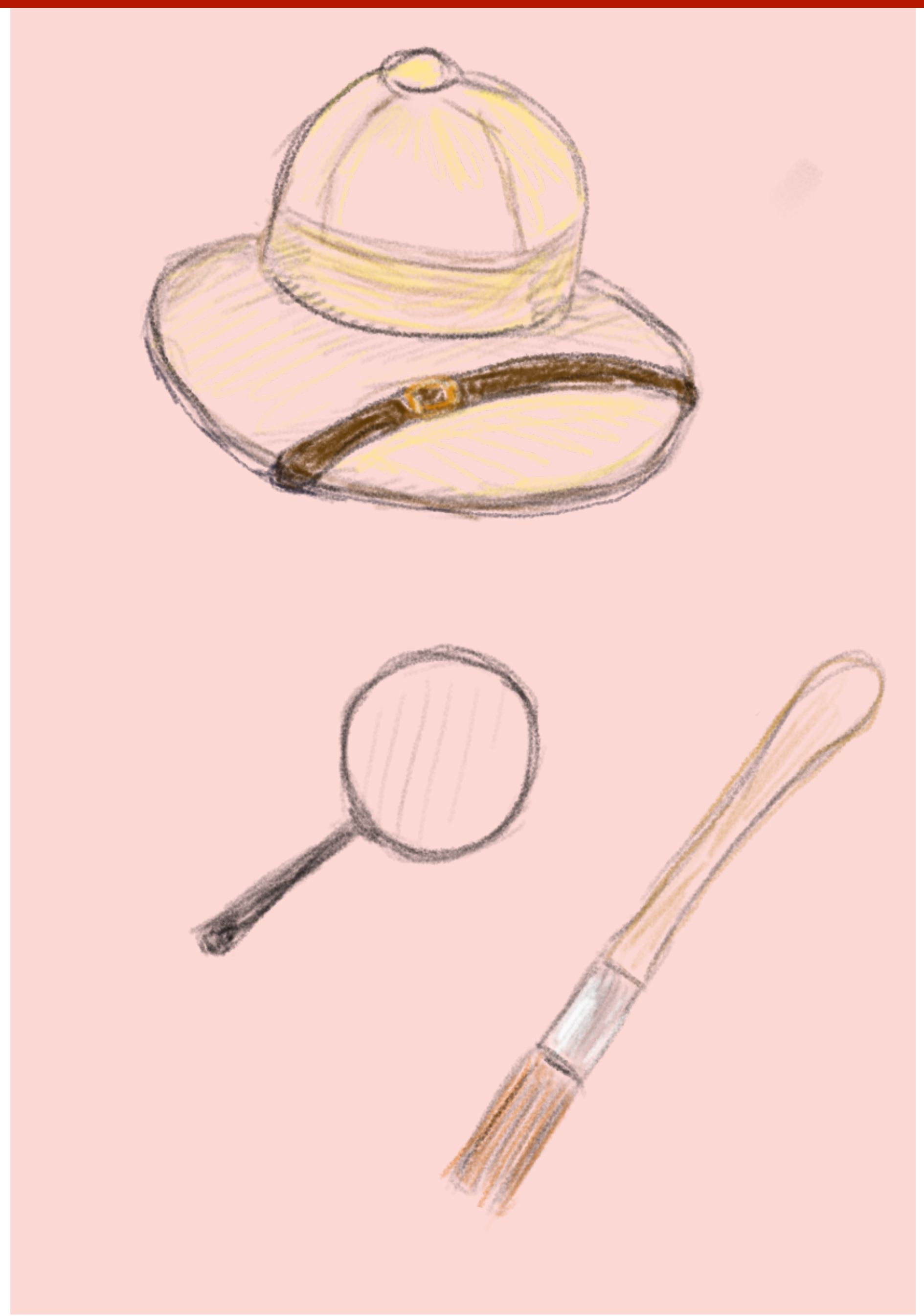
Arrive at useful conclusions

Data Science: Correct?

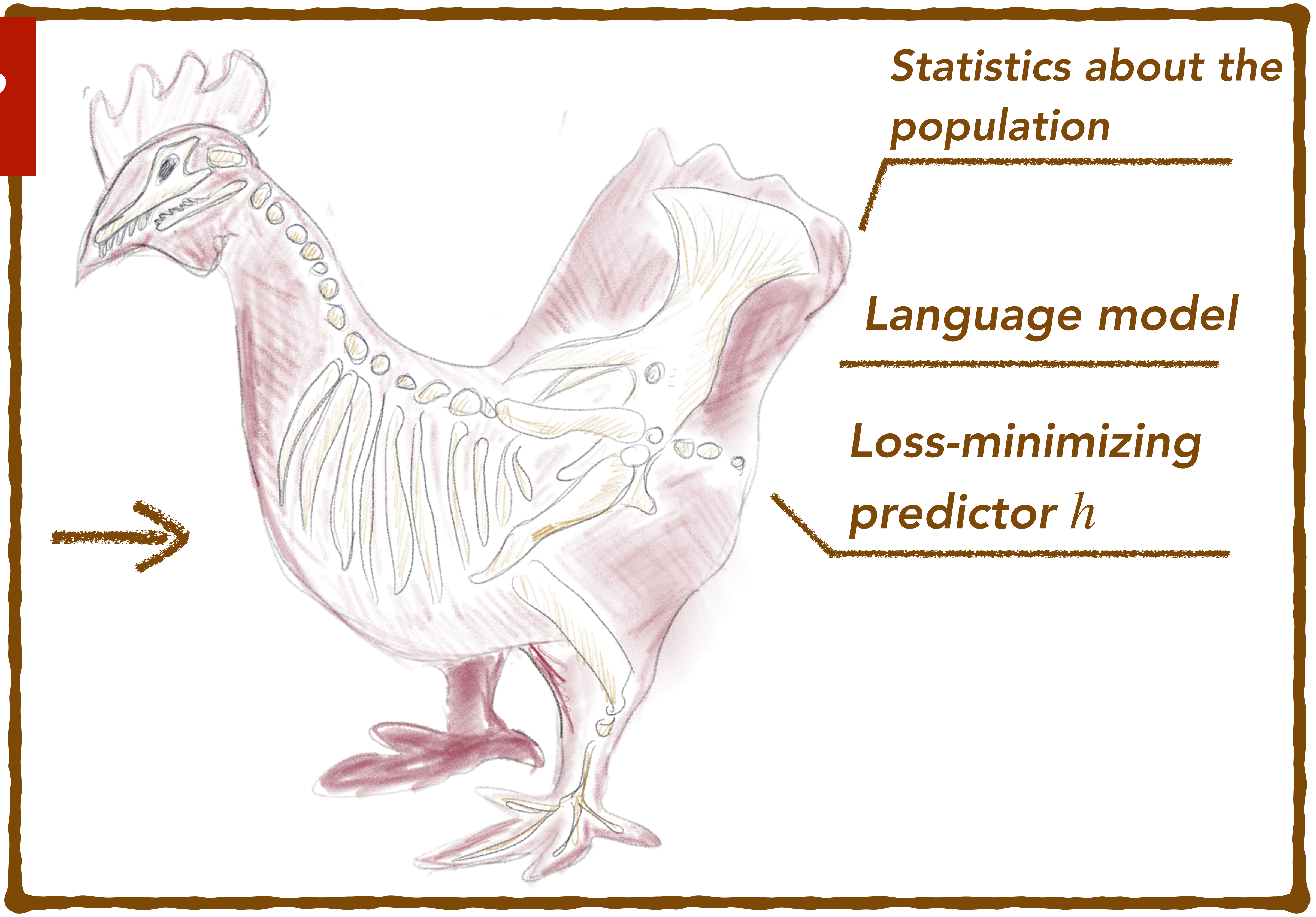
Indepedent sample from correct distribution?

Gather valuable dataset

Was it run correctly?

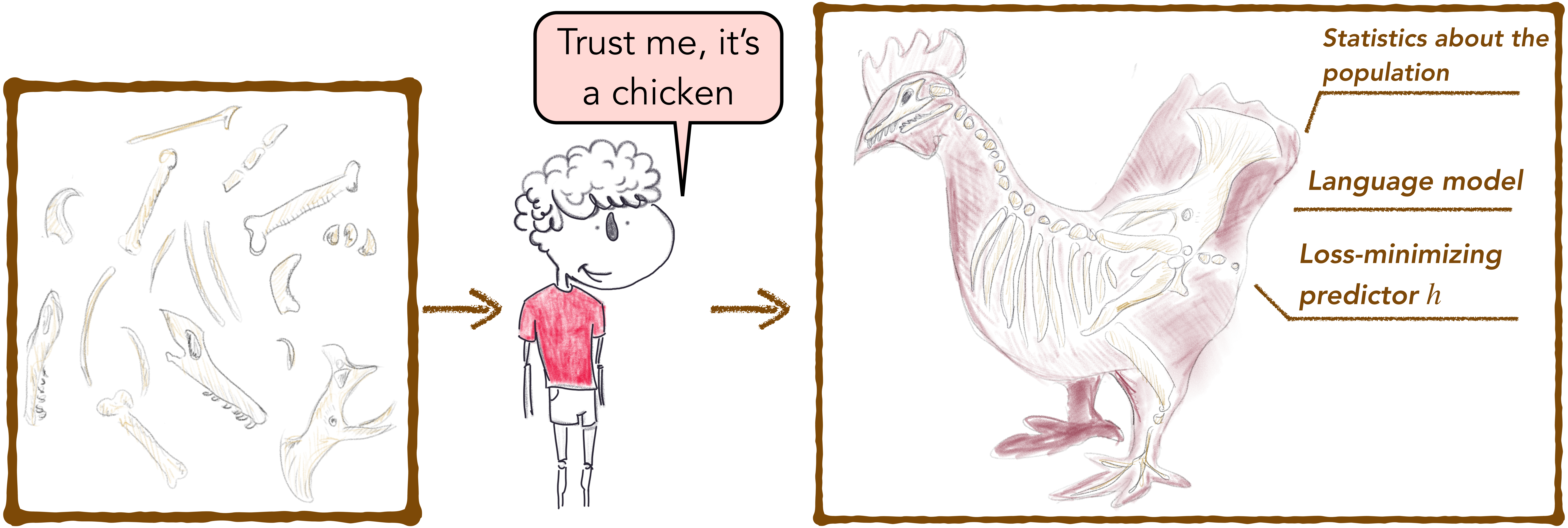


Analyze using sophisticated or extensive algorithm

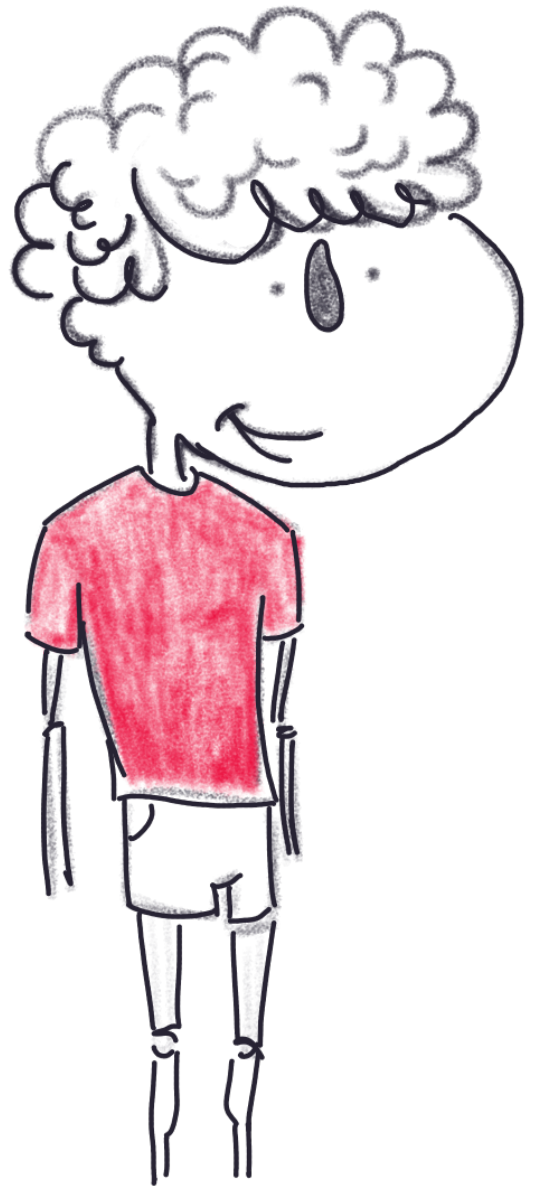


Arrive at useful conclusions

Data Science: Correct?



Can we <u>verify</u> the output was <u>correct</u>, given samples from correct distribution?

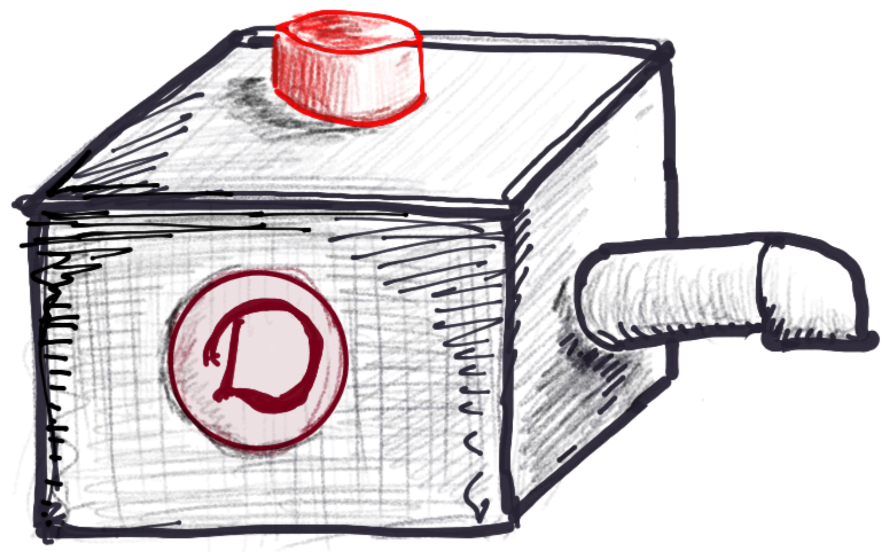


Running the algorithm $\mathscr{A}$ on many samples from D, yields a chicken

Can we *verify* claims about D, *without replication*?

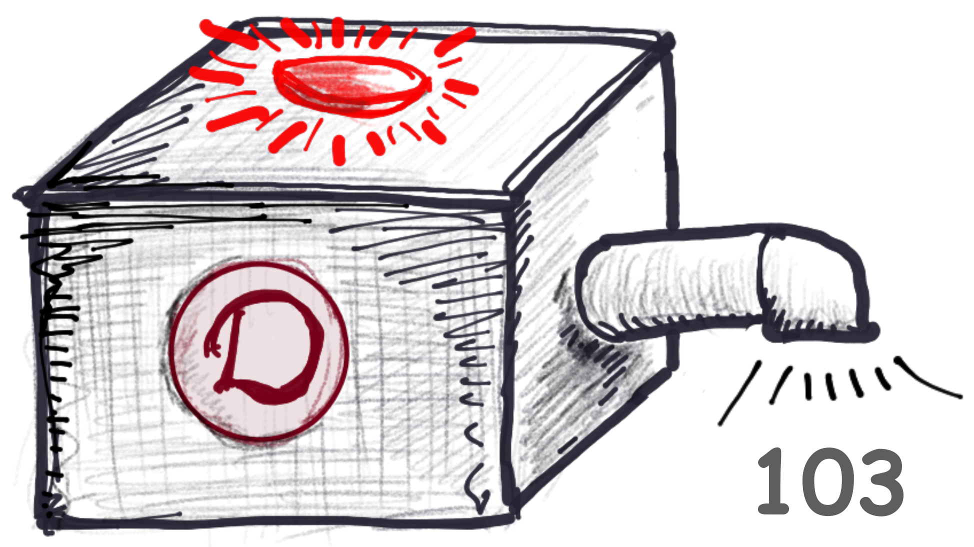
Distribution Testing [GGR'96, BFRSW'00]

Distribution D *over* $[N]$, $\varepsilon \in (0,1]$



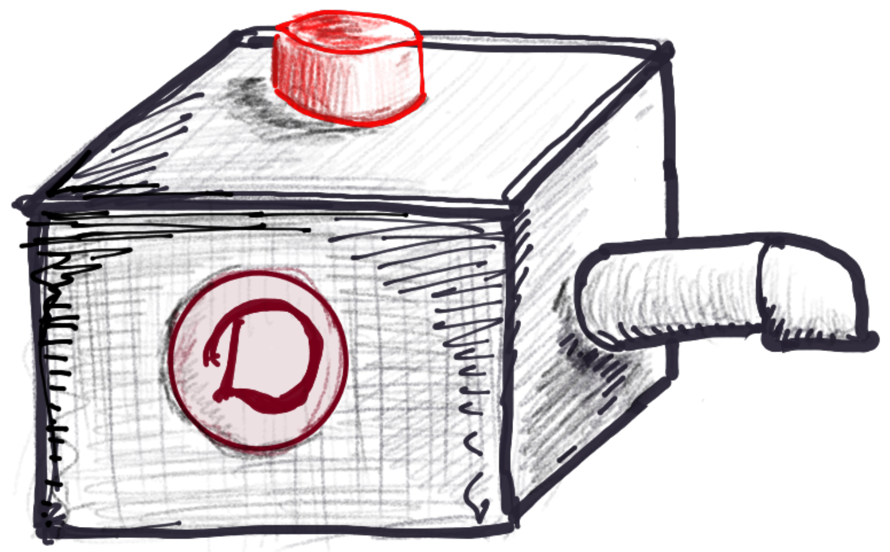
Distribution Testing [GGR'96, BFRSW'00]

Distribution D *over* $[N]$, $\varepsilon \in (0,1]$



Distribution Testing [GGR'96, BFRSW'00]

Distribution D *over* $[N]$, $\varepsilon \in (0,1]$

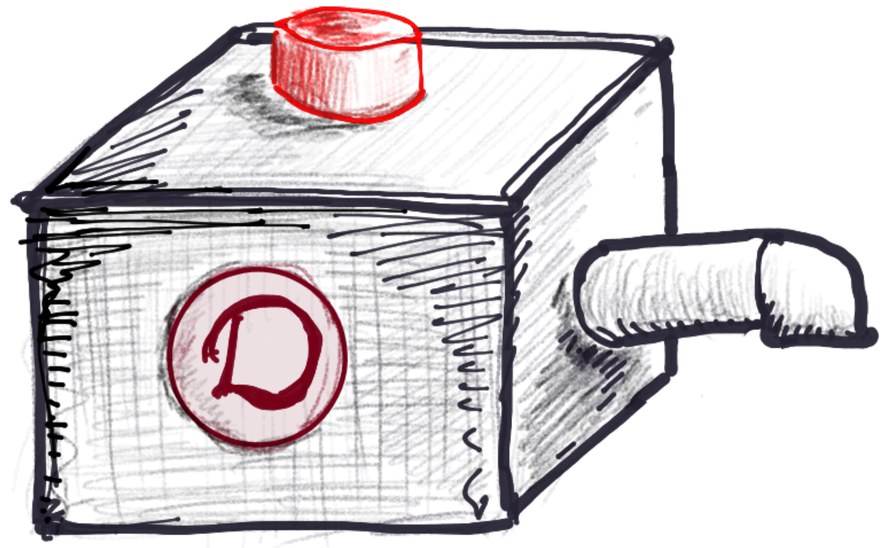


Distribution property $\mathscr{P}$ (same role as language for decision problems).

Accept if $D \in \mathscr{P}$;
reject if D is ε-far

Distribution Testing [GGR'96, BFRSW'00]

Distribution D *over* $[N]$, $\varepsilon \in (0,1]$



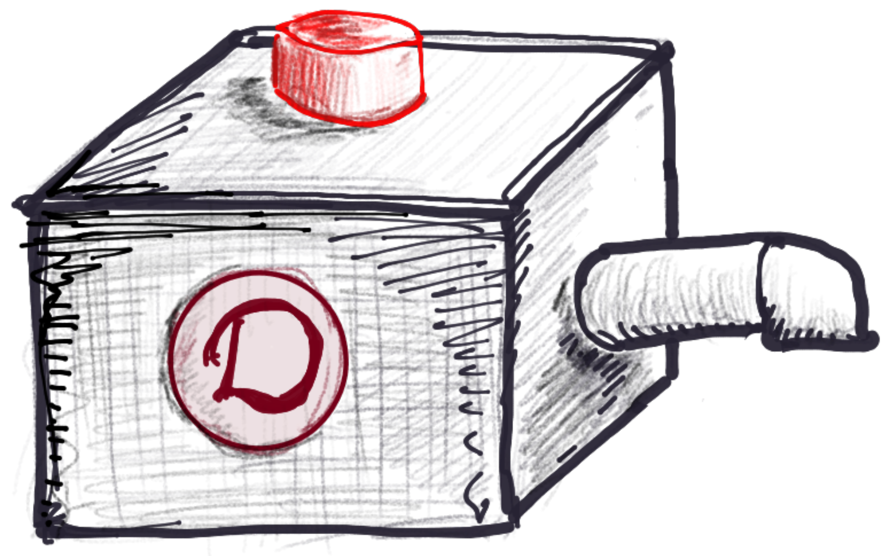
Accept if $D \in \mathscr{P}$;
reject if D is ε-far

Distribution property $\mathscr{P}$ (same role as language for decision problems). Examples:

Label invariant properties (does D have *Shannon Entropy* k? *Support size* M? *Distance* δ *from* U_N?), ***General properties*** *(can a predictor from class* $\mathscr{H}$ *have loss better than* α *over* D*).*

Distribution Testing [GGR'96, BFRSW'00]

Distribution D over $[N]$, $\varepsilon \in (0,1]$



Accept if $D \in \mathscr{P}$;
reject if D is ε-far

Distribution property $\mathscr{P}$ (same role as language for decision problems). Examples:

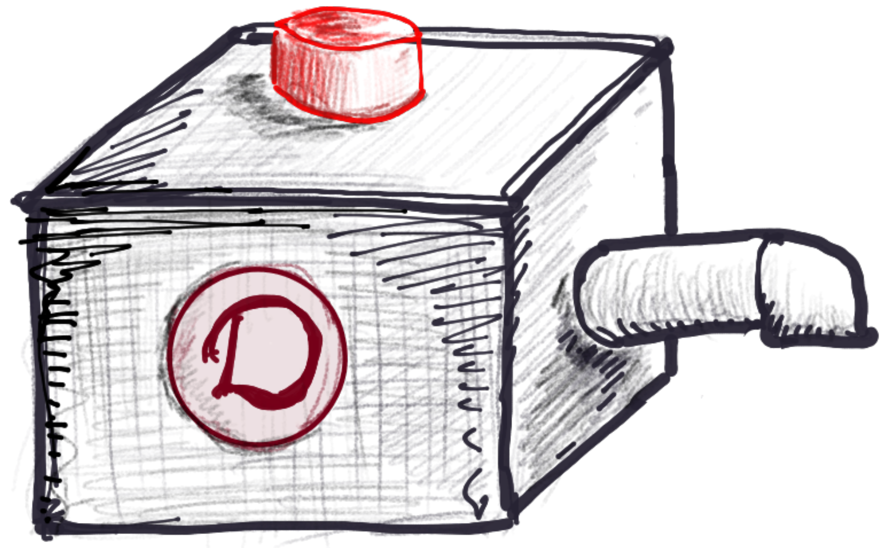
Label invariant properties (does D have *Shannon Entropy k? Support size M? Distance δ from U_N*?), ***General properties*** *(can a predictor from class $\mathscr{H}$ have loss better than α over D).*

Sample access, i.e. $x \sim D$, is **very restrictive.**

A full description would be $(x, D(x))$ for every $x \in [N]$.

Distribution Testing [GGR'96, BFRSW'00]

Distribution D *over* $[N]$, $\varepsilon \in (0,1]$



Accept if $D \in \mathscr{P}$;
reject if D is ε-far

Distribution property $\mathscr{P}$ (same role as language for decision problems). Examples:

Label invariant properties (does D have *Shannon Entropy* k? *Support size* M? *Distance* δ *from* U_N?), ***General properties*** *(can a predictor from class* $\mathscr{H}$ *have loss better than* α *over* D*).*

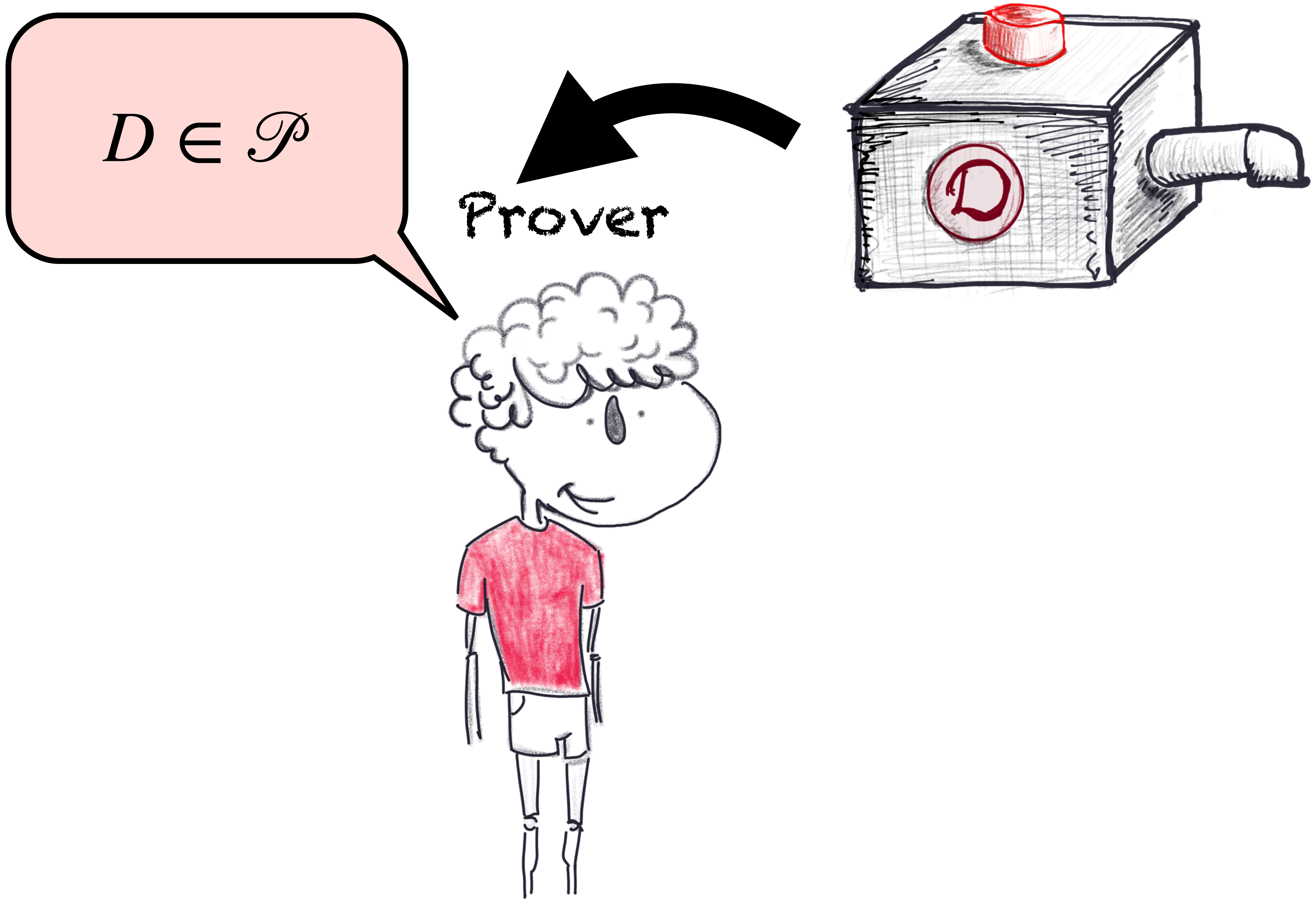
Sample access, i.e. $x \sim D$, is **very restrictive.**

A full description would be $(x, D(x))$ ***for every*** $x \in [N]$.

Indeed, testing via samples alone might be very hard, label invariant properties require $\Theta(N/\log N)$ [RRSS07, VV10]

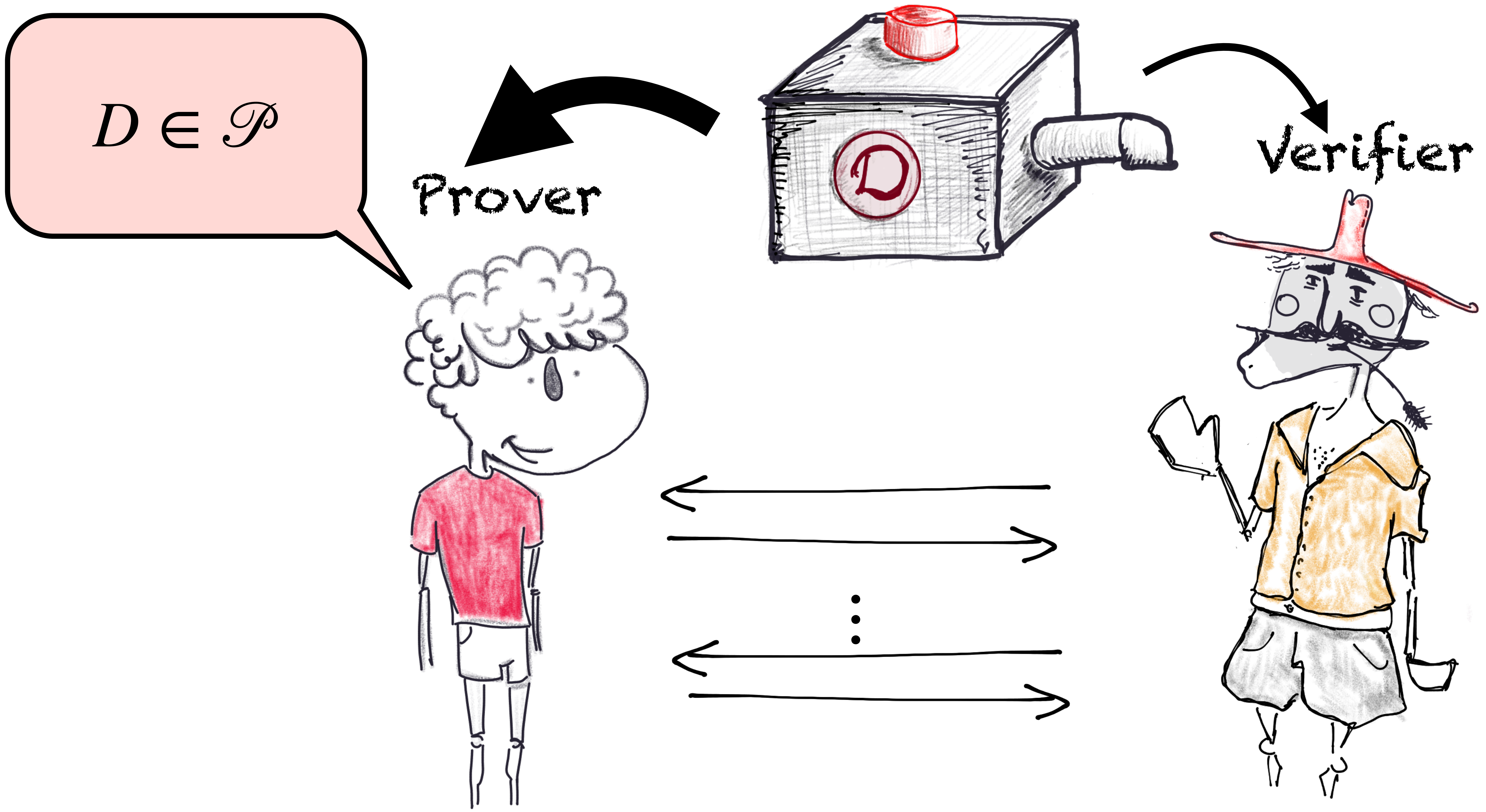
Verifying Properties of Distributions[CG'17, GMR'85]

Distribution D *over* $[N]$*,* $\varepsilon \in (0,1]$



Verifying Properties of Distributions[CG'17, GMR'85]

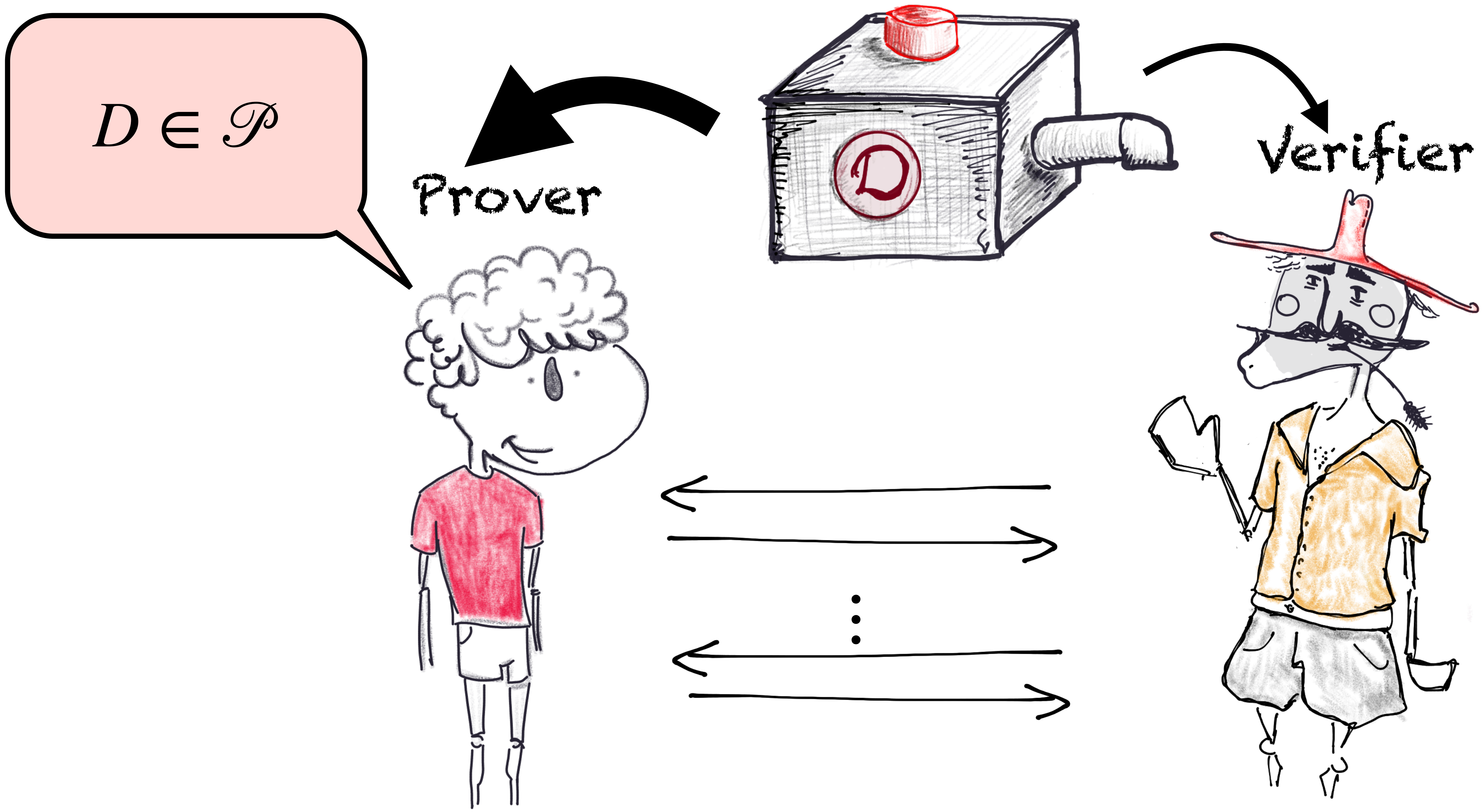
Distribution D *over* $[N]$, $\varepsilon \in (0,1]$



Accept / **Reject**

Verifying Properties of Distributions[CG'17, GMR'85]

Distribution D *over* $[N]$, $\varepsilon \in (0,1]$



Accept / ***Reject***

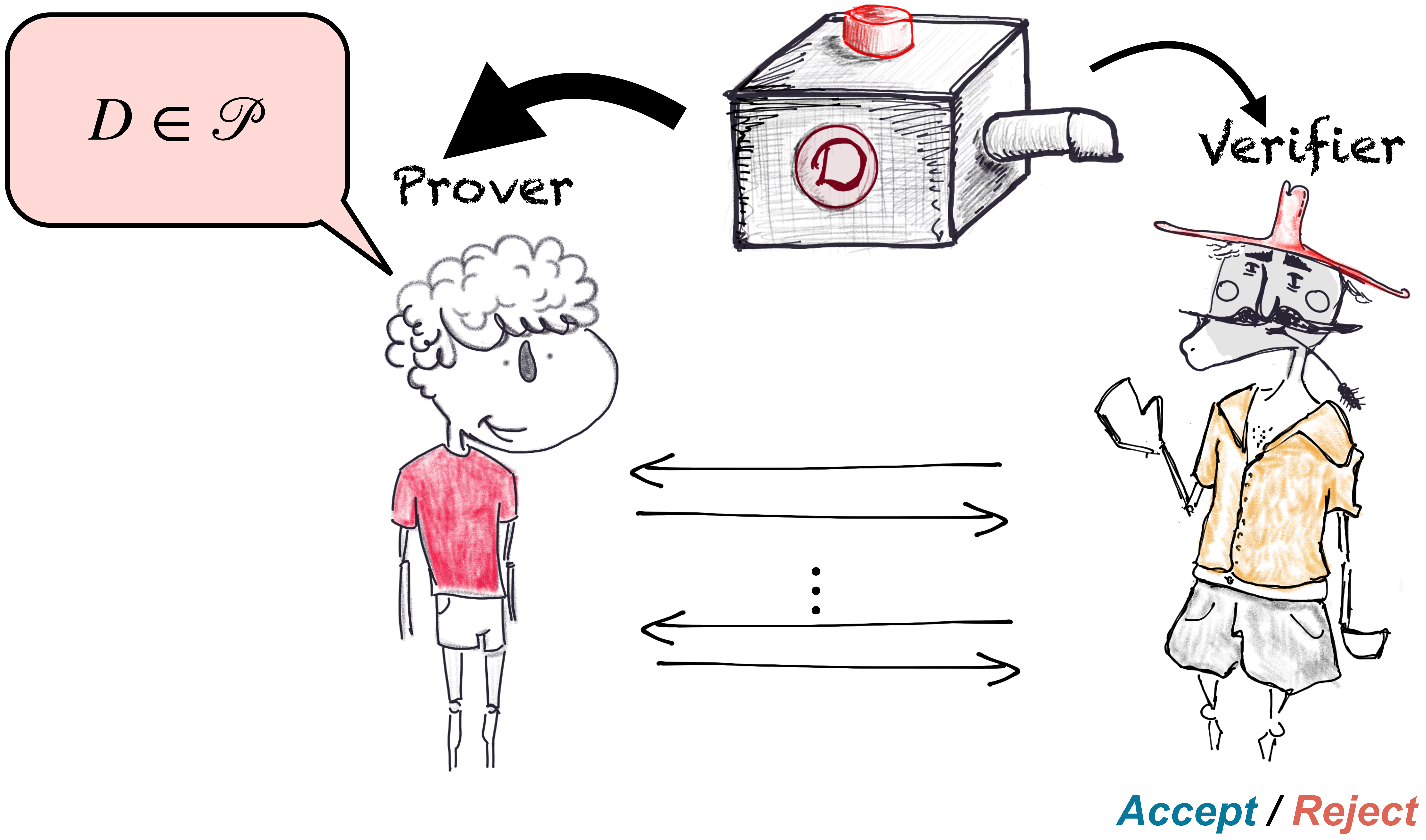
Completeness: $D \in \mathscr{P}$, V accepts w.h.p.

Soundness: if D is ε-far from satisfying $\mathscr{P}$, ∀cheating prover P^*, V rejects whp.

Tolerant verification: approximate distance to property up to ε

Verifying Properties of Distributions[CG'17, GMR'85]

Distribution D over $[N]$, $\varepsilon \in (0,1]$



Completeness: $D \in \mathscr{P}$, V accepts w.h.p.

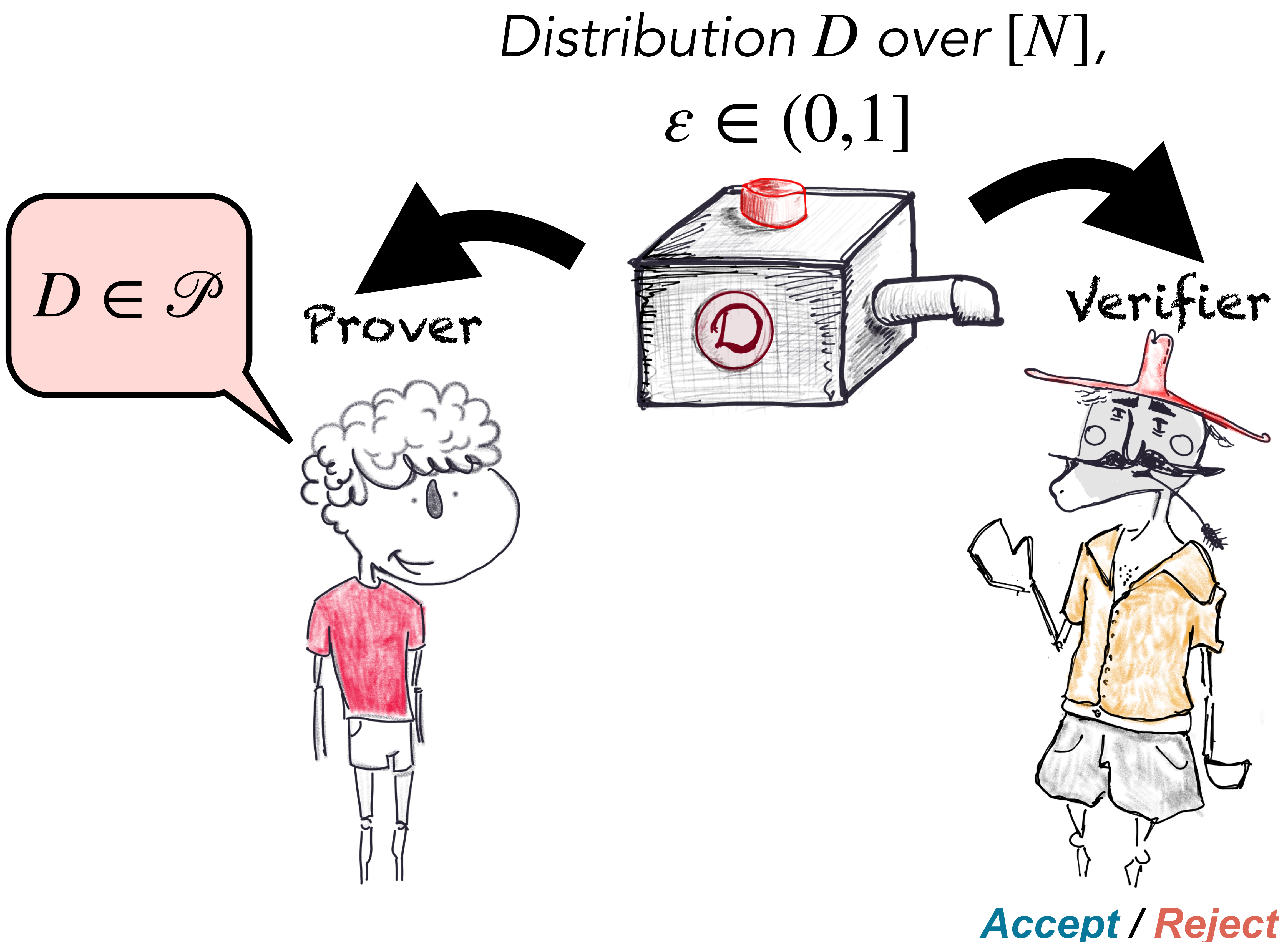
Soundness: if D is ε-far from satisfying $\mathscr{P}$, ∀cheating prover $P*$, V rejects whp.

Tolerant verification: approximate distance to property up to ε

Efficient verification: V's samples & runtime, comm., #of rounds.

Efficient proof: P's samples, P's runtime.

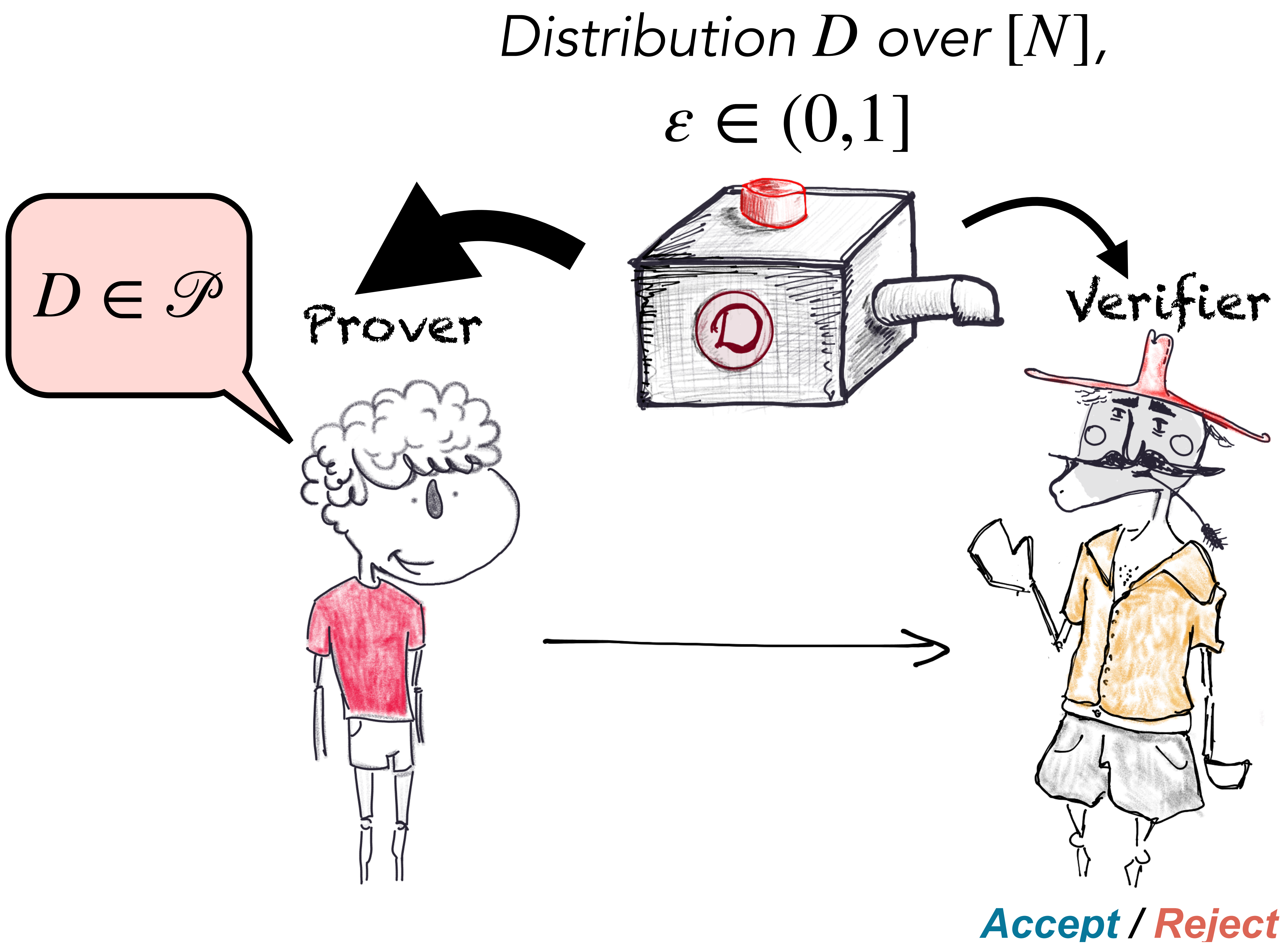
Verifying Properties of Distributions[CG'17, GMR'85]



Trivial Solutions:

1. **No communication (replication).** Verifier ignores prover, *repeats* computation or *learns* D (using $O\left(N \cdot \varepsilon^{-2}\right)$ samples.)

Verifying Properties of Distributions[CG'17, GMR'85]



Trivial Solutions:

1. **No communication (replication).** Verifier ignores prover, *repeats* computation or *learns* D (using $O\left(N \cdot \varepsilon^{-2}\right)$ samples.)

2. **High communication complexity:** P sends $(x, D(x))$ $\forall x \in [N]$, verifier "identity tests", using $O\left(\sqrt{N} \cdot \varepsilon^{-2}\right)$ samples.
   Communication Complexity: $\widetilde{\Omega}(N)$.

Results Overview (joint work with Guy Rothblum)

Ignoring poly(ε^{-1}) factors

Property Family	V Sample Complexity	Comm. Complexity	# of Messages	Honest P Sample Complexity*	Notes
Label-invariant	$\widetilde{O}\left(N^{1/2} \cdot \varepsilon^{-4}\right)$	$\widetilde{O}\left(N^{1/2} \cdot \varepsilon^{-4}\right)$	2	$\widetilde{O}(N)$	

Results Overview (joint work with Guy Rothblum)

*Ignoring poly(ε^{-1}) factors

Property Family	V Sample Complexity	Comm. Complexity	# of Messages	Honest P Sample Complexity*	Notes
Label-invariant	$\widetilde{O}\left(N^{1/2} \cdot \varepsilon^{-4}\right)$	$\widetilde{O}\left(N^{1/2} \cdot \varepsilon^{-4}\right)$	2	$\widetilde{O}(N)$	
Label-invariant	$\widetilde{O}\left(N^{2/3} \cdot \varepsilon^{-6}\right)$	$\widetilde{O}\left(N^{2/3} \cdot \varepsilon^{-6}\right)$	2	$\widetilde{O}(N)$	*Public-coin*

Results Overview (joint work with Guy Rothblum)

*Ignoring poly(ε^{-1}) factors

Property Family	V Sample Complexity	Comm. Complexity	# of Messages	Honest P Sample Complexity*	Notes
Label-invariant	$\widetilde{O}\left(N^{1/2} \cdot \varepsilon^{-4}\right)$	$\widetilde{O}\left(N^{1/2} \cdot \varepsilon^{-4}\right)$	*2*	$\widetilde{O}\left(N\right)$	
Label-invariant	$\widetilde{O}\left(N^{2/3} \cdot \varepsilon^{-6}\right)$	$\widetilde{O}\left(N^{2/3} \cdot \varepsilon^{-6}\right)$	*2*	$\widetilde{O}\left(N\right)$	***Public-coin***
Given complete description $(x, D(x))$, TV($D, \mathscr{P}$) approximated by ***low depth circuit / low space TM***	$\widetilde{O}\left(N^{0.9}\right) \cdot \text{poly}(\varepsilon^{-1})$	$\widetilde{O}\left(N^{0.95}\right) \cdot \text{poly}(\varepsilon^{-1})$	*polylog(N)*	$\widetilde{O}\left(N^{1.1}\right)$	

Results Overview (joint work with Guy Rothblum)

*Ignoring poly(ε^{-1}) factors

Property Family	V Sample Complexity	Comm. Complexity	# of Messages	Honest P Sample Complexity*	Notes
Label-invariant	$\widetilde{O}\left(N^{1/2} \cdot \varepsilon^{-4}\right)$	$\widetilde{O}\left(N^{1/2} \cdot \varepsilon^{-4}\right)$	2	$\widetilde{O}\left(N\right)$	
Label-invariant	$\widetilde{O}\left(N^{2/3} \cdot \varepsilon^{-6}\right)$	$\widetilde{O}\left(N^{2/3} \cdot \varepsilon^{-6}\right)$	2	$\widetilde{O}\left(N\right)$	**Public-coin**
*Given complete description $(x, D(x))$, TV$(D, \mathscr{P})$ approximated by **low depth circuit / low space TM***	$\widetilde{O}\left(N^{0.9}\right) \cdot \text{poly}(\varepsilon^{-1})$	$\widetilde{O}\left(N^{0.95}\right) \cdot \text{poly}(\varepsilon^{-1})$	*polylog(N)*	$\widetilde{O}\left(N^{1.1}\right)$	
*Given complete description $(x, D(x))$, TV$(D, \mathscr{P})$ approximated by **poly-time algorithm***	$O\left(N^{1/2} \cdot \varepsilon^{-2}\right)$	$\widetilde{O}\left(N^{1/2} \cdot \varepsilon^{-2}\right)$	4	$\widetilde{O}\left(N\right)$	**Assuming CRH**

Results Overview (joint work with Guy Rothblum)

*Ignoring poly(ε^{-1}) factors

Property Family	V Sample Complexity	Comm. Complexity	# of Messages	Honest P Sample Complexity*	Notes
Label-invariant	$\widetilde{O}\left(N^{1/2}\cdot\varepsilon^{-4}\right)$	$\widetilde{O}\left(N^{1/2}\cdot\varepsilon^{-4}\right)$	2	$\widetilde{O}(N)$	
Label-invariant	$\widetilde{O}\left(N^{2/3}\cdot\varepsilon^{-6}\right)$	$\widetilde{O}\left(N^{2/3}\cdot\varepsilon^{-6}\right)$	2	$\widetilde{O}(N)$	***Public-coin***
*Given complete description $(x, D(x))$, TV$(D,\mathscr{P})$ approximated by **low depth circuit / low space TM***	$\widetilde{O}\left(N^{0.9}\right)\cdot\text{poly}(\varepsilon^{-1})$	$\widetilde{O}\left(N^{0.95}\right)\cdot\text{poly}(\varepsilon^{-1})$	polylog(N)	$\widetilde{O}\left(N^{1.1}\right)$	
*Given complete description $(x, D(x))$, TV$(D,\mathscr{P})$ approximated by **poly-time algorithm***	$O\left(N^{1/2}\cdot\varepsilon^{-2}\right)$	$\widetilde{O}\left(N^{1/2}\cdot\varepsilon^{-2}\right)$	4	$\widetilde{O}(N)$	***Assuming CRH***

Results Overview (joint work with Guy Rothblum)

**Ignoring poly(ε^{-1}) factors*

Property Family	V Sample Complexity	Comm. Complexity	# of Messages	Honest P Sample Complexity*	Notes
Given c TV(D dep					
Given c TV(D,					

My goals for the first hour:

Show that with sample access ***and*** *communication with a prover we can do many unexpected things!*

Demonstrate tools and ideas.

questions to have in mind: *what assumptions over the distribution can help us? What settings might this capture?*

Verified Tagged Sample Protocol

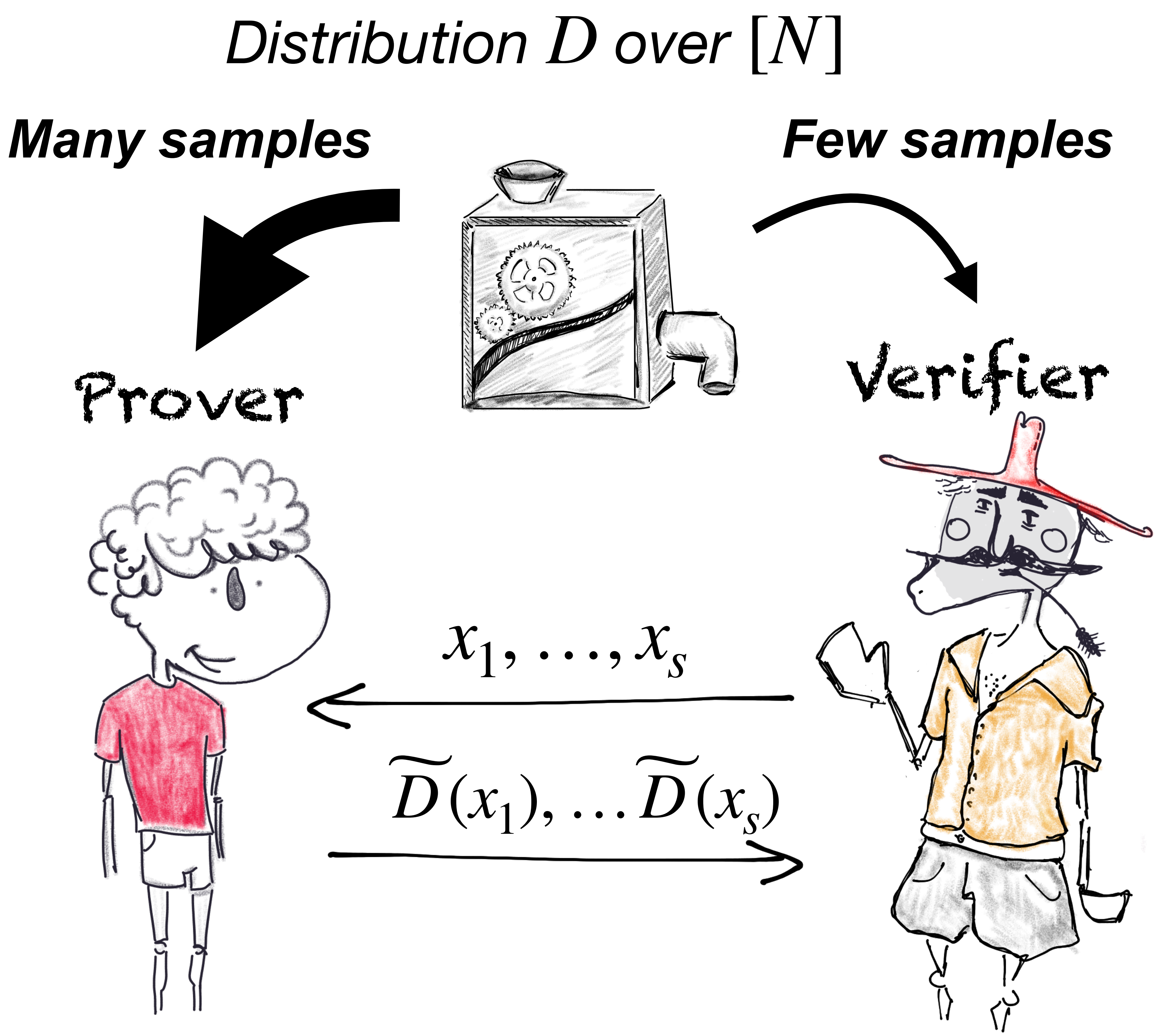
Goal: *verifiably obtain a tagged sample, i.e.* $(x, D(x))$ *where* $x \sim D$ *i.i.d. , while requiring* $\ll N$ samples.

Verified Tagged Sample Protocol

Goal: *verifiably obtain a tagged sample, i.e.* $(x, D(x))$ *where* $x \sim D$ *i.i.d. , while requiring* $\ll N$ samples.

- **Very useful:**
  - Approximate the probability histogram. *E.g. half the samples have probability* $2/N \implies$ *≈0.5 mass of* D *on elements w.p,* $2/N \implies$ *there are (approx.)* $N/4$ *elements with probability* $2/N$.
  - *Approximate distance from fixed distribution* Q *given explicitly.*
  - *More…*

Verified Tagged Sample Protocol



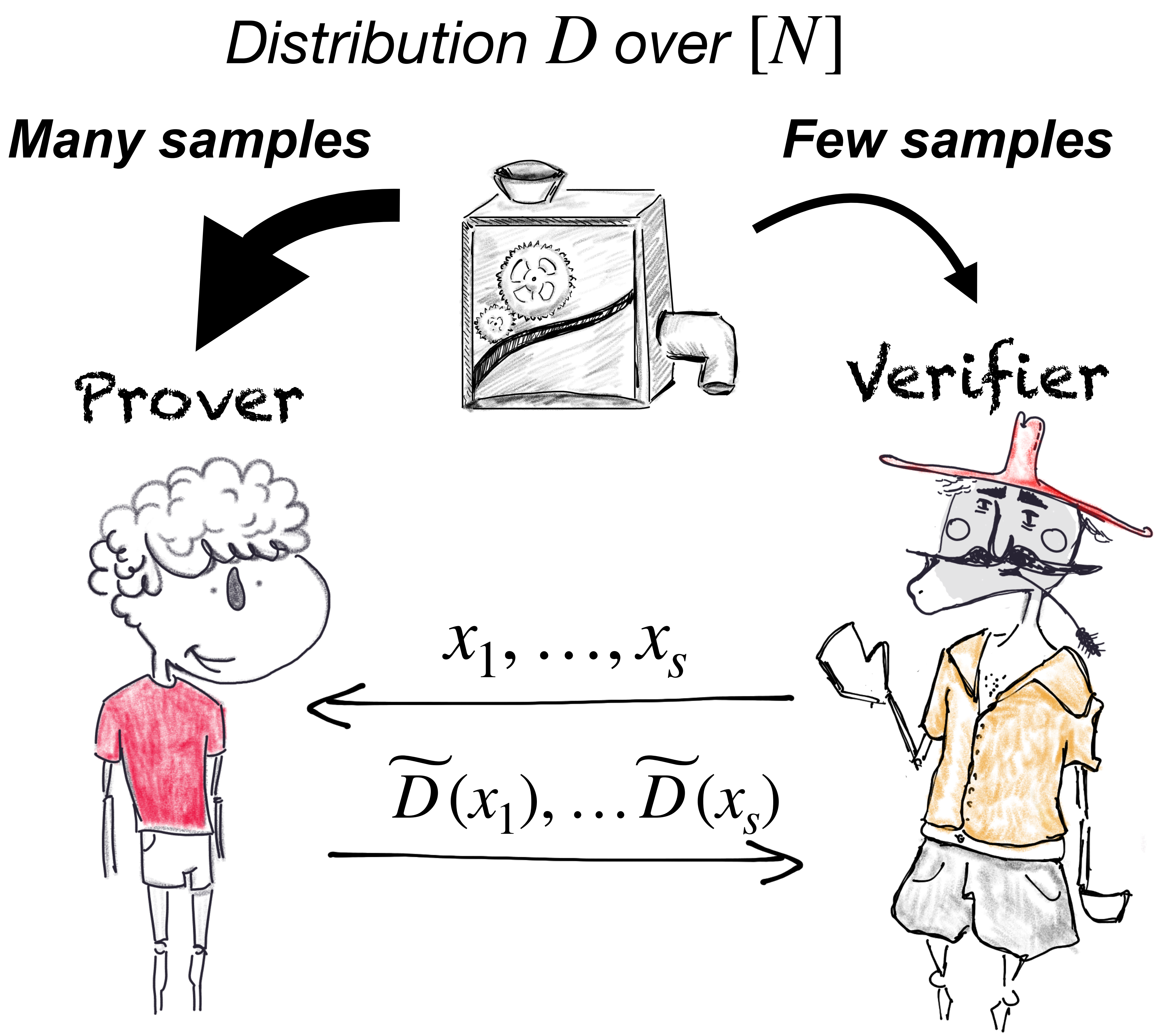
Goal:

Completeness: if $\widetilde{D}(x_i) = D(x_i) \rightarrow$ w.h.p. V accepts.

Soundness: $\forall P^*$ w.h.p. either V rejects or outputs a collection of (x, π_x), where $\pi_x \approx D(x)$ for almost all samples x.

Verifier samp. complexity, runtime $\widetilde{O}\left(N^{1/2}\right)$.

Verified Tagged Sample Protocol



Goal:

Completeness: if $\widetilde{D}(x_i) = D(x_i) \rightarrow$ w.h.p. V accepts.

Soundness: $\forall P^*$ w.h.p. either V rejects or outputs a collection of (x, π_x), where $\pi_x \approx D(x)$ for almost all samples x.

Verifier samp. complexity, runtime $\widetilde{O}\left(N^{1/2}\right)$.

Simplifying assumption: $D(x) \in \left\{\frac{1}{N}, \frac{2}{N}, \ldots, \frac{100}{N}\right\}$

Obtaining Verified *Tagged Sample*

Obtaining Verified *Tagged Sample*

Set $s = \widetilde{\theta}\left(\sqrt{N}\right)$***, repeat*** $2s$ ***times:***

1. Flip fair coin b*.*

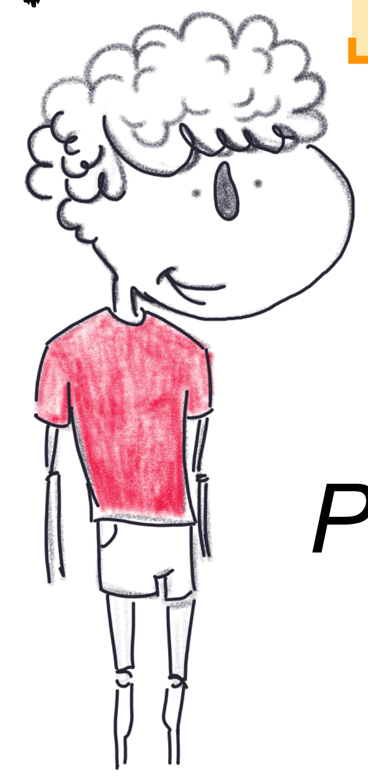
2. If $b = 0$*, draw* $x \sim D$*; else* $x \sim U_{[N]}$*.*

Obtaining Verified *Tagged Sample*

V's secret

$\{i : b_i = 0\},\ \{i : b_i = 1\}$

P



V

$(x_1, x_2, \ldots x_{2s})$ $\longleftarrow$

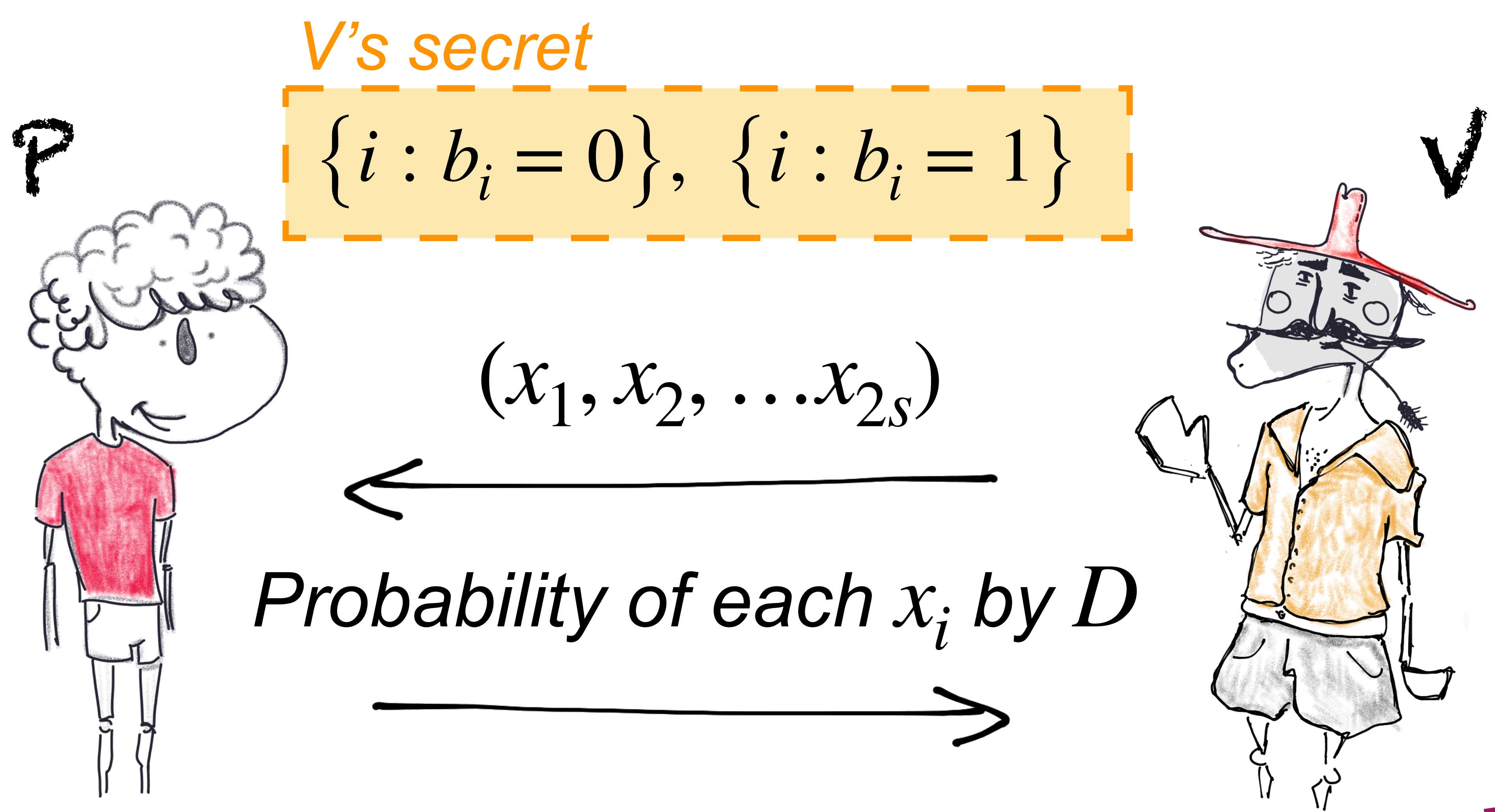
Probability of each x_i by D $\longrightarrow$

Set $s = \widetilde{\theta}\left(\sqrt{N}\right)$, repeat $2s$ times:

1. Flip fair coin b.

2. If $b = 0$, draw $x \sim D$; else $x \sim U_{[N]}$.

Obtaining Verified *Tagged Sample*

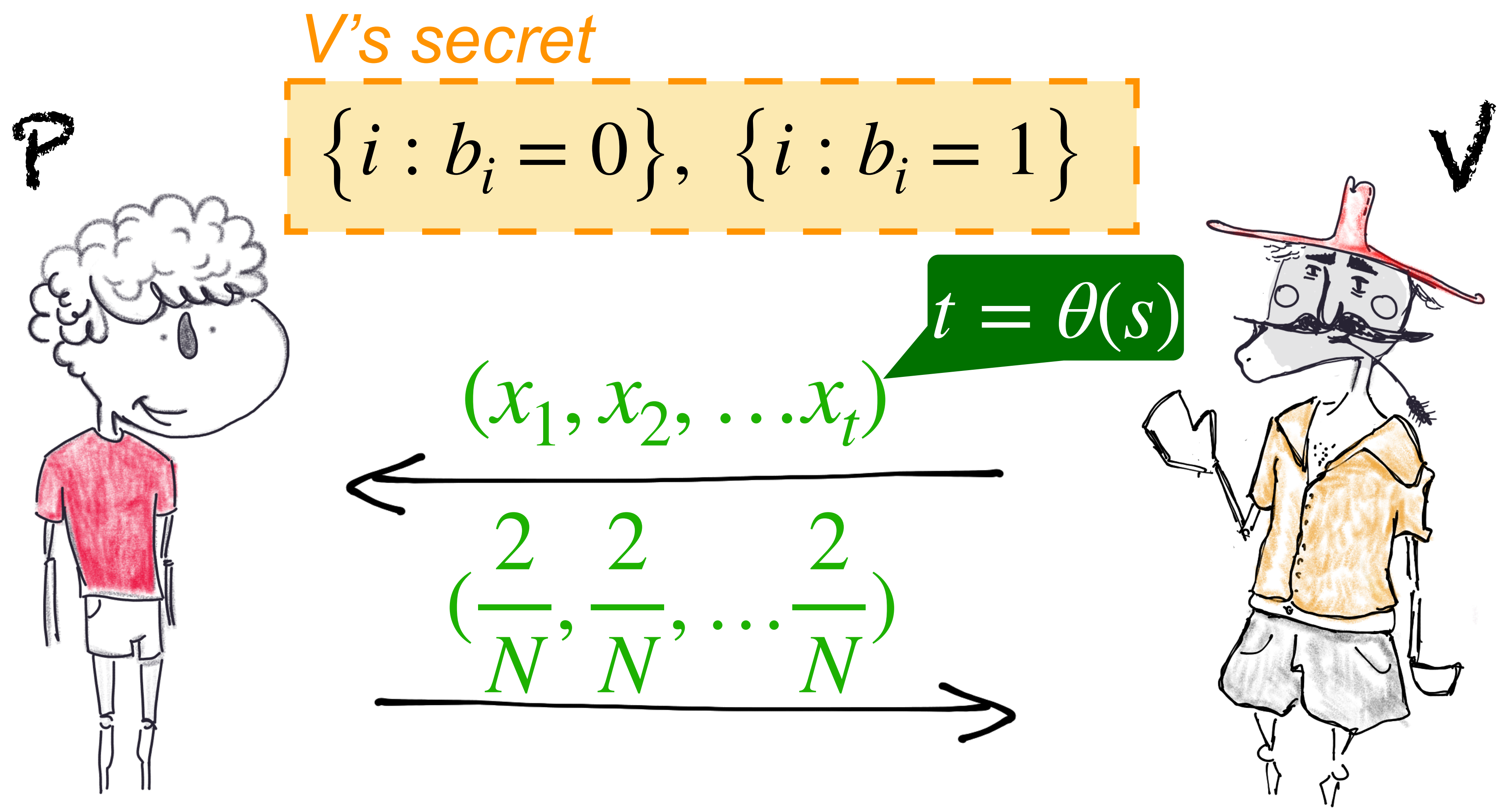


Set $s = \widetilde{\theta}\left(\sqrt{N}\right)$**, repeat** $2s$ **times:**

1. Flip fair coin b.
2. If $b = 0$, draw $x \sim D$; else $x \sim U_{[N]}$.

If P is honest, V obtained a correct tagged sample

Obtaining Verified *Tagged Sample*



Set $s = \widetilde{\theta}\left(\sqrt{N}\right)$***, repeat*** $2s$ ***times:***

1. *Flip fair coin* b.
2. *If* $b = 0$, *draw* $x \sim D$; *else* $x \sim U_{[N]}$.

Obtaining Verified *Tagged Sample*

V's secret

$\{i : b_i = 0\},\ \{i : b_i = 1\}$

$(x_1, x_2, \ldots x_t)$

$(\frac{2}{N}, \frac{2}{N}, \ldots \frac{2}{N})$

Set** $s = \tilde{\theta}\left(\sqrt{N}\right)$**, repeat** $2s$ **times:

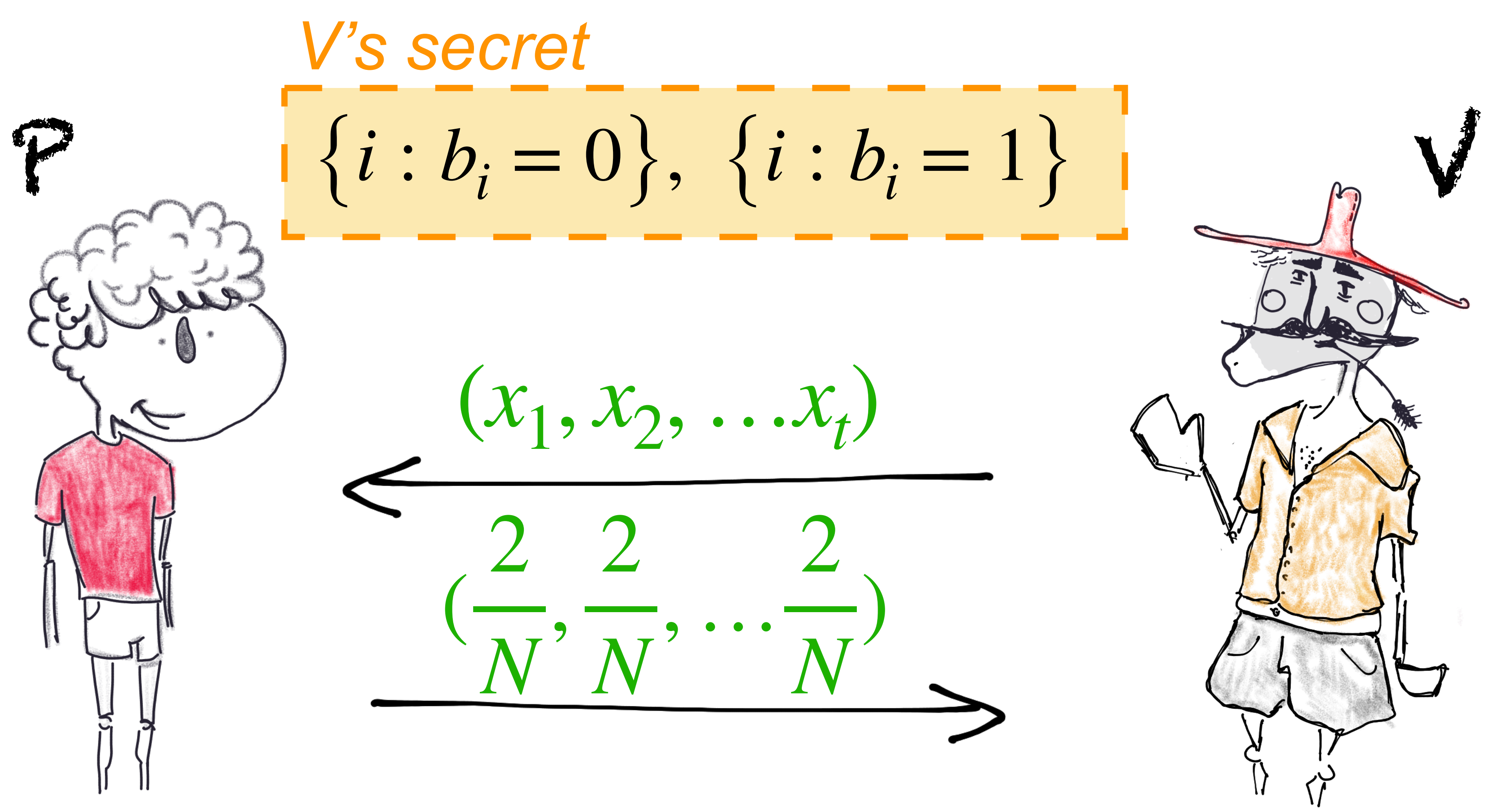
1. *Flip fair coin b.*
2. *If $b = 0$, draw $x \sim D$; else $x \sim U_{[N]}$.*

Verifier tests

Look e.g. at the samples with alleged prob.** $2/N$**:

1. ***Consistency:** (# of D-samples)* = 2 · *(# of U-samples)*
2. ***Correct avg. probability:** total mass is* $\frac{2}{N}$ · *(total # of samples)*

Obtaining Verified *Tagged Sample*



Set $s = \widetilde{\theta}\left(\sqrt{N}\right)$***, repeat*** $2s$ ***times:***

1. Flip fair coin b*.*

2. If $b = 0$*, draw* $x \sim D$*; else* $x \sim U_{[N]}$*.*

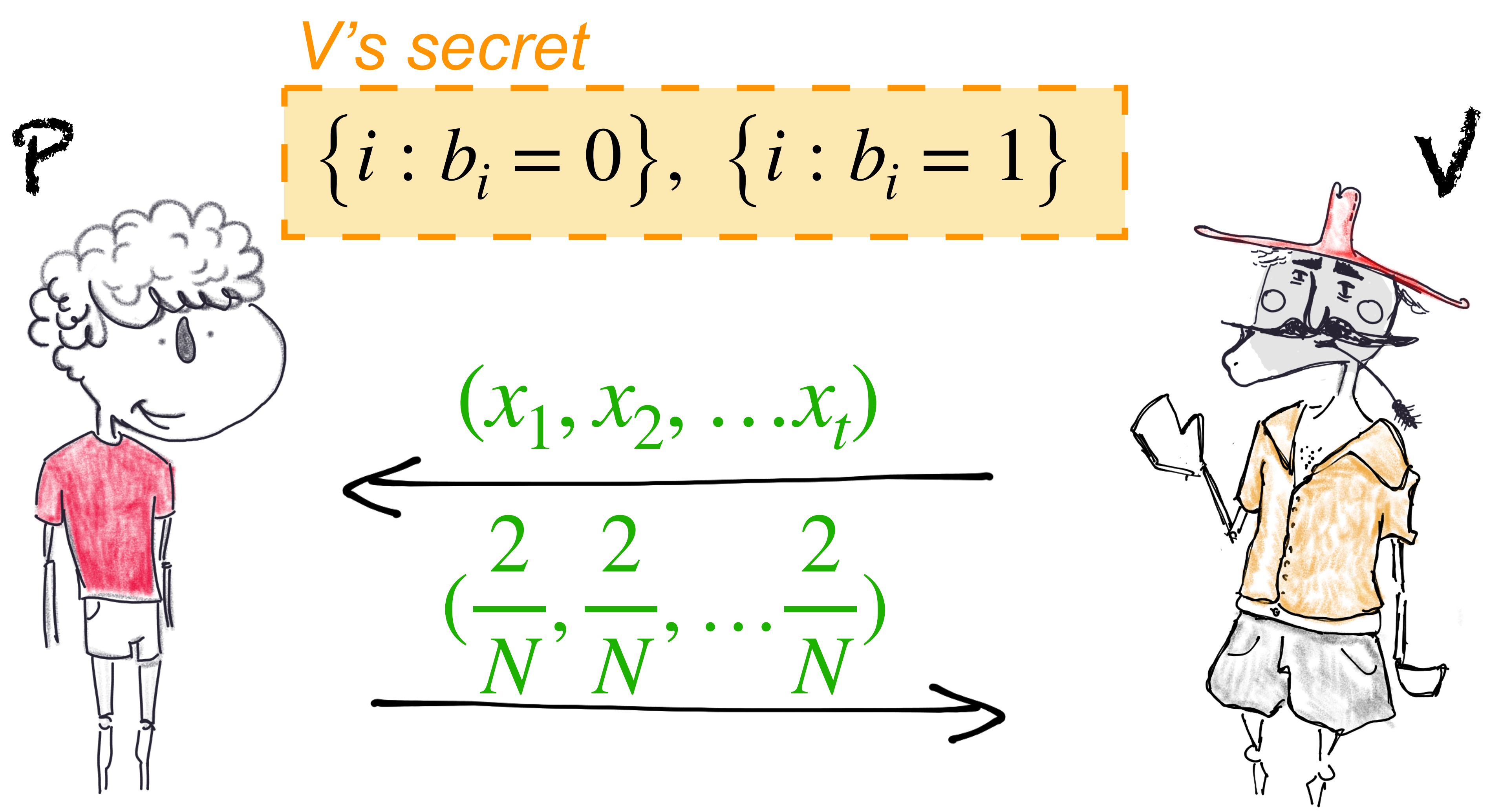
Verifier tests

Look e.g. at the samples with alleged prob. $2/N$**:**

1. ***Consistency:*** *(# of D-samples)* = 2 · *(# of U-samples)* t

2. ***Correct avg. probability:*** *total mass is* $\frac{2}{N}$ · *(total # of samples)*

Obtaining Verified *Tagged Sample*



Set $s = \tilde{\theta}\left(\sqrt{N}\right)$, repeat $2s$ times:

1. Flip fair coin b.

2. If $b = 0$, draw $x \sim D$; else $x \sim U_{[N]}$.

Verifier tests

Look e.g. at the samples with alleged prob. $2/N$:

1. ***Consistency:*** *(# of D-samples)* = 2 · *(# of U-samples)*

2. ***Correct avg. probability:*** *total mass is* $\frac{2}{N}$ · *(total # of samples)*

Draw s fresh D-samples, and check $s\left(t \cdot \frac{2}{N}\right)$ of them have alleged prob. $2/N$.

Obtaining Verified *Tagged Sample*

V's secret

$\{i : b_i = 0\},\ \{i : b_i = 1\}$

$(x_1, x_2, \ldots x_t)$

$(\frac{2}{N}, \frac{2}{N}, \ldots \frac{2}{N})$

Set** $s = \tilde{\theta}\left(\sqrt{N}\right)$**, repeat** $2s$ **times:

1. Flip fair coin b.

2. If $b = 0$, draw $x \sim D$; else $x \sim U_{[N]}$.

Verifier tests

Look e.g. at the samples with alleged prob.** $2/N$**:

1. ***Consistency:*** *(# of D-samples)* = 2 · *(# of U-samples)*

2. ***Correct avg. probability:*** *total mass is* $\frac{2}{N} \cdot$ *(total # of samples)*

Draw s fresh D-samples, and check $s\left(t \cdot \frac{2}{N}\right)$ of them have alleged prob. $2/N$.

Do this for all $\left\{\frac{1}{N}, \frac{2}{N}, \ldots, \frac{100}{N}\right\}$

Obtaining Verified *Tagged Sample*

V's secret

$\{i : b_i = 0\},\ \{i : b_i = 1\}$

$(x_1, x_2, \ldots x_t)$

$(\frac{2}{N}, \frac{2}{N}, \ldots \frac{2}{N})$

Set $s = \tilde{\theta}\left(\sqrt{N}\right)$***, repeat*** $2s$ ***times:***

1. Flip fair coin b.

2. If $b = 0$*, draw* $x \sim D$*; else* $x \sim U_{[N]}$.

Completeness

Verifier tests

Look e.g. at the samples with alleged prob. $2/N$***:***

1. ***Consistency:*** *(# of D-samples)* = $2 \cdot$ *(# of U-samples)*

2. ***Correct avg. probability:*** *total mass is* $\frac{2}{N} \cdot$ *(total # of samples)*

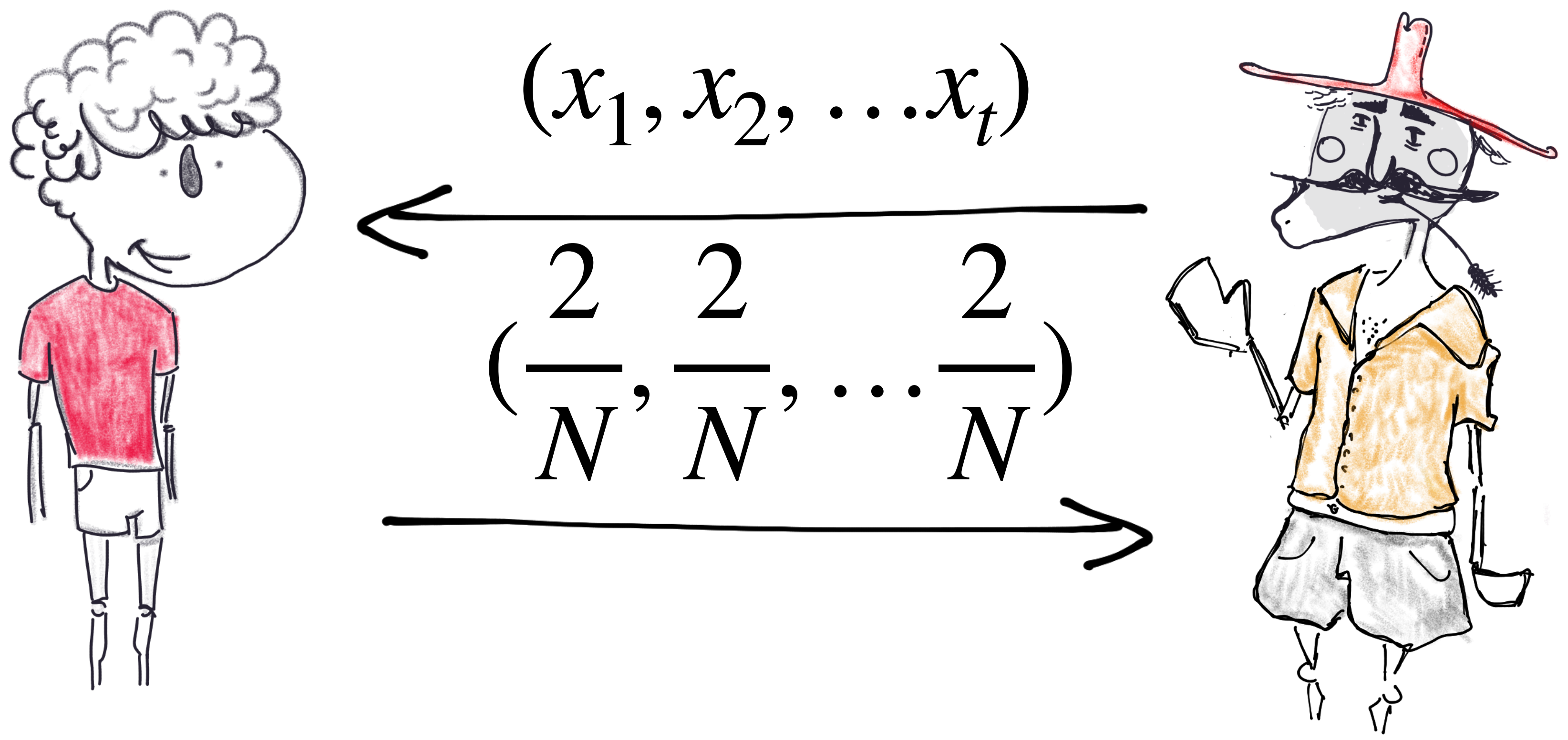
Draw s fresh D-samples, and check $s\left(t \cdot \frac{2}{N}\right)$ of them have alleged prob. $2/N$.

Soundness Analysis

Look at the t ***samples tagged*** $2/N$***:***

1. ***Consistency:*** *(# of D-samples) = 2 · (# of U-samples)*

2. ***Correct avg. probability:*** *total mass is* $\frac{2}{N}$ *· (total # of samples)*

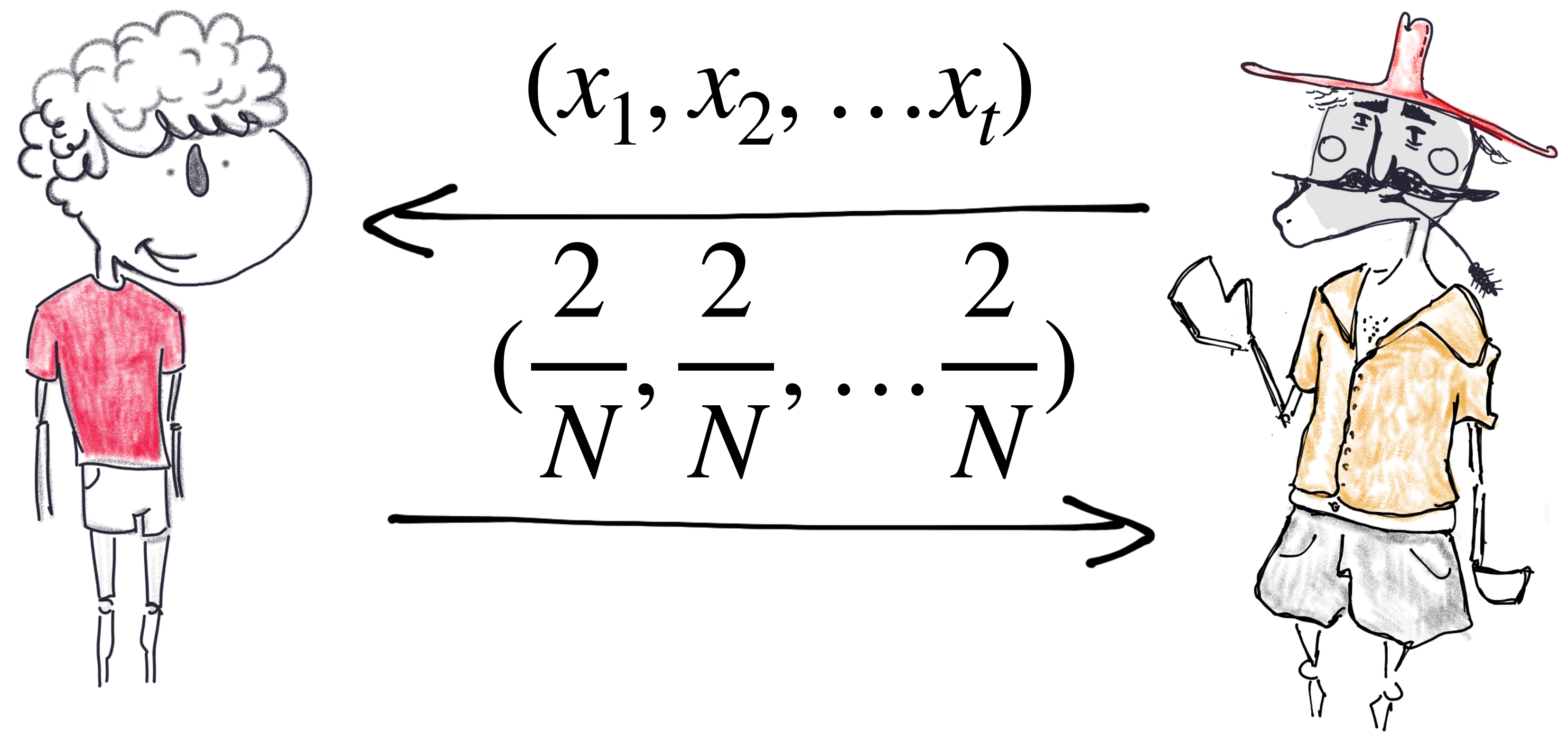


How to Produce $((b_1, x_1), \ldots (b_{2s}, x_{2s}))$***?***

Soundness Analysis

Look at the t samples tagged $2/N$:

1. **Consistency:** *(# of D-samples)* = 2 · *(# of U-samples)*
2. **Correct avg. probability:** *total mass is* $\frac{2}{N}$ · *(total # of samples)*



How to Produce $((b_1, x_1), \ldots (b_{2s}, x_{2s}))$**?**

I, "b then x"

1. *Draw* b
2. *If* $b = 0$, $x \sim D$; *o.w.* $x \sim U_{[N]}$

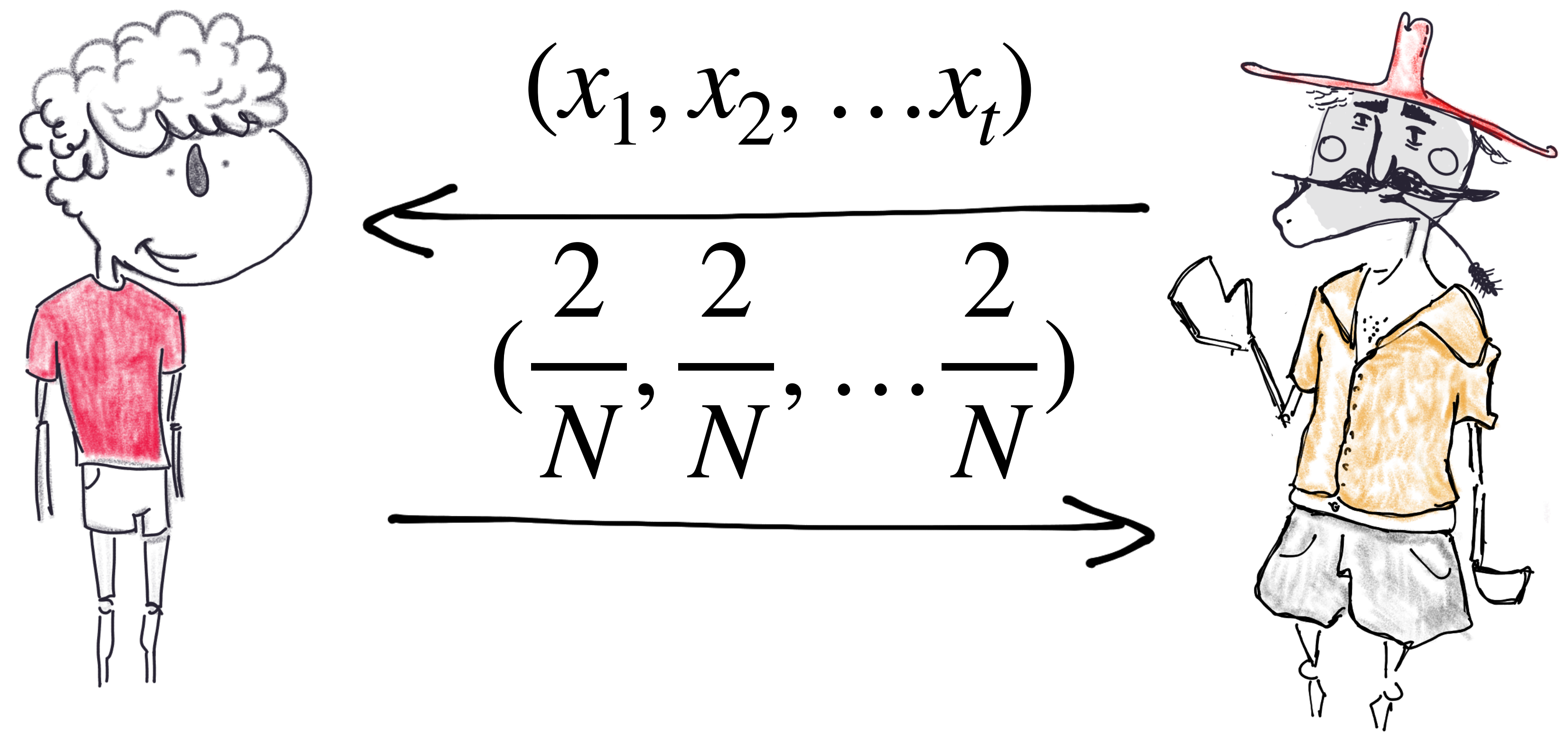
II, "x then b"

1. *Draw* $x \sim \frac{1}{2}D + \frac{1}{2}U_{[N]}$
2. $\forall x$, $b|_x = 0$, *w.p.* $\frac{D(x)}{D(x) + 1/N}$; *o.w.* $b|_x = 1$

Soundness Analysis

Look at the t samples tagged $2/N$:

1. **Consistency:** (# of D-samples) = 2 · (# of U-samples)
2. **Correct avg. probability:** total mass is $\frac{2}{N}$ · (total # of samples)



How to Produce $((b_1, x_1), \ldots (b_{2s}, x_{2s}))$***?***

I, "b then x"

1. *Draw* b
2. *If* $b = 0$, $x \sim D$; *o.w.* $x \sim U_{[N]}$

$\equiv$

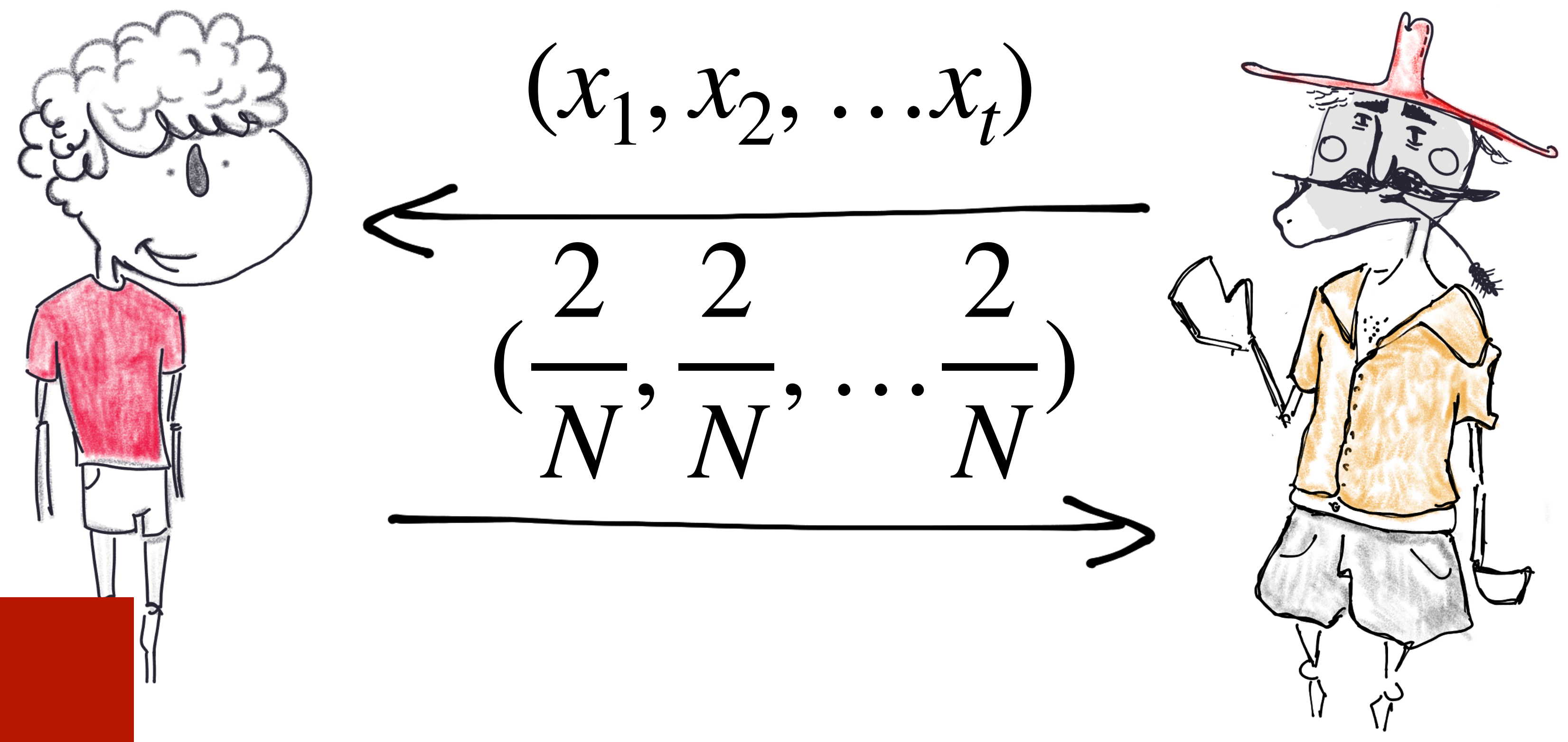
II, "x then b"

1. *Draw* $x \sim \frac{1}{2}D + \frac{1}{2}U_{[N]}$
2. $\forall x$, $b|_x = 0$, *w.p.* $\frac{D(x)}{D(x) + 1/N}$; *o.w.* $b|_x = 1$

Soundness Analysis

Look at the t samples tagged $2/N$:

1. **Consistency:** *(# of D-samples)* = 2 · *(# of U-samples)*
2. **Correct avg. probability:** *total mass is* $\frac{2}{N}$ · *(total # of samples)*



P's perspective: V decides (b_i) AFTER P's message.

I, "b then x"

1. *Draw* b
2. *If* $b = 0$, $x \sim D$; *o.w.* $x \sim U_{[N]}$

$\equiv$

II, "x then b"

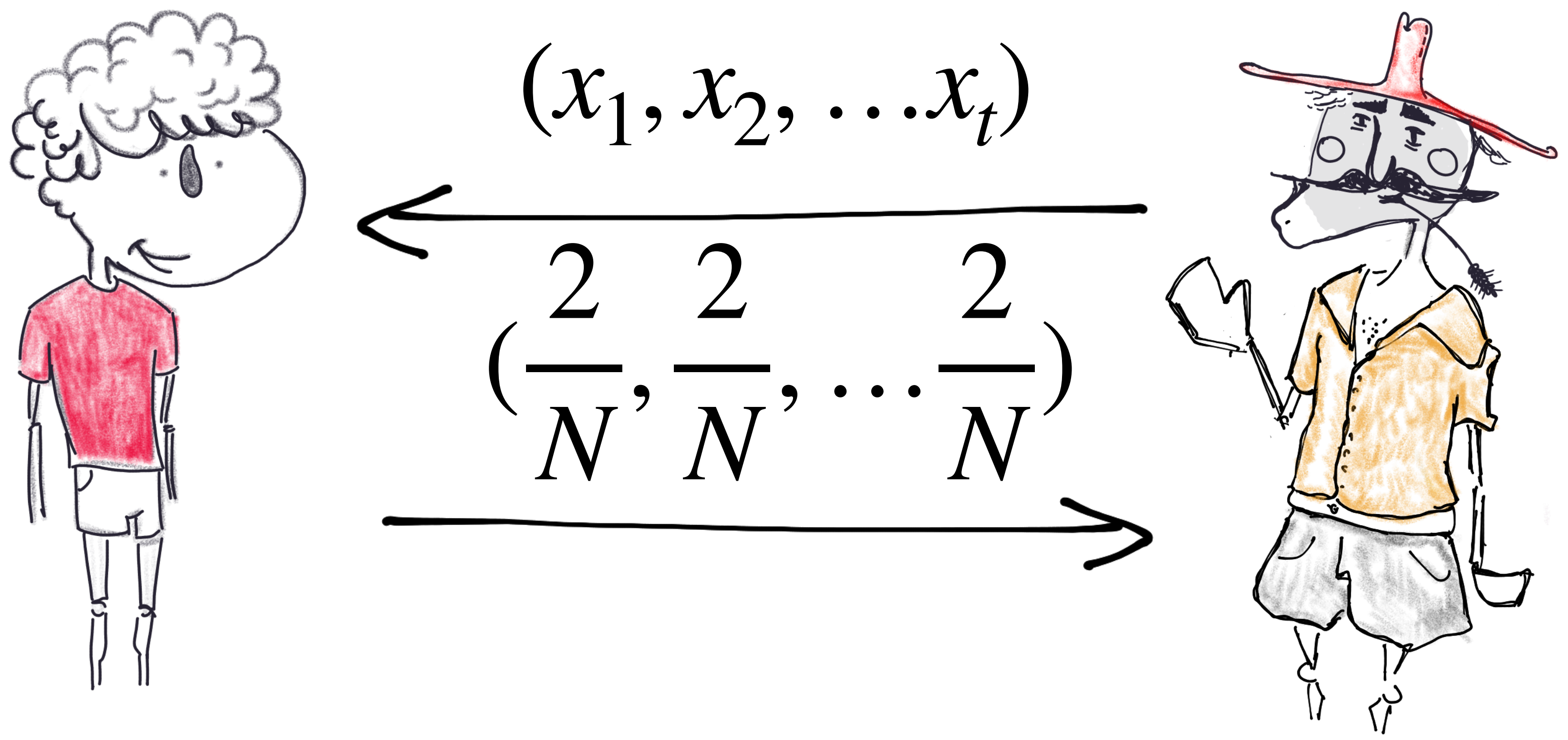
1. *Draw* $x \sim \frac{1}{2}D + \frac{1}{2}U_{[N]}$
2. $\forall x$, $b|_x = 0$, *w.p.* $\frac{D(x)}{D(x) + 1/N}$; *o.w.* $b|_x = 1$

Soundness Analysis

Look at the t samples tagged $2/N$:

1. **Consistency:** (# of D-samples) = 2 · (# of U-samples)

2. **Correct avg. probability:** total mass is $\frac{2}{N}$ · (total # of samples)



(b_i) ***determined AFTER prover's message:***

II, "x then b"

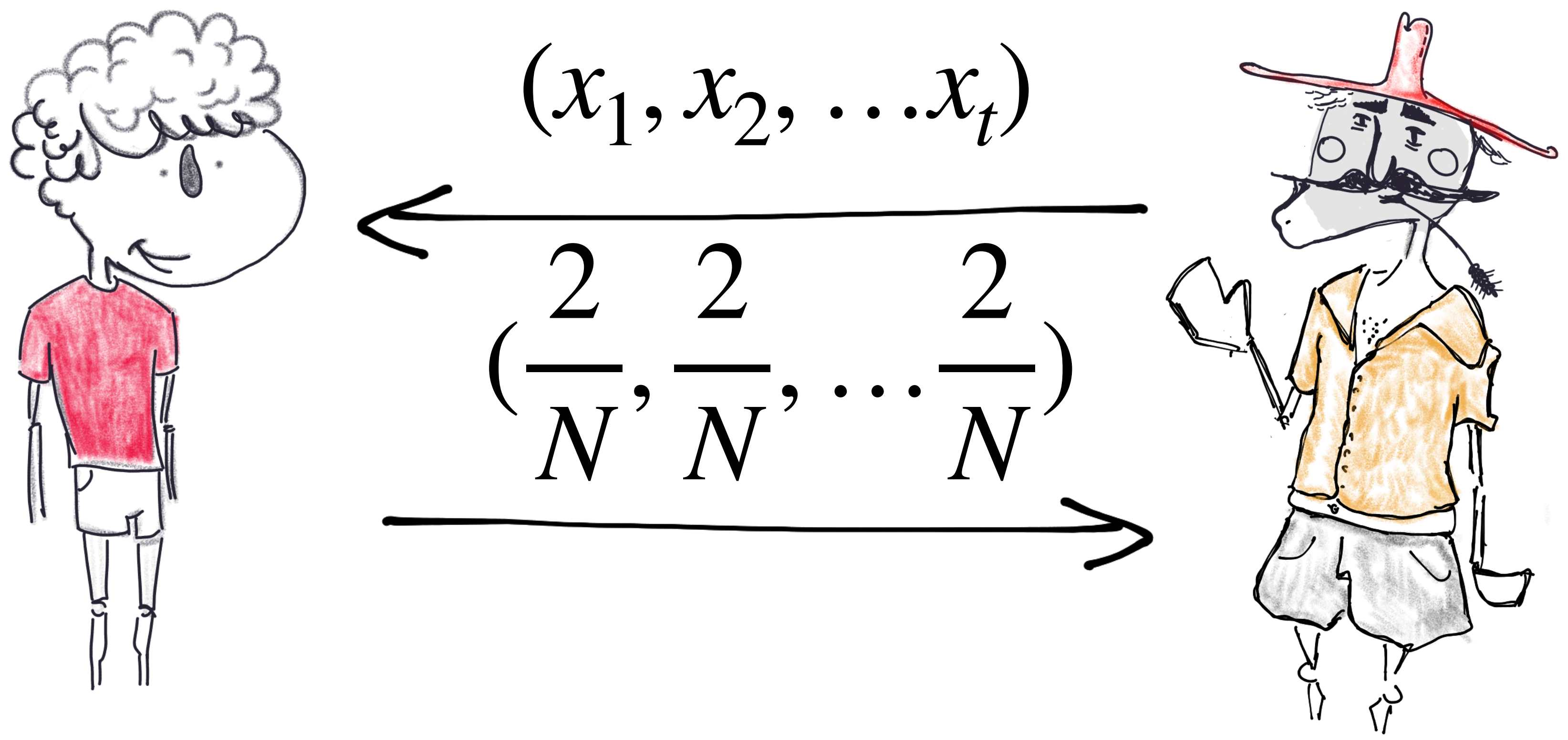
1. Draw $x \sim \frac{1}{2}D + \frac{1}{2}U_{[N]}$

2. $\forall x$, $b|_x = 0$, w.p. $\dfrac{D(x)}{D(x)+1/N}$; o.w. $b|_x = 1$

Soundness Analysis

Look at the t samples tagged $2/N$:

1. **Consistency:** (# of D-samples) = 2 · (# of U-samples)
2. **Correct avg. probability:** total mass is $\dfrac{2}{N}$ · (total # of samples)



(b_i) determined AFTER prover's message:

Test 1 passed:

$$\mathbb{E}_{x \sim_u \{x_1,\ldots,x_t\}} \left[\frac{\Pr\left(b\,|_x = 1\right)}{\Pr\left(b\,|_x = 0\right)} \right] = \frac{1}{2}$$

Test 2 passed:

$$\mathbb{E}_{x \sim_u \{x_1,\ldots,x_t\}} \left[D(x) \right] = \frac{2}{N}$$

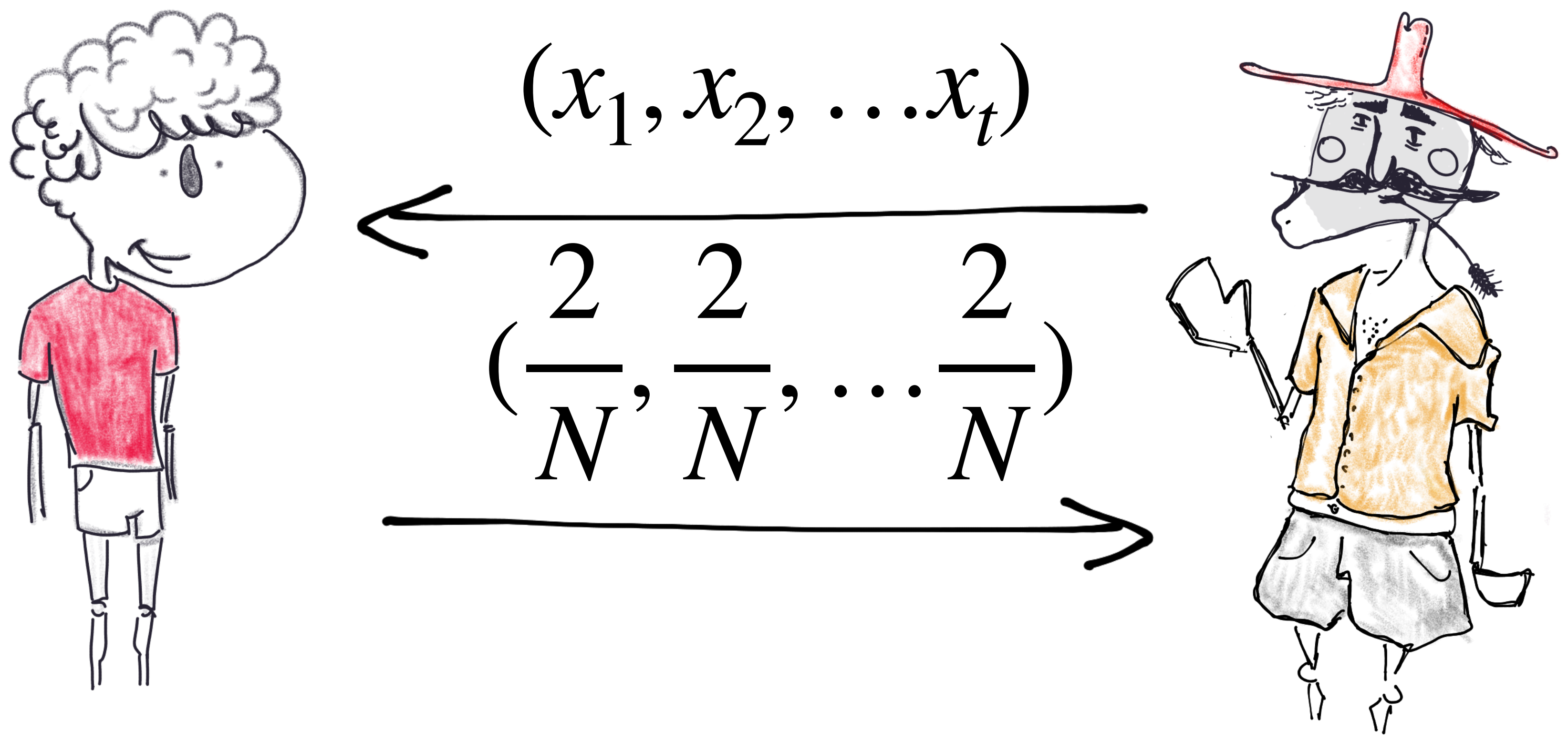
II, “x then b”

1. Draw $x \sim \frac{1}{2}D + \frac{1}{2}U_{[N]}$
2. $\forall x$, $b\,|_x = 0$, w.p. $\dfrac{D(x)}{D(x) + 1/N}$; o.w. $b\,|_x = 1$

Soundness Analysis

Look at the t samples tagged $2/N$:

1. **Consistency:** (# of D-samples) = 2 · (# of U-samples)
2. **Correct avg. probability:** total mass is $\frac{2}{N} \cdot$ (total # of samples)



(b_i) ***determined AFTER prover's***

$$\frac{1/N}{1/N + D(x)}$$

Test 1 passed:

$$\mathbb{E}_{x \sim_u \{x_1, \ldots, x_t\}} \left[\frac{\Pr\left(b|_x = 1\right)}{\Pr\left(b|_x = 0\right)} \right] = \frac{1}{2}$$

Test 2 passed:

$$\mathbb{E}_{x \sim_u \{x_1, \ldots, x_t\}} \left[D(x) \right] = \frac{2}{N}$$

$$\frac{D(x)}{1/N + D(x)}$$

II, "x then b"

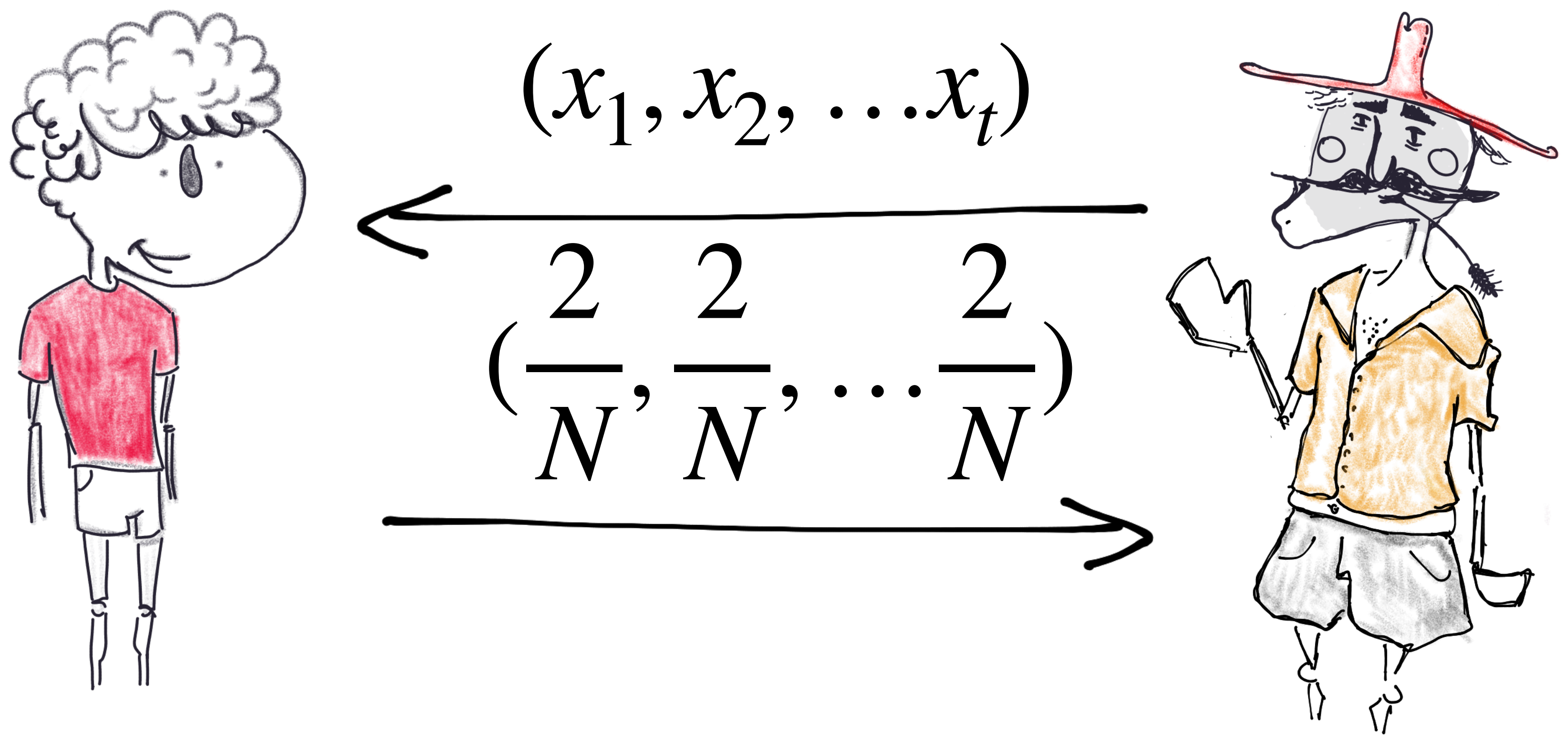
1. Draw $x \sim \frac{1}{2} D + \frac{1}{2} U_{[N]}$

$\forall x$, $b|_x = 0$, w.p. $\frac{D(x)}{D(x) + 1/N}$; o.w. $b|_x = 1$

Soundness Analysis

Look at the t samples tagged $2/N$:

1. ***Consistency:*** *(# of D-samples)* = *2* · *(# of U-samples)*
2. ***Correct avg. probability:*** *total mass is* $\frac{2}{N}$ · *(total # of samples)*



(b_i) ***determined AFTER prover's message:***

Test 1 passed:

$$\mathbb{E}_{x \sim_u \{x_1, \ldots, x_t\}} \left[\frac{1/N}{D(x)} \right] = \frac{1}{2}$$

Test 2 passed:

$$\mathbb{E}_{x \sim_u \{x_1, \ldots, x_t\}} \left[D(x) \right] = \frac{2}{N}$$

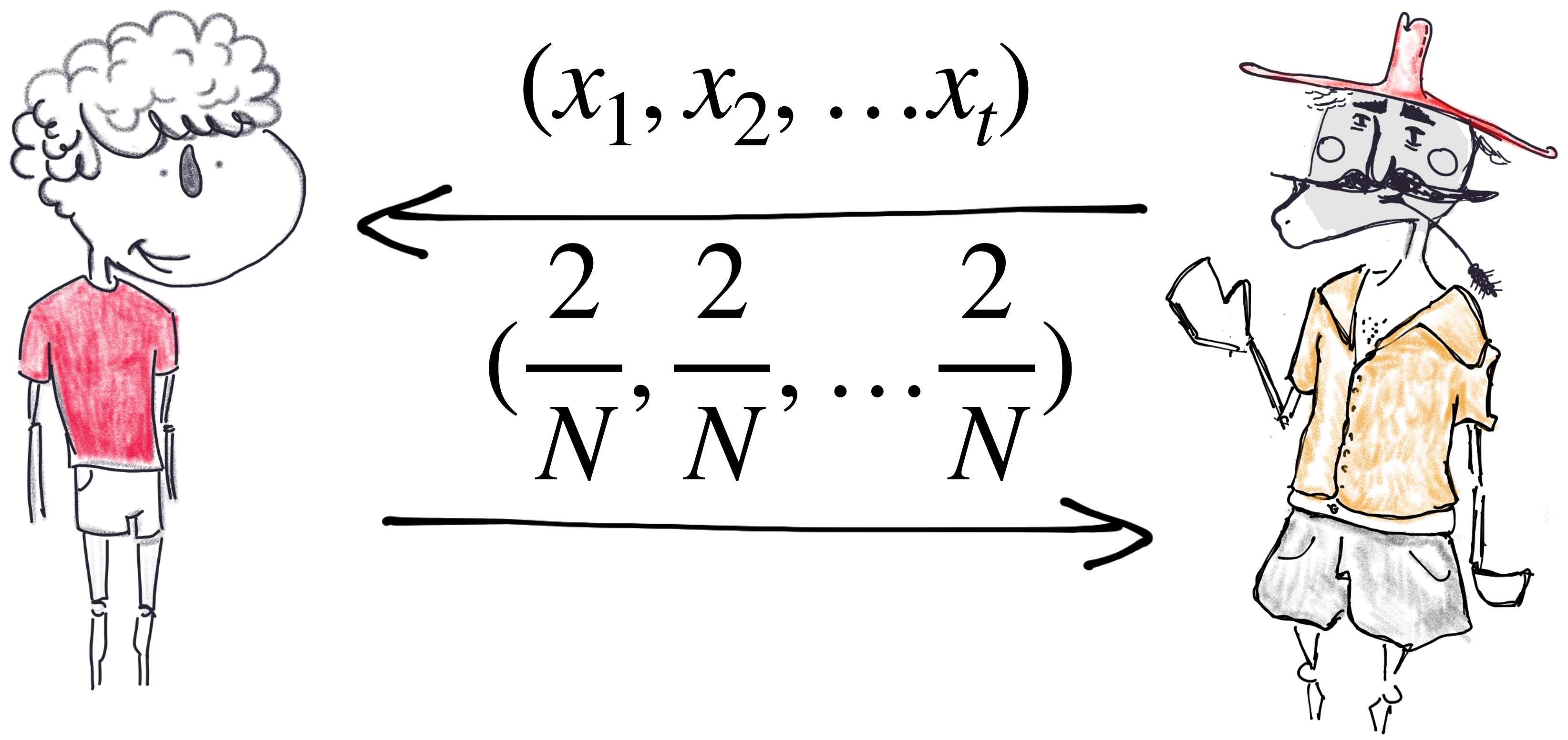
II, “x then b”

1. *Draw* $x \sim \frac{1}{2}D + \frac{1}{2}U_{[N]}$
2. $\forall x$, $b|_x = 0$, *w.p.* $\frac{D(x)}{D(x) + 1/N}$; *o.w.* $b|_x = 1$

Soundness Analysis

Look at the t samples tagged $2/N$:

1. **Consistency:** (# of D-samples) = 2 · (# of U-samples)
2. **Correct avg. probability:** *total mass is* $\dfrac{2}{N}$ · *(total # of samples)*



(b_i) ***determined AFTER prover's message:***

Test 1 passed:

$$\mathbb{E}_{x\sim_u\{x_1,\ldots,x_t\}}\left[\frac{1}{D(x)}\right] = \frac{N}{2}$$

Test 2 passed:

$$\mathbb{E}_{x\sim_u\{x_1,\ldots,x_t\}}\left[D(x)\right] = \frac{2}{N}$$

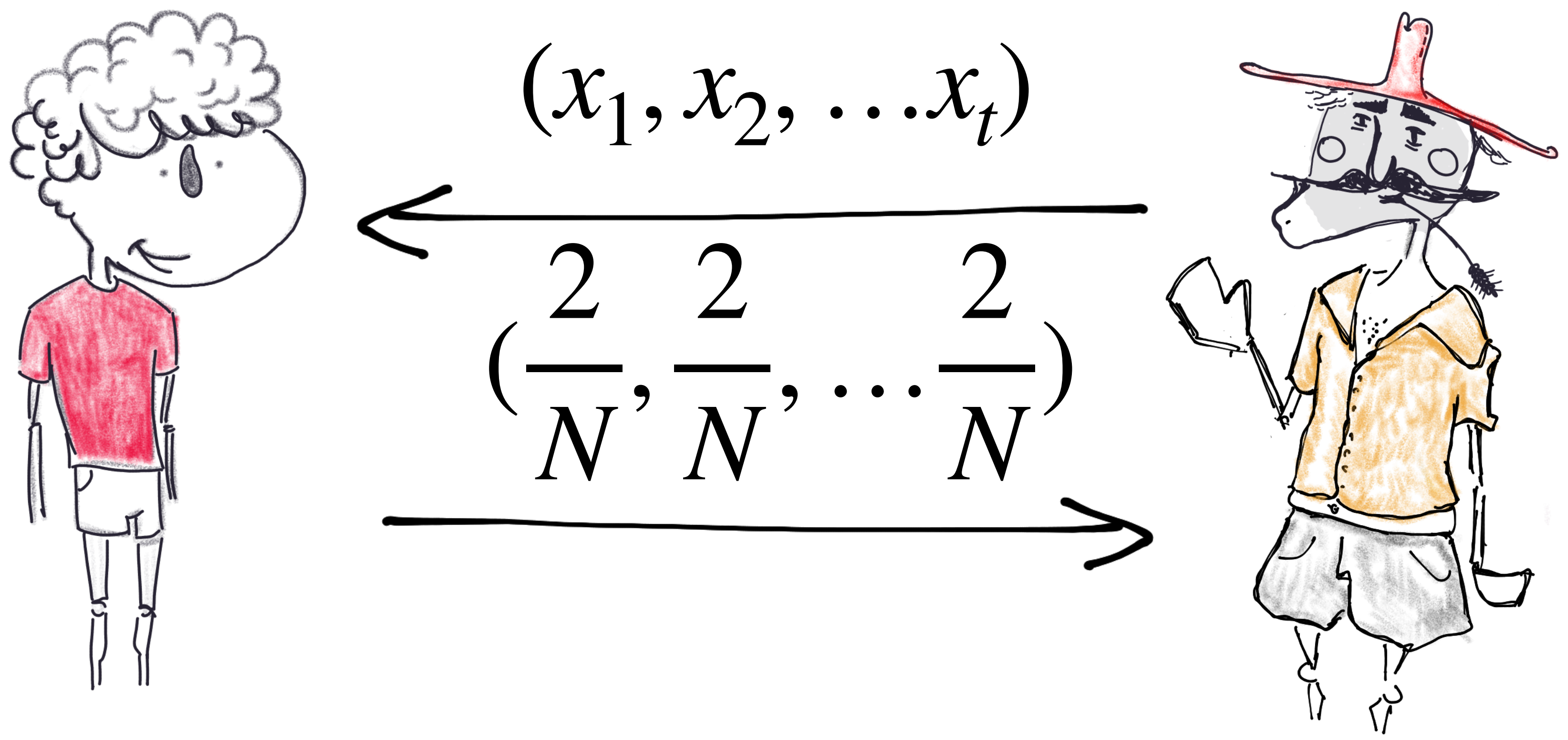
II, "x then b"

1. *Draw* $x \sim \frac{1}{2}D + \frac{1}{2}U_{[N]}$
2. $\forall x,\ b|_x = 0$, *w.p.* $\dfrac{D(x)}{D(x) + 1/N}$; *o.w.* $b|_x = 1$

Soundness Analysis

Look at the t samples tagged $2/N$:

1. **Consistency:** (# of D-samples) = 2 · (# of U-samples)
2. **Correct avg. probability:** *total mass is* $\frac{2}{N}$ · *(total # of samples)*



(b_i) ***determined AFTER prover's message:***

Test 1 passed:

$$\mathbb{E}_{x \sim_u \{x_1, \ldots, x_t\}} \left[\frac{1}{D(x)} \right] = \frac{N}{2}$$

Test 2 passed:

$$\mathbb{E}_{x \sim_u \{x_1, \ldots, x_t\}} \left[D(x) \right] = \frac{2}{N}$$

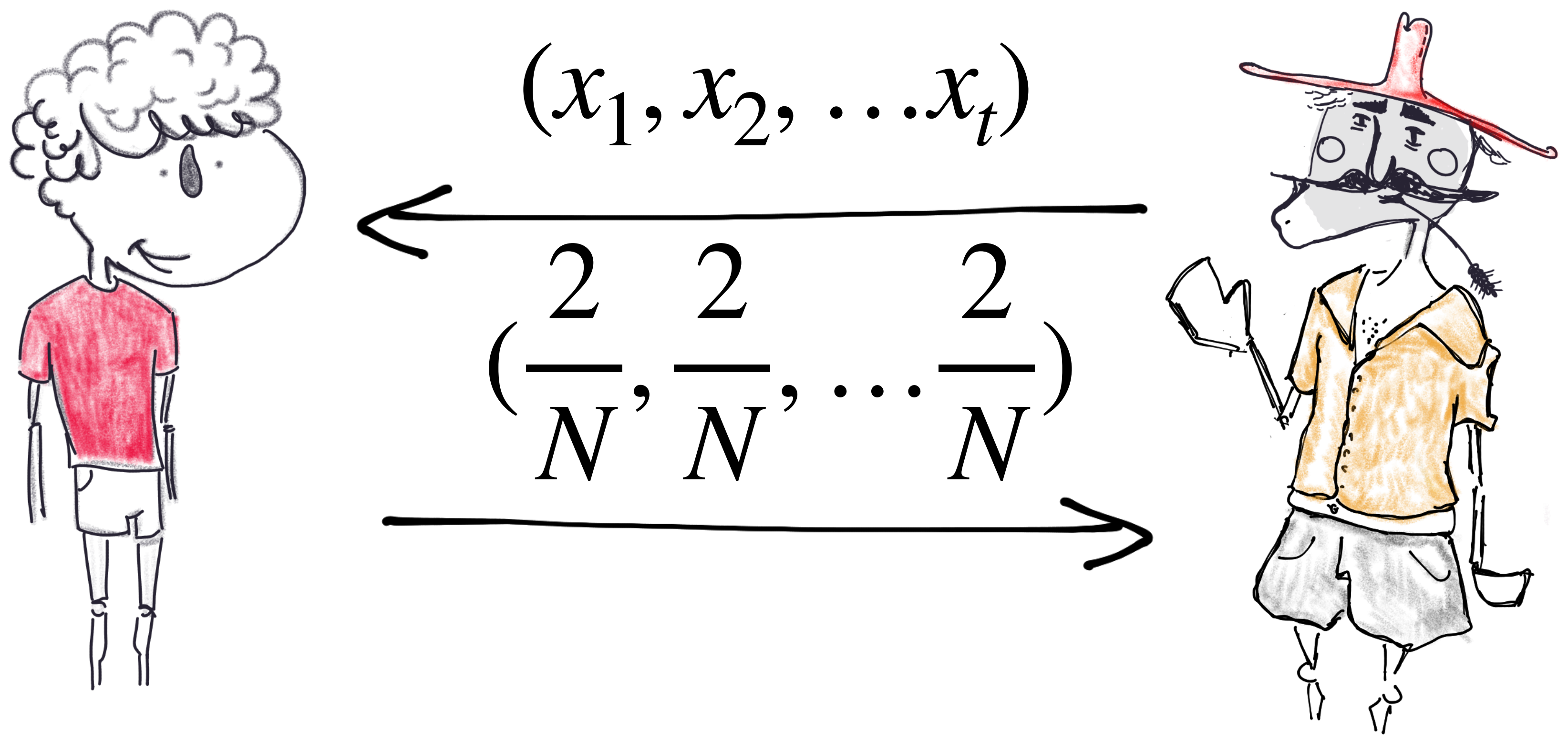
Jensen's Inequality $\Longrightarrow D(X)$ *is a constant r.v.*

$$\forall x, D(x) = \mathbb{E}[D(X)] = 2/N$$

Soundness Analysis

Look at the t samples tagged $2/N$:

1. ***Consistency:*** *(# of D-samples) = 2 · (# of U-samples)*
2. ***Correct avg. probability:*** *total mass is* $\frac{2}{N}$ · *(total # of samples)*



(b_i) ***determined AFTER prover's message:***

Test 1 passed:

$$\mathbb{E}_{x \sim_u \{x_1, \ldots, x_t\}} \left[\frac{1}{D(x)} \right] = \frac{N}{2}$$

Test 2 passed:

$$\mathbb{E}_{x \sim_u \{x_1, \ldots, x_t\}} \left[D(x) \right] = \frac{2}{N}$$

Jensen's Inequality $\Longrightarrow D(X)$ *is a constant r.v.*

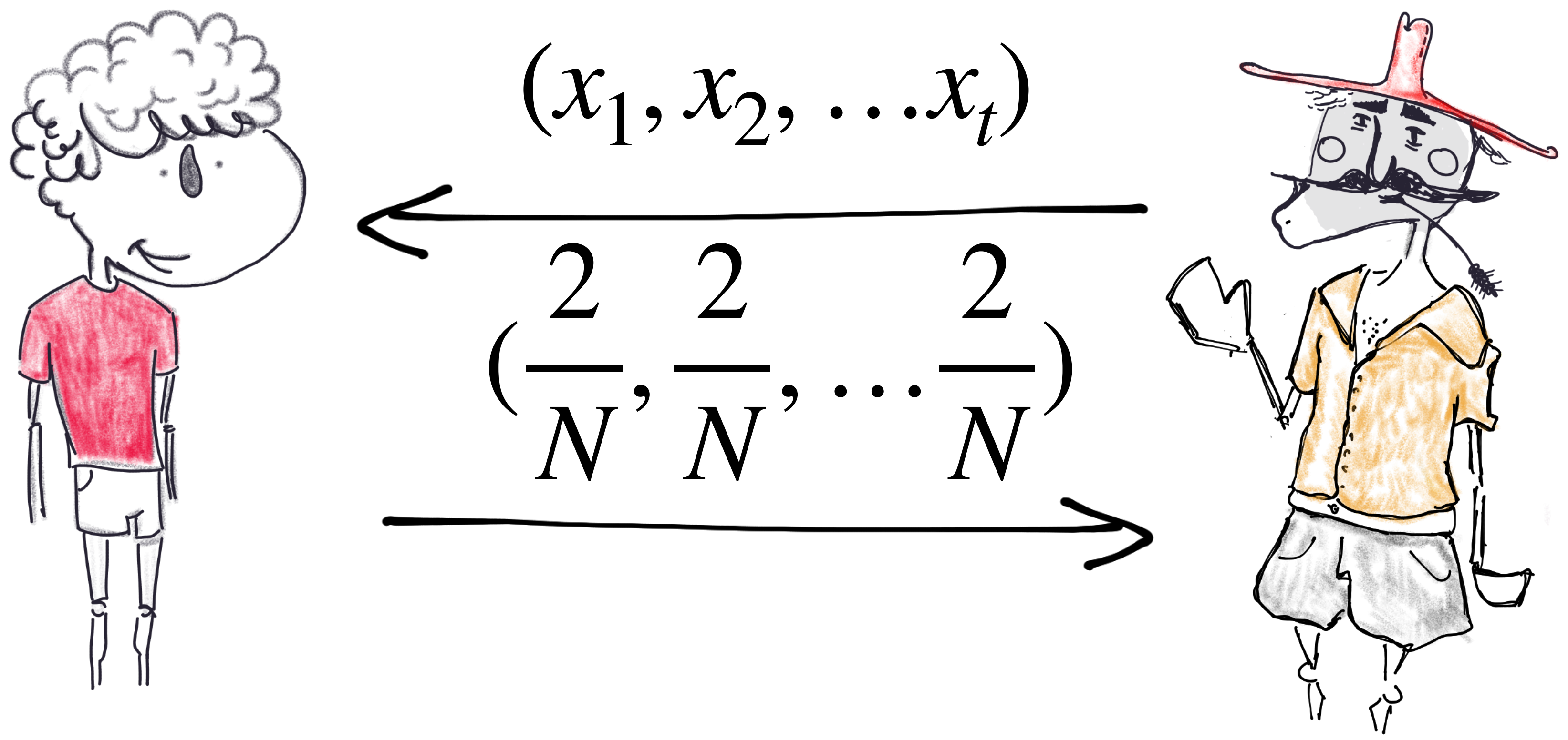
$$\forall x, D(x) = \mathbb{E}[D(X)] = 2/N$$

Equations hold only if
True probability = Alleged probability

Soundness Analysis

Look at the t samples tagged $2/N$:

1. ***Consistency:*** *(# of D-samples)* = 2 · *(# of U-samples)*
2. ***Correct avg. probability:*** *total mass is* $\frac{2}{N}$ · *(total # of samples)*



(b_i) ***determined AFTER prover's message:***

Test 1 passed:

$$\mathbb{E}_{x \sim_u \{x_1,\ldots,x_t\}}\left[\frac{1}{D(x)}\right] \approx \frac{N}{2}$$

Test 2 passed:

$$\mathbb{E}_{x \sim_u \{x_1,\ldots,x_t\}}\left[D(x)\right] \approx \frac{2}{N}$$

More accurately

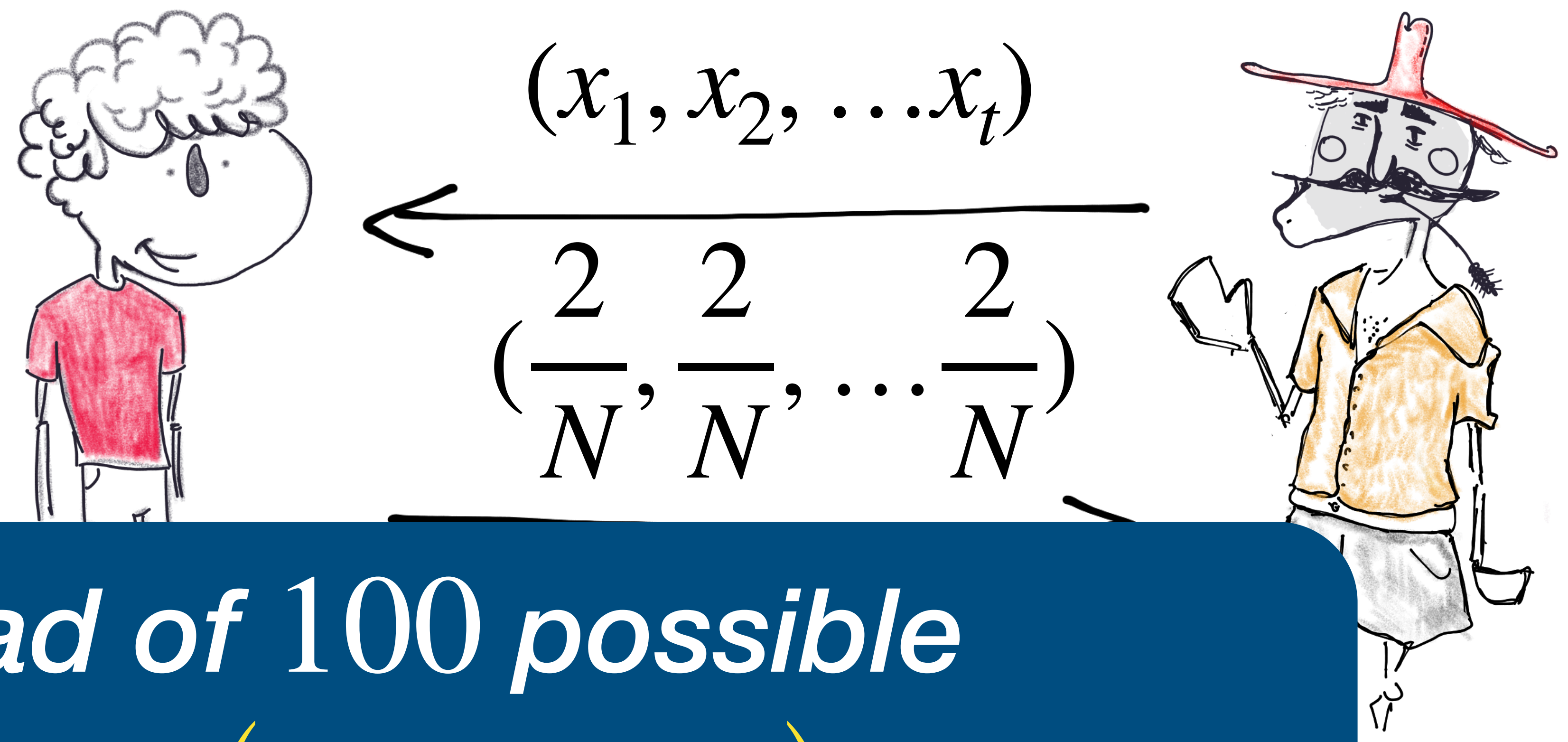
1. *Samples tagged* $\left[2/N, (1+\varepsilon)\cdot 2/N\right]$.
2. *Equations **approximately** hold.*

Soundness Analysis

Look at the t samples tagged $2/N$:

1. ***Consistency:*** *(# of D-samples)* = 2 · *(# of U-samples)*

2. ***Correct avg. probability:*** *total mass is* $\frac{2}{N}$ · *(total # of s*



Instead of 100 possible probabilities, $O\left(\log\left(N/\varepsilon\right)\right)$ possible probabilities

(b_i) ***determined AFTER prover's message:***

Test 1 passed:

$$\mathbb{E}_{x \sim_u \{x_1, \ldots, x_t\}}\left[\frac{1}{D(x)}\right] \approx \frac{N}{2}$$

Test 2 passed:

$$\mathbb{E}_{x \sim_u \{x_1, \ldots, x_t\}}\left[D(x)\right] \approx \frac{2}{N}$$

More accurately

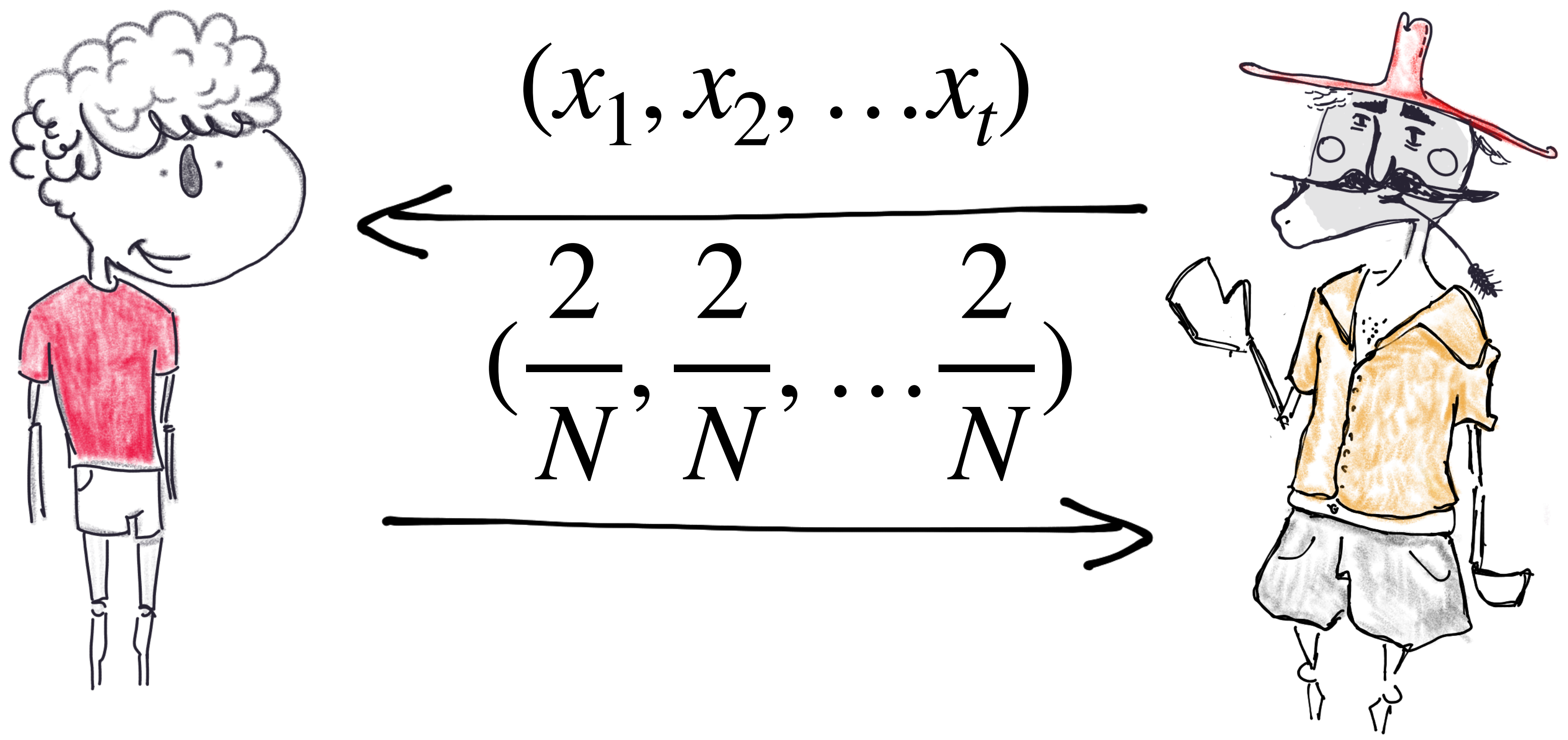
1. *Samples tagged* $\left[2/N, (1+\varepsilon) \cdot 2/N\right]$.
2. *Equations* ***approximately*** *hold.*

Soundness Analysis

Look at the t samples tagged $2/N$:

1. **Consistency:** *(# of D-samples) = 2 · (# of U-samples)*

2. **Correct avg. probability:** *total mass is* $\frac{2}{N}$ · *(total # of samples)*



(b_i) ***determined AFTER prover's message:***

Test 1 passed:

$$\mathbb{E}_{x \sim_u \{x_1, \ldots, x_t\}} \left[\frac{1}{D(x)}\right] \approx \frac{N}{2}$$

Test 2 passed:

$$\mathbb{E}_{x \sim_u \{x_1, \ldots, x_t\}} \left[D(x)\right] \approx \frac{2}{N}$$

Claim:

Alleged probabilities ***far*** *from correct* $\Longrightarrow$
Equations don't hold ***even approximately.***

Recap

- Via sample access, we can verifiably obtain $(x, D(x))$ for many $x \sim D$.

Recap

- Via sample access, we can verifiably obtain $(x, D(x))$ for many $x \sim D$.
- Immediate corollary: verify any label-invariant property.

Recap

- Via sample access, we can verifiably obtain $(x, D(x))$ for many $x \sim D$.
- Immediate corollary: verify any label-invariant property.
- Next - leverage to obtain:

Theorem **[HR'24]:** ***very rich family of properties*** has an interactive proof with:

- **Verifier** sample complexity $\widetilde{O}\left(N^{0.9}\right)$ poly(ε^{-1}), runtime, and communication $\widetilde{O}\left(N^{0.95}\right)$ poly(ε^{-1}).
- **Honest Prover** time $\mathrm{poly}(N)$, sample complexity $\widetilde{O}\left(N^{1.1}\right)$ poly(ε^{-1}).

Recap

- Via sample access, we [...] many $x \sim D$.
- Immediate corollary: ve[...]
- Next - leverage to obtai[...]

Given complete description
$\forall x, (x, D(x))$, *TV*$(D, \mathcal{P})$
approximated by ***low depth circuit / low space TM***

Theorem **[HR'24]:** ***very rich family of properties*** has an interactive proof with:

- **Verifier** sample complexity $\widetilde{O}\left(N^{0.9}\right)$ poly(ε^{-1}), runtime, and communication $\widetilde{O}\left(N^{0.95}\right)$ poly(ε^{-1}).
- **Honest Prover** time $\operatorname{poly}(N)$, sample complexity $\widetilde{O}\left(N^{1.1}\right)$ poly(ε^{-1}).

Summary

Main takeaway - useful tool: given only ***sample access,*** *many claims* can be ***verified*** much more efficiently than repeating computation.

Summary

Main takeaway - useful tool: given only ***sample access,*** *many claims* can be ***verified*** much more efficiently than repeating computation.

Not discussed:

- ***Lower bounds*** **-** sample lower bounds; *proving* is *harder than testing.*
- ***Crypto assumptions*** **-** *reduce sample (what about communication?)*
- ***Super-fast protocols for specific problems*** *(e.g. verifying distributions are far).*

Questions for Discussion

Moving forward, to the context of AI:

1. *Where do we find distributions accessible by samples (and not queries)?*

   Choosing training set? Prompt distribution? Output distribution?

2. What might we want to *verify* about distributions? (Alignment? Compliance? Diversity of data?) Delegation? Can it be a property of the distribution?

3. Extending the model to accommodate *more settings* (better parameter regime):
   a. ***New access models:*** what type of access do we expect to have (cond. sampling?)?
   b. ***More assumptions:*** what if distribution admits some structure (e.g. is uniform, over metric space, We are given some other *advice*)? Can we get *super-fast protocols?*

Theorem[H, Rothblum '25]:

Any interactive proof that a ***distribution is uniform over*** $N/2$ ***elements***, has either:

- The verifier sample complexity is $\Omega\left(N^{2/3}\right)$, or -
- The honest prover sample complexity is $\Omega(N)$.

Theorem[H, Rothblum '25]:

For any constant $c \in \mathbb{R}^+$, and $k \in \mathbb{N}$, ***any*** interactive proof that a ***distribution*** D ***satisfies*** $\|D\|_k \leq \frac{c}{N^{1-1/k}}$, or is ε-far from any such distribution, has either:

- The verifier sample complexity is $\Omega\left(N^{1-1/k}\right)$, or -
- The honest prover sample complexity is $\Omega(N)$.