COMS W4236 Lecture 7: Randomness

Professor: Toni Pitassi Guest Lecturer: Conlan Olson

23 October 2026

When does randomness help?

RP, BPP: we can't rule out the case where these classes are no more powerful than P.

Crypto, privacy: randomness is provably necessary.

Table of Contents

Differential Privacy

Stability and the Reusable Holdout

Zero Knowledge Proofs

Table of Contents

Differential Privacy

Stability and the Reusable Holdout

Zero Knowledge Proofs

Privacy motivation

Motivation: The US Census Bureau wants to release statistics without revealing any individual information.

Naïve idea: Remove personally identifiable information.

Example:

Voter Registration

Name	Birthdate	ZIP
Conlan	01/01/2001	10001
Toni	02/02/2002	11211
Rain	03/03/2003	11377

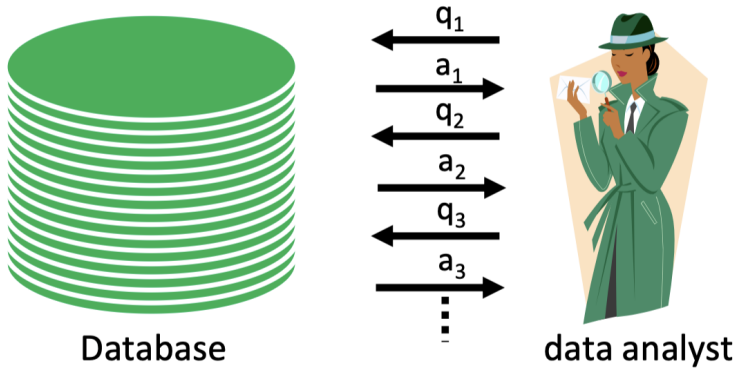
Census

ZIP	HH Size	Age	Avg. Income
10001	1	0-10	0
10001	1	10-20	35,000
10001	1	20-30	50,000

Famous examples: Reidentification of medical records [Sweeney'96], Netflix reviews [Narayanan, Shmatikov'08].

Idea of differential privacy

Setting:



In words: It should be hard to tell whether I am in the database or not.

Note

Is the following a privacy violation? *A health insurance company studies health data, and finds that smokers are more likely to get lung cancer. Because I am a smoker, they set a higher rate for me.*

Typically, the answer is no. General findings are allowed.

Differential privacy

We view a database as a vector in $\mathbb{N}^{|\mathcal{X}|}$, where $\mathcal{X}$ are types of individuals.

Definition ([DMNS'06])

An algorithm $\mathcal{M} : \mathbb{N}^{|\mathcal{X}|} \to \mathcal{Y}$ is ε-**differentially private** if, for all $S \in \text{Range}(\mathcal{M})$ and for all $D_1, D_2 \in \mathbb{N}^{|\mathcal{X}|}$ such that $\|D_1 - D_2\|_1 \leq 1$,

$$\Pr[\mathcal{M}(D_1) \in S] \leq e^{\varepsilon} \Pr[\mathcal{M}(D_2) \in S].$$

at most one bit changes by 1

$\Pr[M(D_1) \in S] / \Pr[M(D_2) \in S] \leq e^{\varepsilon}$

reverse holds by symmetry

Note

Today, we'll usually think of $\mathcal{Y}$ as being in $\mathbb{R}$ or $[0, 1]$. However, we can ask for DP on algorithms with arbitrary outputs, e.g. learning algorithms.

No privacy without randomness

Proposition

If $\mathcal{M}$ is deterministic and nonconstant, it is not ε-DP for any ε.

Proof.

let D_1 be s.t. $M(D_1) = y_1$ and D_2 s.t. $M(D_2) = y_2$

$D_1 \to D_1' \to D_2' \to \cdots \to D_2$ $\Rightarrow$ WLOG $\|D_1 - D_2\| \le 1$

$y_1 \quad y \quad y' \quad y_2$

$S = \{y_1\}$ □

Fundamental Law of Information Recovery [Dinur, Nissim'03]: "Overly accurate answers to too many queries leads to catastrophic privacy loss."

$\Pr[M(D_1) \in S] = 1$

$\Pr[M(D_2) \in S] = 0$

$1 \not\le e^{\varepsilon} \cdot 0$ $\forall \varepsilon$

Randomized response

"Did you vote in the last election?"

1. Flip a coin. If heads, answer truthfully.
2. If tails, flip another coin and answer "yes" if heads and "no" if tails.

Privacy: D_1 = voted D_2 = not voted

$\Pr[RR(D_1) = \text{no}] = \frac{1}{4}$ $\Pr[RR(D_2) = \text{no}] = \frac{3}{4}$

DP if: $\frac{3}{4} \leq e^{\varepsilon} \cdot \frac{1}{4}$

$\varepsilon = \ln(3)$

Laplace mechanism

count + Lap($\frac{1}{\varepsilon}$)

"How many voting-age Asian people live in Census block 1234 in households of size 2?"

1,000,000 queries $\varepsilon = 10$

$\varepsilon' = \varepsilon / 1{,}000{,}000$

Definition

The L_1 **sensitivity** of a function $f : \mathbb{N}^{|\mathcal{X}|} \to \mathcal{Y}$ is

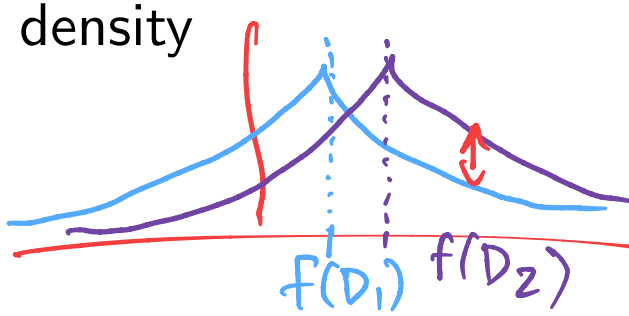
$$\Delta(f) := \max_{\|D_1 - D_2\|_1} |f(D_1) - f(D_2)|.$$

$\Delta(f)$ is 1 for counting queries

Laplace mechanism for real-valued queries: Output $f(D) + \nu$ where $\nu \sim \text{Lap}(\Delta(f)/\varepsilon)$.

The **Laplace distribution** with parameter b has density

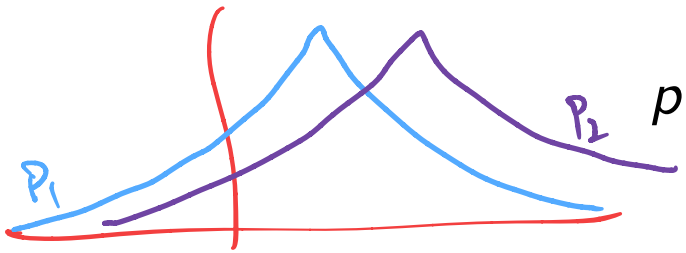
$$p(x) := \frac{1}{2b} \exp(-|x|/b).$$



Laplace mechanism

Laplace mechanism for real-valued queries: Output $f(D) + \nu$ where $\nu \sim \mathrm{Lap}(\Delta(f)/\varepsilon)$.

The **Laplace distribution** with parameter b has density



$$p(x) := \frac{1}{2b} \exp(-|x|/b).$$

Proposition

The Laplace mechanism gives ε-DP.

Proof.

let $p_1(x)$ be pdf of $M(D_1)$

$p_2(x)$ " $M(D_2)$

$$\frac{p_1(x)}{p_2(x)} = \frac{p(x - f(D_1))}{p(x - f(D_2))} \leq \frac{\exp(-|x - f(D_1)|/b)}{\exp(-|x - f(D_2)|/b)}$$

$$\exp\left(\frac{|x - f(D_2)| - |x - f(D_1)|}{b}\right)$$

$$\leq \exp\left(\frac{|f(D_1) - f(D_2)|}{b}\right)$$

$$\leq \exp\left(\frac{\Delta f}{b}\right)$$

$$\leq \exp(\varepsilon)$$

□

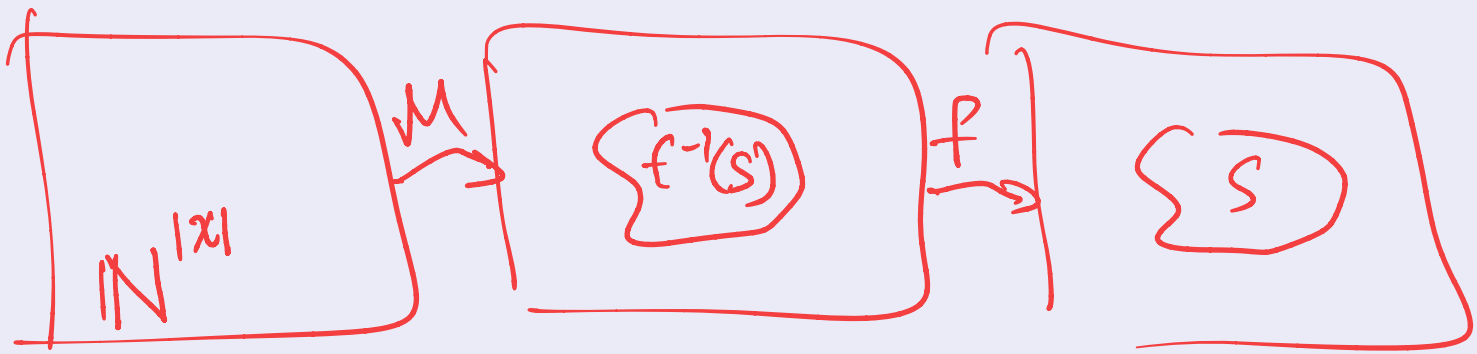
DP is robust to post-processing

Proposition

Let $\mathcal{M} : \mathbb{N}^{|\mathcal{X}|} \to \mathcal{Y}$ be ε-DP and let $f : \mathcal{Y} \to \mathcal{Y}'$ be any function (possibly randomized). Then, $f \circ \mathcal{M}$ is ε-DP.

Proof.

$$\Pr[f(M(D)) \in S] = \Pr[M(D_1) \in f^{-1}(S)]$$



□

Composition of DP

Proposition

Let $\mathcal{M}_1 : \mathbb{N}^{|\mathcal{X}|} \to \mathcal{Y}_1$ be ε_1-DP and let $\mathcal{M}_2 : \mathbb{N}^{|\mathcal{X}|} \to \mathcal{Y}_2$ be ε_2-DP. Then, the function $\mathcal{M}_1 \times \mathcal{M}_2 : \mathbb{N}^{|\mathcal{X}|} \to \mathcal{Y}_1 \times \mathcal{Y}_2$ given by $D \mapsto (\mathcal{M}_1(D), \mathcal{M}_2(D))$ is $(\varepsilon_1 + \varepsilon_2)$-DP

Proof.

$$\Pr[(M_1(D), M_2(D)) = (y_1, y_2)] =$$

$$\Pr[M_1(D) = y_1 \wedge M_2(D) = y_2] =$$

$$\Pr[M_1(D) = y_1] \cdot \Pr[M_2(D) = y_2] \leq$$

$$e^{\varepsilon} \Pr[M_1(D') = y_1] \cdot e^{\varepsilon} \Pr[M_2(D') = y_2] =$$

$$e^{\varepsilon+\varepsilon} \Pr[(M_1(D'), M_2(D')) = (y_1, y_2)]$$

□

Why use DP for the Census?

1. Post-processing: people may analyze Census data and combine it with other data in arbitrary ways.
2. Composition: the Census releases many counts and wants to bound the cumulative privacy loss.
3. Allows for principled "privacy budgeting."

Not without controversy...

Table of Contents

Differential Privacy

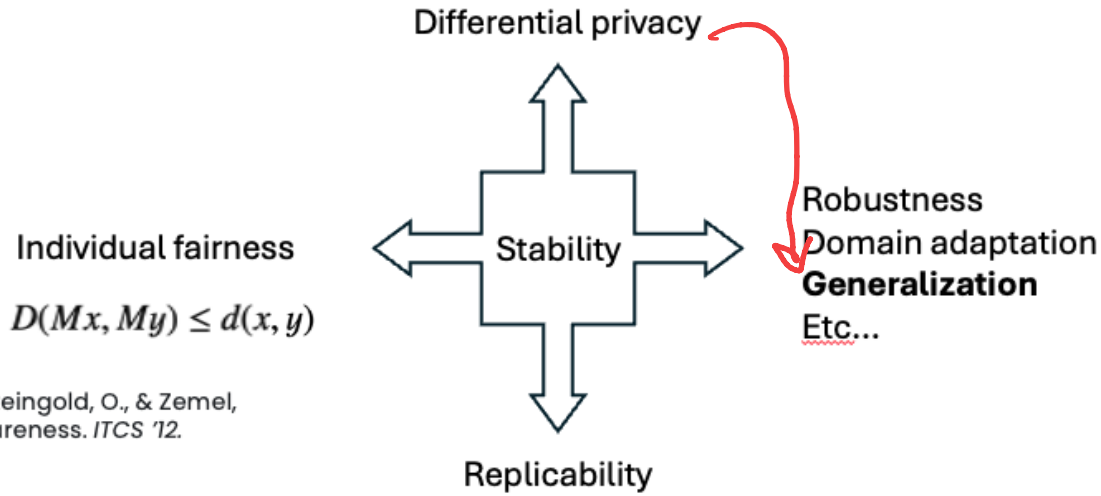
Stability and the Reusable Holdout

Zero Knowledge Proofs

Stability

Dwork, C., McSherry, F., Nissim, K., Smith, A. (2006). Calibrating Noise to Sensitivity in Private Data Analysis. *Theory of Cryptography*. TCC 2006.

$$\mathbb{P}[\mathcal{M}(D) \in S] \leq e^{\epsilon}\mathbb{P}[\mathcal{M}(D') \in S] + \delta$$



Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2012). Fairness through awareness. *ITCS '12*.

$$\mathbf{Pr}_{S,S',r}[\mathcal{A}(S;r) = \mathcal{A}(S';r)] \geqslant 1 - \rho$$

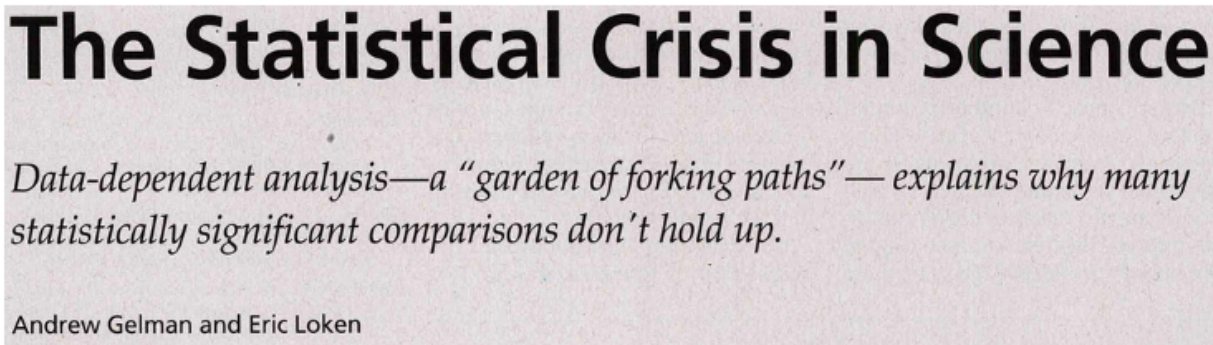
Bun, M., Gaboardi, M., Hopkins, M., Impagliazzo, R., Lei, R., Pitassi, T., ... & Sorrell, J. (2023, June). Stability is stable: Connections between replicability, privacy, and adaptive generalization. *STOC 2023*.

The problem of validity

2005:

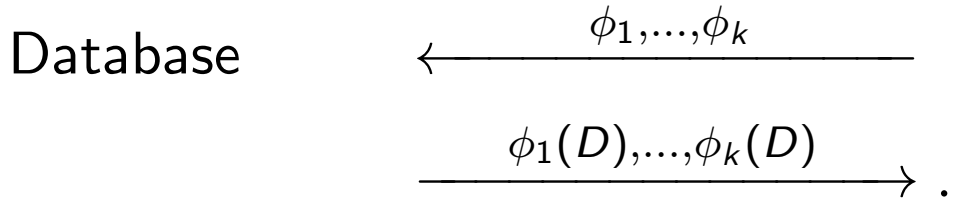


2014:



Statistical query model

Let a **statistical query** be a function $\phi : \mathcal{X} \to [0, 1]$. Let $\phi(D)$ be $\frac{1}{|D|} \sum_{x \in D} \phi(x)$.

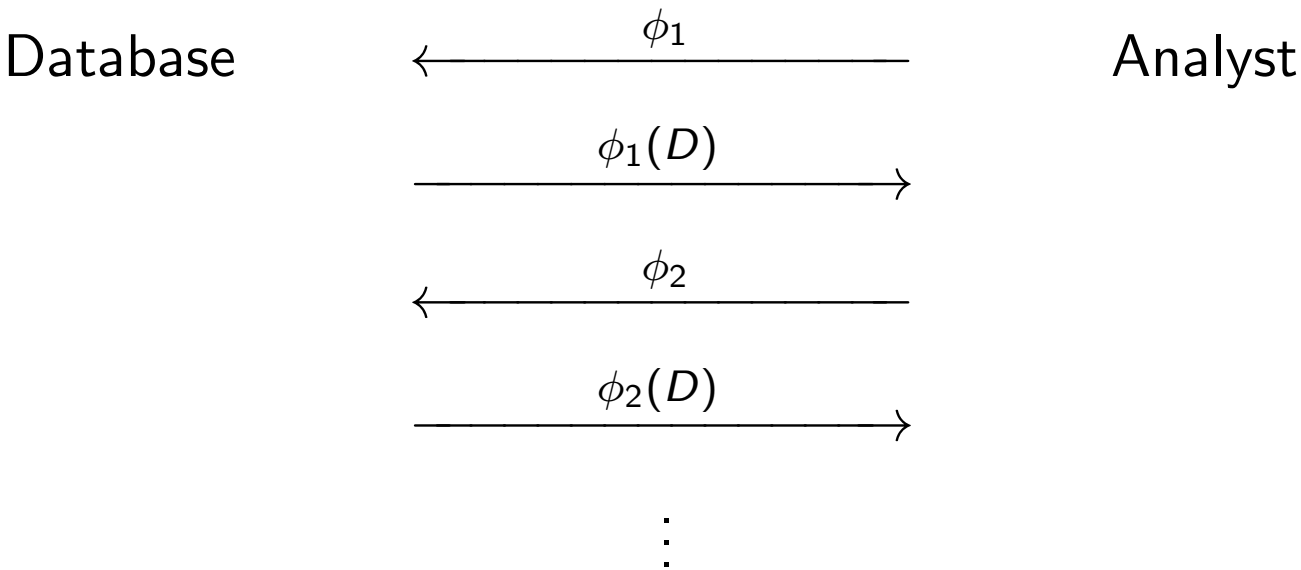


Hoeffding's inequality: If $X_1, \ldots, X_n$ are i.i.d. random variables with values in $[0, 1]$ and mean μ, then

$$\Pr\left[\left|\sum_{i=1}^{n} X_i - \mu\right| \geq t\right] \leq 2e^{-2t^2/n}.$$

Probability that all k queries are within ε of their mean:

What actually happens



Example

Let individuals have features $(v_1, \ldots, v_{100})$ and labels y, all random coin flips.

For $i = 1, \ldots, 100$, $\phi_i(x) = \text{cor}(v_i, y)$.

Then, ϕ_{101} is the correlation of the best coordinate.

$\mathbb{E}[\phi_{101}(x)] = 0$, but on a dataset D the empirical correlation will be large.

Solution

Stability → Generalization (DP)

"If we only allow private access to a dataset, we can avoid overfitting." - The Reusable Holdout [DFHPRR'15]

Intuition: let $D = (x_1, \ldots, x_n)$ and $D^{(i)} = (x_1, \ldots, x_i', \ldots, x_n)$.
Then, if $\mathcal{M}$ is a ε-DP function that outputs functions $\mathcal{X} \to [0,1]$,

$$\left|\mathbb{E}[\mathcal{M}(D)(x)] - \mathbb{E}[\mathcal{M}(D^{(i)})(x)]\right| \leq e^{\varepsilon} - 1.$$

Then

What scientist sees

learned hypothesis's

$$\begin{aligned}
\mathop{\mathbb{E}}_{D}\left[\frac{1}{|D|}\sum_{i=1}^{n}\mathcal{M}(D)(x_i)\right] &= \frac{1}{|D|}\sum_{i=1}^{n}\mathop{\mathbb{E}}_{D}[\mathcal{M}(D)(x_i)] \\
&= \frac{1}{|D|}\sum_{x\in D}\mathop{\mathbb{E}}_{D,x'}[\mathcal{M}(D^{(i)})(x_i')] \\
&\approx \frac{1}{|D|}\sum_{x\in D}\mathop{\mathbb{E}}_{D,x'}[\mathcal{M}(D)(x_i')] \qquad \text{DP} \\
&= \mathop{\mathbb{E}}_{x'}[\mathcal{M}(D)(x_i')]
\end{aligned}$$

Shalev-Schwartz

Table of Contents

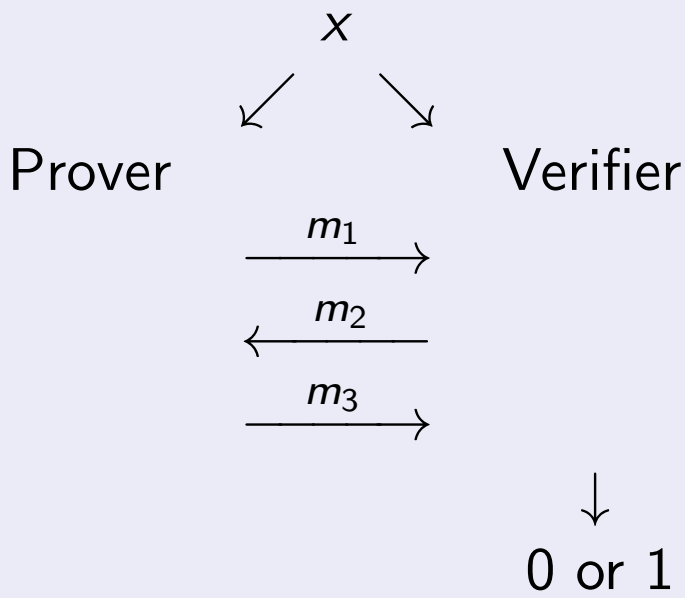
Differential Privacy

Stability and the Reusable Holdout

Zero Knowledge Proofs

Proof systems

Model



A proof system (P, V) for a language L must satisfy

1. Completeness: If $x \in L$, then after V interacts with P, it outputs 1 with probability $\geq 2/3$.
2. Soundness: If $x \notin L$, then, for any prover P', after V interacts with P', it outputs 1 with probability $< 1/3$.

easy to convince if $x \in L$

hard to trick

Intuition of zero knowledge proofs

if $x \in L$

Idea: we don't want the verifier to learn anything except for that $x \in L$.

Toy example: Alice is a tetrachromat, so she can detect a difference between a red/green mixture and an orange that, to Bob, look identical. She has two pieces of paper that she claims are different and she wants to prove this to Bob, without telling him which one is the red/green mixture and which one is the orange.

The proof will be that, for k rounds, Bob will have Alice look away and either switch or not switch the papers. Then, he asks Alice to tell him what he did. Bob outputs 1 if Alice is correct all k times.

Computational indistinguishability

Definition

Sequences of distributions $\{\mathcal{D}_1, \mathcal{D}_2, \ldots\}$ and $\{\mathcal{D}'_1, \mathcal{D}'_2, \ldots\}$ are **polytime indistinguishable** if, for all polynomials $p(n)$ and all sufficiently large n, any algorithm A_n running in $p(n)$ time satisfies

$$\left| \Pr_{x \sim \mathcal{D}_n} [A_n(x) = 1] - \Pr_{x \sim \mathcal{D}'_n} [A_n(x) = 1] \right| \leq 1/p(n).$$

Proposition (No indistinguishability without randomness)

If $\{\mathcal{D}_1, \mathcal{D}_2, \ldots\}$ and $\{\mathcal{D}_1, \mathcal{D}_2, \ldots\}$ are degenerate (constant) and non-equal, then they are not indistinguishable.

Proof.

Defining zero knowledge

Definition

A proof system (P, V) is **zero knowledge** if, for every polytime V', there exists a polytime "simulator" S such that, for all $x \in L$, the transcript of the interaction between P and V' on input x is polytime indistinguishable from the output of S on input x.

A zero knowledge proof for graph isomorphism

$$\text{GRAPHISO} = \{(G_0, G_1) \mid \exists \pi,\ \ G_1 = \pi(G_0)\}.$$

Protocol: Alice knows a permutation π such that $G_1 = \pi(G_0)$.

1. Alice chooses a random permutation σ and flips a coin b. She computes $H := \sigma(G_b)$ and sends H.
2. Bob flips a coin b' and sends b'.
3. Alice computes

$$\tau := \begin{cases} \sigma & b = b' \\ \sigma\pi^{-1} & b = 0, b' = 1 \\ \sigma\pi & b = 1, b' = 0 \end{cases}$$

   and sends τ.
4. Bob checks whether $H = \tau(G_{b'})$ and outputs 1 if it does.

A zero knowledge proof for graph isomorphism

Completeness: Check cases.

Soundness: Suppose there exists σ such that $H = \sigma(G_b)$ for some b. Then, with probability $1/2$, $b' \neq b$. If Bob accepts, this means that there exist permutations σ, τ such that $\sigma(G_b) = \tau(G_{b'})$. This would imply $G_b, G_{b'}$ are isomorphic for $b \neq b'$.
Now, suppose there does not exist σ such that $H = \sigma(G_b)$ for any b. Then, Bob cannot accept.

A zero knowledge proof for graph isomorphism

Zero knowledge: Fix a verifier strategy V'. We give a simulator that outputs H, b', τ.

1. Pick a random permutation σ and a random bit b and set $H := \sigma(G_b)$.
2. Set b' to be whatever V' would say when given G_0, G_1, H.
3. If $b = b'$, then output (H, b', σ). Otherwise, restart.

If G_0 and G_1 are isomorphic, the distribution of $\sigma(G_0)$ is identical to the distribution of $\sigma(G_1)$. Therefore, the distribution of b' is independent of b, so the probability that $b' = b$ is $1/2$. Therfore step 3 succeeds with probability $1/2$.

Since b' is independent of b, the distribution of H conditional on S halting is the same as the distribution of H. b' and σ are exactly what would be sent in the real protocol.

Interactive proofs and complexity

IP is the class of problems decidable by an interactive proof system where the number of bits exchanged is polynomial, V is polynomial time, and P is unbounded.

- $\mathsf{P} \subseteq \mathsf{BPP} \subseteq \mathsf{IP}$
- $\mathsf{NP} \subseteq \mathsf{IP}$, but deterministic IP gives NP and perfectly sound IP also gives NP.
- Languages in P and BPP also have zero-knowledge proofs.
- Languages in NP have zero-knowledge proofs if we assume some crypto primitives.