Introduction to Differential Privacy

Daniel Kifer

NBER Summer Institute

July 17, 2020

Outline

Survey Data

- Goal: collect surveys and publish statistics about the U.S. population.

Survey Data

- Goal: collect surveys and publish statistics about the U.S. population.
- Challenge: protect the confidentiality of respondents.
  - Linking attacks (commercial databases and other external datasets)
  - Billions of statistics about millions of people
  - Advanced data science reconstruction algorithms [DN03] + computation
  - Title 13 of the US Code:
    - Penalties for disclosure of private information: $\leq$ 5 years in prison, $\leq$ \$250,000 fine.
    - If it crosses paths with IRS data (Title 26): $\leq$ 22 years in prison.

Survey Data

- Goal: collect surveys and publish statistics about the U.S. population.
- Challenge: protect the confidentiality of respondents.
  - Linking attacks (commercial databases and other external datasets)
  - Billions of statistics about millions of people
  - Advanced data science reconstruction algorithms [DN03] + computation
  - Title 13 of the US Code:
    - Penalties for disclosure of private information: $\leq$ 5 years in prison, $\leq$ \$250,000 fine.
    - If it crosses paths with IRS data (Title 26): $\leq$ 22 years in prison.
- 2008 OnTheMap (https://onthemap.ces.census.gov/): first large-scale deployment of differential privacy.
- 2020 Decennial Census of Populaton and Housing.
- "Central" model.

How do people use their browsers?

- Goal: understand Chrome browser usage through browser settings.
  - Homepages
  - Plugins

How do people use their browsers?

- Goal: understand Chrome browser usage through browser settings.
  - Homepages
  - Plugins
- Challenge: massive invasion of privacy
  - Needs user consent
  - Low levels of opt-in data collection

How do people use their browsers?

- Goal: understand Chrome browser usage through browser settings.
  - Homepages
  - Plugins
- Challenge: massive invasion of privacy
  - Needs user consent
  - Low levels of opt-in data collection
- 2014 RAPPOR: Chrome browser sends noisy bits to Google.
  - Noise protects your information from Google
  - Aggregating across users reveals population statistics (popular plugins, etc.)
  - "Local" model.

How do people use their devices?

- Goal: understand mobile device usage.
  - Websites people visit (for link recommendation)
  - Words people type (for predictive keyboards)
  - Emojis people use 😊 (for when you are at a loss for words 🤔)

How do people use their devices?

- Goal: understand mobile device usage.
  - Websites people visit (for link recommendation)
  - Words people type (for predictive keyboards)
  - Emojis people use 😊 (for when you are at a loss for words 🤔)
- Challenge: keyloggers ⌨ are malware 🙈😠😈

How do people use their devices?

- Goal: understand mobile device usage.
  - Websites people visit (for link recommendation)
  - Words people type (for predictive keyboards)
  - Emojis people use 😊 (for when you are at a loss for words 🤔)
- Challenge: keyloggers ⌨ are malware 🙈😠😈
- 2016 Apple announces their plan to deploy differential privacy.
  - Usage information converted to bits
  - Bits are randomly perturbed
  - Noise protects your information from Apple
  - Aggregating across users reveals population statistics
  - "Local" model.

Et Alia

- ... Microsoft telemetry, Facebook url shares, Samsung, Uber, etc.
- Differential privacy enables study of data that is otherwise inaccessible.
- Trust models
  - Central Model
    - trusted data collector
    - data perturbed after collection
  - Local Model:
    - untrusted data collector
    - data perturbed before collection
    - lower accuracy of published statistics
- Transparency: access to source code does not increase privacy risk.
  - Source code can be released.

Et Alia

- ... Microsoft telemetry, Facebook url shares, Samsung, Uber, etc.
- Differential privacy enables study of data that is otherwise inaccessible.
- Trust models
  - Central Model
    - trusted data collector
    - data perturbed after collection
  - Local Model:
    - untrusted data collector
    - data perturbed before collection
    - lower accuracy of published statistics
- Transparency: access to source code does not increase privacy risk.
  - Source code ~~can~~ should be released.
  - Census end-to-end test: https://github.com/uscensusbureau/census2020-das-e2e
  - Google RAPPOR: https://github.com/google/rappor
  - OpenDP: https://github.com/opendifferentialprivacy
  - Many others ...

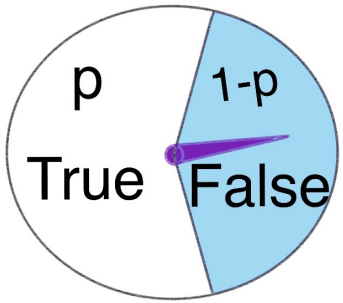
Outline

Meanwhile, in 1965 ...

- Differential privacy officially invented in 2006 [DMNS06]
- Mechanisms for differential privacy existed in 1965 [War65]
- Face-to-face survey: "have you ever engaged in insider trading?"
  - Respondents are likely to lie
  - or withhold information

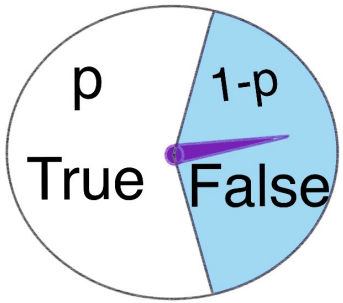
Meanwhile, in 1965 ...

- Differential privacy officially invented in 2006 [DMNS06]
- Mechanisms for differential privacy existed in 1965 [War65]
- Face-to-face survey: "have you ever engaged in insider trading?"
  - Respondents are likely to lie
  - or withhold information
- Warner's Spinner:
  - Only the respondent sees spinner.
  - $P(True) = p > \frac{1}{2}$
  - If arrow lands on "True", answer truthfully
  - If arrow lands on "False", lie



Meanwhile, in 1965 ...

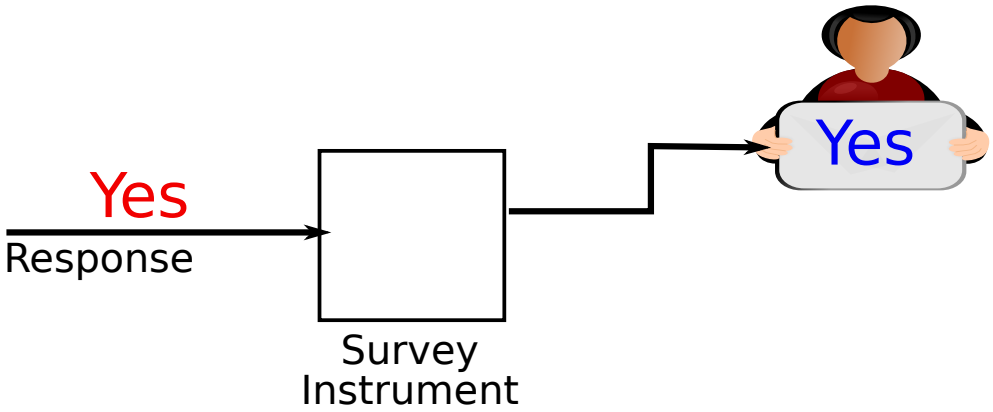
- Differential privacy officially invented in 2006 [DMNS06]
- Mechanisms for differential privacy existed in 1965 [War65]
- Face-to-face survey: "have you ever engaged in insider trading?"
  - Respondents are likely to lie
  - or withhold information
- Warner's Spinner:
  - Only the respondent sees spinner.
  - $P(True) = p > \frac{1}{2}$
  - If arrow lands on "True", answer truthfully
  - If arrow lands on "False", lie



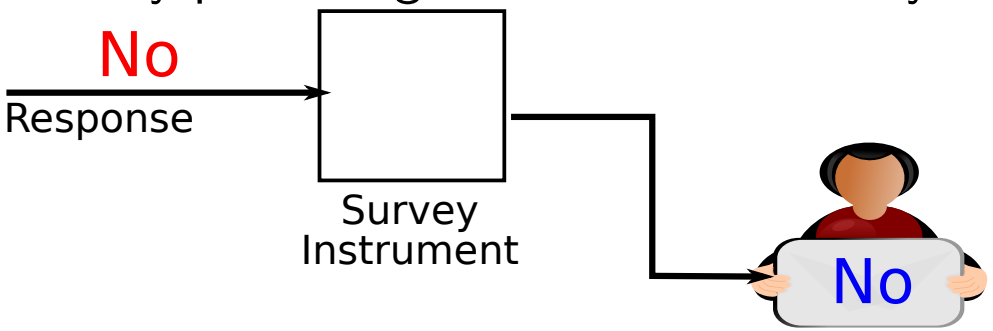
- If respondents use this mechanism, their guarantee is:
  - Their information is protected almost as well as if they lied strategically.
  - Protection only relies on randomness in mechanism.
  - Protection does not rely on prior beliefs.
- Note: mechanism is public, randomness is not.

A Tale of Two Mechanisms

- Pre-Warner face-to-face surveys.
- Suppose real status is True: engaged in insider trading.
- Two Options:
  - Factual World: submit correct response into survey instrument.



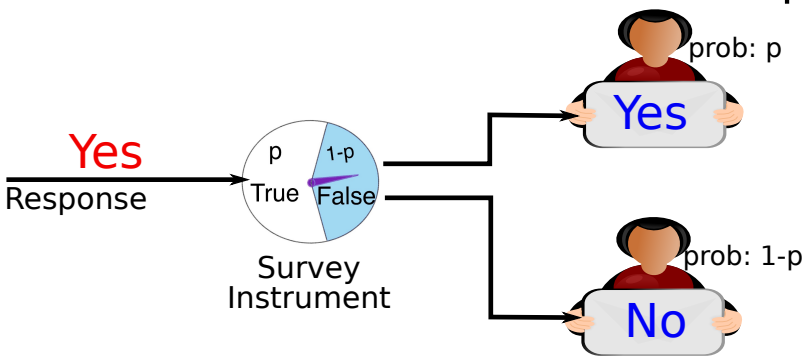
  - Privacy-preserving counterfactual: always submit denial.



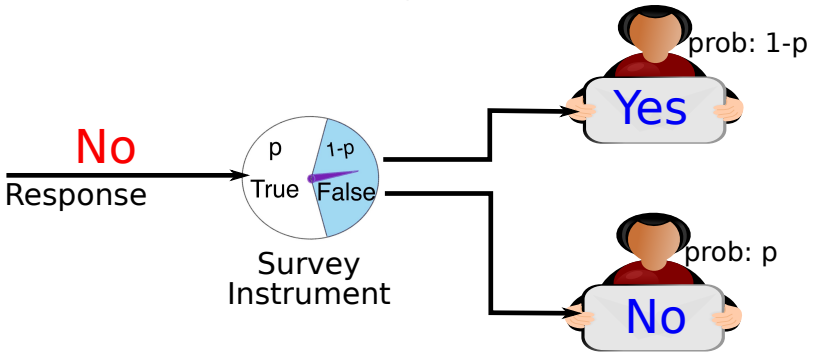
- Conclusion: strong incentive against being factual.

A Tale of Two Mechanisms

- Post-Warner face-to-face surveys.
- Suppose real status is True: engaged in insider trading.
- Two Options:
  - Factual World: submit correct response into survey instrument.



  - Privacy-preserving counterfactual: always submit denial.



- Is there still an incentive to lie?

Privacy of Randomized Response

- Spinner M, $p > \frac{1}{2}$
- Factual World W_f: submit correct response into survey method.
  - Input: Yes
  - Output: Yes with prob p, No with prob 1-p.
- Privacy-preserving counterfactual world W_p: always deny.
  - Input: No
  - Output: Yes with prob 1-p, No with prob p.

Privacy of Randomized Response

- Spinner M, $p > \frac{1}{2}$
- Factual World W_f: submit correct response into survey method.
  - Input: Yes
  - Output: Yes with prob p, No with prob 1-p.
- Privacy-preserving counterfactual world W_p: always deny.
  - Input: No
  - Output: Yes with prob 1-p, No with prob p.
- What is the difference? Compared to privacy-preserving world:
  - Probabilities of Yes increase by a factor at most $p/(1-p)$.
  - Probabilities of No decrease by a factor at most $(1-p)/p$.

Privacy of Randomized Response

- Spinner M, $p > \frac{1}{2}$
- Factual World W_f: submit correct response into survey method.
  - Input: Yes
  - Output: Yes with prob p, No with prob 1-p.
- Privacy-preserving counterfactual world W_p: always deny.
  - Input: No
  - Output: Yes with prob 1-p, No with prob p.
- What is the difference? Compared to privacy-preserving world:
  - Probabilities of Yes increase by a factor at most $p/(1-p)$.
  - Probabilities of No decrease by a factor at most $(1-p)/p$.
  - p close to $\frac{1}{2} \Rightarrow$ more privacy.
  - p far from $\frac{1}{2} \Rightarrow$ less privacy.

Privacy of Randomized Response

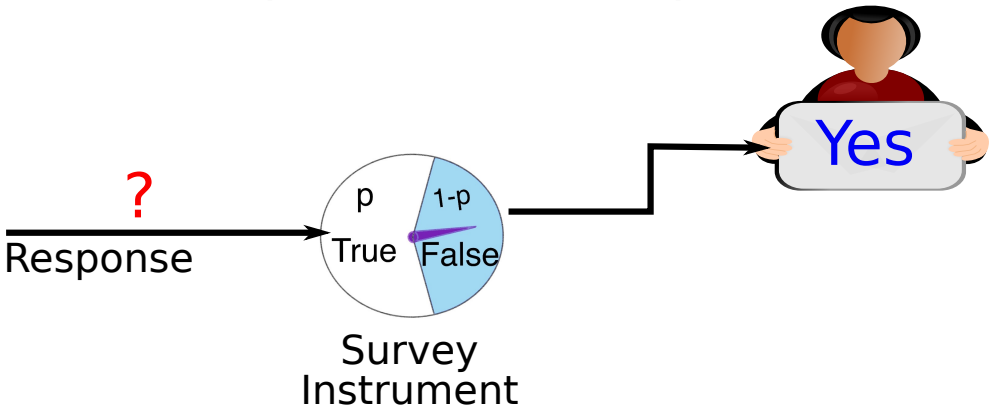
- Spinner M, $p > \frac{1}{2}$
- Factual World W_f: submit correct response into survey method.
  - Input: Yes
  - Output: Yes with prob p, No with prob 1-p.
- Privacy-preserving counterfactual world W_p: always deny.
  - Input: No
  - Output: Yes with prob 1-p, No with prob p.
- What is the difference? Compared to privacy-preserving world:
  - Probabilities of Yes increase by a factor at most $p/(1-p)$.
  - Probabilities of No decrease by a factor at most $(1-p)/p$.
- For any event E (output = True or output = False):

$$\frac{1-p}{p} P(M(W_p) = E) \leq P(M(W_f) = E) \leq \frac{p}{1-p} P(M(W_p) = E)$$

- Let's set $\epsilon = \text{natural}\log \frac{p}{1-p}$ (i.e., ϵ is the log odds).

Interpretation I (Bayesian)

- $\epsilon = \log \frac{p}{1-p}$
- Data snooper's prior belief that Bob participated in insider trading: q.
- Data snooper observes output "Yes"

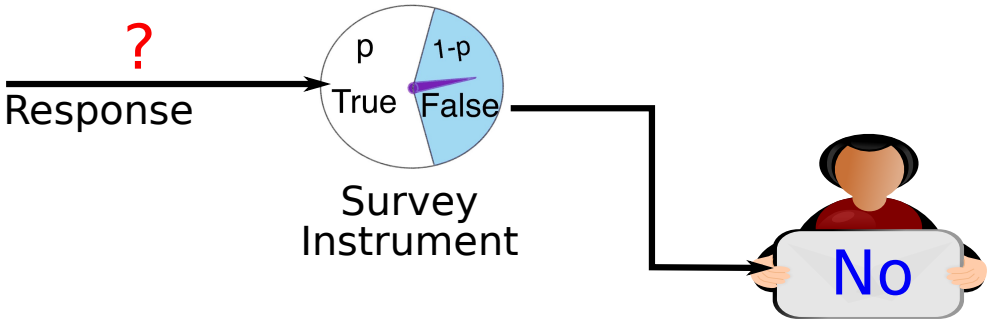


- Spinner guarantees posterior odds are similar to prior odds:

$$
\begin{aligned}
e^{-\epsilon} &= \frac{1-p}{p} \\
&\leq \frac{P(response = Yes \mid output = Yes)}{P(response = No \mid output = Yes)} \Big/ \frac{P(response = Yes)}{P(response = No)} \\
&\leq \frac{p}{1-p} = e^{\epsilon}
\end{aligned}
$$

Interpretation I (Bayesian)

- $\epsilon = \log \frac{p}{1-p}$
- Data snooper's prior belief that Bob participated in insider trading: q.
- Data snooper observes output "No"

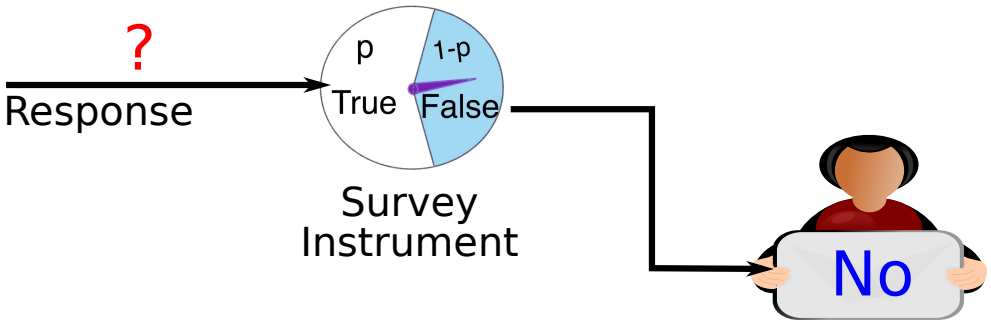


- Spinner guarantees posterior odds are similar to prior odds:

$$
\begin{aligned}
e^{-\epsilon} &= \frac{1-p}{p} \\
&\leq \frac{P(response = Yes \mid output = No)}{P(response = No \mid output = No)} \Big/ \frac{P(response = Yes)}{P(response = No)} \\
&\leq \frac{p}{1-p} = e^{\epsilon}
\end{aligned}
$$

Interpretation I (Bayesian)

- $\epsilon = \log \frac{p}{1-p}$
- Data snooper's prior belief that Bob participated in insider trading: q.
- Data snooper observes output "No"



- Spinner guarantees posterior odds are similar to prior odds:

$$
\begin{aligned}
e^{-\epsilon} &= \frac{1-p}{p} \\
&\leq \frac{P(response = Yes \mid output = No)}{P(response = No \mid output = No)} \Big/ \frac{P(response = Yes)}{P(response = No)} \\
&\leq \frac{p}{1-p} = e^{\epsilon}
\end{aligned}
$$

- No matter what happens, odds change by factor at most e^{ϵ}.

$$e^{1} \approx 2.72 \qquad e^{0.5} \approx 1.65 \qquad e^{0.1} \approx 1.11$$

Interpretation II (Frequentist)

- $p = 0.55$
  - $\frac{p}{1-p} \approx 1.22$
  - $\epsilon = \log \frac{p}{1-p} \approx 0.2$
- Null hypothesis: Bob did not engage in insider trading.
- Alternative hypothesis: Bob is guilty
- P(observed "Yes" | Null hypothesis) = 0.45
- P(observed "No" | Null hypothesis) = 0.55
- No matter what snooper observed, not enough evidence to reject null hypothesis at reasonable levels.

Interpretation II (Frequentist)

- $p = 0.55$
  - $\frac{p}{1-p} \approx 1.22$
  - $\epsilon = \log \frac{p}{1-p} \approx 0.2$
- Null hypothesis: Bob did not engage in insider trading.
- Alternative hypothesis: Bob is guilty
- P(observed "Yes" | Null hypothesis) = 0.45
- P(observed "No" | Null hypothesis) = 0.55
- No matter what snooper observed, not enough evidence to reject null hypothesis at reasonable levels.
- Composition: what happens when multiple information sources are combined.
  - Suppose snooper has evidence about Bob.
  - Later snooper observes result of randomized response.
  - Randomized response causes $\frac{\text{power}}{\text{type I error}}$ to change by at most e^{ϵ}

Interpretation III (For the Math Phobic)

- Suppose Bob is in a survey with 100 other people.
- Suppose Bob is the only one to trade stocks based on non-public information.
- Survey uses randomized response with $p = 0.55$.
  - Almost even chance that Bob's true response is changed to "No"
  - Even if it remains unaltered, ≈ 45 other people's response changed to "Yes"
  - Bob is in a crowd of about 45 other people
  - Even if interviewer knew 1 person is guilty, can't pick out which one

Interpretation III (For the Math Phobic)

- Suppose Bob is in a survey with 100 other people.
- Suppose Bob is the only one to trade stocks based on non-public information.
- Survey uses randomized response with $p = 0.55$.
  - Almost even chance that Bob's true response is changed to "No"
  - Even if it remains unaltered, ≈ 45 other people's response changed to "Yes"
  - Bob is in a crowd of about 45 other people
  - Even if interviewer knew 1 person is guilty, can't pick out which one
- Here we used uncertainty about the data to show randomized response preserves uncertainty.
- Randomized response guarantees can be strengthened by considering data uncertainty.
- Randomized response still protects you even without resorting to data uncertainty.

Applications to Politics

- Newly introduced S9999: CONTROVERSY Act
- How many senators truly think it is a good idea?
- Public statements may differ from private beliefs.
  - Worries about re-election.
  - Prior deals.

Applications to Politics

- Newly introduced S9999: CONTROVERSY Act
- How many senators truly think it is a good idea?
- Public statements may differ from private beliefs.
  - Worries about re-election.
  - Prior deals.
  - Other exogenous concerns.

Applications to Politics

- Newly introduced S9999: CONTROVERSY Act
- How many senators truly think it is a good idea?
- Public statements may differ from private beliefs.
  - Worries about re-election.
  - Prior deals.
  - Other exogenous concerns.

- With randomized response
  - Disincentives for factual response: slight information leakage.
  - Incentives: curiosity about overall senate perception.

Applications to Politics

- Newly introduced S9999: CONTROVERSY Act
- How many senators truly think it is a good idea?
- Public statements may differ from private beliefs.
  - Worries about re-election.
  - Prior deals.
  - Other exogenous concerns.

- With randomized response
  - Disincentives for factual response: slight information leakage.
  - Incentives: curiosity about overall senate perception.
- π_s: (unknown) proportion of senators privately supporting the bill
  - Suppose we receive Y randomized response reports equal to "yes"

Applications to Politics

- Newly introduced S9999: CONTROVERSY Act
- How many senators truly think it is a good idea?
- Public statements may differ from private beliefs.
  - Worries about re-election.
  - Prior deals.
  - Other exogenous concerns.

- With randomized response
  - Disincentives for factual response: slight information leakage.
  - Incentives: curiosity about overall senate perception.
- π_s: (unknown) proportion of senators privately supporting the bill
  - Suppose we receive Y randomized response reports equal to "yes"
  - Do not treat $Y/100$ as the estimate of π_s

Applications to Politics

- Newly introduced S9999: CONTROVERSY Act
- How many senators truly think it is a good idea?
- Public statements may differ from private beliefs.
  - Worries about re-election.
  - Prior deals.
  - Other exogenous concerns.

- With randomized response
  - Disincentives for factual response: slight information leakage.
  - Incentives: curiosity about overall senate perception.
- π_s: (unknown) proportion of senators privately supporting the bill
  - Suppose we receive Y randomized response reports equal to "yes"
  - How do we estimate π_s? [War65]
    - $\hat{\pi}_s = \frac{p-1}{2p-1} + \frac{Y}{100(2p-1)}$, unbiased
    - $std(\hat{\pi}_s) = \frac{1}{10}\sqrt{\frac{1}{16(p-1/2)^2} - (\pi_s - 1/2)^2)}$
    - For $p = 0.6$, $std(\hat{\pi}_s) \approx 0.25$. (disappointing, can we do better?)

The House?

- Senate population size: 100
- House of Representatives: 435
- π_r: (unknown) prop. of representatives privately supporting the bill

The House?

- Senate population size: 100
- House of Representatives: 435
- π_r: (unknown) prop. of representatives privately supporting the bill
  - Suppose we receive Y randomized response reports equal to "yes"
  - $\hat{\pi}_r = \frac{p-1}{2p-1} + \frac{Y}{435(2p-1)}$, unbiased
  - $std(\hat{\pi}_r) = \frac{1}{\sqrt{435}} \sqrt{\frac{1}{16(p-1/2)^2} - (\pi_r - 1/2)^2)}$
  - For $p = 0.6$, $std(\hat{\pi}_r) \approx 0.117$.
- Standard Deviation decreases like $\frac{1}{\sqrt{n}}$

The House?

- Senate population size: 100
- House of Representatives: 435
- π_r: (unknown) prop. of representatives privately supporting the bill
  - Suppose we receive Y randomized response reports equal to "yes"
  - $\hat{\pi}_r = \frac{p-1}{2p-1} + \frac{Y}{435(2p-1)}$, unbiased
  - $std(\hat{\pi}_r) = \frac{1}{\sqrt{435}} \sqrt{\frac{1}{16(p-1/2)^2} - (\pi_r - 1/2)^2)}$
  - For $p = 0.6$, $std(\hat{\pi}_r) \approx 0.117$.
- Standard Deviation decreases like $\frac{1}{\sqrt{n}}$
- Can we do better?
  - Same privacy guarantees
  - More accuracy

The House?

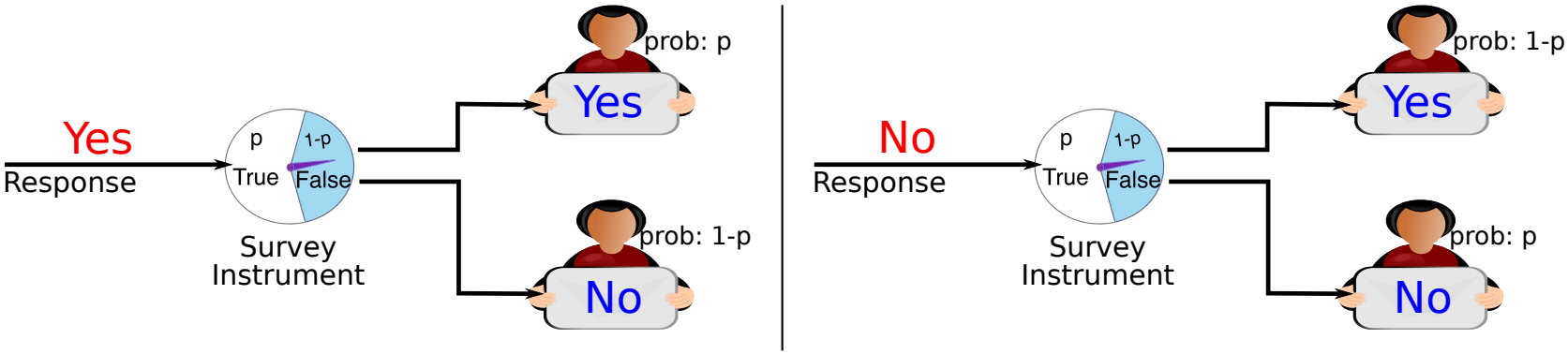
- Senate population size: 100
- House of Representatives: 435
- π_r: (unknown) prop. of representatives privately supporting the bill
  - Suppose we receive Y randomized response reports equal to "yes"
  - $\hat{\pi}_r = \frac{p-1}{2p-1} + \frac{Y}{435(2p-1)}$, unbiased
  - $std(\hat{\pi}_r) = \frac{1}{\sqrt{435}} \sqrt{\frac{1}{16(p-1/2)^2} - (\pi_r - 1/2)^2)}$
  - For $p = 0.6$, $std(\hat{\pi}_r) \approx 0.117$.
- Standard Deviation decreases like $\frac{1}{\sqrt{n}}$
- Can we do better?
  - Same privacy guarantees
  - More accuracy
- Yes, if we have a trusted data collector.
- Previously:
  - We were in the local model.
  - Respondents did not trust the data collector.

A Trusted Data Collector

- Suppose we have a trusted data collector.
  - Senators submit truthful responses to the collector.
  - Collector counts up "yes" responses, publishes "information" about this.
  - Note: the data collector cannot release the exact count.
- Desired guarantee:
  - Even if 99 senators collude ...
  - remaining senator's data as well protected as with randomized response.

A Trusted Data Collector

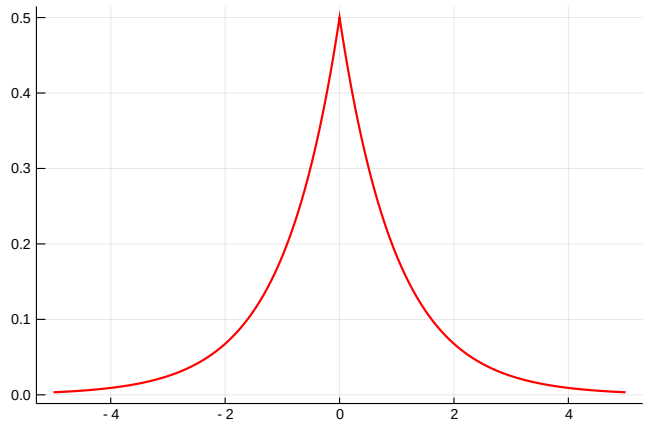
- Suppose we have a trusted data collector.
  - Senators submit truthful responses to the collector.
  - Collector counts up "yes" responses, publishes "information" about this.
  - Note: the data collector cannot release the exact count.
- Desired guarantee:
  - Even if 99 senators collude ...
  - remaining senator's data as well protected as with randomized response.
- Recall randomized response:



  - Probability of any output event increases or decreases by factor between $(1-p)/p$ and $p/(1-p)$ if a person changes their response.
  - $\epsilon = \log \frac{p}{1-p}$.

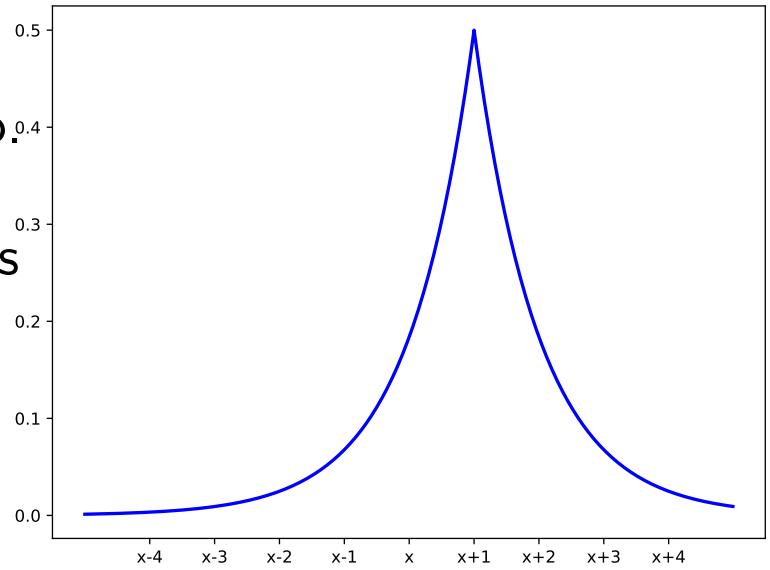
A Trusted Data Collector

- Suppose we have a trusted data collector.
  - Senators submit truthful responses to the collector.
  - Collector counts up "yes" responses, publishes "information" about this.
  - Note: the data collector cannot release the exact count.
- Desired guarantee:
  - Even if 99 senators collude ...
  - remaining senator's data as well protected as with randomized response.
- So data collector adds Laplace$(1/\epsilon)$ noise to # of yes responses [DMNS06].
  - $f(x; 1/\epsilon) = \frac{\epsilon}{2} e^{-\epsilon|x|}$
  - Variance$= 2/\epsilon^2$
- Does it maintain Privacy?
- Is it more accurate?

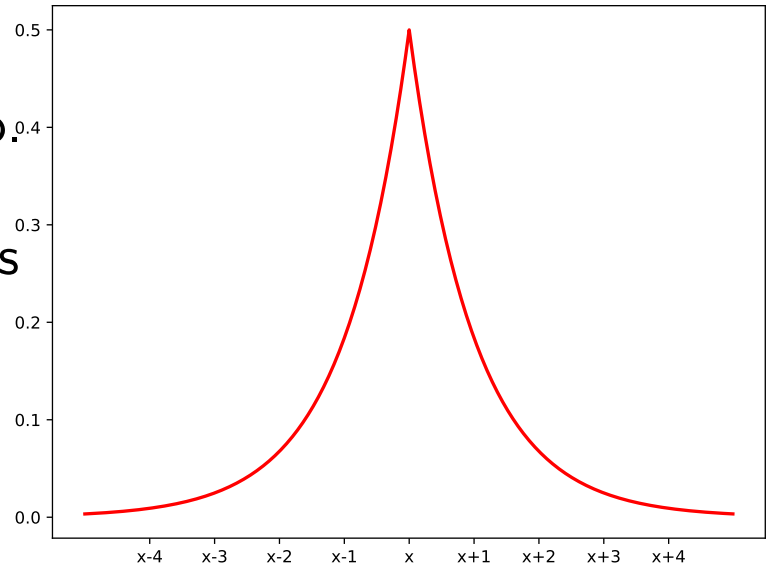


The Mood in the Senate

- Case 1:
  - 99 senators: x yes and 99-x no.
  - Senator 100: "yes"
  - Trusted data collector publishes $x+1+\text{Laplace}(1/\epsilon)$

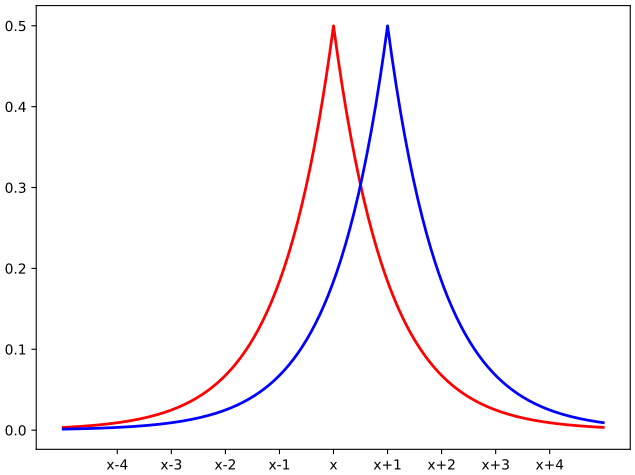


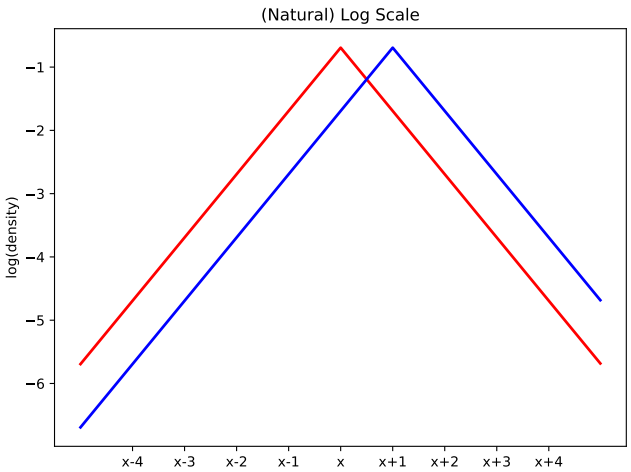
- Case 2:
  - 99 senators: x yes and 99-x no.
  - Senator 100: "no"
  - Trusted data collector publishes $x+\text{Laplace}(1/\epsilon)$



Comparing Densities

- The 2 densities (Senator 100 responding yes vs. no):

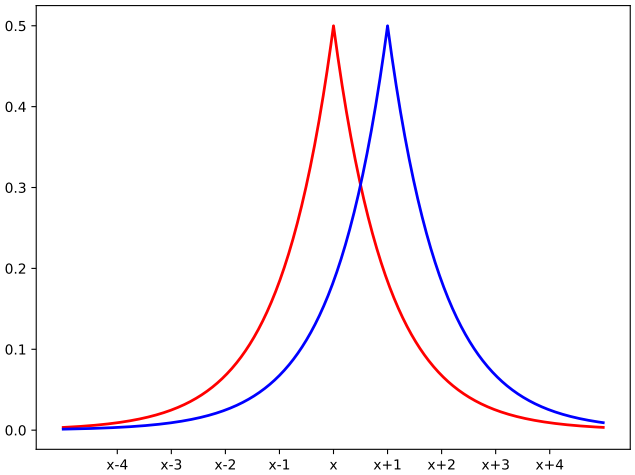


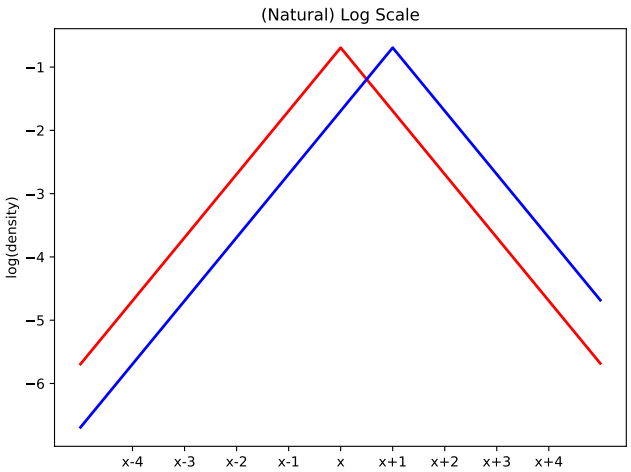


- The ratio of densities is bounded between $e^{-\epsilon}$ and e^{ϵ}.
- Senator 100 changing yes response to no response increases/decreases density by factor of at most e^{ϵ}.

Comparing Densities

- The 2 densities (Senator 100 responding yes vs. no):



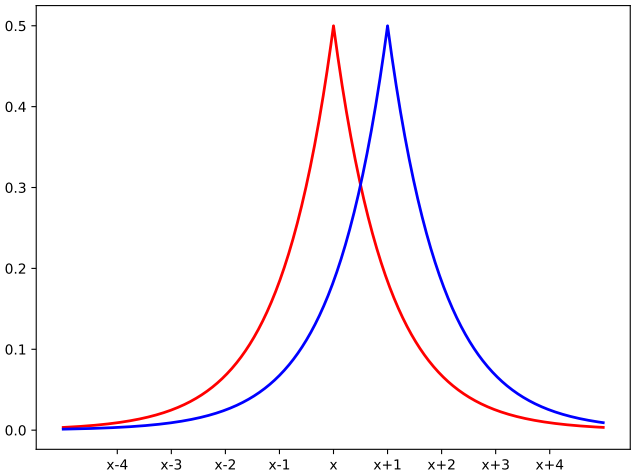


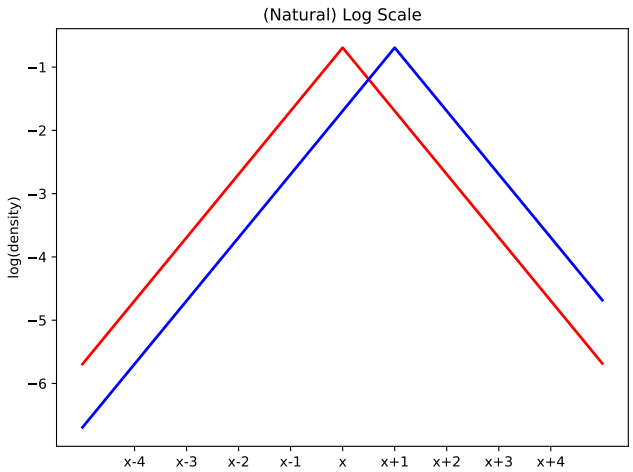
- The ratio of densities is bounded between $e^{-\epsilon}$ and e^{ϵ}.
- Senator 100 changing yes response to no response increases/decreases density by factor of at most e^{ϵ}.
- Thus, for any set E:

$$P(output \in E \mid Senator\ 100 = yes) \leq e^{\epsilon} P(output \in E \mid Senator\ 100 = no)$$
$$P(output \in E \mid Senator\ 100 = no) \leq e^{\epsilon} P(output \in E \mid Senator\ 100 = yes)$$

Comparing Densities

- The 2 densities (Senator 100 responding yes vs. no):





- The ratio of densities is bounded between $e^{-\epsilon}$ and e^{ϵ}.
- Senator 100 changing yes response to no response increases/decreases density by factor of at most e^{ϵ}.
- Thus, for any set E:

$$P(output \in E \mid Senator\ 100 = yes) \leq e^{\epsilon} P(output \in E \mid Senator\ 100 = no)$$

$$P(output \in E \mid Senator\ 100 = no) \leq e^{\epsilon} P(output \in E \mid Senator\ 100 = yes)$$

- Same guarantees as for randomized response with $\epsilon = \log \frac{p}{1-p}$.
  - Low power in distinguishing between Senator 100 = yes vs. Senator 100 = no.

Accuracy

- $p = 0.6$ and $\epsilon = \log p/(1-p) \approx 0.405$
- Senate:
  - π_s: proportions of senators privately supporting bill
  - standard deviation of estimate:
    - Under randomized response: ≈ 0.25
    - Under the Laplace mechanism: ≈ 0.035
- House:
  - π_r: proportion of representatives privately supporting bill
  - standard deviation of estimate:
    - Under randomized response: ≈ 0.117
    - Under the Laplace mechanism: ≈ 0.008

Accuracy

- $p = 0.6$ and $\epsilon = \log p/(1-p) \approx 0.405$
- Senate:
  - π_s: proportions of senators privately supporting bill
  - standard deviation of estimate:
    - Under randomized response: ≈ 0.25
    - Under the Laplace mechanism: ≈ 0.035
- House:
  - π_r: proportion of representatives privately supporting bill
  - standard deviation of estimate:
    - Under randomized response: ≈ 0.117
    - Under the Laplace mechanism: ≈ 0.008
- Asymptotically, standard deviation due to privacy decreases like
  - $1/\sqrt{n}$ under randomized response.
  - $1/n$ under the Laplace mechanism.
- Transparency! In both cases:
  - Full details of how mechanism works can be made public.
  - Only the randomness must be kept secret.
  - Allows inferences to be adjusted.

Outline

Differential Privacy

- Two databases D_1 and D_2 are neighbors if:
  - D_1 is the result of changing one record in D_2.
  - Number of respondents n remains the same (we will revisit this).

Differential Privacy

- Two databases D_1 and D_2 are neighbors if:
  - D_1 is the result of changing one record in D_2.
  - Number of respondents n remains the same (we will revisit this).

Definition (Differential Privacy [DMNS06])

Given a privacy loss budget $\epsilon > 0$, an randomized algorithm M satisfies ϵ-differential privacy if for all $E \subset \text{range}(M)$ and all pairs of databases D_1, D_2 that are neighbors of each other,

$$P(M(D_1) \in E) \leq e^{\epsilon} P(M(D_2) \in E)$$

Differential Privacy

- Two databases D_1 and D_2 are neighbors if:
  - D_1 is the result of changing one record in D_2.
  - Number of respondents n remains the same (we will revisit this).

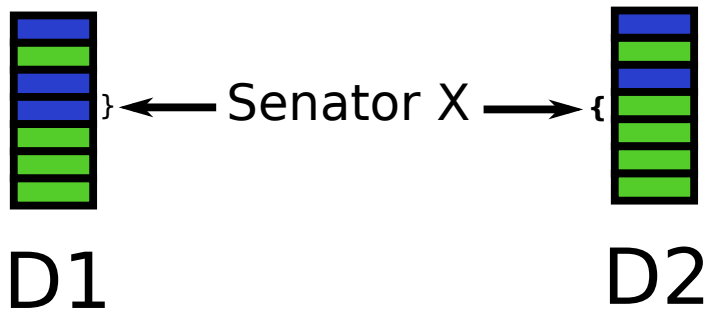
Definition (Differential Privacy [DMNS06])

Given a privacy loss budget $\epsilon > 0$, an randomized algorithm M satisfies ϵ-differential privacy if for all $E \subset \text{range}(M)$ and all pairs of databases D_1, D_2 that are neighbors of each other,

$$P(M(D_1) \in E) \leq e^{\epsilon} P(M(D_2) \in E)$$

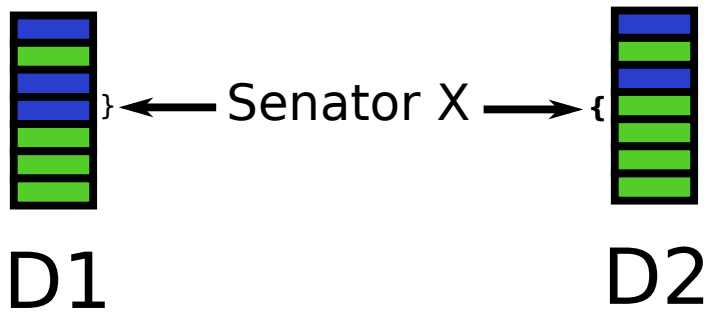
- All pairs of neighbors, not just neighbors of current database.
- M must be a randomized algorithm, e.g.,
  - Randomized Response ✓
  - Laplace Mechanism ✓
  - Publish true count ✗
- Probability is over randomness in M only (not uncertainty in data).
  - Privacy is a function of the mechanism, not the data.

Bounded Neighbors



- D_1 and D_2 are neighbors:
  - so $P(M(D_1) \in E) \leq e^{\epsilon} P(M(D_2) \in E)$ for all E
- D_2 and D_1 are neighbors:
  - so $P(M(D_2) \in E) \leq e^{\epsilon} P(M(D_1) \in E)$ for all E
- Regardless of what other senators do,
  - Noise masks senator X's response (blue vs green).

Bounded Neighbors



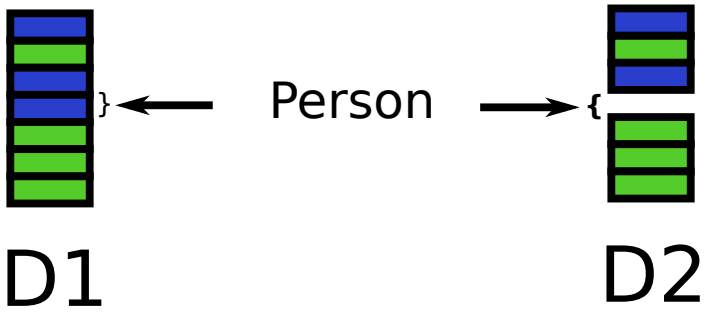
- D_1 and D_2 are neighbors:
  - so $P(M(D_1) \in E) \leq e^{\epsilon} P(M(D_2) \in E)$ for all E
- D_2 and D_1 are neighbors:
  - so $P(M(D_2) \in E) \leq e^{\epsilon} P(M(D_1) \in E)$ for all E
- Regardless of what other senators do,
  - Noise masks senator X's response (blue vs green).
- Neighbors that differ on the response of
  - Senator 1 protect the privacy of Senator 1
    - There are 2^{100} such pairs of neighbors.
  - Senator 2 protect the privacy of Senator 2
    - There are 2^{100} such pairs pairs of neighbors.
  - And so on. Each senator is protected.
- The differential privacy equations must hold for all of these pairs.

Unbounded Neighbors

- But what if I want to hide not just responses, but also hide participation?
  - e.g., in statistical uses of IRS data, even fact of filing is protected.
  - e.g., participation in STD study.

Unbounded Neighbors

- But what if I want to hide not just responses, but also hide participation?
  - e.g., in statistical uses of IRS data, even fact of filing is protected.
  - e.g., participation in STD study.
- Redefine neighbors!
- D_1 and D_2 are neighbors if either:
  - D_1 is obtained from D_2 by removing a record.
  - D_2 is obtained from D_1 by removing a record.



- Whether you opt-in or not:
  - Inference about your participation is protected.
  - Inference about your record is protected.

Action-level Neighbors

- Database 1:

Customer ID	Purchase History
1	{shrimp, lobster, crab, mussel, ...}
2	{beer, wine, bourbon, ...}
3	{cookie, cookie, cookie, ...}
⋮	{...}

- Database 2:

Customer ID	Purchase History
1	{shrimp, lobster, mussel, ...}
2	{beer, wine, bourbon, ...}
3	{cookie, cookie, cookie, ...}
⋮	{...}

- D_1 and D_2 are neighbors if they differ on one action by one person.
- Users with many actions have higher privacy loss.
  - Protected from inference: customer 1 bought crab on 7/17/2020 (single action).
  - Not protected from inference: customer 1 likes seafood (result of many actions).

Neighbors Summary

- Bounded neighbors:
  - D_1 and D_2 differ on value of one record.
  - Number of respondents n can be released without noise.
  - Only use when n is public.

Neighbors Summary

- Bounded neighbors:
  - D_1 and D_2 differ on value of one record.
  - Number of respondents n can be released without noise.
  - Only use when n is public.
  - Why? Due to fragmentation.
    - One dataset about 32 year-old Hispanic women with cancer.
    - Another dataset about 32 year-old non-Hispanic women with flu.
    - Another dataset about 32 year-old Hispanic males with cancer.
    - Another dataset about 32 year-old non-Hispanic males with flu.
    - etc.
    - Allowing the # of respondents in each dataset to be revealed degrades privacy.

Neighbors Summary

- Bounded neighbors:
  - D_1 and D_2 differ on value of one record.
  - Number of respondents n can be released without noise.
  - Only use when n is public.
- Unbounded neighbors:
  - D_1 and D_2 differ on the presence/absence of one individual's data.
  - n cannot be released without noise.
  - Most recommended choice of neighbors.
- Action-level neighbors
  - D_1 and D_2 differ on one action of one person.
    - Protects inference about an item (e.g., specific purchase).
    - Privacy degrades for users with multiple purchases (allows inference about customer interests).

DP Summary

Definition (Differential Privacy [DMNS06])

Let $\epsilon > 0$. An randomized algorithm M satisfies ϵ-differential privacy if for all $E \subset \text{range}(M)$ and all pairs of databases D_1, D_2 that are neighbors of each other,

$$P(M(D_1) \in E) \leq e^{\epsilon} P(M(D_2) \in E)$$

- Protects individual records, reveals estimates of population properties.
- Resists linking attacks and other background knowledge attacks.
- Transparency – source code can be released.
- Privacy only depends on randomness in the mechanism.
- Noise is introduced to mask the effect of any particular individual.
- Central and Local models.

Thank You

Outline

https://www.nber.org/papers/w15703

Inaccurate age and sex data in the Census PUMS files:

Evidence and Implications

J. Trent Alexander
Minnesota Population Center, University of Minnesota

Michael Davern
National Opinion Research Center, University of Chicago

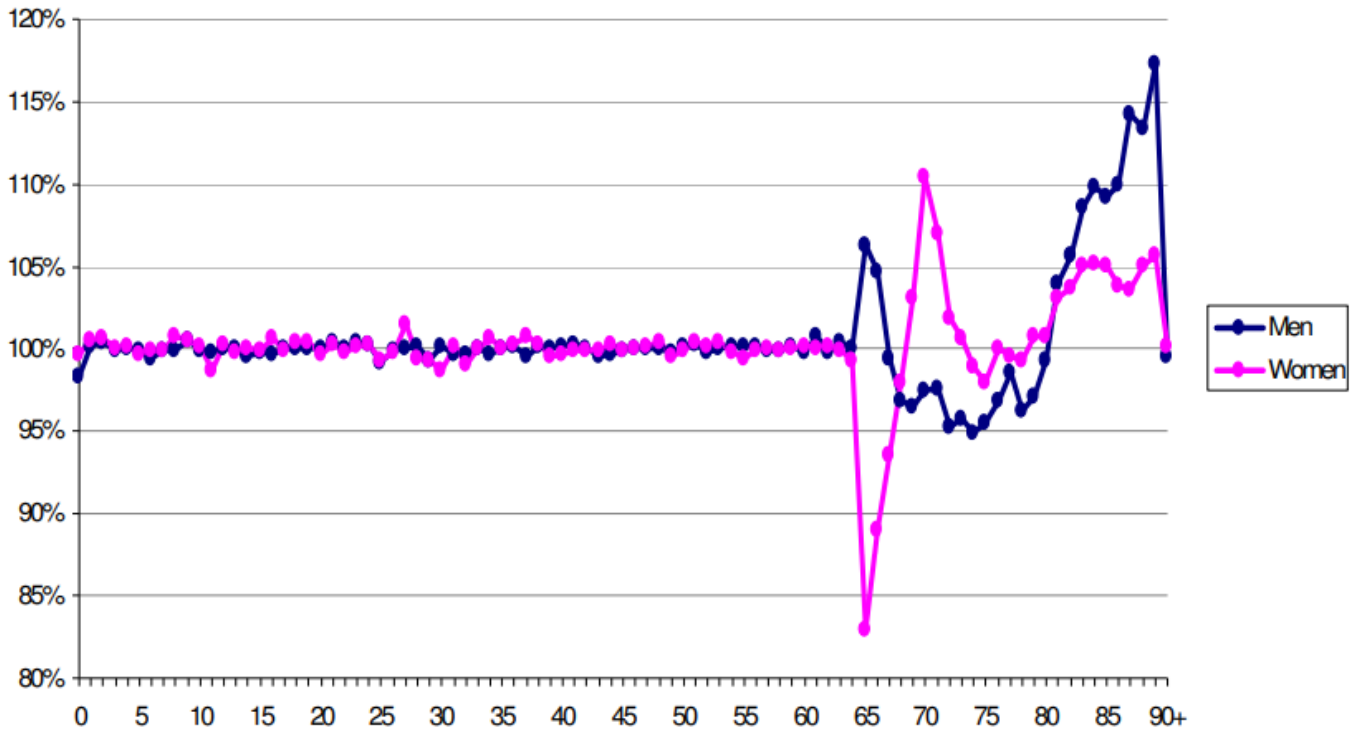
Betsey Stevenson
The Wharton School, University of Pennsylvania, CESifo, and NBER

Abstract

We discover and document errors in public use microdata samples ("PUMS files") of the 2000 Census, the 2003-2006 American Community Survey, and the 2004-2009 Current Population Survey. For women and men ages 65 and older, age- and sex-specific population estimates generated from the PUMS files differ by as much as 15% from counts in published data tables. Moreover, an analysis of labor force participation and marriage rates suggests the PUMS samples are not representative of the population at individual ages for those ages 65 and over. PUMS files substantially underestimate labor force participation of those near retirement ages and overestimate labor force participation rates of those at older ages. These problems were an unintentional by-product of the misapplication of a newer generation of disclosure avoidance procedures carried out on the data. The resulting errors in the public use data could significantly

Alexander, Davern, Stevenson, 2010

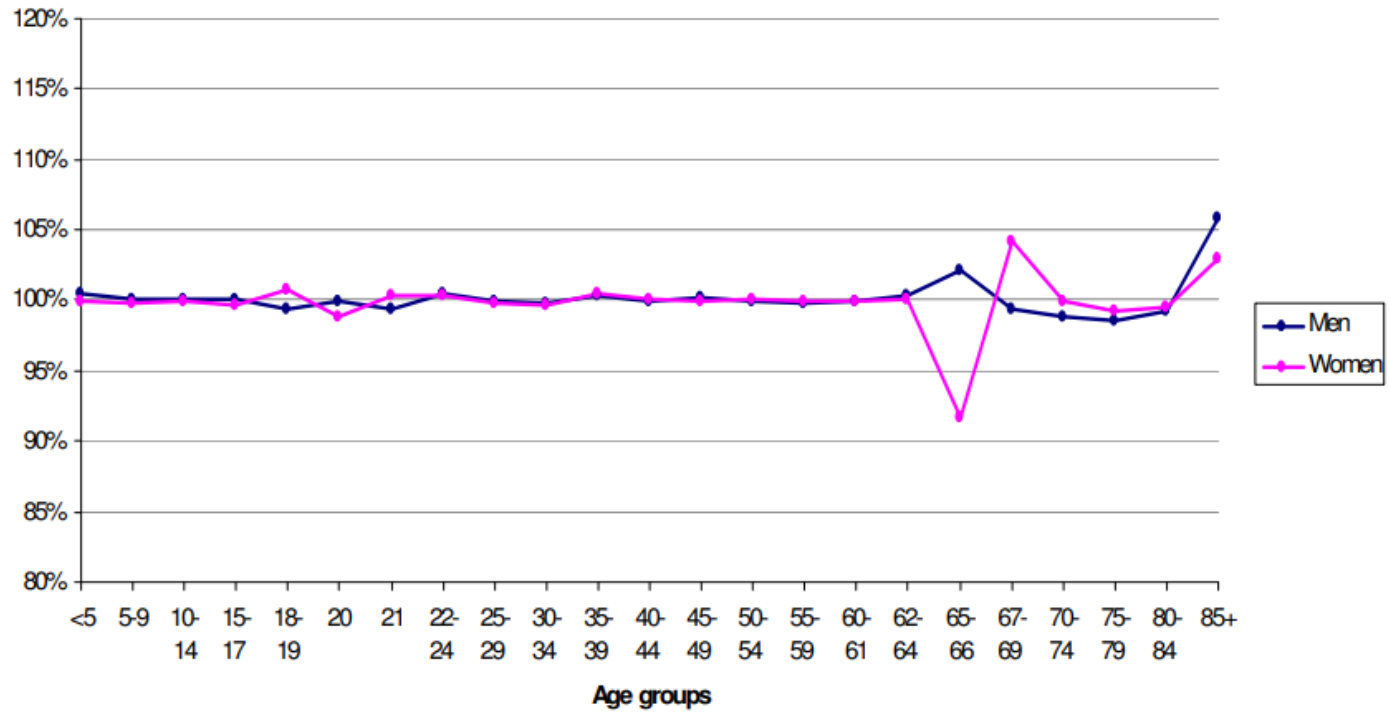
Figure 1. Population estimates from 2000 5% Census PUMS as a percentage of published data



Sources: Census 2000 Summary File 4, Table PCT3 (http://factfinder.census.gov); Census 2000 5% sample, IPUMS-USA (http://usa.ipums.org/).

Alexander, Davern, Stevenson, 2010

Figure 2. ACS 2006: Population estimates from PUMS as a percentage of published data



Sources: 2006 ACS Table B01001 (http://factfinder.census.gov); 2006 ACS PUMS, IPUMS-USA (http://usa.ipums.org/).

Transparency

- Legacy methods hide their source code and parameters.
  - Errors are hard to detect.
  - Inferences cannot be adjusted for disclosure avoidance control.
- Differential privacy allows source code to be published.

Practical Considerations

- Randomness
  - Differential privacy relies on non-predictable sources of randomness.
  - Random number generators used for statistics (defaults in R, python, etc.) are not secure enough.
- Floating point math
  - Much of differential privacy theory assumes real-valued numbers $x \in \mathbb{R}$.
  - Mismatch: computers use floating point representations [Mir12].
- These issues are exploitable.
- Ok for prototyping, but use secure system for deployment.
- OpenDP: https://github.com/opendifferentialprivacy/

Approximate Differential Privacy

- Differential Privacy relies heavily on Laplace noise.
- Is there a way to use Gaussian Noise with Differential Privacy?
  - Yes, but privacy guarantees are random [Mir17, DKM$^+$06].

Definition (Approximate Differential Privacy [DKM$^+$06])

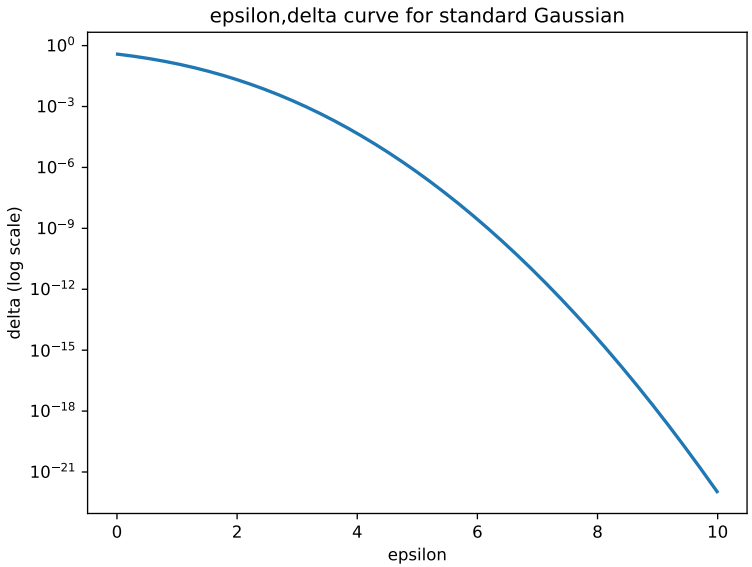
Given a $\epsilon > 0$ and $\delta \in [0, 1]$, an randomized algorithm M satisfies (ϵ, δ)-approximate differential privacy if for all $E \subset \text{range}(M)$ and all pairs of databases D_1, D_2 that are neighbors of each other,

$$P(M(D_1) \in E) \leq e^{\epsilon} P(M(D_2) \in E) + \delta$$

- δ is interpreted as probability ϵ-differential privacy guarantees may fail.
- Best examined through the ϵ, δ curve.

ϵ, δ curve for Gaussian Noise

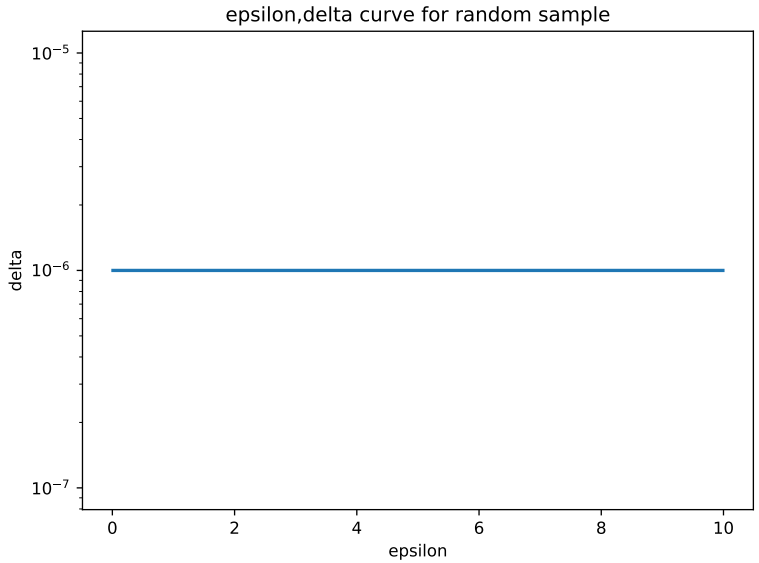
- Mechanism: number of yes responses in Senate + Gaussian noise ($\mu = 0, \sigma^2 = 1$).
- This mechanism satisfies ϵ, δ-approximate differential privacy for every ϵ, δ pair on this curve:



- Better composition properties: slower depletion of privacy budget.
- Privacy analysis must consider entire curve.

Use with Care

- Random sampling also satisfies ϵ-differential privacy.
- Mechanism: return 1 randomly selected record from database of size n.
- Satisfies approximate differential privacy with $\epsilon = 0$ and $\delta = 10^{-6}$
  - Yet someone's privacy is always violated.



References I

Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor.
Our data, ourselves: Privacy via distributed noise generation.
In EUROCRYPT, pages 486–503, 2006.

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith.
Calibrating noise to sensitivity in private data analysis.
In TCC, 2006.

Irit Dinur and Kobbi Nissim.
Revealing information while preserving privacy.
In PODS, 2003.

Ilya Mironov.
On significance of the least significant bits for differential privacy.
In CCS, 2012.

References II

Ilya Mironov.
Rényi differential privacy.
In 30th IEEE Computer Security Foundations Symposium, CSF 2017, Santa Barbara, CA, USA, August 21-25, 2017, pages 263–275, 2017.

S. L. Warner.
Randomized response: A survey technique for eliminating evasive answer bias.
Journal of the American Statistical Association, 1965.