Locally Differentially Private Frequency Estimation via Joint Randomized Response

Authors: Ye Zheng, Shafizur Rahman Seeam, Yidan Hu, Rui Zhang, Yanchao Zhang

RIT | Rochester Institute of Technology

University of Delaware

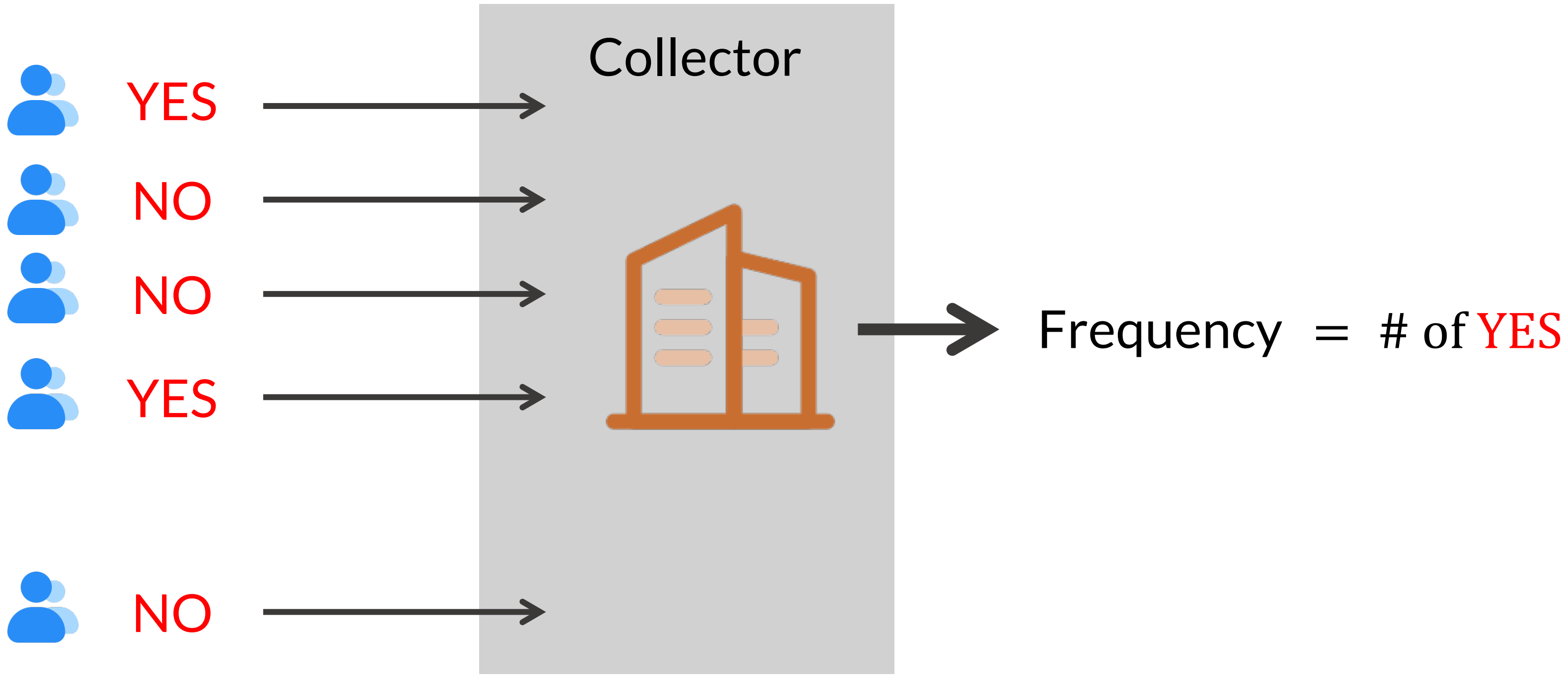
ASU Arizona State University

Frequency Estimation

- Social scientists: <u>How many people engage in tax evasion?</u>

  - ask one person if they had evaded tax

  - the person answers YES or NO



Randomized Response for Privacy

- People have privacy concerns on sensitive/embarrassing question

  - i.e. don't want to let the collector know

- A privacy mechanism $\mathcal{M}$ satisfies LDP if

For any truth x_1, x_2,
and randomized answer y:

$$\max \boxed{\frac{\Pr[\mathcal{M}(x_1) = y]}{\Pr[\mathcal{M}(x_2) = y]}} \leq e^{\varepsilon}$$

Distinguishability of x_1 (YES) and x_2 (NO)
from y (randomized answer)

Randomized Response for Privacy

- People have privacy concerns on sensitive/embarrassing question

  - i.e. don't want to let the collector know

- A privacy mechanism $\mathcal{M}$ satisfies LDP if

For any truth x_1, x_2,
and randomized answer y:

$$\max \frac{\Pr[\mathcal{M}(x_1) = y]}{\Pr[\mathcal{M}(x_2) = y]} \le e^{\varepsilon}$$

Distinguishability of x_1 (YES) and x_2 (NO)
from y (randomized answer)

- **quantifiable hardness** to distinguish x_1 (YES) and x_2 (NO) from the randomized answer y

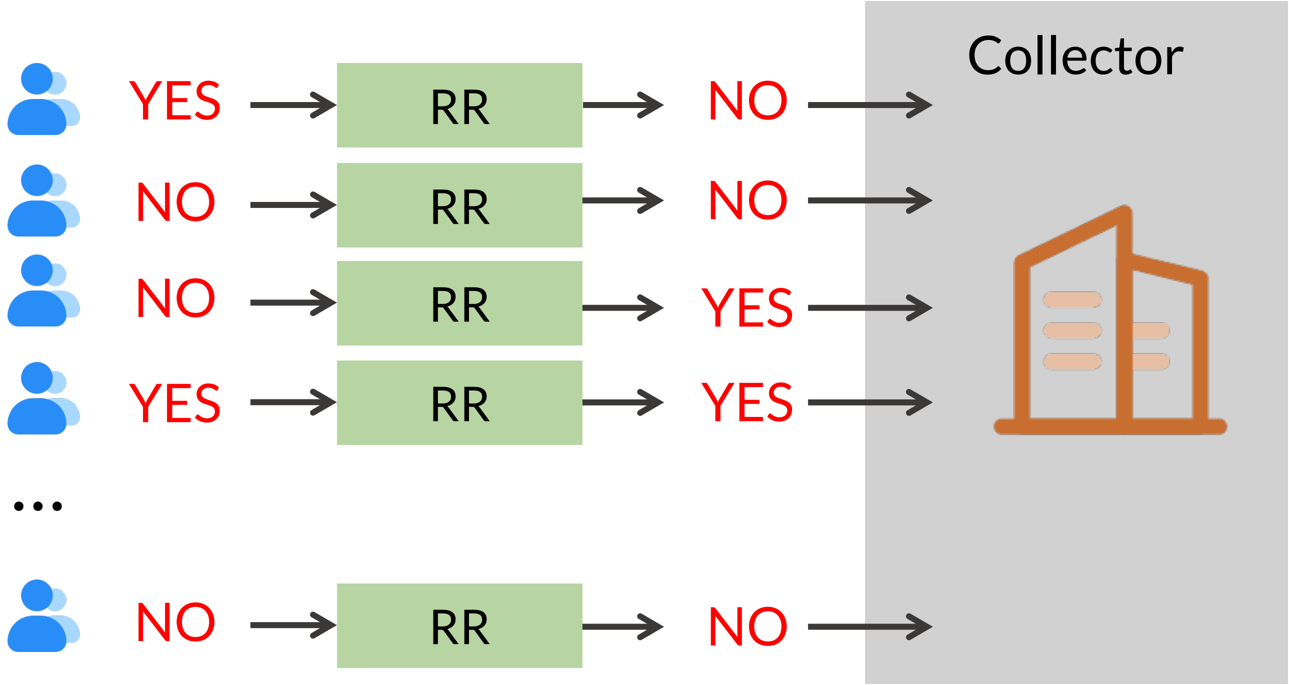
- against inference from data collectors 🏢 or adversaries 👹

Randomized Response for Privacy

- People have privacy concerns on sensitive/embarrassing question

  - i.e. don't want to let the collector know

- Randomized Response: Randomize the truth before answering the collector



Randomized Response for Privacy

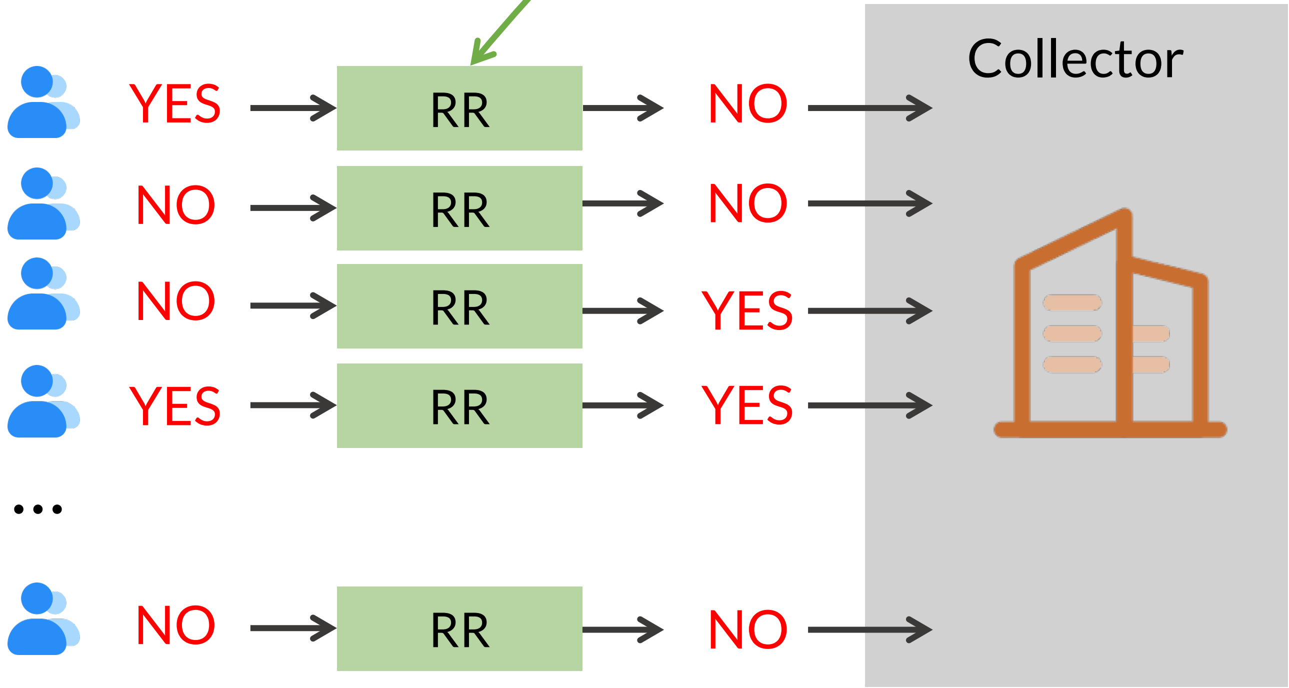
$$\max \frac{\Pr[\mathbf{RR}(x_1) = y]}{\Pr[\mathbf{RR}(x_2) = y]} \leq e^{\ln\frac{p}{1-p}}$$

- People have privacy concerns on sensitive/embarrassing questi
  - i.e. don't want to let the collector know
- Randomized Response: Randomize the truth before answering

Private

RR: [Warner, 1965]
answer truth with probability p

$$\mathrm{RR}(x) = \begin{cases} x & \text{w.p. } p \\ \neg x & \text{w.p. } 1-p \end{cases}$$



Randomized Response for Privacy

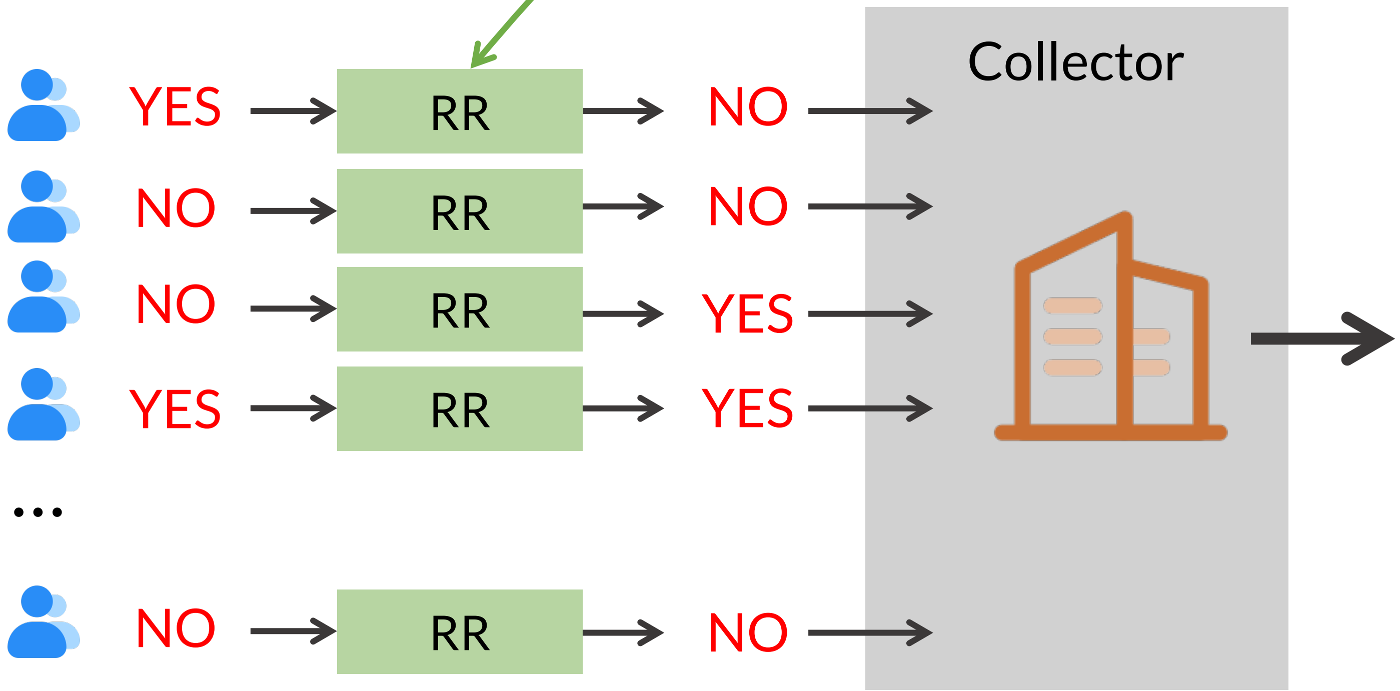
$$\max \frac{\Pr[\mathbf{RR}(x_1) = y]}{\Pr[\mathbf{RR}(x_2) = y]} \le e^{\ln\frac{p}{1-p}}$$

- People have privacy concerns on sensitive/embarrassing questi
  - i.e. don't want to let the collector know
- Randomized Response: Randomize the truth before answering

Private

RR: [Warner, 1965]
answer truth with probability p

$$\mathrm{RR}(x) = \begin{cases} x & \text{w.p. } p \\ \neg x & \text{w.p. } 1-p \end{cases}$$



estimated frequency

$$= \frac{\#\text{ of YES} - \# \times q}{p-q}$$

Unbiased:
expectation = truth

Utility: RR's Variance

- Randomization reduces data utility

$$\mathrm{Var}\left[\frac{\#\text{ of }\text{YES} - \# \times q}{p-q}\right] = \frac{\mathrm{Var}[\#\text{ of YES}]}{(p-q)^2} = \frac{npq}{(p-q)^2}$$

\- summation of variance from n **independent** randomization

Utility: RR's Variance

- Randomization reduces data utility

$$\mathrm{Var}\left[\frac{\#\text{ of YES} - \# \times q}{p - q}\right] = \frac{\mathrm{Var}[\#\text{ of YES}]}{(p-q)^2} = \frac{npq}{(p-q)^2}$$

- summation of variance from n **independent** randomization

- larger $p \in (0.5, 1]$ → lower variance → larger privacy parameter ε

↑ data utility  ↓ privacy

Utility: RR's Variance

- Randomization reduces data utility

$$\mathrm{Var}\left[\frac{\#\text{ of YES} - \# \times q}{p-q}\right] = \frac{\mathrm{Var}[\#\text{ of YES}]}{(p-q)^2} = \frac{npq}{(p-q)^2}$$

\- summation of variance from n **independent** randomization

\- larger $p \in (0.5, 1]$ → lower variance → larger privacy parameter ε

↑ data utility ↓ privacy

- Q: Can we improve this privacy-utility tradeoff?

Utility: RR's Variance

- Randomization reduces data utility

$$\mathrm{Var}\left[\frac{\#\text{ of YES} - \# \times q}{p-q}\right] = \frac{\mathrm{Var}[\#\text{ of YES}]}{(p-q)^2} = \frac{npq}{(p-q)^2}$$

- summation of variance from n **independent** randomization

- larger $p \in (0.5, 1]$ → lower variance → larger privacy parameter ε

↑ data utility  ↓ privacy

- Q: Can we improve this privacy-utility tradeoff?

- **yes, by correlated (joint) randomization**

This Paper: Joint RR (JRR)

- JRR: Better data utility by joint randomization

This Paper: Joint RR (JRR)

- JRR: Better data utility by joint randomization
- **Example:** 2-person ($x_1 = \text{YES}$ and $x_2 = \text{YES}$) with $p = 0.8$ ($\mathrm{P}[T = 1] = 0.8$)

RR: Joint distribution

	$T_1 = 1$	$T_1 = 0$
$T_2 = 1$	0.64 ($= p^2$)	0.16 ($= pq$)
$T_2 = 0$	0.16 ($= pq$)	0.04 ($= q^2$)

Truthfulness of x_1

Truthfulness of x_2

This Paper: Joint RR (JRR)

- JRR: Better data utility by joint randomization
- **Example:** 2-person (x_1 = YES and x_2 = YES) with $p = 0.8$ ($\mathrm{P}[T = 1] = 0.8$)

RR: Joint distribution

	$T_1 = 1$	$T_1 = 0$
$T_2 = 1$	0.64 ($= p^2$)	0.16 ($= pq$)
$T_2 = 0$	0.16 ($= pq$)	0.04 ($= q^2$)

Truthfulness of x_1

Truthfulness of x_2

Independent T_1 and T_2 ($\mathbf{P}[T_1 \cap T_2] = \mathbf{P}[T_1] \cdot \mathbf{P}[T_2]$)

Joint probability = Π of marginal probabilities

This Paper: Joint RR (JRR)

- JRR: Better data utility by joint randomization
- **Example:** 2-person ($x_1 = \text{YES}$ and $x_2 = \text{YES}$) with $p = 0.8$ ($\text{P}[T = 1] = 0.8$)

RR: Joint distribution

	$T_1 = 1$	$T_1 = 0$
$T_2 = 1$	0.64 ($= p^2$)	0.16 ($= pq$)
$T_2 = 0$	0.16 ($= pq$)	0.04 ($= q^2$)

Truthfulness of x_2

JRR: Joint distribution

	$T_1 = 1$	$T_1 = 0$
$T_2 = 1$	0.6 ($= p^2 + \rho pq$)	0.2 ($= pq - \rho pq$)
$T_2 = 0$	0.2 ($= pq - \rho pq$)	0 ($= q^2 + \rho pq$)

Independent T_1 and T_2 ($\mathbf{P}[T_1 \cap T_2] = \mathbf{P}[T_1] \cdot \mathbf{P}[T_2]$)

Joint probability = Π of marginal probabilities

This Paper: Joint RR (JRR)

- JRR: Better data utility by joint randomization
- **Example:** 2-person (x_1 = YES and x_2 = YES) with $p = 0.8$ ($\mathrm{P}[T = 1] = 0.8$)

Same marginal prob for each person

$\mathrm{P}[T_1 = 1] = 0.8$

RR: Joint distribution

	$T_1 = 1$	$T_1 = 0$
$T_2 = 1$	0.64 ($= p^2$)	0.16 ($= pq$)
$T_2 = 0$	0.16 ($= pq$)	0.04 ($= q^2$)

Truthfulness of x_2

JRR: Joint distrib

	$T_1 = 1$	$T_1 = 0$
$T_2 = 1$	0.6 ($= p^2 + \rho pq$)	0.2 ($= pq - \rho pq$)
$T_2 = 0$	0.2 ($= pq - \rho pq$)	0 ($= q^2 + \rho pq$)

Independent T_1 and T_2 ($\mathbf{P}[T_1 \cap T_2] = \mathbf{P}[T_1] \cdot \mathbf{P}[T_2]$)

Joint probability = Π of marginal probabilities

This Paper: Joint RR (JRR)

- JRR: Better data utility by joint randomization
- **Example:** 2-person $(x_1 = \text{YES}$ and $x_2 = \text{YES})$ with $p = 0.8$ $(\mathrm{P}[T = 1] = 0.8)$

Same marginal prob for each person

RR: Joint distribution

	$T_1 = 1$	$T_1 = 0$
$T_2 = 1$	0.64 $(= p^2)$	0.16 $(= pq)$
$T_2 = 0$	0.16 $(= pq)$	0.04 $(= q^2)$

Truthfulness of x_2

Independent T_1 and T_2 $(\mathrm{P}[T_1 \cap T_2] = \mathrm{P}[T_1] \cdot \mathrm{P}[T_2])$

Joint probability = Π of marginal probabilities

JRR: Joint distri

$\mathrm{P}[T_1 = 1] = 0.8$

	$T_1 = 1$	$T_1 = 0$
$T_2 = 1$	0.6 $(= p^2 + \rho pq)$	0.2 $(= pq - \rho pq)$
$T_2 = 0$	0.2 $(= pq - \rho pq)$	0 $(= q^2 + \rho pq)$

$\mathbf{P}[T_1 = 0 \cap T_2 = 0] = \mathbf{0} \neq \mathbf{P}[T_1 = 0] \cdot \mathbf{P}[T_2 = 0] = \mathbf{0.04}$

NOT independent T_1 and T_2

Joint probability $\neq$ Π of marginal probabilities

Utility: JRR's Variance

- Same estimator as RR

$$\mathrm{E}[\#\text{ of YES}] = \sum_{i=1}^{\#} \mathrm{P}[y_i = \text{YES}] = n_{\text{YES}} \cdot p + (\# - n_{\text{YES}}) \cdot q$$

Expectation

Ground truth

Utility: JRR's Variance

- Same estimator as RR

$$\mathrm{E}[\#\text{ of }\mathrm{YES}] = \sum_{i=1}^{\#} \mathrm{P}[y_i = \mathrm{YES}] \;= n_{\mathrm{YES}} \cdot p + (\# - n_{\mathrm{YES}}) \cdot q$$

→ Unbiased estimator $\hat{n}_{\mathrm{YES}} = \boxed{\dfrac{\#\text{ of }\mathrm{YES} - 2q}{p-q}}$

Identical to RR

Utility: JRR’s Variance

- Variance: (# = 2, $p = 0.8$)

$$\mathrm{Var}[\hat{n}_{\mathrm{YES}}] = \frac{\mathrm{Var}[\#\text{ of YES}]}{(0.8-0.2)^2}$$

Utility: JRR's Variance

- Variance: (# = 2, $p = 0.8$)

$$\mathrm{Var}[\hat{n}_{\mathrm{YES}}] = \frac{\mathrm{Var}[\#\text{ of YES}]}{(0.8 - 0.2)^2}$$

Randomized

YES → RR / JRR → YES / NO

YES → RR / JRR → YES / NO

Utility: JRR's Variance

- Variance: (# = 2, $p = 0.8$)

$$\mathrm{Var}[\hat{n}_{\mathrm{YES}}] = \frac{\mathrm{Var}[\#\text{ of YES}]}{(0.8-0.2)^2}$$

Randomized

YES → RR / JRR → YES / NO

YES → RR / JRR → YES / NO

- Distribution table:

RR

# of YES	0	1	2
Probability	0.04	0.16 + 0.16	0.64

$$\mathrm{Var}[\#\text{ of YES}] = E[(X-\mu)^2] = \mathbf{0.32}$$

$$= \sum_{X=0,1,2} (X-1.6)^2 \cdot \Pr[X] \approx \mathbf{0.1 + 0.12 + 0.1}$$

JRR

# of YES	0	1	2
Probability	0	0.2 + 0.2	0.6

$$\mathrm{Var}[\#\text{ of YES}] = E[(X-\mu)^2] = \mathbf{0.24}$$

$$= \sum_{X=0,1,2} (X-1.6)^2 \cdot \Pr[X] = \mathbf{0 + 0.14 + 0.1}$$

Utility: JRR's Variance

- Variance: (# = 2, $p = 0.8$)

$$\mathrm{Var}[\hat{n}_{\mathrm{YES}}] = \frac{\mathrm{Var}[\#\text{ of YES}]}{(0.8-0.2)^2}$$

Randomized

YES → RR / JRR → YES / NO

YES → RR / JRR → YES / NO

- Distribution table:

Better utility

RR

# of YES	0	1	2
Probability	0.04	0.16 + 0.16	0.64

$\mathrm{Var}[\#\text{ of YES}] = E[(X-\mu)^2] = \mathbf{0.32}$

$$= \sum_{X=0,1,2} (X-1.6)^2 \cdot \Pr[X] \approx \mathbf{0.1 + 0.12 + 0.1}$$

JRR **(near to μ)**

# of YES	0	1	2
Probability	0	0.2 + 0.2	0.6

$\mathrm{Var}[\#\text{ of YES}] = E[(X-\mu)^2] = \mathbf{0.24}$

$$= \sum_{X=0,1,2} (X-1.6)^2 \cdot \Pr[X] = \mathbf{0 + 0.14 + 0.1}$$

JRR's General Form

- Correlated randomization with 2 persons x_{2i-1} and x_{2i}

JRR: Joint distribution

	$T_{2i-1} = 1$	$T_{2i-1} = 0$
$T_{2i} = 1$	$p^2 + \rho pq$	$(1 - \rho)pq$
$T_{2i} = 0$	$(1 - \rho)pq$	$q^2 + \rho pq$

$\rho \in [-1,1]$: correlated coefficient

- RR is a special case of JRR with $\rho = 0$ (no correlation)

JRR's General Form

- Correlated randomization with 2 persons x_{2i-1} and x_{2i}

JRR: Joint distribution

	$T_{2i-1} = 1$	$T_{2i-1} = 0$
$T_{2i} = 1$	$p^2 + \rho pq$	$(1 - \rho)pq$
$T_{2i} = 0$	$(1 - \rho)pq$	$q^2 + \rho pq$

$\rho \in [-1,1]$: correlated coefficient

Utility Theorem. The variance of JRR's estimator $\hat{\boldsymbol{n}}_{\boldsymbol{v}}$ is

$$\mathrm{Var}[\hat{\boldsymbol{n}}_{\boldsymbol{v}}] = \frac{pq}{(p-q)^2} \cdot \left(n + \frac{\rho\left((2n_{\mathrm{YES}} - n)^2 - n\right)}{n-1} \right).$$

Privacy: NOT as Simple as RR

- If any person can be an adversary

JRR: Joint distribution

	$T_1 = 1$	$T_1 = 0$
$T_2 = 1$	0.6	0.2
$T_2 = 0$	0.2	0

T_1: I am an adversary 😈

When I report untruthfully $(T_1 = 0)$,
My partner will report truthfully $(T_2 = 1)$

Privacy: NOT as Simple as RR

- If any person can be an adversary

T_1: I am an adversary 😈

JRR: Joint distribution

	$T_1 = 1$	$T_1 = 0$
$T_2 = 1$	0.6	0.2
$T_2 = 0$	0.2	0

When I report untruthfully $(T_1 = 0)$,
My partner will report truthfully $(T_2 = 1)$

- Correlation results in privacy leakage

JRR – Privacy Model in This Paper

- Randomly groups into form 2-person groups for correlated randomization

JRR – Privacy Model in This Paper

- Randomly groups into form 2-person groups for correlated randomization

- **Threat model:**

  - any person can be an adversary 👹

  - if a group contains an adversary, the adversary knows who is their partner (after random grouping)

JRR – Privacy Model in This Paper

- Randomly groups into form 2-person groups for correlated randomization

- **Threat model:**

  - any person can be an adversary 👹

  - if a group contains an adversary, the adversary knows who is their partner (after random grouping)

A person

Partner is an 👹

Partner is NOT an adversary

JRR – Privacy Model in This Paper

- Randomly groups into form 2-person groups for correlated randomization

- **Threat model:**

  - any person can be an adversary 😈

  - if a group contains an adversary, the adversary knows who is their partner (after random grouping)

$$\Pr = \frac{\#\text{😈}}{\#\text{👥}}$$

A person

Partner is an 😈

Partner is NOT an adversary

JRR – Privacy Model in This Paper

- Randomly groups into form 2-person groups for correlated randomization

- **Threat model:**

  - if a group contains an adversary, the adversary knows who is their partner (after random grouping)

  - the adversary cannot control randomness, but can infer their partner's

$$\Pr = \frac{\#\text{👿}}{\#\text{👤}}$$

A person

Partner is an 👿

$$\Pr[\text{person} \mid \text{👿}] = \frac{(1-\rho)pq}{p}$$ Higher confidence

Partner is NOT an adversary

JRR – Formal Privacy & Utility

Privacy Theorem. Assume a set of data contributors $\mathcal{T}_m$ whose reporting truthfulness is known to the adversary. For any data contributor i, the JRR mechanism satisfies:

$$\frac{\Pr[\,\mathrm{JRR}(x_i)\mid \mathcal{T}_m]}{\Pr[\,\mathrm{JRR}(x_i')\mid \mathcal{T}_m]} \leq e^{\varepsilon}, \quad \text{where} \quad \varepsilon = \ln \frac{mp_{\max} + (n-m-1)p}{mp_{\min} + (n-m-1)q}.$$

$p_{\max} = \max\{(1-\rho)p, p + \rho q\}$: confidence of adversaries inferring a specific value

Privacy affected by

m	# adversaries 👹
n	# of persons 👥
ρ	Correlated coefficient

JRR – Formal Privacy & Utility

Privacy Theorem. Assume a set of data contributors $\mathcal{T}_m$ whose reporting truthfulness is known to the adversary. For any data contributor i, the JRR mechanism satisfies:

$$\frac{\Pr[\,\mathrm{JRR}(x_i)\,|\mathcal{T}_m]}{\Pr[\,\mathrm{JRR}(x_i')\,|\mathcal{T}_m]} \le e^{\varepsilon}, \quad \text{where} \quad \varepsilon = \ln \frac{mp_{\max} + (n-m-1)p}{mp_{\min} + (n-m-1)q}.$$

Utility Theorem. The variance of JRR's estimator $\hat{\boldsymbol{n}}_{\boldsymbol{v}}$ is

$$\mathrm{Var}[\hat{\boldsymbol{n}}_{\boldsymbol{v}}] = \frac{pq}{(p-q)^2} \cdot \left(n + \frac{\rho\big((2n_{\mathrm{YES}} - n)^2 - n\big)}{n-1}\right).$$

JRR – Formal Privacy & Utility

Privacy Theorem. Assume a set of data contributors $\mathcal{T}_m$ whose reporting truthfulness is known to the adversary. For any data contributor i, the JRR mechanism satisfies:

privacy constraint

$$\varepsilon = \ln \frac{mp_{\max} + (n - m - 1)p}{mp_{\min} + (n - m - 1)q}.$$

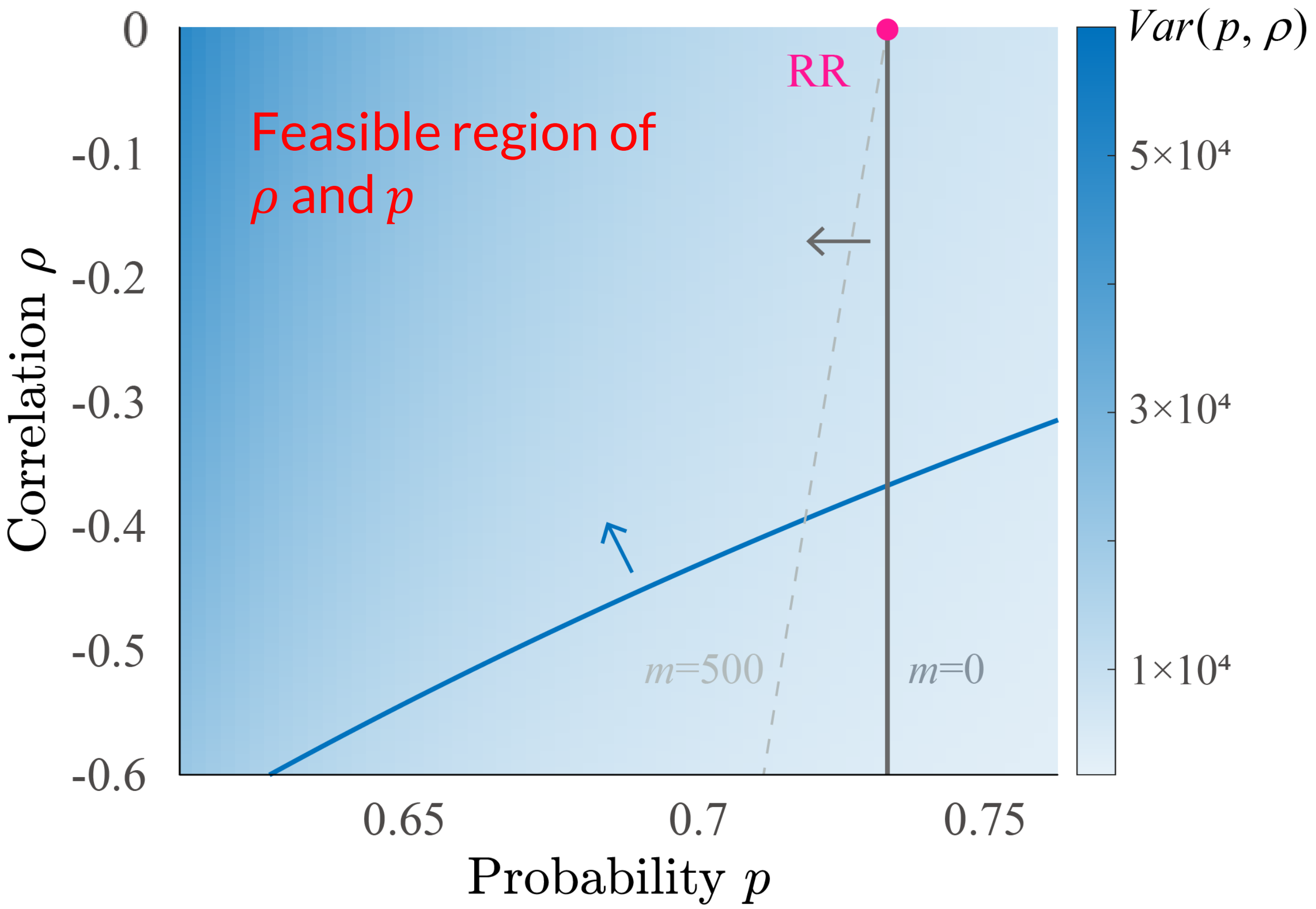
Utility Theorem. The variance of JRR's estimator $\hat{\boldsymbol{n}}_v$ is

minimize

$$\mathrm{Var}[\hat{\boldsymbol{n}}_v] = \frac{pq}{(p-q)^2} \cdot \left(n + \frac{\rho\left((2n_{\mathrm{YES}} - n)^2 - n\right)}{n - 1} \right).$$

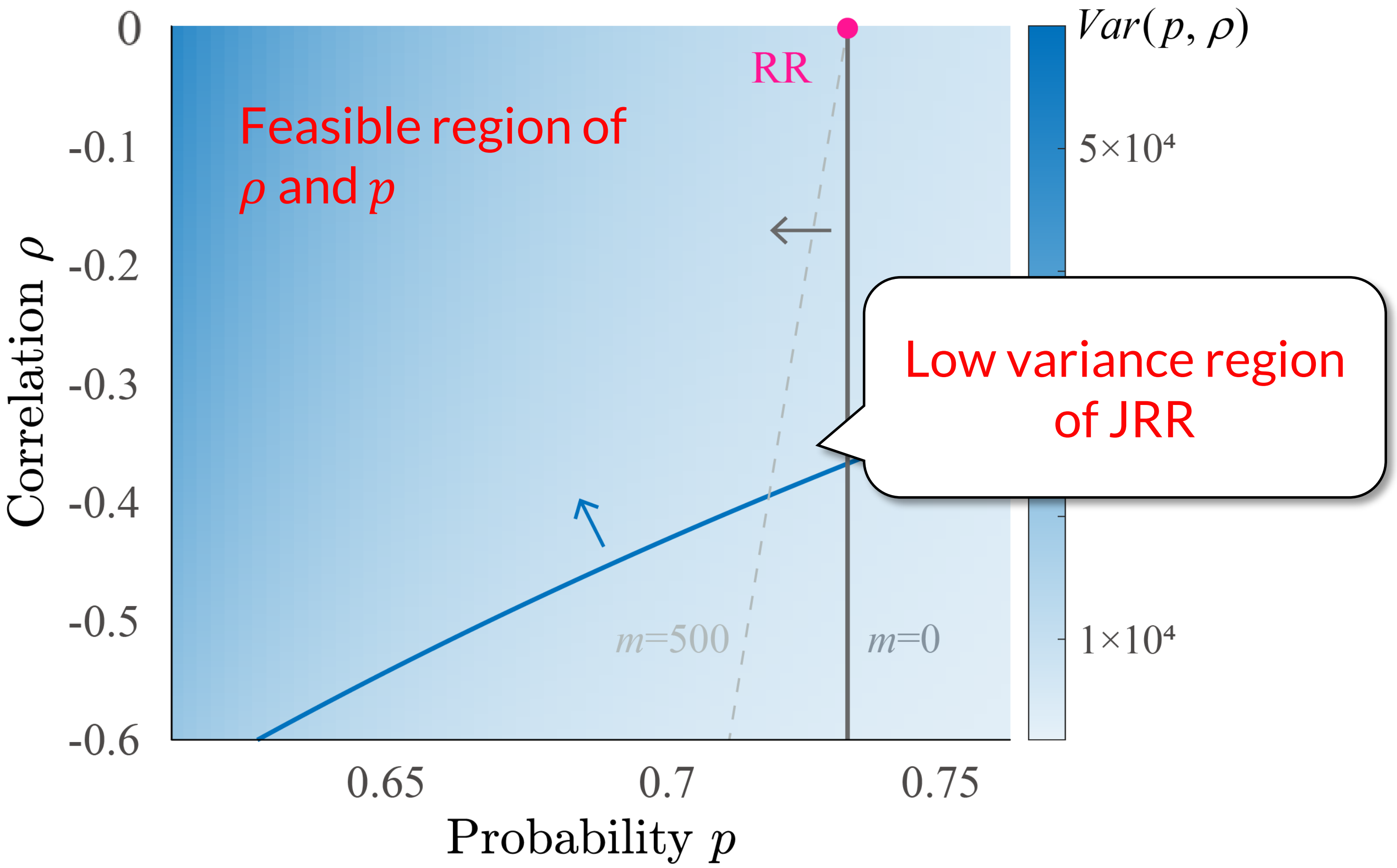
JRR – Variance Heatmap

- Effect of ρ and p (when $\varepsilon = 1, n = 10^4, n_{\text{Yes}} = 200$, and $m = 0$ & 500)



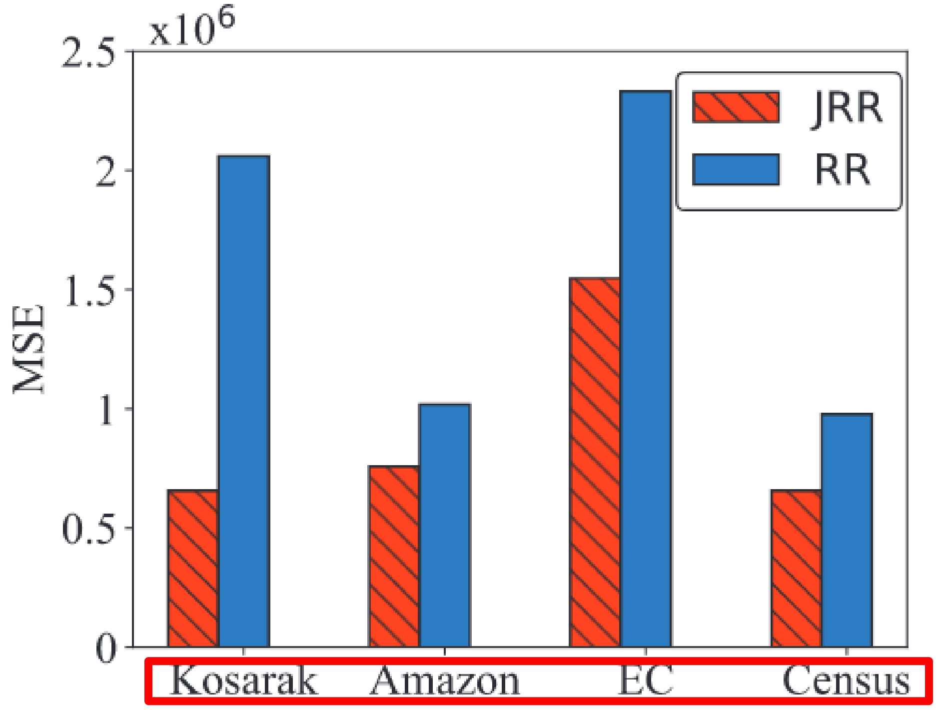
JRR – Variance Heatmap

- Effect of ρ and p (when $\varepsilon = 1, n = 10^4, n_{\text{Yes}} = 200$, and $m = 0$ & 500)

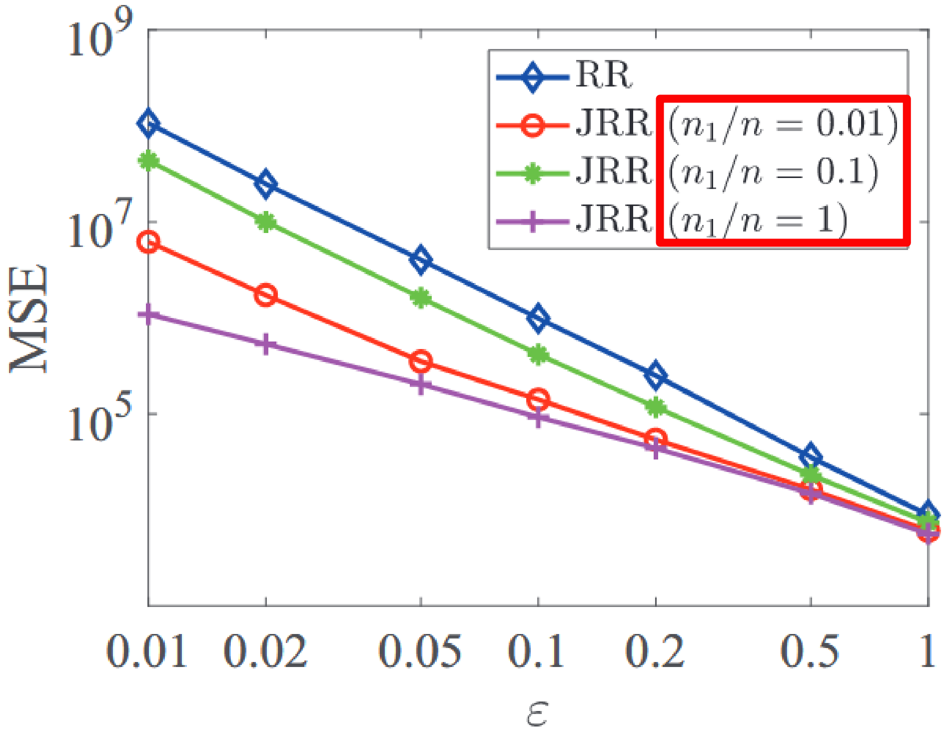


Experiments

- Comparison with RR under the same privacy level - JRR: $\varepsilon(n, m, \rho, p)$, RR: $\varepsilon(p)$



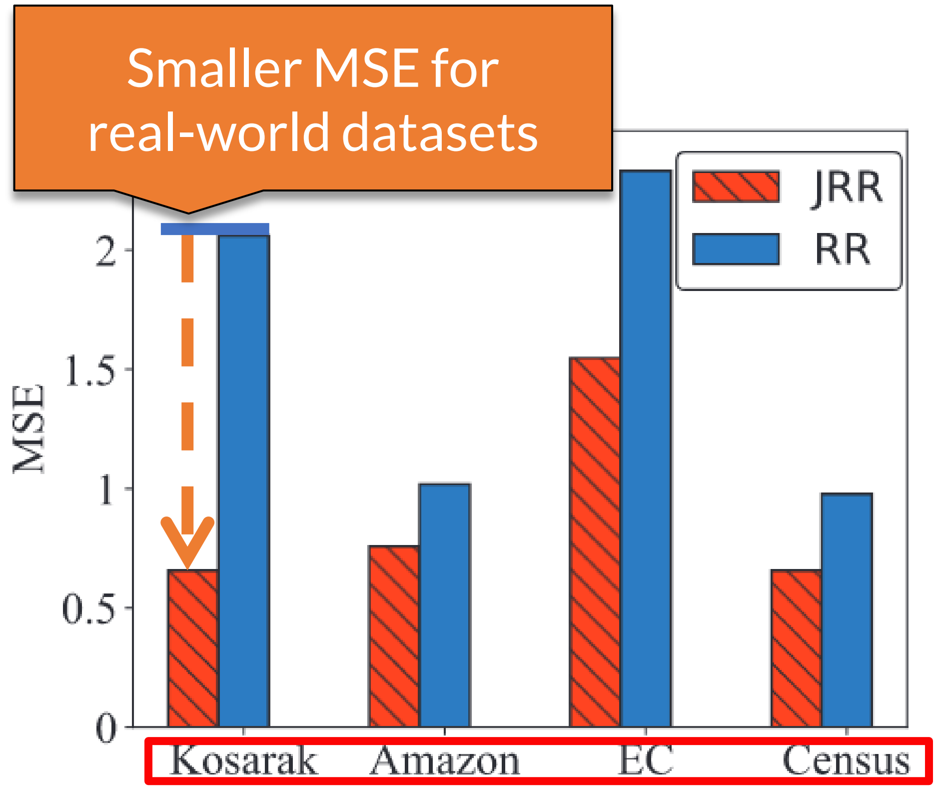
Real-world datasets ($\varepsilon = 0.1$)



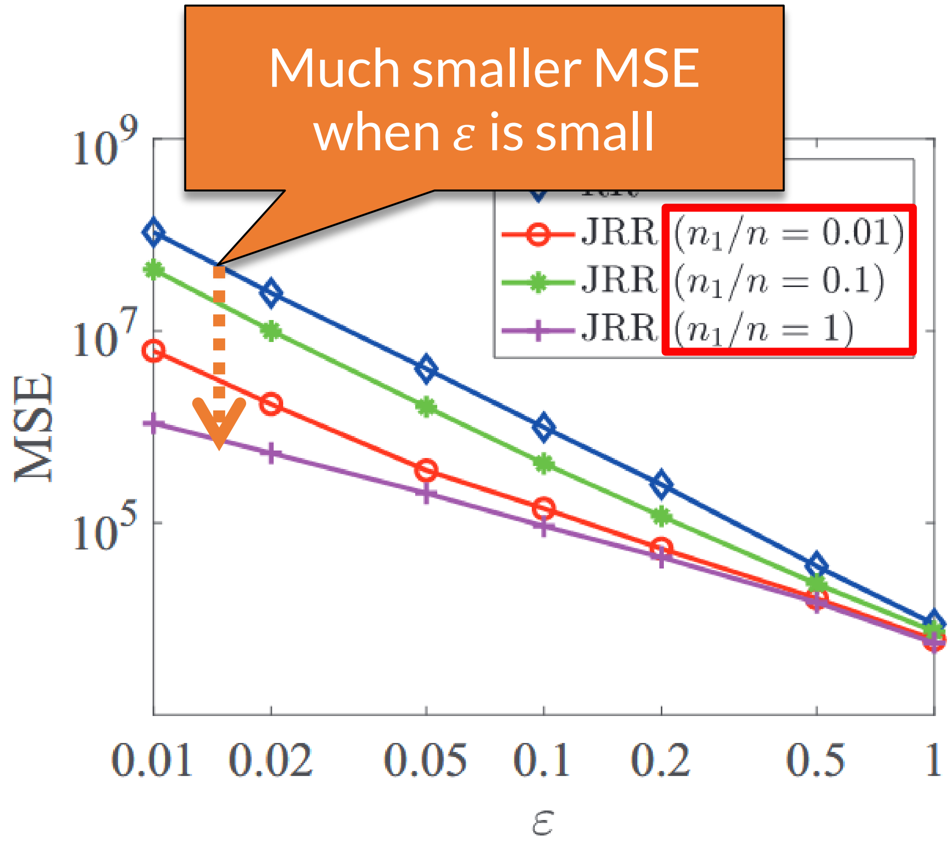
Synthetic datasets ($n = 10^4$)

Experiments

- Comparison with RR under the same privacy level - JRR: $\varepsilon(n, m, \rho, p)$, RR: $\varepsilon(p)$



Real-world datasets ($\varepsilon = 0.1$)



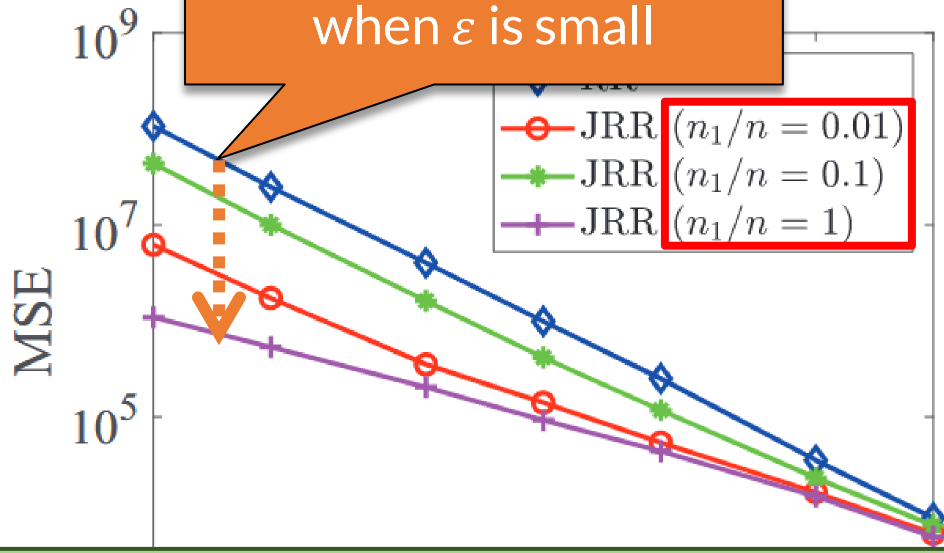
Synthetic datasets ($n = 10^4$)

Experiments

- Comparison with RR under the same privacy level - JRR: $\varepsilon(n, m, \rho, p)$, RR: $\varepsilon(p)$

Smaller MSE for real-world datasets

Much smaller MSE when ε is small



- Correlated randomization can improve the data utility of frequency estimation
- **JRR: Privacy & utility model for correlated randomization**

Locally Differentially Private Frequency Estimation via Joint Randomized Response

Randomized Response for Privacy

$\max \frac{\Pr[\mathbf{RR}(x_1)=y]}{\Pr[\mathbf{RR}(x_2)=y]} \le e^{\ln\frac{p}{1-p}}$

- People have privacy concerns on sensitive/embarrassing questi
  - i.e. don't want to let the collector know
- Randomized Response: Randomize the truth before answering

Private

RR: [Warner, 1965]
answer truth with probability p

$\mathrm{RR}(x) = \begin{cases} x & \text{w. p. } p \\ \neg x & \text{w. p. } 1-p \end{cases}$

YES →	RR	→ NO →	Collector
NO →	RR	→ NO →	
NO →	RR	→ YES →	
YES →	RR	→ YES →	
…			
NO →	RR	→ NO →	

estimated frequency

$= \frac{\#\text{ of YES} - \# \times q}{p - q}$

Unbiased:
expectation = truth

This Paper: Joint RR (JRR)

- JRR: Better data utility by joint randomization
- **Example:** 2-person (x_1 = YES and x_2 = YES) with $p = 0.8$ ($\mathrm{P}[T=1] = 0.8$)

Same marginal prob for each person

$\mathrm{P}[T_1 = 1] = 0.8$

RR: Joint distribution

	$T_1 = 1$	$T_1 = 0$
$T_2 = 1$	0.64 ($= p^2$)	0.16 ($= pq$)
$T_2 = 0$	0.16 ($= pq$)	0.04 ($= q^2$)

Truthfulness of x_2

Independent T_1 and T_2 ($\mathrm{P}[T_1 \cap T_2] = \mathrm{P}[T_1] \cdot \mathrm{P}[T_2]$)

Joint probability = Π of marginal probabilities

JRR: Joint distri

	$T_1 = 1$	$T_1 = 0$
$T_2 = 1$	0.6 ($= p^2 + \rho pq$)	0.2 ($= pq - \rho pq$)
$T_2 = 0$	0.2 ($= pq - \rho pq$)	0 ($= q^2 + \rho pq$)

$\mathrm{P}[T_1 = 0 \cap T_2 = 0] = \mathbf{0} \neq \mathrm{P}[T_1 = 0] \cdot \mathrm{P}[T_2 = 0] = 0.04$

NOT independent T_1 and T_2

Joint probability ≠ Π of marginal probabilities

Utility: JRR's Variance

- Variance: (# = 2, $p = 0.8$)

$\mathrm{Var}[\hat{n}_{\mathrm{YES}}] = \frac{\mathrm{Var}[\#\text{ of YES}]}{(0.8-0.2)^2}$

Randomized

YES → RR / JRR → YES / NO
YES → YES / NO

- Distribution table:

Better utility

RR

# of YES	0	1	2
Probability	0.04	0.16 + 0.16	0.64

$\mathrm{Var}[\#\text{ of YES}] = E[(X-\mu)^2] = \mathbf{0.32}$

$= \sum_{X=0,1,2} (X-1.6)^2 \cdot \Pr[X] \approx \mathbf{0.1 + 0.12 + 0.1}$

JRR (near to μ)

# of YES	0	1	2
Probability	0	0.2 + 0.2	0.6

$\mathrm{Var}[\#\text{ of YES}] = E[(X-\mu)^2] = \mathbf{0.24}$

$= \sum_{X=0,1,2} (X-1.6)^2 \cdot \Pr[X] = \mathbf{0 + 0.14 + 0.1}$

JRR's General Form

- Correlated randomization with 2 persons x_{2i-1} and x_{2i}

JRR: Joint distribution

	$T_{2i-1} = 1$	$T_{2i-1} = 0$
$T_{2i} = 1$	$p^2 + \rho pq$	$(1-\rho)pq$
$T_{2i} = 0$	$(1-\rho)pq$	$q^2 + \rho pq$

$\rho \in [-1,1]$: correlated coefficient

- RR is a special case of JRR with $\rho = 0$ (no correlation)

JRR – Privacy Model in This Paper

- Randomly groups into form 2-person groups for correlated randomization
- **Threat model:**
  - if a group contains an adversary, the adversary knows who is their partner (after random grouping)
  - the adversary cannot control randomness, but can infer their partner's

$\Pr = \frac{\#}{\#}$

A person

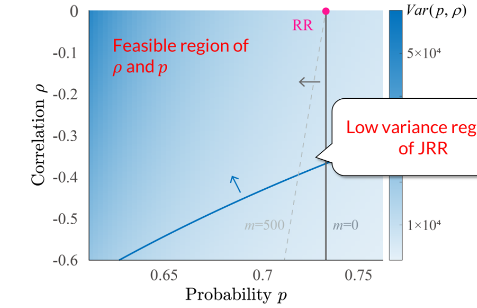
Partner is an

Partner is NOT an adversary

→ $\Pr[\text{person} \mid \ ] = \frac{(1-\rho)pq}{p}$ Higher confidence

JRR – Variance Heatmap

- Effect of ρ and p (when $\varepsilon = 1, n = 10^4, n_{\mathrm{Yes}} = 200$, and $m = 0$ & 500)



Thank you!

Q&A

Privacy Model

- No need for securing shuffling:

  - when one person hold multiple items