Data & Privacy

- Data is fundamental for companies, drug developers, scientists, and politics.
- Unregulated access to data creates privacy risk for individuals.
- Differential privacy is the de facto tool for data analysis with mathematical guarantees.
- In practice, utility of data greatly diminishes with the use of differential privacy.
- Researchers introduce relaxations of differential privacy.
- Implications on privacy are not completely understood.

- This poster: Label differential privacy.

Label Differential Privacy

- Users may have public information (gender, zip code, age, ...).
- User has a sensitive attribute (disease, income, ...).
- Researcher wants to train a model to predict sensitive attributes without learning information about individual users.
- Ideally: Noise public information and sensitive label.
- In practice, very low utility.
- Proposal: Noise only sensitive attribute.
- How private is this?

Randomized response

Thought experiment:
- Do a study on the incidence of lung cancer.
- Every person is asked a question: Do you have lung cancer?
- Respondent flips a coin (probability of heads = p).
- If heads: answer truthfully.
- If tails: say yes or no uniformly at random.
- For moderate values of p, respondent information is protected.
- Data collector can get accurate aggregate information about incidence of lung cancer.

Inverting randomized response

- Assume data collector knows \( f(\text{smoking}) = P(\text{lung cancer} | \text{smoking}) \).
- Easy to estimate \( P(\text{lung cancer} | \text{report}) \).
- Theorem: \( f(\text{smoking}) \) is close to 0 or 1 simply ignore report and infer lung cancer status based on f.
- Theorem: If \( f(\text{smoking}) \) is not close to 0 or 1, randomized response provides privacy protections.
- If data collector already knows \( f(\text{smoking}) \) then this isn't really a privacy violation.

Learning to invert randomized response

- With enough data it is possible to estimate \( f(\text{smoking}) \) accurately by debiasing reports.
- Learning trade-off:
  - If \( f(\text{smoking}) \) is close to 0 or 1. Then experiment will leak information about user.
  - If \( f(\text{smoking}) \) is not close to 0 or 1. Then experiment does not leak particular information about a user but also likely to be a bad predictor.

General scenario

- Public feature vectors \( X \).
- True sensitive label \( Y \).
- Randomized response of sensitive label \( Y' \).
- What can we say about the privacy protection of users?
  - Using the regression function \( n(x) = P(Y = 1 | X = x) \), "true" privacy leakage increases by \( \log(n(x)/(1-n(x))) \).
  - Do all users get the same protection (or do users that are easier to classify have more risk of leakage)?
  - Depends on the regression function.
  - Can we estimate the true privacy risk of label randomized response?
    - Yes, using nearest neighbor estimators.
  - Can attackers estimate the regression function?
    - Yes, using nearest neighbor estimators.

Conclusion

- Differential privacy is a powerful tool for data analysis.
- Several relaxations of differential privacy have been proposed to make it more practical.
- We demonstrated that label differential privacy has higher privacy leakage risks than expected.