Lukas Steinberger:

Statistical Efficiency in Local Differential Privacy

We develop a theory of asymptotically efficient estimation in regular parametric models when data confidentiality is ensured by local differential privacy (LDP). The idea of LDP is that individual data owners should be able to release an anonymized or sanitized version Z_i of their possibly sensitive information X_i by drawing Z_i from a pre-specified conditional distribution Q that satisfies the formal α-differential privacy constraint. The problem is now to identify a randomization mechanism Q, generating Z_i, and an estimator $\hat{\theta}$, that uses the sanitized data to estimate the population parameter, with minimal vari-ance among all data-generation and estimation schemes satisfying the privacy constraint. Starting from a regular parametric model for the iid unobserved sensitive data $X_1, \ldots, X_n$, we establish local asymptotic mixed normality (along subsequences) of the model describing the sanitized observations $Z_1, \ldots, Z_n$. This result readily implies convolution and local asymptotic minimax theorems. In case $p = 1$, the optimal asymptotic variance is found to be the inverse of the supremal Fisher-Information, where the supremum runs over all α-differentially private (marginal) privacy mechanisms. We present a numerical algorithm for finding a (nearly) optimal privacy mechanism and an estimator based on the corresponding sanitized data that achieves this asymptotically optimal variance under mild assumptions. In special cases, such as the Gaussian location model, our theory also enables us to identify exact closed form expressions of efficient privacy mechanism and estimators.