Testing Algorithms for Private Linear Regression

A. Antuca, H. Wang, J. Honaker, V. Karwa, O. Sheffet

Privacy Tools for Sharing Data Project, Summer REU

Privacy Tools for Sharing Research Data

A National Science Foundation Secure and Trustworthy Cyberspace Project

with additional support from the Sloan Foundation and Google, Inc.

DIFFERENTIALLY PRIVATE LINEAR REGRESSION

Background:

Existing techniques [Chaudhuri et al 11, Bassily et al 14] give a differentially private version of linear regression by formulating it as a convex optimization problem, with two main drawbacks;

1. A single specific problem: these works deal with one regression problem (predict the label y as a linear function of a given p-tuple features $\boldsymbol{x}$), while a d-dimensional table can lead for $\exp(d)$ different regression problems.
2. Ignore statistical inference: compare the differentially private predictor with the standard, non-private predictor, and not with some true model that generates as data --- as opposed to the statistical approach.

“Analyze Gauss” algorithm of Dwork et al 14 is an algorithm that approximates the 2^{nd}-moment matrix of the data, from which any linear regression can be estimated, yet it was never analyzed for the context of linear regression but rather for PCA.

All of these approaches work by adding random noise, and have no interpretation in classic literature of linear-regression.

Our Contribution #1: New Algorithms

We give 3 additional algorithms, all approximating the 2^{nd} moment matrix of the data with a PSD matrix.

(Input: a $(n{\times}d)$-matrix D; output: a PSD matrix approximating $D^T D$)

All corresponding to techniques in existing literature on linear regression.

1. Additive Wishart noise:

- Adds to the data $O(1/\epsilon^2)$ new completely random datapoints.
- Corresponds to regularization / shrinkage

2. Johnson Lindenstrauss using Ridge Regularization:

- Blocki et al 12 have shown that the Johnson-Lindenstrauss transform preserves privacy is the data is sufficiently well spread (all singular values are large).
- Our algorithm tests for the least-signular value privately, then adds a matrix $w{\cdot}I$ to the data to make it well-spread.
- Translates to approximating a l_2 -regularized version of linear regression called ridge regression with normalization factor of w^2.

3. Sampling from an Inverse-Wishart distribution:

- The JL-algorithm is equivalent to sampling from a Wishart-distribution. We could also sample the inverse of the matrix from an inverse-Wishart distribution – if the data is well spread.
- Corresponds to sampling from a Bayesian posterior belief on the 2^{nd}-moment matrix given a spread-enough prior.

Our contribution #2: Statistical Inference

We analyze the JL algorithm for the purposes of statistical inference.

- Case 1: no regularization needed. We can show that by projecting the n-datapoints given in D using a $(r{\times}n)$-matrix R, we effectively reduce the sample-size to r. (We can do the same inference, up to an approximation factor of $\exp(\pm r/(n\text{-}p))$).
- Case 2: regularization needed. We can give coordinate based confidence intervals, but they are based on the problem parameters.

References

- Bassily, Smith, Thakurta. Private empirical risk minimization: Efficient algorithms and tight error bounds. FOCS 2014.
- Blocki, Blum, Datta, Sheffet. The Johnson-Lindenstrauss Transform itself preserves differential privacy. FOCS 2012.
- Chaudhuri, Monteleoni, Sarwate. Differentially private empirical risk minimization. JMLR 2011.
- Dwork, Talwar, Thakurta, Zhang. Analyze gauss – optimal bounds for privacy preserving principal component analysis. STOC 2014.

TESTING FRAMEWORK

We created a testing framework focused on the statistical significance of regression coefficients. That is, we wanted to test the performance of private algorithms with respect to rejecting the null hypothesis that a given regression coefficient equals 0.

Our testing framework can evaluate the following 4 algorithms:

- Analyze Gauss
- Additive Wishart
- Johnson-Lindenstrauss transform
- Johnson-Lindenstrauss transform with Inverse Wishart distribution

The testing framework does this by generating a dataset with parameters as specified by the researcher and then iterating a chosen differentially private algorithm over the dataset. This allows us to test both the randomness of the differentially private algorithm as well as error introduced from the process of generating data—equivalently, the process of sampling from a population.

The framework accepts the following parameters as rows of a matrix. Any of these parameters can be varied.

- T: number of times to run differentially private algorithm on current row of matrix
- *n*: number of observations in the dataset
- *d*: number of variables (or columns/dimensionality) in the dataset
- means: vector of means for each column of the dataset to be generated
- sigma: the variance-covariance matrix of the population of the dataset to be generated
- ranges: gives upper and lower bounds for each column of the dataset
- choice of whether we truncate or censor synthetically generated values outside our allowed ranges of data
- value of epsilon as in differential privacy
- value of delta as in differential privacy
- identity of the algorithm to be tested
- indication of which variables in the dataset are dependent or independent

The framework then outputs three vectors of length $d + 1$ inside a matrix:

- the vector of regression coefficients
- the vector of standard errors for each regression coefficient
- the vector of t-values, which are computed directly by dividing a regression coefficient by its corresponding standard error

These t-values are necessary to perform statistical testing to verify the significance of the regression coefficients. We hoped to use the testing framework to examine the following questions:

- How do the performances of the algorithms compare to each other?
- Where is each algorithm most effective?

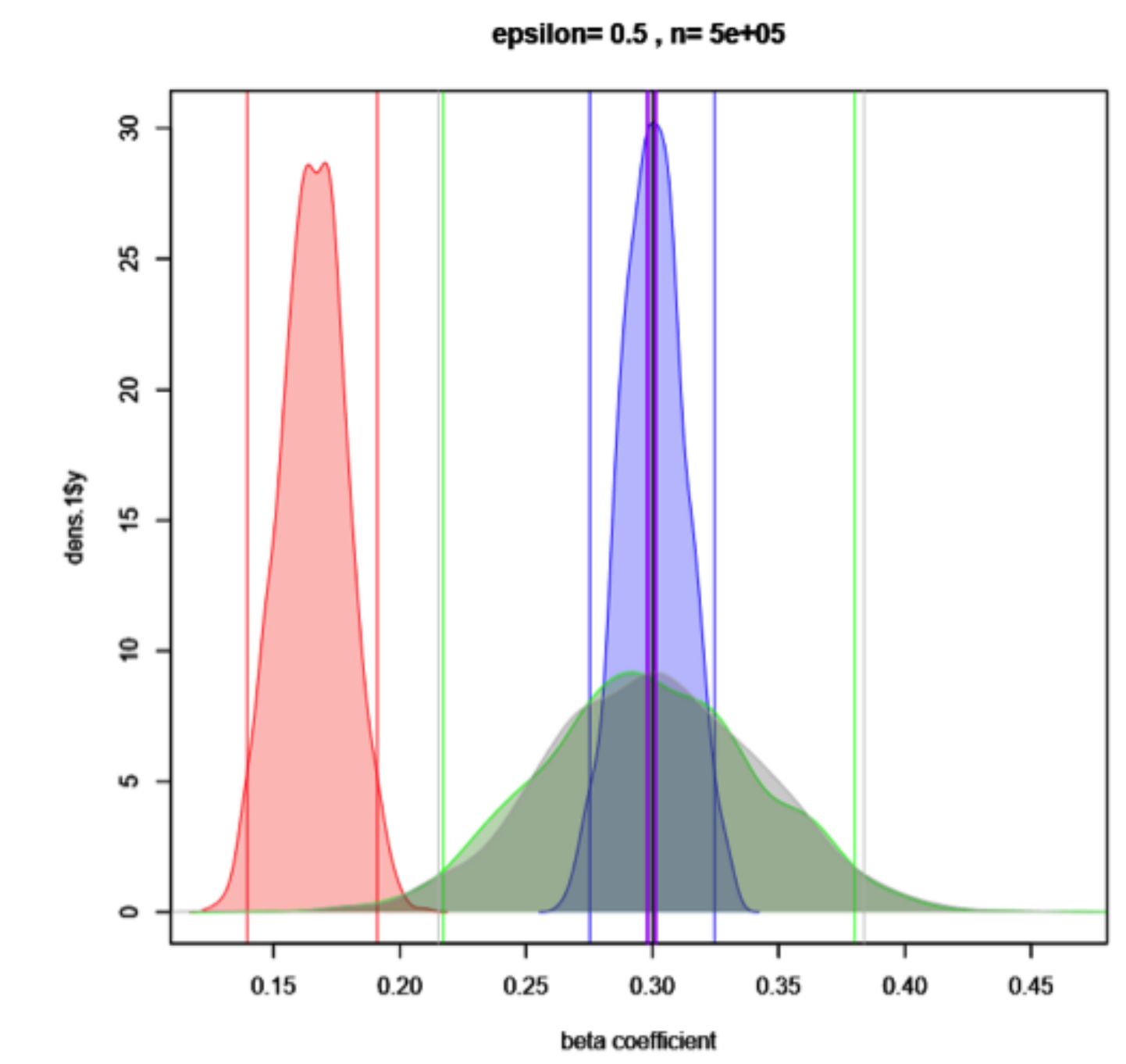
PRELIMINARY RESULTS

We began by initially running each of our algorithms according to the parameters as follows. We generate data with 500000 counts, dimensionality of 5, means of 0 for each row, and the following variance-covariance matrix:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0.3 \\ 0 & 1 & 0 & 0 & 0.1 \\ 0 & 0 & 1 & 0 & -0.1 \\ 0 & 0 & 0 & 1 & 0 \\ 0.3 & 0.1 & -0.1 & 0 & 1 \end{pmatrix}$$

We restrict data to the range [-4, 4] for each column and censor rather than truncate any observations that lie outside this range. For each algorithm, we use 0.5 as our value of epsilon and exp(-10) as our value of delta. We perform regressions taking the first 4 variables in the matrix above as independent and the last one as dependent.

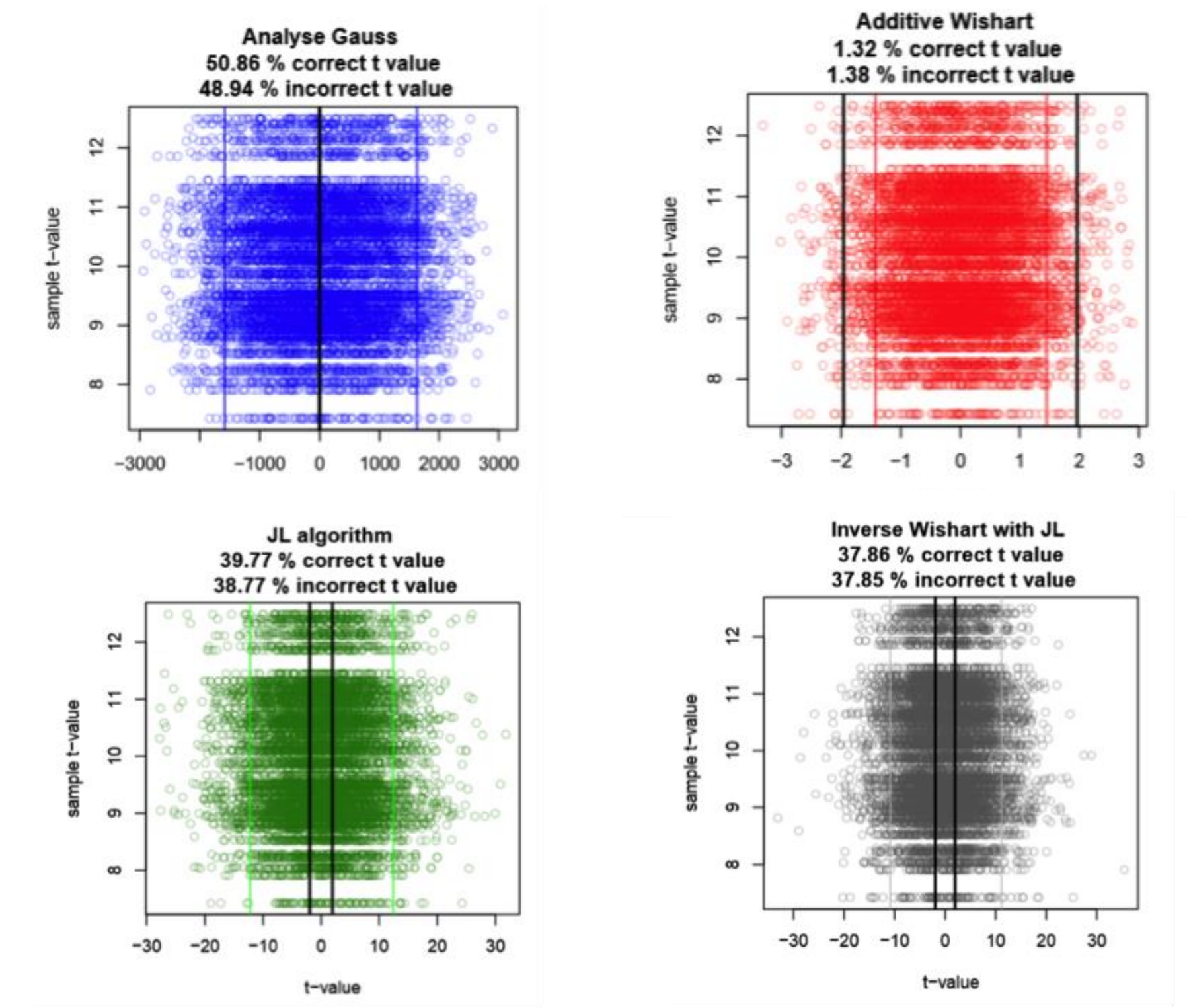
The below graphs give us a comparison of the four algorithms, for the regression coefficient between the first and fifth variables:



Key: Analyze Gauss (blue), Additive Wishart (green), Johnson-Lindenstrauss (green), Johnson-Lindenstrauss with Inverse Wishart (gray).

The differentially private estimates of standard error for each of these algorithms are 0.004, 0.004, 0.18, and 0.18, respectively. A brief glance at t-values here shows that in general, we will reject the null hypothesis that this coefficient is 0, which is correct, as the population coefficient is 0.3.

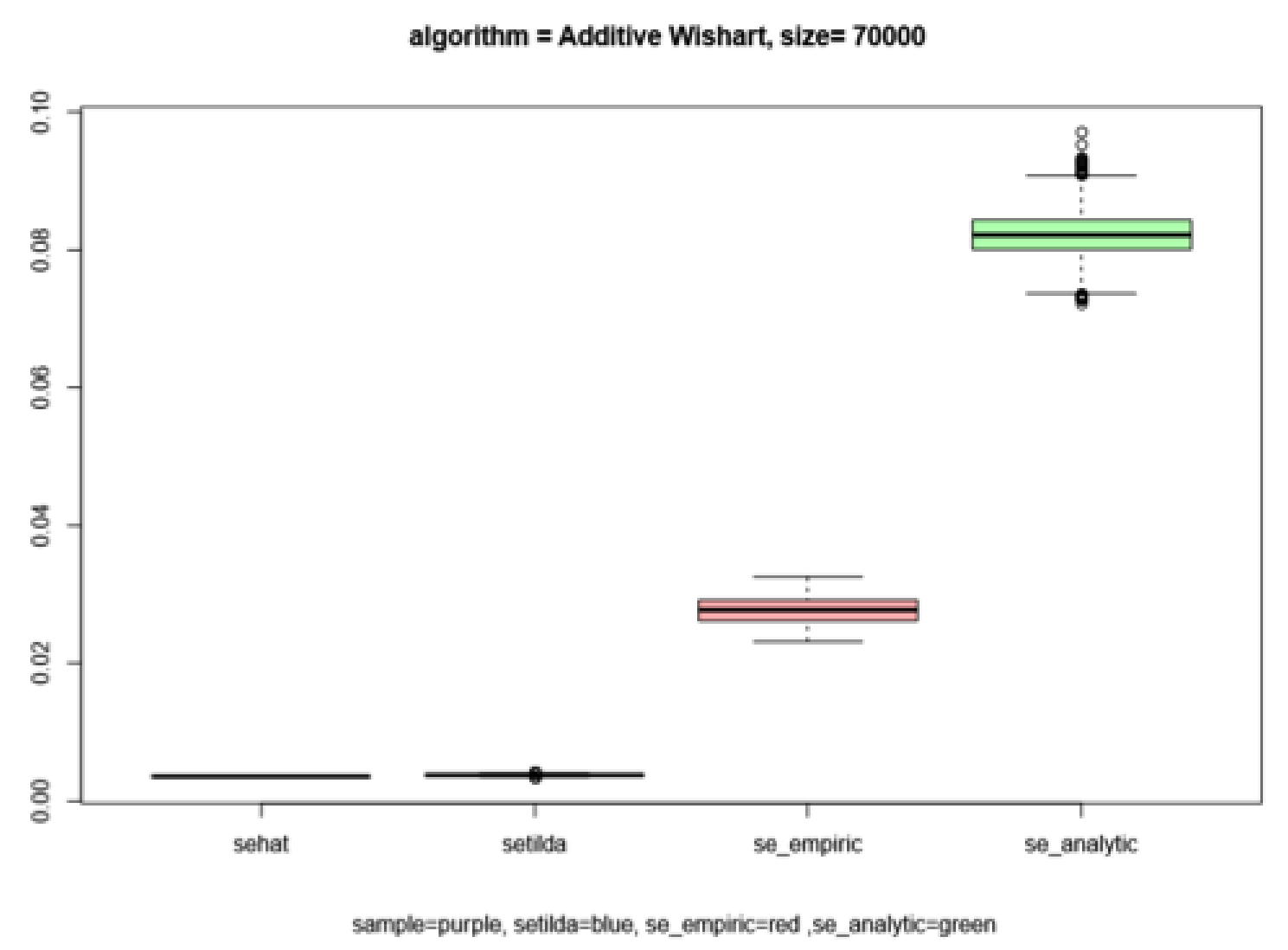
What happens for datasets of other sizes? We can examine t-values for *n* = 1000. These are seen in the graphs below:

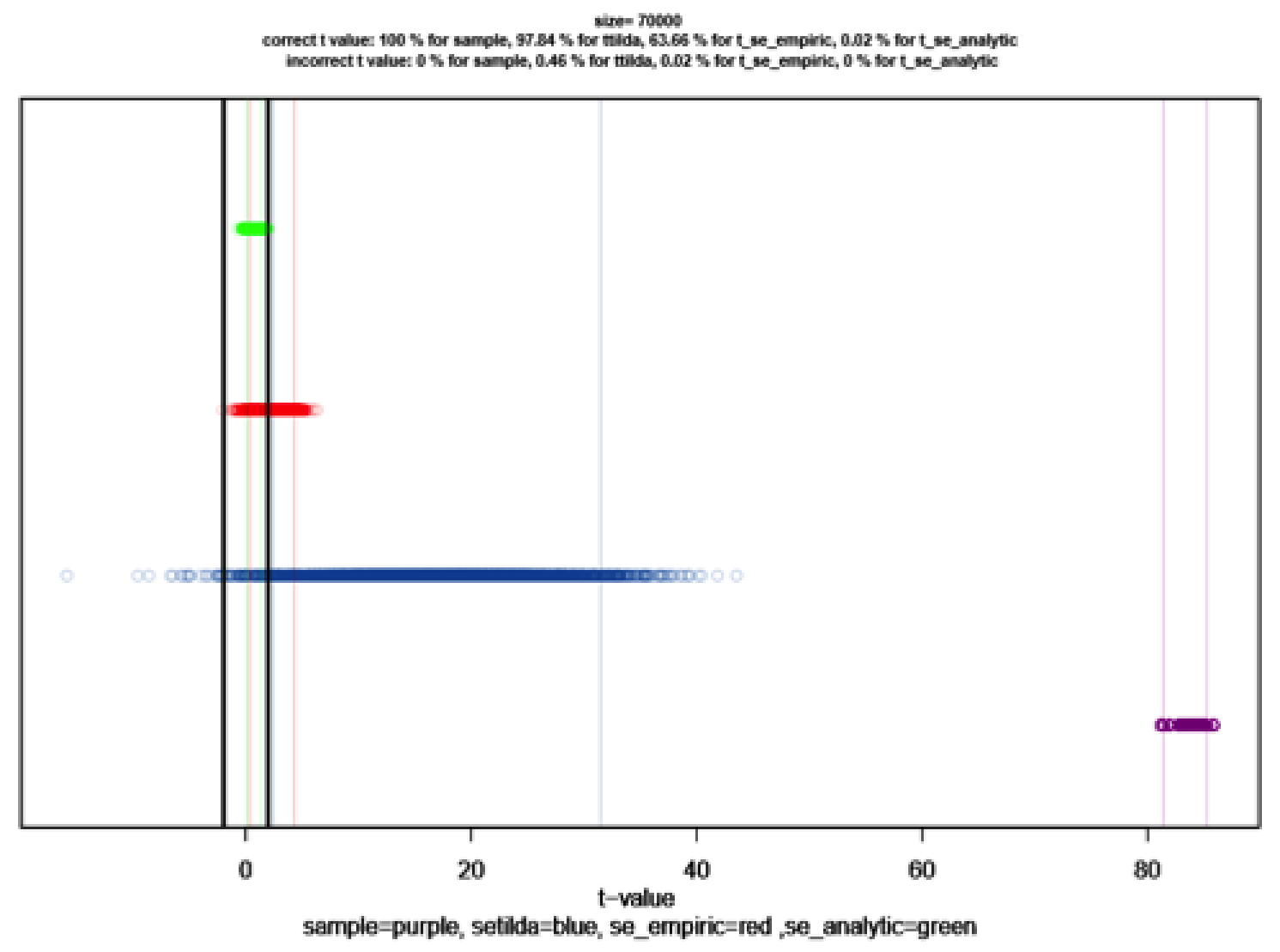


MORE ON T-VALUES

Note that our t-values for small *n* are problematic in a lot of cases: We have many t-values that suggest that we have high confidence in an incorrect result, although our regression coefficient estimates are generally correct.

We would like to find different ways of computing standard error for our algorithms. We demonstrate this using the Additive Wishart algorithm: We compute standard error 4 different ways (sample, differentially private estimate, empirical through computing many differentially private regression coefficients, and an analytical upper bound) and plot the resulting standard errors and t-values (for *n* = 7000 and all other parameters as previously stated):





We make the following observations regarding standard error:

- The analytical bound for standard error is very loose and gives far too conservative estimates.
- The differentially private estimate for standard error gives t-values that may be blatantly incorrect from an inferential perspective. For example, we sometimes get negative t-values of large magnitude when we should get positive ones.
- We would like to find a way to get a standard error near the empirical standard error, which takes into account randomness in both the sampling and differentially private algorithm., but have no way of calculating one.

FUTURE

In addition to obtaining these results that we already have, the testing framework could be used to further investigate these algorithms and in particular methods for performing inference on the results of these algorithms. As seen in the poor t-values, there is improvement needed in releasing a standard error consistent with the randomness that occurs when computing regression coefficients.

CRCS Center for Research on Computation and Society at Harvard School of Engineering and Applied Sciences

BERKMAN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

The Institute for Quantitative Social Science at Harvard University

DATA PRIVACY LAB