Statistically Valid Inferences from Differentially Private Data Releases, II: Extensions to Nonlinear Transformations*

Georgina Evans† Gary King‡

First version: August 8, 2020
This version: March 6, 2021

Abstract

We extend Evans and King (Forthcoming, 2021) to nonlinear transformations, using proportions and weighted averages as our running examples.

*All information necessary to replicate the results in this paper is available at Evans and King (2020).

†Ph.D. Candidate, Department of Government, Harvard University, 1737 Cambridge Street Cambridge, MA 02138; Georgina-Evans.com, GeorginaEvans@g.harvard.edu.

‡Albert J. Weatherhead III University Professor, Institute for Quantitative Social Science, 1737 Cambridge Street, Harvard University, Cambridge MA 02138; GaryKing.org, King@Harvard.edu.

0.1 Introduction

Evans and King (Forthcoming, 2021) develops valid statistical methods for analyzing "differentially private" data, where specially calibrated random noise has been added to provide formal privacy guarantees. This random noise creates measurement error which, if ignored in the analysis stage, induces statistical bias. The approach adapts methods developed to correct for naturally occurring measurement error, with special attention to computational efficiency for large datasets. The result is statistically consistent and approximately unbiased linear regression estimates, including interaction terms, quadratics, and some descriptive statistics.

Evans and King (Forthcoming, 2021) corrects regression coefficients when mean-zero gaussian noise is added directly to covariates. However nonlinear transformations of variables included in the regression and measured with error require different statistical procedures to avoid bias. In this paper, we thus develop a more general framework for incorporating nonlinear transformed variables. The simple insight is that the original estimator corrects least squares (LS) coefficients by adjusting the observed moments of the noisy variables according to the known error variance. We show here that, even if mean-zero noise is not added to the transformed variables directly (which is required in the original article), the estimator can still be applied if, firstly, the bias of the noisy transformed variable is approximately zero and, secondly, a reasonable estimator for the noise variance can be constructed using our knowledge of the error in the untransformed variables. We provide simulations demonstrating the effectiveness of this approach in the case of noisy proportions and noisy weighted averages.

1 Strategy and Notation

Following the notation in Evans and King (Forthcoming, 2021), let Z be an $n \times K$ data matrix that we do not have access to for privacy reasons. Instead we observe $X = Z + v$, where $v \sim \mathcal{N}(0, S^2)$. I.e., v is an $n \times K$ error matrix, and S^2 is a $K \times K$ diagonal matrix where diagonal elements represents the variance of the independent noise in each of the K private variables. Let Y be an outcome of interest which we observe either with or without error. Evans and King (Forthcoming, 2021) proposed the following estimator for β where $Y = Z\beta + \epsilon$:

$$\tilde{\beta} = \left(\frac{X'X}{n} - S^2\right)^{-1} \frac{X'y}{n} \tag{1}$$

where they show $\text{plim}(\tilde{\beta}) = \beta$. They key idea here is to correct the moments of $X'X$ by accounting for the known measurement error variance, S^2.

We now consider the extension where design matrix includes non-linear transformations of noisy variables. We focus on the univariate setting for expositional purposes but the method immediately extends to models with multiple covariates (transformed or untransformed) with independent error. Thus, let $Y = \gamma_0 + \gamma_1 g(Z) + u$, where $g(\cdot)$ is a non-linear function. Knowledge of S^2 is no longer sufficient to apply the estimator in (1) since the variance of a transformed variable is not the same as the variance of the untransformed variables, i.e., $\text{Var}(g(X)|Z) \neq S^2$. Moreover, non-linear transformations may introduce considerable bias: $\mathbb{E}[X|Z] = Z$ does not imply that $\mathbb{E}(g(X)|Z] = g(Z)$. In the presence of such bias, correcting only for the moment-inflation from the variance is not sufficient to recover a consistent estimate of γ; bias will remain. The converse implication is that if we are able to demonstrate that the transformed variable is approximately unbiased and quantify the variance, $\text{Var}(g(X)|Z)$, then we can apply estimator (1). We will demonstrate this in the remainder of the paper.

2 Noisy Proportions

To make this concrete, consider the case where the Z_j, for $j \in 1...K$, are counts, and $g()$ is a function that calculates the *proportion* of those counts in the first category. We denote this proportion by p where $p \equiv \frac{Z_1}{\sum_{j=1}^K Z_j}$. We then denote the same function applied to X instead by $r \equiv \frac{X_1}{\sum_{j=1}^K X_j}$. Our goal is to evaluate the bias of r. We start by approximating the expectation of r using a second-order Taylor approximation (we will implicitly condition on Z throughout):

$$\mathbb{E}[r] = \mathbb{E}\left[\frac{X_1}{\sum_{j=1}^K X_j}\right]$$

$$\approx \frac{\mathbb{E}[X_1]}{E[\sum_{j=1}^K X_j]} - \frac{\text{Cov}\left(X_1, \sum_{j=1}^K X_j\right)}{\mathbb{E}[\sum_{j=1}^K X_j]^2} + \frac{\mathbb{E}[X_1]}{\mathbb{E}[\sum_{j=1}^K X_j]^3}\text{Var}\left(\sum_{j=1}^K X_j\right)$$

$$= \frac{X_1}{\sum_{j=1}^K Z_j} - \frac{S_1^2}{\left(\sum_{j=1}^K Z_j\right)^2} + \frac{Z_1}{\left(\sum_{j=1}^K Z_j\right)^3}\left(\sum_{j=1}^K S_j^2\right) \tag{2}$$

The bias of r as an estimate of p is hence:

$$\approx \frac{Z_1}{\left(\sum_{j=1}^K Z_j\right)^3}\left(\sum_{j=1}^K S_j^2\right) - \frac{S_1^2}{\left(\sum_{j=1}^K Z_j\right)^2}$$

For intuition, suppose $S_1 = S_j$ and $z_j > 0$ for all j, then the *maximum* the bias could be is $\left(\frac{S_1^2}{Z_1^2}\right) \cdot (K-1)$. Hence, this suggests that for contexts in which the noise to signal ratio is moderate, we can generally approximate the bias as 0. We will show that this is reasonable via simulation.

We can also approximate the variance of r using a Taylor expansion. Denote the variance of r by S_r^2:

$$S_r^2 = \text{Var}\left(\frac{X}{\sum_{j=1}^K X_j}\right)$$

$$\approx \frac{\text{Var}(X_1)}{\mathbb{E}[\sum_{j=1}^K X_j]^2} - \frac{2E[X_1]}{\mathbb{E}[\sum_{j=1}^K X_j]^3}\text{Cov}\left(X_1, \sum_{j=1}^K X_j\right) + \frac{\mathbb{E}[X_1]^2}{E[\sum_{j=1}^K X_j]^4}\text{Var}\left(\sum_{j=1}^K X_j\right)$$

$$S_r^2 = \frac{S_1^2}{\left(\sum_{j=1}^K Z_j\right)^2} - \frac{2Z_1}{\left(\sum_{j=1}^K Z_j\right)^3}S_1^2 + \frac{Z_1^2}{\left(\sum_{j=1}^K Z_j\right)^4}\left(\sum_{j=1}^K S_j^2\right) \tag{3}$$

We do not observed the Z's, so for our estimator we can plug in the x's, which are unbiased estimates of the z's, to yield our estimator for the variance, $\hat{S}_r$.

3 Noisy Weighted Averages

We next consider the natural extension of a noisy proportion – a noise weighted average. In particular, consider the proportion $w \equiv \sum_{j=1}^{J} c_j p_j$ and $p_k \equiv \frac{X_k}{\sum_{j=1}^{J} X_j}$, where c_j is some known constant. As before, we do not observe the weights directly and hence can only calculate a noisy version $a \equiv \sum_{j=1}^{J} c_j r_j$, where $r_k \equiv \frac{X_k}{\sum_{j=1}^{J} X_j}$.

Like in the previous section, a reasonable approach to correcting OLS coefficients estimated with a rather than w is to estimate the variance of the error in a and adjust accordingly using the estimator in Evans and King (Forthcoming, 2021). Note that since a weighted average is a linear function of the ratios, if the ratios are unbiased (meaning the error is mean 0) then the weighted average is also unbiased. We demonstrated that this is a reasonable assumption if the noise to signal ratio is moderate.

All that is left, then, is to estimate the variance, which we do via the multivariate delta method. The multivariate delta method tells us that:

$$\text{Var}(w) \approx \sum_{j=1}^{J} \left(\frac{\partial w}{\partial X_j} \right)^2 \text{Var}(X_j)$$

Which can be written as:

$$= \sum_{j=1}^{J} \left(\sum_{k=1}^{J} \frac{\partial c_k r_k}{\partial X_j} \right)^2 \text{Var}(X_j)$$

Where:

$$\frac{\partial c_k r_k}{\partial X_\ell} = \begin{cases} \dfrac{c_k \sum_{j' \neq \ell} X_{j'}}{(\sum_j X_j)^2} & \text{if} \quad k = \ell \\ \dfrac{-c_k X_k}{(\sum_j X_j)^2} & \text{if} \quad k \neq \ell \end{cases}$$

$$\implies \text{Var}(w) \approx \sum_j \left[\left(\sum_{j' \neq j} \frac{-c_{j'} X_{j'}}{(\sum_j X_j)^2} \right) + \frac{c_j \sum_{j'} X_j}{(\sum_j X_j)^2} \right]^2 S_j^2$$

Therefore our final variance estimator is:

$$S_a^2 = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{J} \left[\left(\sum_{j' \neq j} \frac{-c_{j'} X_{j'}}{(\sum_j X_j)^2} \right) + \frac{c_j \sum_{j'} X_{j'}}{(\sum_j X_j)^2} \right]^2 S_j^2 \quad (4)$$

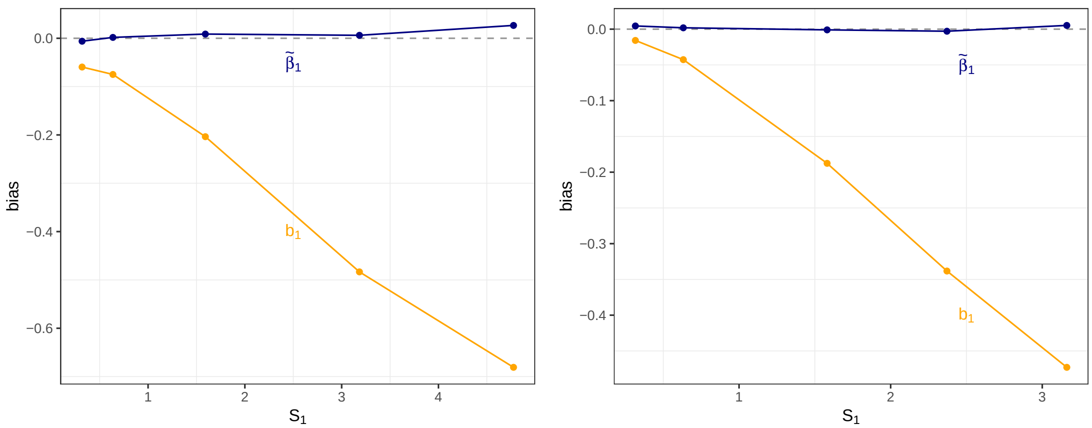

Figure 1: Statistical bias for noisy proportion (left panel) and noisy weighted average (right panel). In both panels, results for least square coefficients are in orange and for our estimator in blue.

4 Simulation

We now demonstrate the finite sample properties of the proposed approach. We begin by setting $n = 100,000$, $Z_1 \sim \text{Pois}(10)$ and $Z_2 \sim \text{Pois}(50)$. We draw $Y_i = 1 - \beta(\frac{Z1}{Z1+Z2}) + \epsilon_i$ where $\epsilon_i \sim \mathcal{N}(0, 1)$. We fix the noise in X_2 to $S_2 = \frac{\sqrt{50}}{2}$ and vary the noise variance in X_1. For each level of S_1 we run 500 simulations. We use the same underlying Z's and X's for our weighted average model, but now draw $Y_i = 1 - \beta w_i + u_i$ where $w_i = c_1 \frac{Z_{1i}}{Z_{1i}+Z_{2i}} + c_2 \frac{Z_{2i}}{Z_{1i}+Z_{2i}}$. We set $c_1 = 1$ and $c_2 = 2$ and draw $u_i \sim \mathcal{N}(0, 1)$. We fix $\beta = 1$ across both simulation settings.

In Figure 1, we give results for point estimates averaged over our 500 simulations. In the left planel, we plot the statistical bias (vertically) for our model estimated with a noisy proportion as a covariate where the standard deviation of the noise in the untransformed variable, X_1, is increasing along the horizontal axis. The right panel is analogous but our covariate is now a noisy weighted-average. Im both cases, we see that naively running an OLS regression with the noisy transformed variables introduces increasing bias and the underlying noise increases. In contrast, our alternative estimator, which explicitly accounts for the variance in the *transformed* variables is always approximately unbiased in this noise range. The far right of the figures is a context where the noise in X_1 is equal to the standard deviation of Z_1.

5 Concluding Remarks

The results here should extend to other nonlinear transformations, such as logs, square roots, etc., although in every case they should be checked by following the analytical and Monte Carlo methods introduced in this paper.

References

Evans, Georgina and Gary King (Forthcoming, 2021): "Statistically Valid Inferences from Differentially Private Data Releases, with Application to the Facebook URLs Dataset". In: *Political Analysis*. URL: GaryKing.org/dpd.

— (2020): *Replication Data for: Statistically Valid Inferences from Differentially Private Data Releases, with Application to the Facebook URLs Dataset*. Version DRAFT VERSION. DOI: 10.7910/DVN/UDFZJD. URL: https://doi.org/10.7910/DVN/UDFZJD.