

Classification with Measurement Error

Sarah Shy, Jonathan Hehir

April 21, 2020

(this slide is blank)

Introduction

What is measurement error?

In many statistical scenarios, we assume that our variables are perfect measurements from some sample.

What is measurement error?

In many statistical scenarios, we assume that our variables are perfect measurements from some sample.

This is often **false**. In many cases, we know our measurements are inaccurate.

What is measurement error?

In many statistical scenarios, we assume that our variables are perfect measurements from some sample.

This is often **false**. In many cases, we know our measurements are inaccurate.

In especially nice cases, we know **how inaccurate** they are.

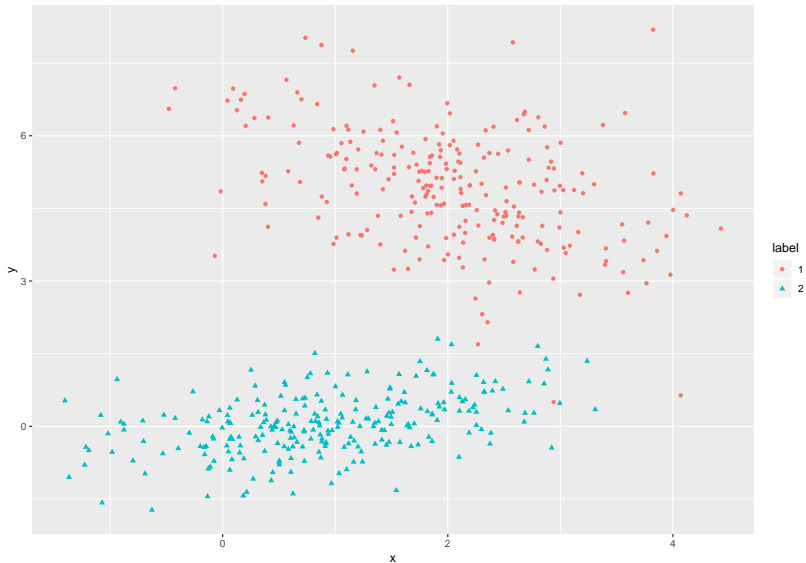
Applications of Measurement Error

Situations where we know our measurement error include:

- ▶ Astronomy data
 - ▶ DETAILS NEEDED HERE
- ▶ Privacy applications
 - ▶ Intentionally adding noise to observations is a common method for achieving privacy. Here we know the exact distribution of the real data.
- ▶ Earthquake data?
 - ▶ Would need to ask our friend Ian Lee about this.

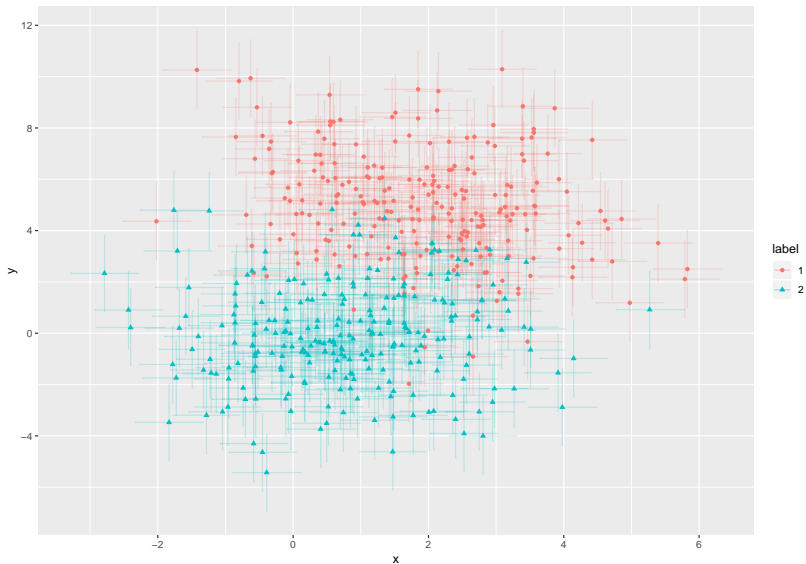
Example: A Clean Sample

This is an example of a particularly nice sample.



Example: Noisy Sample

This is the exact same sample, but with some error added to each point.



Example classifier performance

Fitting a basic linear SVM to the above yields the following accuracies:

| Original Sample | Noisy Sample |
|-----------------|--------------|
| 100.0% | 93.0% |

Example classifier performance

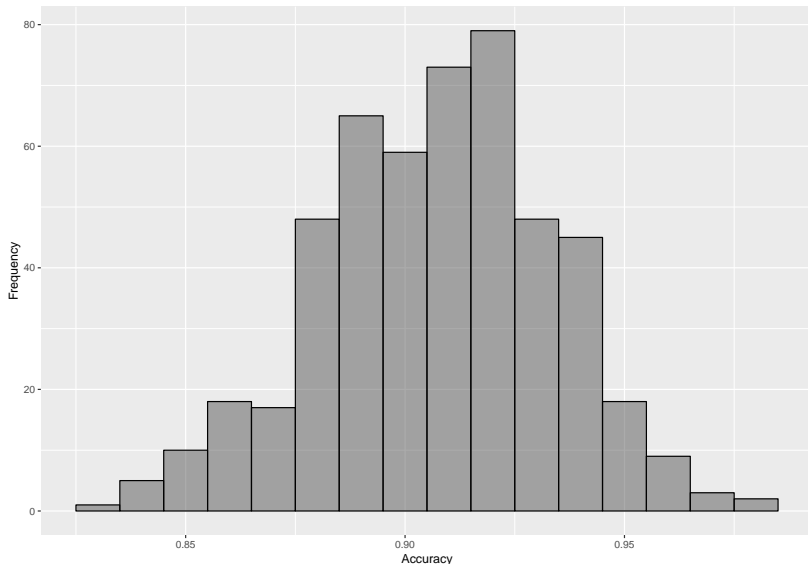
Fitting a basic linear SVM to the above yields the following accuracies:

| Original Sample | Noisy Sample |
|-----------------|--------------|
| 100.0% | 93.0% |

Can we trust this accuracy estimate?

Simulating Measurement Error

Using that same original sample, we repeat this process 500 times to obtain the following distribution of classifier accuracy:



Can we estimate that distribution without access to the clean sample?

Ideas

What do we want?

- ▶ Want a method for estimating uncertainty in classification performance that can be used for any classifier.
- ▶ etc.

How do we do it?

Insert pipeline here.

Real Data

Astronomy

Here's some data from astronomers (Ogden et al. 2014).

Astronomy

Here's some data from astronomers (Ogden et al. 2014).

There are literally rocket scientists.

Astronomy

Here's some data from astronomers (Ogden et al. 2014).

There are literally rocket scientists.

Just kidding. They don't have rockets.

Astronomy

Here's some data from astronomers (Ogden et al. 2014).

There are literally rocket scientists.

Just kidding. They don't have rockets.

But I bet they wish they did.

Results

In our testing, our method performs amazingly. It is so good.

The End

Bibliography

Ogden, C. L., Carroll, M. D., Kit, B. K., and Flegal, K. M. (2014), "Prevalence of childhood and adult obesity in the united states, 2011-2012," *Jama*, American Medical Association, 311, 806–814.