To Err is Machine: Measurement Error and DUI Law

1 Introduction

Every year in North Carolina, thousands of people are arrested for driving under the influence. Police officers stop drivers who show signs of impairment and administer several field tests to determine sobriety. If a suspect fails these tests probable cause has been established, and the individual can be taken into police custody for further testing. The most common test administered at the station is a breath analysis using a machine (these machines are often called breathalyzers). Results of breath analysis tests are frequently used as evidence of in DUI court cases. But how accurate are these machines? Manufacturers rarely produce estimates of error and a good, skeptical statistician might wish to conduct their own analysis, a quintessential motivating application in statistics classrooms. Statisticians use the term measurement error to refer to the deviations between measurements made by a device and the underlying quantity the device is attempting to measure. The current legal framework fails to incorporate measurement error when weighing the evidence provided by breathalyzers, which may lead to convictions of innocent persons, for most cases defined as an individual with a BAC (blood alcohol content) of less than .08%.

A naive approach to calculating measurement error would involve calculating the sample variance via the standard formula. However, we will see our data has a structure that makes those assumptions unreasonable. If we had a model we could determine the value via a technique called maximum likelihood estimation. We believe it is reasonable to model true BAC values as draws from a continuous distribution. Furthermore, we assume that instead of a true BAC value, what the machine measures is the true value plus some measurement error due to the machine. We will assume that the measurement error is uncorrelated with the population mean. If we assume normally distributed BAC values and measurement error terms, we can write this as $\mu_i \sim N(\theta, \sigma^2)$ $\epsilon_{ij} \sim N(0, \tau^2)$. We can then write the value of a breathalyzer measurement as $X_{ij} = \mu_i + \epsilon_{ij} \sim N(\theta, \sigma^2 + \tau^2)$ (the subscript i represents different individuals and the subscript j represents possible replicates on a person). A histogram of the averages of the two readings suggests that normally distributed BAC values are reasonable.

2 Analysis

When a breathalyzer is administered, the test is performed twice per suspect. The readings actually reported by breathalyzers provide a couple of issues (here we make a distinction between the measurements that the breathalyzer calculates and the readings which are output by the machine). The readings are the values of the measurements truncated and rounded down to the nearest hundredth. So the value of the reading looks like $Y_{ij} = \lfloor 100 \times X_{ij} \rfloor / 100$ where $\lfloor \rfloor$ is the floor function, which always rounds down to the nearest integer. This presents a challenge for maximum likelihood, because the Y_{ij} data we observe are no longer normally distributed. Mathematically, we observe

$$\begin{pmatrix} Y_{i1} \\ Y_{i2} \end{pmatrix} = \begin{pmatrix} \lfloor X_{i1} \rfloor \\ \lfloor X_{i2} \rfloor \end{pmatrix}$$

$$\begin{pmatrix} X_{i1} \\ X_{i2} \end{pmatrix} = \begin{pmatrix} \mu_i + \epsilon_{i1} \\ \mu_i + \epsilon_{i2} \end{pmatrix} \sim N\left(\begin{pmatrix} \theta \\ \theta \end{pmatrix}, \begin{bmatrix} \sigma^2 + \tau^2 & \sigma^2 \\ \sigma^2 & \sigma^2 + \tau^2 \end{bmatrix} \right)$$

To derive the distribution of $\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$,

$$\begin{aligned} & P(Y_1 = y_1, Y_2 = y_2) \\ = & (\lfloor X_1 \rfloor = y_1, \lfloor X_2 \rfloor = y_2) \\ = & P(y_1 \leq X_1 < y_1 + 1, y_2 \leq X_2 < y_2 + 1) \end{aligned}$$

This finally equation can be evaluated using the bivariate normal CDF.

Now that we have a model, we are interested in estimating the value of τ^2, the variance of the measurement error terms. One of the tools that statisticians use to estimate unknown parameters is called maximum likelihood estimation. The intuition behind maximum likelihood estimation is to consider several models as candidates, and to pick the model that best fits the data (the model under which the observed data has the highest probability). This technique will look at all bivariate normal distributions with parameters τ, σ, and θ unknown and to pick the parameters that give us the best value.

3 Results

Our data come from 3 different machine models, and we estimate the measurement error separately for each model. Our model contains an assumption that there is no correlation between the duplicate measurement error terms the machine calculates for each person.

This table contains parameter values for machine 8856 under the assumption that the correlation between measurements is 0.

Parameter	*Estimate*
τ	0.0083
σ	0.0480
θ	0.1420

4 Conclusions

Using the parameter estimates calculated above, we perform a test of the hypothesis that the defendant in a DUI case had a true BAC of .079%, which would be below the legal limit. The probability of observing a pair of values under the null hypothesis depends on the model of the machine and the assumed correlation between the measurement error terms. For a pair of measurements at 0.08, the wrongful conviction rate could be as high as 25% in the case that the machine reported .08 for both of the defendant's measurements. The figure below describes whether an individual would be found guilty and the probability of wrongful conviction under the model for reported BAC values. This table corresponds to the case where we assume no correlation between first and second reading.

1st \ 2nd	0.06	0.07	0.08	0.09	0.1	0.11
0.06						
0.07						
0.08			0.204	0.042	0.003	
0.09			0.042	0.009	0.001	0
0.1			0.003	0.001	0	0
0.11				0	0	0

Not Guilty | Guilty, p-value < .05 | Guilty, p-value > .05

If we assume the correlation between the readings is 0.3, the false conviction rates change.

1st \ 2nd	0.06	0.07	0.08	0.09	0.1	0.11
0.06						
0.07						
0.08			0.257	0.078	0.009	
0.09			0.078	0.028	0.004	0
0.1			0.009	0.004	0	0
0.11				0	0	0

Not Guilty | Guilty, p-value < .05 | Guilty, p-value > .05

For many real-world applications, we want the false conviction (more generally, the false positive) rate to be less than 5%. Among a random sample of NC DUI cases, approximately 3% of suspects had readings in the red region in the first table (high false positive rate), but would be convicted under a straightforward application of DUI law. If we look at the second table, that proportion rises to 5.6%. In 2013 there were 41,247 DUI arrests in North Carolina. If we extrapolate the proportion in the red region to our entire sample we have 1253 and 2337 arrests for people who were in the red region, and we believe it is reasonable to assume many of these arrests may have led to convictions.

We propose 3 policies to reduce the false positive rate. DUI law could be reworked to exclude convictions among people in the red region. This has the benefit of avoiding changes to police procedure or breathalyzer machines, but would require the passage of new legislation. Alternately, if police took 3 measurements instead of 2 it would increase the power of the hypothesis test and reduce the false positive rate. The most direct solution would be to have breathalyzers report all of their digits instead of truncating the readings, which would allow for precise confidence intervals and improve estimation of the measurement error component. Each of these policies has the added benefit of reducing public spending on DUI court cases. If defendants who are close to the legal limit are more likely to go to trial, the reduction in public legal resources could be considerable.

Given the current duplicate measurements per person, the correlation cannot be estimated in a model. This analysis did not account for time between breathalyzer readings as that information was not recorded. Further work on breathalyzer measurement error is needed to improve the justice system and reduce public expenditures on hard-to-win court cases.