Calibration Experiments for a Computer Vision Oyster Volume Estimation System

G. Andy Chang
Youngstown State University

G. Jay Kerns
Youngstown State University

D. J. Lee
Brigham Young University

Gary L. Stanek
Youngstown State University

Key Words: Calibration; Linear regression; Inverse regression.

Abstract

Calibration is a technique that is commonly used in science and engineering research that requires calibrating measurement tools for obtaining more accurate measurements. It is an important technique in various industries. In many situations, calibration is an application of linear regression, and is a good topic to be included when explaining and learning the concepts of linear regression. However, calibration is not often mentioned in the introductory statistics textbooks or in the introductory statistics classrooms. The goal of this paper is to share with instructors an example with real data for simple linear regression and its application in calibration. It can be used as a lecture example, a class project, or a lab activity.

1. Introduction

Correlation between quantitative variables and simple linear regression are topics that are covered in most high school and college-level introductory statistics courses. The standard lecture usually introduces Pearson's correlation coefficient and discusses least square estimation for the parameters in a linear regression equation, and then studies the relationship using this regression equation. Data provided in this paper can be used for demonstrating all the concepts mentioned above. However, the main focus of the use of this data set is on calibration.

Calibration is a procedure that is used for adjusting readings or output from a measurement tool to agree with an applied standard and for achieving better accuracy in the measurement process. It can also be used to convert

measurement data obtained from a different format or scale to achieve a measurement result. Statistical calibration usually involves an inverse estimation and it is sometimes called inverse regression. The data story in this article is a good example of calibration. The two common approaches to calibration using regression are the classical and inverse methods. The classical estimation method was first investigated by Eisenhart (1939). Krutchkoff (1967) derived the inverse estimation method which caused considerable controversy at the time. Many studies (Berkson, 1969; Krutchkoff, 1971; Shukla, 1972; Osborne, 1991; Brown, 1993) have been done to describe the properties of inverse regression for estimation, and to examine the advantages and disadvantages of these two approaches. The ideas behind these two methods will be explained and discussed in this paper. Since this paper emphasizes common practice with the objective of introducing a simple example for an introductory class, the Bayesian, nonparametric, and bootstrapping approaches to calibration will not be discussed.

Most introductory statistics textbooks do not specifically cover inverse regression or calibration. Some textbooks (Peck and Devore, 2008; Brase and Brase, 2009; and DeVeaux, Velleman, and Bock, 2006) warn readers against estimations of the independent variable, x, using the regression line established from regressing the dependent variable, y, on x. However, the main objective in calibration problems is in estimating the value of the independent variable corresponding to a measured value of the dependent variable as described in this data story. In this special situation the inverting of the regression line to find the estimates for x is necessary and allowed. Although there are confidence interval estimation formulas for calibration problems (Ott and Longnecker, 2001), in many applications, people do calibration only for point estimation without involving confidence intervals, such as the story in this article. In this way, calibration can be considered as an application of simple linear regression. Calibration is a useful technique, an interesting application of simple linear regression, and many students use it in their college science and engineering courses.

The data set in this article is from a calibration experiment related to an oyster volume estimation (OVE) system. Instructors may use it for lecture material or a class activity for teaching or learning basic techniques in simple linear regression, such as Pearson's correlation coefficient, least squares regression, and estimation.

2. The Story – Developing an Oyster Volume Estimation (OVE) System

The data set in this paper came from a real situation in the food industry. It was from the joint work by two of the authors of this article in a project for designing a fast and inexpensive computer vision oyster volume estimation system using 3-dimensional (3-D) information. Volume measurement or estimation is an important process in the food industry, since the price of the product is often determined by the size of the product. For instance, the price of oyster meat is determined by its size. Large oyster meat can sell for a higher price than regular or small sizes. Before computer vision sorting systems were available for oyster sorting, human vision was the main measurement tool used for oyster sorting. But manual sorting is prone to errors which result in both profit loss and customer complaints. Parr et al. (1994) implemented the idea of using a 2-dimensional (2-D) digital image of the oyster through a machine vision system to estimate its size. The system takes the digital image of oyster meat that is put on a conveyor belt from a camera directly above the meat looking down. The system binarizes the image taken from the digital camera so that the pixels (i.e., *picture elements*) of the oyster meat are set to 1 and the pixels of the background are set to 0. Sample digital images before and after binarization are shown in Figure 1. The 2-D area of the oyster meat can then be measured by counting the number of pixels that have the value 1. This 2-D area measurement of oyster meat in pixels is then correlated with the actual volume of the oyster meat to find an equation for estimating the oyster meat volume. This is the calibration process.
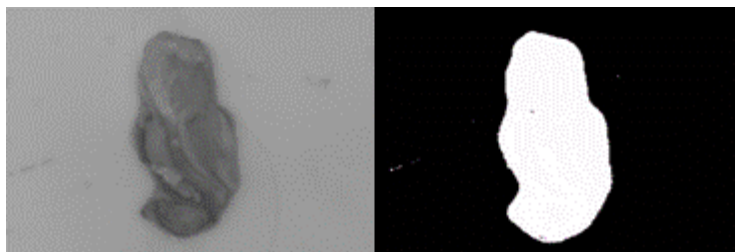
Original Image   Binarized Image

Figure 1. Digital Images for a 2-D system.

Using several laser light beams with a digital camera and a triangulation technique (see the Appendix), a low cost system can be designed to obtain the thickness information of oysters. Combining the 2-D data with the thickness information, the volume of oyster meat can be estimated in terms of number of 3-D pixels. Similar to the 2-D system, the volume estimate in pixels from the 3-D system can be correlated with the actual volume of the oyster meat. The volume estimation can be established using a regression equation relating these two variables.

Instructors who would like to assign a simple project to students can use the 2-D data, and the 3-D data can be used for longer projects or to explore or explain the use of statistics from regression for comparative study.

The actual volume measurements of oyster meat in this study were carefully taken by lab technicians, and the computer vision OVE system was well designed. As a result, the data produced from this study were remarkably well behaved, with no extreme values. This study consequently serves as a good example or project for use in an introductory statistics course, because instructors do not need to worry about outliers, a poor fit, or many of the other issues typically encountered with real-world data. Of course, the issue of outliers in regression modeling is important but it is not the main concern of this paper.

Oyster Data

In this study, thirty oysters of various sizes, weighing from 5 grams to 18 grams, were selected. Oysters were placed in random orientations on the platform of the imaging system. The image information for each oyster was taken simultaneously by utilizing both 2-D and 3-D settings described above to obtain volume information (in number of pixels) calculated by the computer vision volume estimation system. After system calibration, each system produced an estimation formula for estimating the volume of oyster meats. One can use these formulas to evaluate and compare the performance of the two systems.

The experiment, including designing and setting up the system and taking measurements of the oyster meat, was done at an agriculture computer vision technology laboratory. The data collected were the *weight of the oyster meat* in grams (g), and the *volume of the oyster meat* in cubic centimeters (cc) which was measured by an agriculture product measurement specialist using the displacement of water adjusted by a temperature adjustment factor. We recorded the *number of pixels for observed oyster meat from the 2-D system* and the *number of pixels for observed oyster meat from the 3-D system*. And, for this pedagogical article, we rounded the actual weight and volume measurements to two decimal places.

Please note that while the oyster meat weights are included in the data set associated with this article, they do not play a substantive role in the discussion that follows. We have included them because they were an important part of the original data collection, and are strongly correlated with the oyster meat volumes. An instructor may wish to use them for an alternate model of the volume, or perhaps even to discuss multicollinearity concepts in multiple regression. Ultimately, it is hoped that the additional information will give more options to the designer of classroom activities.

3. Pedagogical Uses

Using this data set, students can learn and practice all of the basic elements in simple linear regression such as making a scatter plot, Pearson's correlation, testing of zero correlation, the coefficient of determination, building regression equations and estimation. Evaluation of the aptness of a regression model as suggested by Matson and Huguenard (2007) can be performed with this data set. In actual applications, the main objective for these data is to establish a calibration equation. One simple assignment idea is to use either the 2-D or 3-D data to fit a regression line and build a model for calibration and estimation. Of course, the data can be used to develop assignments with different objectives.

In the following sections, the difference between the classical and inverse approaches to calibration will be discussed. A subset of 20 randomly selected observations from the complete set of 30 cases is utilized for demonstrating the use of the data. Instructors may use this as an example to show the class and use the full data set for assignments or in-class activities. The statistics and graphs in this paper were generated using the R Commander, developed by Fox (2005), which is a graphical user interface to R – the open source statistical environment. It can be downloaded for free from the R Project web site. It is the software that we use in our introductory probability and statistics course for science, engineering, and mathematics majors.

The instructor may set up an exercise for comparing the data obtained from the 2-D and 3-D systems to see which system performed better using various statistics generated from the respective regression analyses. For an in-class activity, the instructor may also divide the class into two groups of students, one group to analyze the 2-D data and the other to analyze the 3-D data. Students can then make comparisons after finishing their parts of the exercise. A sample for class assignments is in Section 5.

Calibration and Estimation

The *calibration problem* in statistics may be simply described as doing linear regression in reverse. The idea is that a researcher is confronted with an independent variable of interest which is difficult or expensive to measure. Luckily, however, (s)he also has a second variable, which depends on the first, but which is relatively easy or inexpensive to measure. Given a suitable (linear) model relating the two, the goal is to use inexpensive future values of the dependent variable to *estimate* the corresponding values of the *independent* variable. Even though it may seem backward, such an approach can generate substantial savings of time and money in applications.

There are two popular methods for attacking the calibration problem: the classical method and the inverse method. The difference between them concerns the respective roles that the two variables play when fitting the regression model. We will discuss both methods in detail below, since they can sometimes be counterintuitive to both students and instructors unfamiliar with the topic.

In this study, the independent variable is the actual volume of oyster meat and the dependent variable is the area in pixels (for 2-D) or volume in pixels (for 3-D) reported from the computer vision volume estimation system. We have identified the pixel count to be the dependent variable because its value *depends* on the actual volume of the oyster meat, the independent variable. The pixel count reported by the 2-D or 3-D volume estimation system is used with the calibration equation to estimate the actual volume of the oyster meat. Visual inspection (with a scatter plot) is usually a preliminary step to fitting a straight line model since in some situations the relationship between the two variables may be non-linear. Figure 2 shows a clear linear relation from the data given by the 2-D system, which indicates that a simple linear regression model might be appropriate for the data (which may be found in an ASCII data file containing a sub-sample of 20 oysters: 20oysters.dat). Given the care with which the data were collected, the assumption of independent and identically distributed (i.i.d.) errors seems plausible (and subsequent residual plots and diagnostic tests confirm this). We can then compute the Pearson correlation coefficient, find the equation of the straight line that best fits the data, and compute the coefficient of determination, R^2.
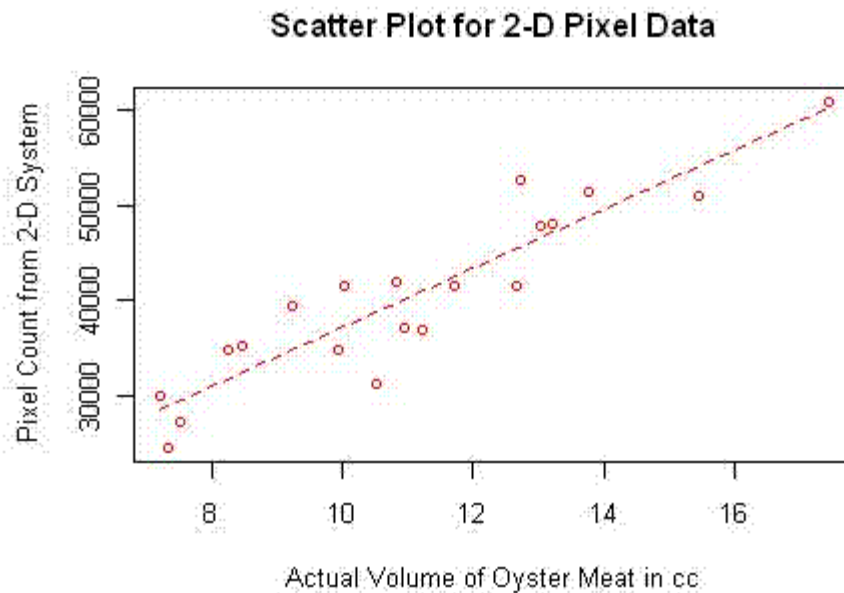
Figure 2: Scatter Plot for Actual Oyster Volume versus Pixel Count from 2-D System.

The Classical Method for Calibration

In the classical method for calibration, the idea is to preserve the dependent and independent roles of the variables, and to fit a standard linear regression line to the observed data. Once the regression equation has been determined, it is inverted to solve for the independent variable and estimates are obtained by substitution into the resulting calibration equation.

Let x_i be the actual volume of the oyster meat measured and y_i be the pixel count from the 2-D system for the observed oyster. The classical simple linear regression model assumes a model of the form

$$y_i = a + b \cdot x_i + e_i, \qquad (1)$$

where the e_i's are independent and identically normally distributed random variables with mean zero and common standard deviation s. Given a set of data points, the least-squares estimates for the parameters in the model are denoted respectively $\hat{\alpha}$ and $\hat{\beta}$. Therefore, the least-squares equation can be written as $\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x$. The calibration equation for the estimation of the volume of oyster meat (x) using pixel count (y) can then be obtained by solving the least-square equation for x,

$$\hat{x} = \frac{y - \hat{\alpha}}{\hat{\beta}}. \qquad (2)$$

A confidence interval for the estimate can be computed at a specified confidence level through inverse estimation. However, in the actual oyster sorting process, the decision is generally made based on the predicted value from equation (2), and a confidence interval is not utilized. The output from using the statistical package R for simple linear regression with the actual oyster meat volume as the independent variable is presented in Figure 3. The parameter

estimates are $\hat{\alpha} = 5863$ and $\hat{\beta} = 3125$. The calibration equation is $\hat{x} = (y - 5863)/3125$. So, if the 2-D system reported an oyster size in pixels of 47907—the pixel count for case number 1 in the data which has an actual volume of 13.04 cc—then the estimated oyster size would be (47907 - 5863)/3125 = 13.45 cc. It follows that the error ("actual" minus "estimated") would be 13.04 – 13.45 = –0.41.
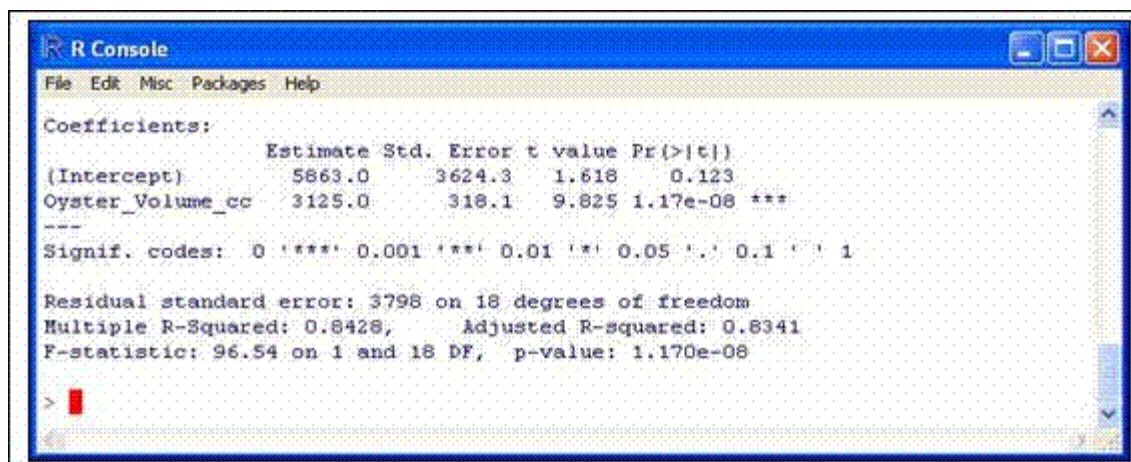

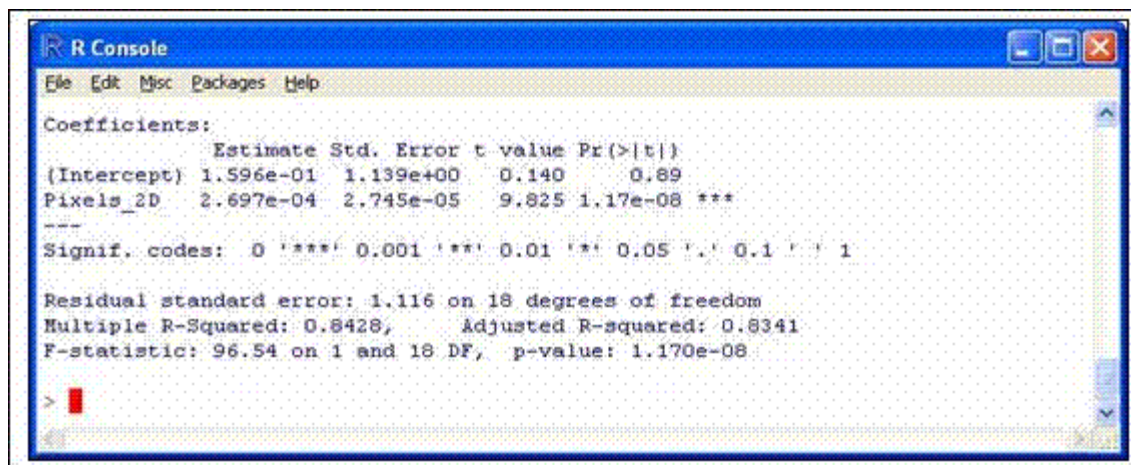**Figure 3:** Output from R for Simple Linear Regression Using Classical Method.

A common problem for students while trying to use the classical method for calibration is the determination of the dependent and independent variables. Some students will choose the actual volume measured as the dependent variable, *y* in equation (1), arguing that, intuitively, it is the variable to be estimated later. Are they wrong for doing this? In fact, their intuitive approach is called the inverse method for calibration.

The Inverse Method for Calibration

The inverse method is also a commonly used approach in industry for system calibration. This approach is straightforward and simple in practice. Since the actual volume is the variable to be estimated, it is tempting to suppose that the actual volume of the oyster meat, x_i, is the dependent variable in the model. In that case, the estimation of the volume of oyster meat could then be done by considering the regression of the oyster meat volume x_i on y_i, the observed pixel count of the oyster meat, and the regression model would have the form

$$x_i = a^* + b^* \cdot y_i + e_i^*, \qquad (3)$$

where e_i^*'s are random and normally distributed random variables with mean 0 and common standard deviation s^*. Continuing, the calibration equation would be written as $\hat{x} = \hat{\alpha}^* + \hat{\beta}^* \cdot y$, where $\hat{\alpha}^*$ and $\hat{\beta}^*$ are the least square estimates for a^* and b^*. In the R output below for the inverse method, the estimates are $\hat{\alpha}^* = 0.1596$ and $\hat{\beta}^* = 0.0002697$. The calibration equation would be $\hat{x} = 0.1596 + 0.0002697 \cdot y$. For an oyster size in pixels of 47907, (recall that this is the pixel count for case number 1 which has an actual volume of 13.04 cc), the estimated oyster size would be 0.1596 + 0.0002697 · (47907) = 13.08 cc. The error (*actual* minus *estimated*) would then be 13.04 – 13.08 = –0.04. Please note that, for the inverse method, what we have called the "error" for case number 1 is simply its *residual* from the inverse regression line. The same is not true for the classical method. Table 1 contains the errors of estimation for both of the methods.

```
R Console

File  Edit  Misc  Packages  Help

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.596e-01  1.139e+00   0.140     0.89
Pixels_2D    2.697e-04  2.745e-05   9.825 1.17e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.116 on 18 degrees of freedom
Multiple R-Squared: 0.8428,      Adjusted R-squared: 0.8341
F-statistic: 96.54 on 1 and 18 DF,  p-value: 1.170e-08

>
```

Figure 4: Output from R for Simple Linear Regression Using Inverse Method.

Further Investigation of Classical versus Inverse Calibration Methods

Among the classical and inverse methods, which one should be used for this data set? Many research studies have been done to better understand the properties of these two methods. Osborne (1991) did a thorough review on methods of statistical calibration and examined the pros and cons of the classical and inverse methods and other methods. Standard statistical textbooks (Ott and Longnecker, 2001; Draper and Smith, 1998; Graybill, 1976) usually focus on the classical method. Many statisticians prefer the classical method for its consistency--in the sense of convergence in probability of the estimator to the true value of the parameter as the sample size increases without bound--and distributional properties. Brown (1993) and Osborne (1991) mentioned that if the values of the independent variable are a random sample from a population, which is usually referred to as natural or random calibration, then the better way for the inverse estimation is to use inverse method. But, if the independent variable values are controlled at fixed levels by design in the training experiments, which is sometimes called controlled calibration, then the classical method has been believed to be better, at least since Eisenhart (1939). Controlled calibration is a popular technique in calibration experiments.

For the calibration experiment in this paper, the observations were made somewhat randomly but were controlled in such a way that they have good spread through a range of oyster sizes. We chose the classical method in the original research project for calibrating the systems. We could have chosen the inverse method; the point estimates would not have differed substantially. There would, however, have been more marked differences in the confidence intervals, but confidence interval estimation was not needed in the original experiment.

From the Classical and Inverse discussions above, the pixel count of 47907 from the 2-D system for the first oyster meat volume listed in the data set has an actual volume of 13.04 cc and has an estimated volume of 13.45 cc (with an error of -0.41 cc) for the classical method and 13.08 cc (with an error of -0.04 cc) for the inverse method. But, these results were only from one case in the data set and should not be used to conclude that the inverse method is better.

In the search for a valid manner to compare the performance of the classical and inverse methods, we might first consider studying the error (actual volume minus estimated volume) variability from each of the two methods. The inverse method simply calculates the least-squares line where oyster volume is the response variable and the 2-D pixel count is the explanatory variable. Thus the resulting sum of squares of the errors is as small as possible, by definition of the least-squares line. The classical method fits a different straight line to the same data; the sum of squares of its errors is by definition going to be larger. It is therefore not possible to compare the variances of the errors from the 20 oysters used to fit the models.

However, comparing the error variability of the two methods on the remaining 10 pieces of oyster offers insight into the performance of the methods. To see whether the difference in variances is statistically significant, we may employ resampling methods (note that the standard F test for comparing sample variances assumes independent samples, which is clearly not satisfied here). A 95% bootstrap percentile interval (see DiCiccio and Efron, 1996) for the ratio of
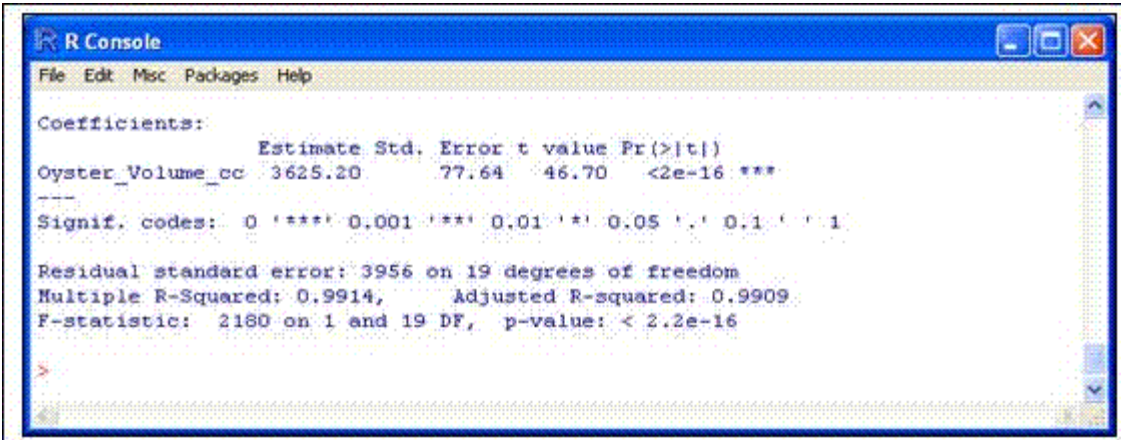
variances is (0.521, 2.600) with 1000 replicates. Since the interval does cover one (1), we see that, for the size 10 sub-sample in the 2D case, there is no significant difference in the variances for the two methods.

Table 1. Errors in Volume Estimation from Classical and Inverse Methods.

Oyster ID	Error = Actual - Estimated	
	Classical_method	Inverse_method
6	2.08	2.47
13	0.72	0.93
15	0.92	1.27
16	-0.13	-0.10
17	-3.69	-2.44
19	0.22	-0.66
22	0.54	-0.03
25	0.57	-0.99
26	-1.18	-0.95
30	-0.80	-0.16
Mean	-0.19	-0.07
Variance	2.41	1.87

In the R output for the model estimations in Figures 3 and 4, the estimates of the intercept of the regression line for both classical and inverse methods are not significantly different from zero, at 10% level of significance. An inevitable question is whether a model without an intercept is appropriate for these data. It is tempting to assert that if an oyster's meat volume were zero then the oyster meat must not exist; therefore, the pixel count data from the volume estimation system would necessarily be zero as well and regression through the origin is appropriate. However, we must be careful. On one hand, the smallest oyster meat in the dataset has volume more than 5 cc, meaning that no data were collected near a volume of 0. Consequently, we have no evidence to support that the linearity assumption holds outside the range of the observed data. Extrapolation to small or large volumes is a dangerous practice, and should not be taken lightly. On the other hand, physical principles suggest that physical volume is linearly related to the pixel count reported from the computer vision OVE system.

Figure 5 has the R output for the classical method with regression through the origin using the sample of 20 oysters studied earlier. This model has a much higher R^2 value (0.99) and the inverse estimation equation is $\hat{x} = y/3625.20$. For an oyster that has a pixel count of 47907 from a 2-D OVE system using the classical method, the estimated volume would be 47907/3625.2 = 13.22 (cc). The error would be 13.04 - 13.22 = -0.18.


```
R Console                                                                    _ □ ✕
File  Edit  Misc  Packages  Help

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
Oyster_Volume_cc  3625.20      77.64   46.70   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3956 on 19 degrees of freedom
Multiple R-Squared: 0.9914,     Adjusted R-squared: 0.9909
F-statistic:  2180 on 1 and 19 DF,  p-value: < 2.2e-16

>
```

Figure 5. Regression through Origin Output from R Using Classical Method.

4. Comparing the 2-D and 3-D Systems

Scatter plots can help us understand whether straight lines are appropriate for both 2-D and 3-D methods. Two scatter plots for graphing pixel count versus actual volume of the oyster meat are shown in Figure 6. The graphs suggest a strong linear relationship for both cases, and the 3-D data seem to have smaller variability. Of course, one should be careful to read the scales used in the graphs before attempting to judge the performance of the two methods in terms of variability. The scatter plot is not the best method to visualize the precision of the two methods.
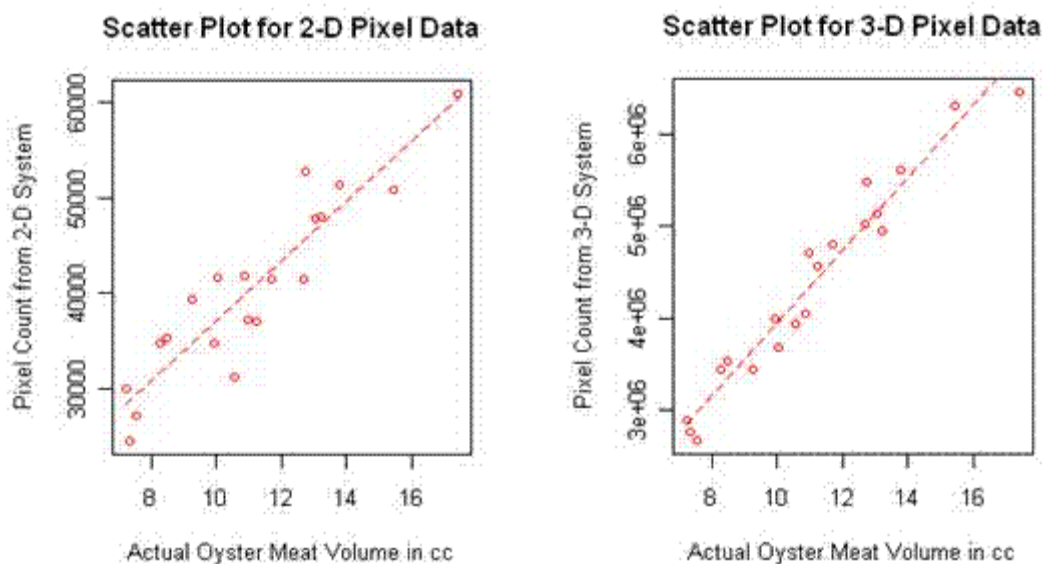

Figure 6. Scatter Plots for Comparing 2-D and 3-D Systems.

There are other items that can be used in the decision making process, such as the Pearson correlation coefficient, R, the coefficient of determination, R^2, and plot of errors or standardized errors calculated from using the estimation equations. When students were asked to select a statistic for this comparison in class, most of them chose R^2. One should be careful in interpreting R^2. It is not necessarily the best statistic for explaining how well a regression equation fits the data. Its value is affected by the spread of the data and, therefore, one should also refer to a scatter plot to help in making a judgment. However, in most of the calibration experiments, the observations are controlled so that the data are usually evenly spread and R^2 becomes more meaningful to be used for comparison. If the classical method is used for comparison in the 20 case data set, the R^2 for the 3-D system (0.9518) is higher than the R^2 for the 2-D system (0.8428).

Using the Variance of the Errors

Examining the estimation errors is important for comparing the two systems. For the 3-D data, the regression equation using the classical method is $\hat{y} = -9887 + 395518\,x$; see the parameter estimates in Figure 7. The inverted estimation equation is $\hat{x} = (y + 9887)/395518$. The 3-D pixel count for case number 1 is 5136699. The estimated volume for this oyster using its 3-D pixel count would be $(5136699 + 9887)/395518 = 13.01$ (cc). The error is $13.04 – 13.01 = 0.03$. The errors can be computed for all the data collected for the 2-D and 3-D systems.

Figure 7. R Output for Simple Linear Regression Using Classical Method for 3-D data.

Using the classical method, for the 3-D system, the sample variance of errors (not shown, but computed in the same way as displayed in [Table 2]) is 0.3797 which is decidedly less than the sample variance of the errors from the 2-D system (1.40). The smaller error sample variance indicates that the 3-D system may provide better precision, and as before, we may use resampling to obtain a confidence interval for the ratio of variances. A 95% bootstrap percentile confidence interval for the ratio turns out to be (1.97, 6.75), with 1000 replicates. As the interval does not cover 1, the difference in variances of the two methods is statistically significant. Both sets of errors from the 2-D and 3-D systems passed Shapiro-Wilk normality tests with p-values of 0.880 and 0.266, respectively. In summary, there is significant evidence that the 3-D system performed better than the 2-D system in terms of errors using the classical method.

Note the distinction between the comparison made here and the comparison in the subsection "Further Investigation of Classical versus Inverse Calibration Methods" of [Section 3]. Before, we were comparing two lines, the least-square line and one derived from it, and we used the remaining 10 data points for reasons described previously. But in this section, the comparison is between two OVE systems both using the classical method. Therefore, the 20 pieces oyster data used for establishing the calibration models may be utilized for variance comparison.

An error plot is useful for visually comparing the precision between the two systems, and it is more effective than the scatter plots showed earlier. In Figure 8, the errors from the 3-D system fall within 1 cc of zero, whereas the errors from the 2-D system extend past 2 cc. These graphs indicate that the 3-D system has a narrower spread of errors.
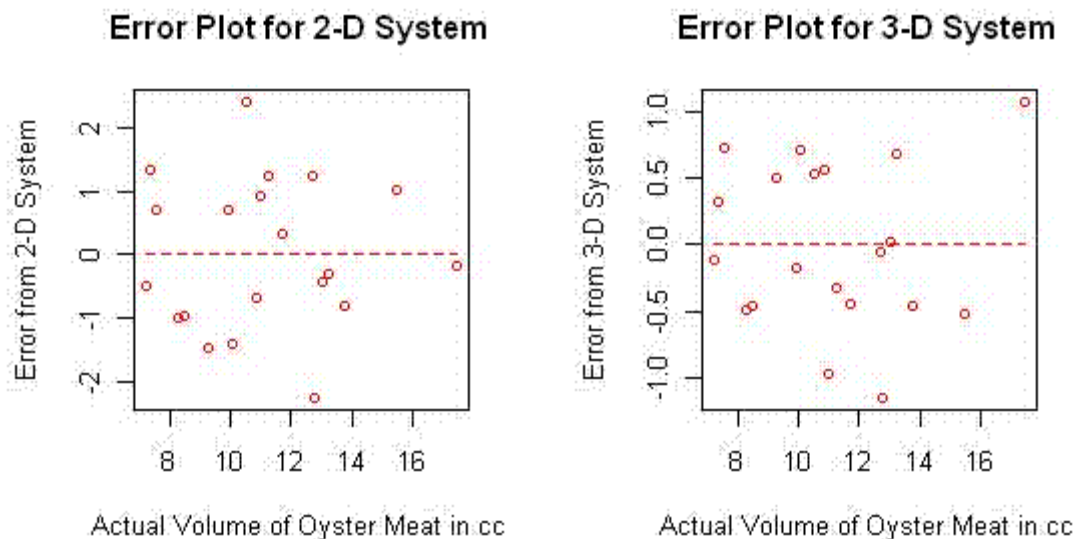
Figure 8. Error Plots from 2-D and 3-D Systems.

A common mistake that students have made, when no guidance was given, was the use of the coefficient of variation (CV) to compare the two systems. Because the pixel counts from the two systems are substantially different in scale (5-digit numbers for 2-D and 7-digit numbers for 3-D), students often have thought that they could just use the CV as a relative measure of system spread. Indeed the CV may be helpful if only one oyster were repeatedly measured. However, since oysters of multiple sizes were used in this experiment, the CV is not a proper statistic for system comparison.

Using the Hit Ratio

The main objective for designing the OVE system was to properly classify the oysters for packaging. The percentage of correct classification (also called the hit ratio) is a popular statistic for comparing classification methods. In this project, the OVE system was designed to be integrated with an automatic sorting machine. The 2-D and 3-D systems can be considered as two different classification techniques for classifying oysters. If one uses an oyster grading scale such as that listed in Table 2, after an oyster's size in cc is determined by the OVE system, the sorting machine will place it into one of the three bins: large, medium, or small, according to its size. For instance, an oyster larger than 13 cc would be classified as a large oyster. A good sorting system should have a high rate of correct classification (that is, a high hit ratio).

Please see Table 3 for a hit ratio comparison of the 2-D and 3-D systems, using the classical method. The rows of the table list the ID, the actual volume, and the estimated volume for each observation in the sub-sample of 20 oysters. The "*" in front of an estimate in Table 3 indicates a misclassification. In the bottom row we see that the 2-D system had a hit ratio of 80%, while the 3-D system had a hit ratio of 75%.

One may ask, "how could the 2-D system have a higher hit ratio if it has lower precision (larger error variance)?" One possible reason is that a small sample increases the sampling variability which decreases the accuracy of the model. Another possible reason relates to the general recommendation that the observed data values for calibration should be spread evenly throughout the range of the data to achieve a better result. In the full data set for this study, there are 10 oysters from each of the three grading scale ranges. When a random sample of 20 is selected, there is potential for the size distribution to become uneven, which increases the probability of an unexpected result like this in hit ratios. Note that there are three oysters (oysters 5, 7, and 24, in Table 3) that are classified correctly by the 2-D method and incorrectly by the 3-D method. The 3-D estimates for two of these oysters (5 and 7) are actually closer to the correct value than the 2-D estimates. Thus the higher hit ratio for the 2-D method is likely just a fluke, arising because some of the volumes are close to the boundary between categories.

To drive this last point home, the reader is encouraged to try the above procedures with the full data set which constitutes a slightly larger sample and see what happens. A larger sample would reduce the sampling variation and provide better estimates from both systems. We would expect the 3-D system to have a higher hit ratio if the same number of oysters from each oyster grade were used. It would be interesting to examine the hit ratios from using different samples of size 20, randomly selected from the full data set, for understanding the sampling variability of the hit ratio statistic. This example of using a sub-sample of 20 oysters highlights the possible problems in using the hit ratio in a comparative study.

Table 2. Oyster Grading Scale.

Oyster Grade	Actual Volume	
Large	(13, ∞)	Above 13 cc
Medium	(10, 13]	Between 10 to 13 cc
Small	(0, 10]	10 cc or below

Table 3. Hit Ratios for the 2-D and 3-D Oyster Volume Estimation Systems.

Oyster ID	Oyster Actual Volume	2-D System	3-D System
1	13.04	13.45	13.01
2	11.71	11.39	12.15
3	17.42	17.61	16.34
4	7.23	7.71	7.35
5	10.03	11.44	*9.31
7	9.94	9.23	*10.11
8	7.53	6.84	6.80
9	12.73	*14.99	*13.88
10	12.66	11.40	12.71
11	10.53	*8.11	*9.99
12	10.84	11.52	10.27
14	8.48	9.43	8.94
18	15.44	14.42	15.96
20	8.26	9.26	8.75
21	10.95	10.01	11.93
23	7.34	5.99	7.02
24	13.21	13.51	*12.53
27	11.22	*9.97	11.55
28	9.25	*10.71	8.75
29	13.75	14.56	14.21
Hit Ratio	-	80%	75%
* Not in the same class with the observed oyster.			

5. A Sample Project Assignment

The following is a sample project assignment that can be used in a statistics class right after covering the least-square estimates in simple linear regression. The "and/or" in the assignment description is for making the document more flexible to use for different formats that instructors would like to follow based on their teaching preferences or time constraints.

Two computer vision metrology systems were developed by engineers. One system uses a 3-D estimation method and the other system uses a 2-D estimation method. The data were recorded in the following ASCII file: 30oysters.dat

In the data file, the first column is the oyster ID, the 2nd column is the weight of the oyster meat in grams (g), the 3rd column is actual volume of the oyster meat in cubic centimeters (cc), the 4th column is the number of pixels recorded for the observed oyster using the 3-D method, and the 5th column is the number of pixels recorded for the observed oyster using the 2-D method. Before the system can be used for sorting oysters, a calibration process needs to be performed to obtain a formula for using the pixel data from the vision system to estimate the volume of the oyster meat.

The goal of this project is to use the linear regression technique to model the relationship between the actual oyster meat volume and the number of pixels recorded from the 2-D (and/or 3-D) imaging system, and to build a model for oyster volume estimation (OVE) and oyster sorting.

1. Make a scatter plot to examine the correlation between the actual volume of the oyster meat and the number of pixels for the oyster in the image taken from the 2-D (and/or 3-D) metrology system, and describe the relationship between the two quantitative variables.
2. If the relationship looks linear, report the Pearson's correlation coefficient between actual oyster meat volume and number of pixels obtained from the 2-D (and/or 3-D) metrology system. Is there a significant correlation between the actual oyster volume and the number of pixels for the oyster in the image taken from the 2-D (and/or 3-D) metrology system? Perform a **zero-correlation test** to see if the correlation is significantly different from zero, at 5% level of significance.
3. Build a simple linear regression equation to model the relationship between the actual oyster meat volume and the number of pixels recorded from the 2-D (and/or 3-D) system by, first, **determining the dependent and the independent variables**, and then **performing simple linear regression** and **writing the regression equation**, and then **interpreting the results and the meaning of the regression coefficients**.
4. Establish the oyster meat **volume estimation equation** (calibration equation) using the regression result above for the 2-D (and/or 3-D) volume estimation system.
5. If an oyster is measured by the 2-D (and/or 3-D) system and the number of pixels given by the system is 58,000 (and/or 5,800,000 from 3-D) what would be the **expected volume of this oyster using the volume estimation system**?
6. A grading scale for the oysters is defined in the following table. Conclude the analysis by examining and commenting on the performance of the system using the **percent of the oysters in this study that would be correctly classified using the model(s) above**. (Based on the actual volume information in these data, there were 10 small, 10 medium, and 10 large oysters, according to the oyster grading system in the following table. If you were to use the data provided in this project, in which the actual volumes are already known, to estimate the volume, what percentage of the oysters would be correctly classified?)

Oyster Grade	Actual Volume	
Large	(13, ∞)	Above 13 cc
Medium	(10, 13]	Between 10 to 13 cc
Small	(0, 10]	10 cc or below

6. Summary

Calibration is an important and practical application of simple linear regression, but is rarely seen in standard introductory statistics textbooks. Yet it is a simple technique and is quite useful in many science and engineering applications.

In this paper, a subset of 20 oysters (taken from a data set of 30) was used to show how to establish calibration equations and to illustrate some of the possible analyses that can be done for understanding the performance of 2-D and 3-D OVE systems. The data file also contains the actual weights of the oysters for additional explorations.

Instructors can use the full data set, a subset, or even several sub-samples for lecture or assignment purposes. When using the classical method for the full data set, the hit ratio from the 3-D system turns out to be higher than the hit ratio from the 2-D system. But as we have seen in this article, when using a sub-sample, the hit ratio from the 3-D system is not always higher than the hit ratio from the 2-D system perhaps due to a smaller sample size or possibly an uneven distribution in oyster sizes.

We strongly advocate giving a calibration exercise as one of the assignments for simple linear regression or using it as an example in a regular lecture. If possible, instructors could also discuss issues related to the design of calibration experiments and suggest further improvements for the study.

7. Getting the Data

The file 30oysters.dat is a text file containing the raw data, and the file 20oysters.dat is the sub-sample used for explaining several analysis techniques in the article. The file 30oysters.txt is a documentation file containing a brief description of the dataset.

Appendix: Triangulation Technique for Thickness Data

Using a digital camera and a multi-line laser (Figure A1), the oyster thickness information can be obtained using a triangulation technique. To implement the triangulation technique, a laser light beam is set to aim at an oyster with a particular angle q. When there is no object on the table, the light beam points at the location o which is directly underneath the camera positioned as shown in Figure A2. When an object is placed on the table, the horizontal displacement d of the point o from its original location is recorded by the digital camera, and the thickness measure h can be calculated using trigonometry.

In the actual process, thickness information from various locations of the object can be collected in real time using the triangulation method with system settings as in Figure A1. By combining the thickness information with the 2-D data acquired in real time, the 3-D volume estimation of the object in pixels can be obtained using mathematical interpolation or estimation techniques. The data provided in this article were generated from such a system.
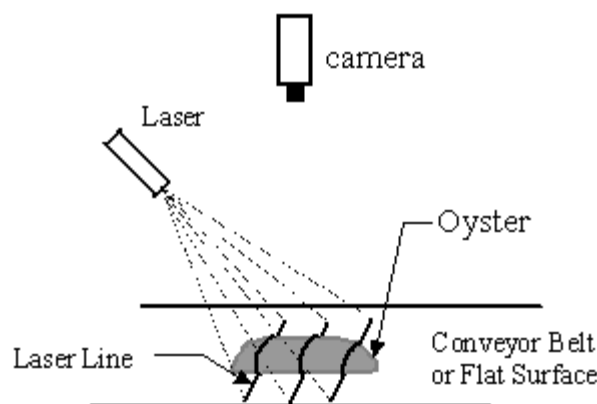

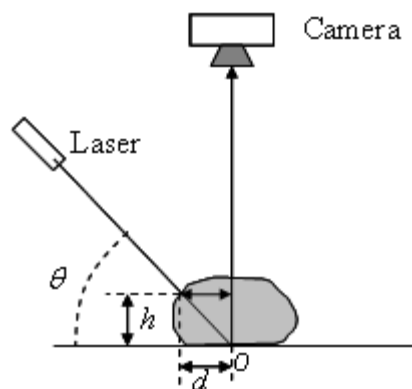
Figure A1. Optical System Design

Figure A2. Triangulation Technique for Obtaining Oyster Thickness (*h*) Data.

Acknowledgments

References

Berkson, J. (1969). "Estimation of a linear function for a calibration line: consideration of a recent proposal," *Technometrics,* 11, 649-660.

Brase, C. H. and Brase, C. P. (2009), *Understandable Statistics*, 9th edition, Houghton Mifflin.

Brown, P. J. (1993), *Measurement, Regression and Calibration*, Clarendon Press, Oxford.

De Veaux, R., Velleman, P., and Bock, D. (2006), *Intro Stats*, 2nd edition, Pearson.

DiCiccio, T. J., and Efron, B. (1996) "Bootstrap Confidence Intervals," *Statistical Science*, 11(3), 189-228.

Draper, N. R., and Smith, H. (1998), *Applied Regression Analysis*, Wiley.

Eisenhart, C. (1939) "The Interpretation of Certain Regression Methods and Their Use in Biological and Industrial Research," *The Annals of Mathematical Statistic*s, 10, 162-186.

Fox, J. (2005) "The R Commander: A Basic-Statistics Graphical User Interface for R," *Journal of Statistical Software*, 11(9).

Graybill, F. (1976), *Theory and Application of the Linear Model*, Duxbury.

Krutchkoff, R. G. (1967), "Classical and Inverse Regression Methods of Calibration," *Technometrics*, 9, 425-40.

Krutchkoff, R. G. (1971), "The Calibration Problems and Closeness," Journal of Statistical Computation and Simulation. 1, 7-95.

Matson, J., and Huguenard, B. (2007), "Evaluating Aptness of a Regression Model," *Journal of Statistics Education*, 15(2). http://www.amstat.org/publications/jse/v15n2/datasets.matson.html.

Osborne, C. (1991), "Statistical Calibration: A Review," *International Statistical Review*, 59(3), 309-336.

Ott, L., and Longnecker, M. (2001), *Introduction to Statistical Methods and Data Analysis*, Duxbury.

Parr, M., Byler, R., Diehl, K., and Hackney, C. (1994), "Machine Vision Based Oyster Meat Grading and Sorting Machine," *Journal of Aquatic Food Product Technology*. 3(4), 5-24.

Peck, R. and Devore, J. (2008), *Statistics: The Exploration and Analysis of Data*, 6th edition, Duxbury.

Shukla, G. K. (1972), "On the Problem of Calibration," *Technometrics*, 14(3), 547-553.

G. Andy Chang
Department of Mathematics and Statistics
Youngstown State University
One University Plaza
Youngstown, OH 44555
gchang@ysu.edu

G. Jay Kerns
Department of Mathematics and Statistics
Youngstown State University
One University Plaza
Youngstown, OH 44555
gkerns@ysu.edu

D. J. Lee
Department of Electrical and Computer Engineering
Brigham Young University, 459 CB
Provo, Utah 84602
djlee@ee.byu.edu

Gary L. Stanek
Department of Mathematics and Statistics
Youngstown State University
One University Plaza
Youngstown, OH 44555
stanek@math.ysu.edu

Volume 17 (2009) | Archive | Index | Data Archive | Resources | Editorial Board | Guidelines for Authors | Guidelines for Data

http://www.amstat.org/publications/jse/v17n2/datasets.changpdf.html[8/25/2009 11:35:25 AM]

Journal of Statistics Education, v17n2: G. Andy Chang, G. Jay Kerns, D. J. Lee, and Gary L. Stanek