film (Kodak); (d) wash in running water for 4 min; and (e) dry with an air dryer.

V. CONCLUSION

The experiment we have described is simple, inexpensive, and effective. It can be easily done without a glass cylindrical lens. It needs only a beaker or a bottle of water, over which was placed a 2-mm-wide slit between the object and film. With this simple setup we can make rainbow holograms that give clear, bright, astigmatic, and large fields

of view to the reconstructed image with a white light point source.

¹⁾ On leave from Physics Department, Tianjin Normal University, Tianjin, People's Republic of China.

²⁾ S. A. Benton, *J. Opt. Soc. Am.* **59**, 1545 (1969).

³⁾ H. Chen, A. Tai, and F. T. S. Yu, *Appl. Opt.* **17**, 1490 (1978).

⁴⁾ H. Chen, *Appl. Opt.* **17**, 3290 (1978).

⁵⁾ H. Chen, Qi-Zhe Shan, and Ming-Yi Chen, *Appl. Opt.* **20**, 3557 (1981).

⁶⁾ H. Chen and Ricardo Pastor, Jr., *Am. J. Phys.* **55**, 623 (1987).

⁷⁾ Francis A. Jenkins and Harry E. White, *Fundamentals of Optics* (McGraw-Hill, New York, 1976), pp. 81-85.

⁸⁾ J. W. Huang, *Am. J. Phys.* **46**, 737 (1978).

Linear regression analysis in a first physics lab

Lawrence Hmurcik,^{a)} Amy Slacik, Hilda Miller, and Sandy Samoncik
Physics Department, University of Bridgeport, Bridgeport, Connecticut 06601

(Received 24 July 1987; accepted for publication 22 March 1988)

Linear regression analysis (least squares) is used in the first physics lab in order to introduce students to computer-aided analysis and to teach data fitting techniques. Application is made to two experiments: Fletcher's trolley and Hooke's law. Least squares will extract information from raw data in a very precise way, and it opens the way for the study of more complicated phenomena than a first lab usually covers. Students who learn least squares in the first semester do much better in future labs, including nonphysics labs.

The main purpose of the first freshman physics lab is to teach students good lab practices, including data analysis. In addition to the usual techniques of data analysis, we introduce linear regression analysis (least squares) in a first lab for physics courses that may be either calculus or noncalculus based. We apply least squares to a series of experiments that serve to teach data fitting via the computer and that allow for the study of more complicated physical phenomena than a freshman lab usually covers.

Using a computer is not a substitute for basic data analysis. It is disturbing that many students who use the computer cannot explain or replicate their work longhand. We insist that, before students use the computer, they plot their data and determine by eyesight if all or part of it can be fit by a straight line. By eyesight, they draw the "best" line through this data, they find the slope m as rise over run, they find the intercept b as the point of intersection of this line and the y axis, and they guess at a correlation coefficient R^2 based on how their data scatter around the line. Here, R^2 represents the confidence that the measured data fit a linear model.¹ If all the data fall on a single line, then $R^2 = 1$, and the confidence that the data fit a linear model is 100%. If $R^2 = 0.70$, the confidence is 70%, and the data scatter around a straight line. As R^2 decreases, the scatter increases.

When the graphical analysis is complete, the data are entered as N pairs of (x, y) data into a computer program that calculates the "best" linear equation as^{1,2}

$$y = mx + b, \quad (1)$$

where

$$m = \frac{\Sigma(xy) - 1/N(\Sigma x)(\Sigma y)}{\Sigma(x^2) - 1/N(\Sigma x)^2}, \quad (2)$$

$$b = 1/N(\Sigma y - m\Sigma x), \quad (3)$$

$$\sigma_b = \sqrt{\frac{\Sigma(y^2) - b\Sigma y - m\Sigma(xy)}{N - 2}}, \quad (4)$$

$$\sigma_m = \frac{\sigma_b}{\sqrt{\Sigma(x^2) - 1/N(\Sigma x)^2}}, \quad (5)$$

and

$$R^2 = \frac{b\Sigma y + m\Sigma(xy) - 1/N(\Sigma y)^2}{\Sigma(y^2) - 1/N(\Sigma y)^2}, \quad (6)$$

with σ as the standard error and with Σx as shorthand for

$$\sum_{i=1}^N x_i,$$

etc. Values for m , b , and R^2 obtained by eye must be equal to those obtained by computer to within one or two significant digits. For example, if $m(\text{graph}) = 2$ and $m(\text{computer}) = 2.35$, then the computer value, which is more accurate, is used for further analysis. If, on the other hand, $m(\text{graph}) = 2$ and $m(\text{computer}) = 17.14$, then both analyses must be rechecked. Either the slope was calculated incorrectly from the graph, or (more likely) the wrong data were fed into the computer.

At the first lab meeting, important aspects of error analysis are covered, such as standard deviation and significant digits. In addition, we give students two "canned" data sets like those shown in Table I. For this lab only, they do least

Table I. Practice data sets used for introductory work on least squares analysis. First data set is extremely linear but hard to calculate by hand. (Slope = 0.134 ± 0.009 ; intercept = 7.96 ± 0.04 ; $R^2 = 98.2\%$). Second set is easy to calculate but not very linear (Slope = 0.371 ± 0.484 ; intercept = 1.87 ± 2.03 ; $R^2 = 12.8\%$).

x	y	x	y
-6.00	7.16	1.00	1.00
-5.00	7.25	2.00	5.00
-4.00	7.43	3.00	2.00
-3.00	7.61	4.00	2.00
-2.00	7.70	5.00	6.00
-1.00	7.79	6.00	3.00

squares on these data sets in three different ways: by using a calculator, a computer, and a graph. For all remaining labs, just graphical and computer analysis are required.

The first data set is very linear. The data are very sensitive to numerical error. When students use a calculator on these data, invariably they get wrong results and must repeat their calculations. With m small, Eqs. (3)–(6) are sensitive to small variations in the data. About six or seven significant digits must be retained through all calculations in order to get the correct results. Only the computer can do this with facility. The second data set, on the other hand, is a poor approximation to linearity. At any point in the calculation, round-off to three significant digits will still produce good results in the end. However, the second data set is more difficult to analyze by graph. Four of the data pairs fall close to a straight line, but the remaining two scatter far from linearity. It requires a greater conceptual effort to do graphical analysis. Generally, the student is forced to perform the computer analysis first. The best-fit line is drawn on the graph by reconciling the scatter on the graph with the computer results. This reinforces the idea that the graph and the computer are *interdependent* tools of analysis.

During the remainder of the course, least squares analysis is used to analyze four of the ten experiments given in the first semester. These are discussed further in Secs. II and III. Experimental details are omitted but can be found in a number of sources.^{3,4}

II. FLETCHER'S TROLLEY

Fletcher's trolley can be used to demonstrate the validity of Newton's second law for linear motion. A cart of mass M_1 rides on a level air track. Mylar tape, which is attached to the cart, passes over a peg at the end of the track. Various masses (M_2) are attached to the other end of the tape, and these cause the cart to accelerate from rest. There are two sources of friction: M_1 riding on the air track and the mylar tape sliding over the peg. Since M_2 is usually much smaller than M_1 , the second source is negligible or, at worst, it causes only small changes in friction for differing values of M_2 . Then

$$(M_1 + M_2)a = M_2g - f, \quad (7)$$

where f is the force of friction acting on the entire system. The acceleration equals $(2s/t^2)$, where s is the distance of travel in time t . By letting $y = (M_1 + M_2)a$ and $x = M_2$, we find $m = g$, $b = -f$, $\sigma_m = \sigma_g$, and $\sigma_b = \sigma_f$. See Fig. 1 and Table II. Graphical analysis and computer analysis are

Fig. 1. A cart of mass M_1 is accelerated by the weight of mass M_2 . A plot of $(M_1 + M_2)a$ vs M_2 produces a straight line with g as the slope and the force of friction as the intercept.

in good agreement. The value 9.93 ± 0.77 encompasses the accepted value of g , i.e., 9.81 m/s^2 . Since Newton's second law is used to formulate Eq. (7), a large confidence (97%) in (7) verifies the application of Newton's law in this system.

The error in friction is greater than the value of the friction itself, but this is expected. Each (x, y) data pair is a function of g and f and the masses used. The masses are measured to great accuracy, with an error of 0.1 g in the worst case. The gravitational field is extremely uniform and does not significantly contribute to measurement errors for each data pair. Friction, however, is an extremely unstable parameter that can vary greatly even if two measurements are made for the same value of M_2 . By plotting data in the form of Eq. (7), differing values of g (slope) are averaged over the entire range of the graph. Since the scatter in data is small, data points can be shifted slightly, and the average slope would change very little. However, friction is only determined at one point. Its value is extremely small. A small shift in data can cause a great change in the value of the y intercept, i.e., friction. Numerical results show the friction to be indeterminate.

The data used to plot Fig. 1 can be used to find friction if the focus of the data is changed. If f equals the slope, then the friction is averaged over the entire range of data as was done for g in Fig. 1. First, rewrite (7) as

$$a(1 + M_1/M_2) = -(1/M_2)f + g. \quad (8)$$

Let $y = a(1 + M_1/M_2)$ and $x = (-1/M_2)$. Then $m = f$, $\sigma_m = \sigma_f$, $b = g$, and $\sigma_b = \sigma_g$. See Fig. 2. Confidence in Eq. (8) is 66%. This would indicate a decreased confidence in the application of Newton's law. Also, the value of g is

Table II. Computer-generated least squares parameters for the data sets plotted in each figure.

Figure	$m + \sigma_m$	$b + \sigma_b$	$R^2(\%)$
1	9.93 ± 0.77	0.0115 ± 0.0203	97.1
2	0.00874 ± 0.00279	9.33 ± 0.41	66.3
3	17.4 ± 0.1	0.0370 ± 0.0400	99.9
4	18.4 ± 0.9	0.00584 ± 0.00885	98.1
5	-10.2 ± 1.1	19.0 ± 0.1	96.5

Fig. 2. Data in Fig. 1 are replotted as $(1 + M_1/M_2)a$ vs $-1/M_2$. Scatter in data is larger, but friction can be found in an unambiguous way.

farther from the accepted value than it was for Eq. (7). Nevertheless, the value of friction has been averaged in an unambiguous way. Variations of the slope (friction) from point to point are large, but these produce an unambiguous range of values with $f = (0.00874 \pm 0.00279)$ N.

Using a different air track, a second set of data was collected. In this case, $f = (0.0171 \pm 0.2405)$ N and $R^2 = 99.6\%$ by using Eq. (7), and $f = (0.00463 \pm 0.00230)$ N and $R^2 = 40.3\%$ by using (8). These results reaffirm the previous ones.

A modified form of Eq. (7) can be applied to an accelerating disk. The frictional torque of the turning mechanism cones from the y intercept, when the applied torque (y axis) is plotted versus the circular acceleration. Values of frictional torque and moment of inertia for a standard lab setup (i.e., using a 2-kg disk) are 0.00522 ± 0.00019 N m and 0.0338 ± 0.0004 kg m² with a confidence of 99.9%. This high confidence would support the application of Newton's second law to circular motion.

III. HOOKE'S LAW

A spring is the simplest physical system that produces a force that varies with distance. According to Hooke's law, the displacement s of a spring is proportional to the force applied. If M is the mass applied and if the spring has a force constant k , then

$$Mg = ks. \quad (9)$$

Using least squares, students study two aspects of the spring system: (a) simple harmonic motion and (b) non-linear displacement.

First, consider that (9) applies to a spring at rest. If M is displaced from equilibrium and released, then the frequency of oscillation (f) is given by

$$M = k / (4\pi^2 f^2) - M_0, \quad (10)$$

where M_0 is that portion of the spring's mass that oscillates with M . Theoretically, M_0 is $\frac{1}{3}$ of the total spring mass. In (9), let $y = Mg$ and $x = s$. Then $m = k$, $\sigma_m = \sigma_k$, and $b = 0$. These results are shown in Fig. 3. Using (10), let $y = M$ and $x = (1/4\pi^2 f^2)$. Then $m = k$, $\sigma_m = \sigma_k$, $b = -M_0$, $\sigma_b = \sigma_{M_0}$. See Fig. 4. Values of m , b , and R^2 determined by eye (graphs) agree well with the computer results (Table II).

Fig. 3. Force Mg versus displacement s for a spring shows a linear variation as predicted by Hooke's law. Slope gives the value of the spring constant.

Since the confidence (R^2) is high for both Eqs. (9) and (10), Hooke's law and the model of simple harmonic motion are good descriptions of the physical behavior of the spring. Also, both k values, found for the spring at rest or in motion, are the same within two standard errors (2σ). For a normal, random distribution, the "true" value of a parameter has a 95% chance of lying in the range of the mean value $\pm 2\sigma$. Since the ranges in k values found by (9) and (10) clearly overlap in 2σ , the two methods of finding k are physically equivalent.

In Eq. (9), the intercept should be zero. Least squares shows a nonzero intercept (0.0370). However, the error in this value (0.0400) is greater than the value itself. Thus the expected value (i.e., zero) falls into this range. Also, the range of the intercept found using (10) encompasses the expected value of $(M_0/3)$, i.e., 0.00290 kg.

Second, Hooke's law applies to a great variety of spring systems. However, when the force applied to a spring is too large or too small, Hooke's law breaks down. For too large

Fig. 4. For a spring in simple harmonic motion, the mass attached to the spring M varies with the frequency of oscillation as $1/f^2$. Slope equals the spring constant, while the intercept equals $\frac{1}{3}$ of the spring's mass.

Fig. 5. Data in Fig. 3 are replotted to show quadratic and higher-order variations of force Mg versus displacement s in a spring. For $s \geq 0.14$ m, a linear variation indicates a Hooke's law dependence. For $0.05 \text{ m} \leq s < 0.14$ m, a quadratic dependence is found. Below 0.05 m, cubic and higher-order variations can be fit to the data.

a force, the spring is plastic rather than elastic. It stretches but does not return to equilibrium when the force is removed. We do not consider this case here.

When a small force is applied to a spring in the static case, the displacement varies with force in a nonlinear manner. The elastic stress in a metal spring is anisotropic.⁵ Initially, the spring stretches a little and the orientation of each wrap of the spring changes greatly. The effective spring constant changes with displacement. Equation (9) can be modified to account for spring anisotropy as

$$Mg = ks + F(s), \quad (11)$$

where F is a nonlinear function. In the simplest case, F can be approximated as a quadratic function of s . Then

$$Mg = ks + (k_1 s + k_2 s^2) \quad (12)$$

and

$$(Mg/s) = k_2 s + (k + k_1). \quad (13)$$

The data in Fig. 3 were fit to the linear model of a spring, i.e., Hooke's law. The linear model is a first-order variation. Variations in the data from linearity (and not just random scatter) are "smeared out" by this linear fit. Using (13), we can see the second-order variation in these data. Let $y = (Mg/s)$ and $x = s$. This is plotted in Fig. 5. Three distinct regions become noticeable. For $s \geq 0.14$ m, $k_1 \approx k_2 \approx 0$, and $k = 17.5$ is found as the intercept of a straight line fit to these data. This value of k is within the ranges of k found previously. Physically, the spring has stretched so far that anisotropic variations in its stretching are negligible. For s less than 0.14 m but greater than 0.05 m, the stretching action can be fit to a quadratic variation, which obeys Eq. (13). These give us $k_2 = -10.2 \pm 1.1$ and $(k_1 + k) = 19.0 \pm 0.1$. For $s \leq 0.05$ m, a cubic or higher-order fit can be made to the data. We did not pursue this, since error in the measurement of s (0.0005 m) has a large effect on the data for $s < 0.05$ m. Another set of measurements was made for a second spring of the same size. The results give $k = 26.4$, $k_2 = -142$, and $k_1 = 21.8$ with $R^2 = 91\%$.

The data used to plot Figs. 3 and 5 are the same. Figure 3 "smears" out all but linear variations of force with displacement. Figure 5 shows quadratic and higher variations. The mixing of data in Fig. 5, which are partially linear and partially quadratic, is a common statistical process.⁶ It applies to many physical systems. The point we wish to emphasize is that not only is the variation of data (y vs x) important, but also the variation of this variation (dy/dx vs x) may contain a great deal of additional information about the system studied. In the former, the system (spring) is isotropic. In the latter, it is anisotropic. Also, with the spring system, the measurement of s is so precise that we can see the second pattern described in Fig. 5. Data are not always this precise. For example, in the case of Fletcher's trolley, random scatter would hide any nonlinear variations in the data.

IV. CONCLUSION

We have used linear regression analysis in the first physics lab. The experiments that deal with basic mechanics are conceptually easy to understand. Attention is paid to the successful application of computer and graphical analysis to data that vary linearly. Experiments in the second and later physics labs cover topics from electricity, optics, thermodynamics, and modern physics. For the student, these labs are harder to visualize, or they require greater precision to obtain good results. As an example, consider Ohm's law ($V = IR$). A simple setup has many wires and meters, which can baffle a student who has never hooked up an electric circuit before. To teach least squares analysis and circuit analysis at the same time is too much for most students to absorb properly. However, students who have used least squares in their first semester have a much easier job. Once they collect data of current I versus voltage V , they can plot these and analyze them via computer to extract the slope (resistance), intercept (generally zero), and confidence in linearity.

One additional advantage of teaching least squares in the first semester was not anticipated. With the popularity of personal computers in home and school, we assumed that students would be computer literate in the first semester. Many were not, and those who had computer experience lacked the necessary analytical skills to use the computer for lab analysis. By doing least squares on the computer in this first semester, students became skilled at computer analysis, and they were able to handle more difficult computer projects in future labs.

ACKNOWLEDGMENTS

We wish to thank Richard W. Jenner and Leslie Klabatz for their contributions to the development of this article.

¹ Present address: Electrical Engineering Department, University of Bridgeport, Bridgeport, CT 06601.

² N. C. Barford, *Experimental Measurements: Precision, Error, and Truth* (Wiley, New York, 1985), p. 60.

³ P. Lancaster and K. Salkauskas, *Curve and Surface Fitting* (Academic, New York, 1986), p. 49.

⁴ D. S. Edmonds, *Cioffari's Experiments in College Physics* (Heath, Lexington, MA, 1983).

⁵ J. D. Wilson, *Physics Laboratory Experiments* (Heath, Lexington, MA, 1986).

⁶ L. H. van Vlack, *Elements of Materials Science and Engineering* (Addison-Wesley, Reading, MA, 1985), p. 172.

⁷ Reference 2, p. 87.