
Christine M. Hansen,* Marc A. Evans† and Terry D. Shultz*2

*Department of Food Science and Human Nutrition and the †Program in Statistics, Washington State University, Pullman, WA 99164-6376

ABSTRACT The bootstrap procedure is a versatile statistical tool for the estimation of standard errors and confidence intervals. It is useful when standard statistical methods are not available or are poorly behaved, e.g., for nonlinear functions or when assumptions of a statistical model have been violated. Inverse regression estimation is an example of a statistical tool with a wide application in human nutrition. In a recent study, inverse regression was used to estimate the vitamin B-6 requirement of young women. In the present statistical application, both standard statistical methods and the bootstrap technique were used to estimate the mean vitamin B-6 requirement, standard errors and 95% confidence intervals for the mean. The bootstrap procedure produced standard error estimates and confidence intervals that were similar to those calculated by using standard statistical estimators. In a Monte Carlo simulation exploring the behavior of the inverse regression estimators, bootstrap standard errors were found to be nearly unbiased, even when the basic assumptions of the regression model were violated. On the other hand, the standard asymptotic estimator was found to behave well when the assumptions of the regression model were met, but behaved poorly when the assumptions were violated. In human metabolic studies, which are often restricted to small sample sizes, or when statistical methods are not available or are poorly behaved, bootstrap estimates for calculating standard errors and confidence intervals may be preferred. Investigators in human nutrition may find that the bootstrap procedure is superior to standard statistical procedures in cases similar to the examples presented in this paper. J. Nutr. 129: 1915–1919, 1999.

KEY WORDS: • Vitamin B-6 requirement • statistics • inverse regression • bootstrap

Statistical methods have become the foundation for objective decision making in much of the scientific world. ANOVA, regression analysis, chi-square tests for independence, t-tests and so on are the champions of statistical inference. Although it is our belief that these utilitarian approaches will continue to dominate the paths to sound statistical inference, several new methodologies are currently available but are relatively unknown to the scientific community. One of the methods to which we are referring is the bootstrap (Efron 1979). This procedure is of recent origin, having arisen, for the most part, because of the advent of powerful computers.

The bootstrap is a statistically elegant procedure. It can be applied in situations where standard statistical tools do not exist or in situations where the usual statistical methods are inappropriate because the underlying assumptions are violated (e.g., ANOVA and regression analysis have assumptions of normality and constant variance). This violation may cause the inferences based on the usual methods of analysis to lead to spurious conclusions. On the other hand, the bootstrap procedure can produce valid inferences for ANOVA and regression analysis when the typical assumptions are violated. Although the bootstrap would appear useful in those aforementioned situations, this is not the principal area of application. The bootstrap procedure is also applicable in those situations where analytic statistical methods are not readily available.

This paper introduces the bootstrap procedure to compute standard errors and confidence intervals for the estimation of the vitamin B-6 requirement by using inverse regression. Bootstrap estimates of the mean, standard error of the mean and confidence intervals will be compared to estimates based on formulas commonly found in statistics textbooks. Furthermore, through Monte Carlo simulation of inverse regression the characteristics of bootstrap estimators compared to the most typically used estimators are explored.

MATERIALS AND METHODS

In a recent controlled-diet study (Huang et al. 1998), after a 9-d adjustment period (1.60 mg vitamin B-6/d), eight young women consumed a low vitamin B-6 diet (0.45 mg/d) for 27 d, followed by repletion with three levels of vitamin B-6 (1.26, 1.66 and 2.06 mg/d) for successive periods of 21, 21 and 14 d, respectively. Six measures of vitamin B-6 status were assessed at the end of each experimental period. One method of determining the vitamin B-6 requirement is to determine the intake at which status indicators are restored to their...
95% confidence limits for the mean were then calculated (the regression analysis is illustrated in inverse prediction or calibration (Kutner et al. 1996). An example of using inverse regression as follows:

the vitamin B-6 requirement for each of the six status indicators by statistics are listed in Appendix A. Equations used for the calculation of both classical and bootstrap

\[ y = \beta_0 + \beta_1 x + \epsilon \]

measure of B-6 status [vitamin B-6 intake estimated regression lines were then solved for vitamin B-6 intake by B-6 requirement, values for each of the six status indicators were baseline values after depletion. To provide estimates of the vitamin B-6 requirement, values for each of the six status indicators were regressed on vitamin B-6 intake by using linear regression, and the estimated regression lines were then solved for vitamin B-6 intake by using the mean from the end of the adjustment period for each measure of B-6 status [vitamin B-6 intake = (mean adjustment period value – intercept)/slope]. This type of regression analysis is known as inverse prediction or calibration (Kutner et al. 1996). An example of the regression analysis is illustrated in Fig. 1. The standard error of the predicted vitamin B-6 requirement was calculated by using an asymptotic estimator (Kutner et al. 1996). The weighted mean of the six estimates (by using the standard error as the weight factor) and 95% confidence limits for the mean were then calculated (Table 1). Equations used for the calculation of both classical and bootstrap statistics are listed in Appendix A. The multivariate bootstrap was also applied to the prediction of the vitamin B-6 requirement for each of the six status indicators by using inverse regression as follows:

1. Because there were measurements at four vitamin B-6 intakes for each of the eight subjects, our observed data consisted of 32 (x, y) pairs for each status indicator. For each vitamin B-6 status indicator, 2,000 independent samples of n = 32 pairs were selected, with replacement, from the observed data. These represent 2,000 bootstrap samples.

2. The simple linear regression equation \( Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \) was calculated for each of the 2,000 bootstrap samples and solved for vitamin B-6 intake \( X_{new} \) at the adjustment mean for that status indicator \( Y_{new} = \hat{X}_{new} = (Y_{new} - \beta_0)/\beta_1 \). (The simple linear regression model was appropriate because repeated observations on the same subject had been found to be uncorrelated.) The average of the 2,000 bootstrap estimates of the vitamin B-6 requirement is shown in Table 1 for each vitamin B-6 status indicator.

3. The bootstrap standard error and t-confidence interval were computed as shown in Appendix A.

4. The bootstrap percentile confidence interval was constructed as follows:

(a) The bootstrap estimates were ordered from smallest to largest.

(b) The 50th and the 1,950th values represent the lower and upper limits of the 95% confidence interval (0.025 × 200 and 0.975 × 200, respectively).

A program using the SAS® IML software (SAS Institute, Cary, NC) for bootstrap estimation of inverse regression standard errors and confidence intervals is provided in Appendix B. This program was used to compute the bootstrap estimation results displayed in Table 1 for the inverse regression example. Included with the estimates of the vitamin B-6 requirement are the estimated standard errors and the lower and upper limits for 95% confidence intervals based on both the asymptotic estimator and the bootstrap estimator.

As shown in Table 1, the standard and bootstrap methods produced inverse regression estimates and standard errors that were nearly identical. However, the bootstrap percentile confidence intervals were asymmetric, unlike the confidence intervals produced by the standard method. This asymmetry is likely the result of a skewness in the data and possibly in the distribution of the underlying population. Thus, the bootstrap confidence intervals may better reflect reality. Unfortunately, one cannot draw such a conclusion from a single experiment. Therefore, a Monte Carlo simulation was undertaken to assess the performance of the bootstrap estimator as applied to inverse regression. Although extensive statistical theory was developed for the bootstrap procedure (Efron and Tibshirani 1993), and the bootstrap procedure was shown to have outstanding performance characteristics, analytical assessment of the quality of bootstrap estimators is not available for all statistical estimators (e.g., inverse regression). Therefore, alternative means of assessing the quality of specific bootstrap applications are required. Monte Carlo simulation provides this alternative means of assessment.

Monte Carlo simulation (Manly 1991) is simple in nature: generate data under a specific set of conditions, compute the statistic(s) of interest and store the results. This process is repeated several thousand times to accumulate information concerning the characteristics of the statistic(s) in question. Because the conditions (i.e., mean, standard deviation, probability distribution) under which the data are generated are known, the properties of the statistic(s) can be compared to the known conditions.

To assess the inverse regression bootstrap estimators, response (Y) and predictor (X) data for a population of size 1,000 were generated under the simple linear regression model: \( Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \), with \( \beta_0 = 0; \beta_1 = 5; 0.0 \leq X_i \leq 50.0 \) in increments of 0.05. Data were generated under two different conditions for sample sizes of n = 10, 20, 40 and 80: (a) regression model with valid assumptions: \( \epsilon_i \) normally distributed random variable with mean 0.0 and standard deviation 5.0; (b) regression model with nonconstant variance: \( \epsilon_i \) normally distributed random variable with mean 0.0 and standard deviation 0.2 × \( X_i \). Specific values of the response \( Y_{new} \) were selected to calculate the inverse regression estimate of the predictor \( X_{new} \): \( Y_{new} = 3.125 \Rightarrow X_{new} = 6.25; Y_{new} = 62.50 \Rightarrow X_{new} = 12.50; Y_{new} = 125.00 \Rightarrow X_{new} = 25.00 \).

The two conditions described above are possible conditions under which linear regression data might occur in nature, i.e., from data that satisfy the usual regression model assumptions of normality and constant variance to data that would seriously violate one of these assumptions. Furthermore, the sample sizes considered range from small to relatively large, and the selected values of \( Y_{new} \) range across the first half of the distribution of \( Y_i \).

Figure 2 is a graphical representation of the population under the usual regression assumptions. This graph also depicts the true values of \( X_{new} \) given each value of \( Y_{new} \). For each of the 24 combinations of sample size, \( Y_{new} \) and distribution conditions, 1,000 samples were generated. The regression coefficient, \( X_{new} \), and the standard error of \( X_{new} \) were estimated for each of the 1,000 samples. For each sample, the 95% classical t-confidence interval for \( X_{new} \) was constructed and checked for coverage of the true value of \( X_{new} \). Also, for each sample, 1,000 bootstrap samples were drawn, and the bootstrap standard error for \( X_{new} \) was computed along with the 95% bootstrap t-confidence intervals and 95% bootstrap percentile confidence intervals. Again, the coverage of the confidence interval was checked against the true

\[ y = 0.0844 + 2.0051 x \]

\[ R^2 = 0.8535 \]
Lastly, the mean of the $X_{new}$ for the 1,000 samples was computed, as was the standard deviation of these 1,000 values. This last value is an empirically derived value for the true standard error of $X_{new}$.

## RESULTS

Table 1 shows the inverse regression estimates, standard error estimates and 95% confidence intervals for the vitamin B-6 requirement based on mean adjustment values for the six vitamin B-6 status indicators, by using standard statistical estimators and bootstrap estimators. The standard and bootstrap methods produced inverse regression estimates and standard errors that were nearly identical. However, the bootstrap percentile confidence intervals were asymmetric, unlike the confidence intervals produced by the standard method, which is likely the result of a skewness in the data and possibly in the distribution of the underlying population. Thus, the bootstrap confidence intervals may better reflect reality. The Monte Carlo simulation was undertaken to assess the precision of bootstrap estimates for inverse regression.

The results of the Monte Carlo simulation are presented in Tables 2, 3 and 4. Table 2 presents the simulated mean values for $X_{new}$. These results clearly show that the inverse regression estimator of $X_{new}$ has little or no bias, even when the regression assumptions are violated. Table 3 shows the empirically derived true standard error, the average estimated asymptotic standard error and the average bootstrap standard error for $Y_{new} = 62.50$. Except for a couple of instances, the bootstrap estimated standard error more closely estimated the true standard error than the asymptotic standard error. In fact, the bootstrap standard error estimate was essentially unbiased for sample sizes of 20 or more, whereas the asymptotic standard error was biased when the assumption of constant variance was violated. Table 4 presents the percent noncoverage of the 95% confidence intervals. These confidence intervals are based on the classical $t$-confidence interval, using the asymptotic standard error; the bootstrap $t$-confidence interval, using the bootstrap standard error; and the bootstrap percentile confidence interval. In general, when the regression assumptions are satisfied, the classical $t$-confidence interval, the bootstrap $t$-confidence interval and the bootstrap percentile confidence intervals cover the true value of $X_{new}$ at the nominal 95% rate (2.5% noncoverage above and below $X_{new}$). However, when the assumptions are violated (e.g., nonconstant variance), the classical $t$-confidence interval produces nearly 100% coverage. This would indicate that the confidence intervals constructed in this manner were excessively broad. Both the bootstrap $t$-confidence interval and the percentile interval came closer to 95% coverage (5% noncoverage) for most conditions of sample size and distribution. Similar results were obtained for $Y_{new} = 31.25$ and 125.00 (data not reported here).

## DISCUSSION

The bootstrap procedure is a computationally intensive statistical tool used in situations where standard techniques for estimating...
Evidence of the superiority of one estimator over another can only be obtained through analytic, mathematical methods that are generally not available or through Monte Carlo computer simulation. The simulation presented in this paper for inverse regression clearly shows the superiority of the bootstrap procedure compared to the asymptotic methods presented in statistics textbooks (Kutner et al. 1996), when the standard assumptions are violated. The mean of the estimated standard error based on the bootstrap procedure is, in nearly all simulations, either of the same magnitude or smaller and closer to the true standard error than the mean of the estimated standard error based on the asymptotic procedure. Furthermore, the bootstrap procedure tends to produce confidence intervals that are either of the same nominal 95% coverage rate as the standard methods or closer to the nominal 95% coverage rate. The lack of confidence interval coverage for a 95% confidence interval should be 2.5% in the region below the confidence interval and 2.5% in the region above the confidence interval. In the inverse regression problem, the classical t-confidence interval based on the asymptotic standard error is conservative when the assumption of constant variance has been violated.

In summary, the simple beauty of the bootstrap procedure is that it can be applied in situations where point estimators, but not the methods for computing standard errors or confidence intervals, have been developed. The bootstrap is a procedure of relatively recent origin; its development was largely based on the availability of high-speed computers. Because of this recent derivation, the bootstrap has not found its way into all scientific disciplines. This conclusion appears to be true in the area of human nutrition, and this paper is an attempt to introduce the bootstrap within this discipline.

### TABLE 2

Monte Carlo simulated mean inverse regression estimates for true \( X_{\text{new}} = 6.25, 12.50 \) and 25.00, based on \( Y_{\text{new}} = 31.25, 62.50 \) and 125.00, respectively

<table>
<thead>
<tr>
<th>Conditions of the simulation</th>
<th>Normality, constant variance</th>
<th>Normality, nonconstant variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size, n</td>
<td>( X_{\text{new}} = 6.25 )</td>
<td>( X_{\text{new}} = 12.50 )</td>
</tr>
<tr>
<td>10</td>
<td>6.221</td>
<td>12.468</td>
</tr>
</tbody>
</table>

1 Data for a population of size 1,000 were generated using the simple linear regression model \( Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \) with \( \beta_0 = 0 \) and \( \beta_1 = 5 \); 0.0 \( \leq X_i \leq 50.0 \) in increments of 0.05, under two different conditions: \( \epsilon_i \) normally distributed random variable with mean 0.0 and standard deviation 5.0 and, \( \epsilon_i \) normally distributed random variable with mean 0.0 and standard deviation 0.2 \( \times X_i \). One thousand bootstrap samples were drawn for each sample size \( n = 10, 20, 40 \) and 80 using \( Y_{\text{new}} = 62.50 \), and the standard deviation of the \( X_{\text{new}} \) for the 1,000 samples was calculated. Abbreviations: ASE, average asymptotic standard error; BSE, average bootstrap standard error.

### TABLE 3

Empirically derived true standard error, the average estimated asymptotic standard error and the average bootstrap standard error based on results of the Monte Carlo simulation for \( Y_{\text{new}} = 62.50 \) and \( X_{\text{new}} = 12.50 \)

<table>
<thead>
<tr>
<th>Conditions of the simulation</th>
<th>Normality, constant variance</th>
<th>Normality, nonconstant variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size, n</td>
<td>True SE</td>
<td>ASE</td>
</tr>
<tr>
<td>10</td>
<td>0.4600</td>
<td>0.4633</td>
</tr>
<tr>
<td>20</td>
<td>0.3212</td>
<td>0.3018</td>
</tr>
<tr>
<td>40</td>
<td>0.1976</td>
<td>0.2255</td>
</tr>
<tr>
<td>80</td>
<td>0.1292</td>
<td>0.1485</td>
</tr>
</tbody>
</table>

1 Data for a population of size 1,000 were generated under the simple linear regression model \( Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \) with \( \beta_0 = 0 \) and \( \beta_1 = 5 \); 0.0 \( \leq X_i \leq 50.0 \) in increments of 0.05, under two different conditions: \( \epsilon_i \) normally distributed random variable with mean 0.0 and standard deviation 0.2 \( \times X_i \). One thousand bootstrap samples were drawn for each sample size \( n = 10, 20, 40 \) and 80 using \( Y_{\text{new}} = 62.50 \), and the standard deviation of the \( X_{\text{new}} \) for the 1,000 samples was calculated. Abbreviations: ASE, average asymptotic standard error; BSE, average bootstrap standard error.
TABLE 4
Percent upper and lower limit noncoverage by 95% confidence intervals based on the classical t-confidence interval, the bootstrap t-confidence interval and the bootstrap percentile confidence interval from the Monte Carlo simulation for Y_{new} = 62.50 and X_{new} = 12.50

<table>
<thead>
<tr>
<th>Conditions of the simulation</th>
<th>Classical t-interval</th>
<th>Bootstrap t-interval</th>
<th>Bootstrap percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size, n</td>
<td>L</td>
<td>U</td>
<td>L</td>
</tr>
<tr>
<td>10</td>
<td>2.60</td>
<td>2.00</td>
<td>2.40</td>
</tr>
<tr>
<td>20</td>
<td>2.20</td>
<td>2.20</td>
<td>3.00</td>
</tr>
<tr>
<td>40</td>
<td>3.60</td>
<td>2.80</td>
<td>3.20</td>
</tr>
<tr>
<td>80</td>
<td>1.90</td>
<td>2.40</td>
<td>2.10</td>
</tr>
</tbody>
</table>

1 Data for a population of size 1,000 were generated under the simple linear regression model Y_i = \beta_0 + \beta_1 X_i + e_i with \beta_0 = 0 and \beta_1 = 5; 0.0 \leq X_i \leq 50.0 in increments of 0.05, under two different conditions: e_i normally distributed random variable with mean 0.0 and standard deviation 5.0 and, e_i normally distributed random variable with mean 0.0 and standard deviation 5.0. One thousand bootstrap samples were drawn for each sample size n = 10, 20, 40 and 80 using Y_{new} = 62.50, and the 95% confidence interval of the \hat{X}_{new} for the 1,000 samples was calculated. Nominal noncoverage rates are 2.5% for lower and upper limits. L, lower; U, upper.

LITERATURE CITED


APPENDIX A

Statistical equations

Let X_i (i = 1, 2, ..., n) be the observed values for a random sample of size n taken from a population with mean \mu and standard deviation \sigma. Let \hat{\theta}(x) represent an estimator of a parametric function, \theta(x) and se(\hat{\theta}(x)) the standard error of \hat{\theta}(x). A simple example of this is the sample mean \bar{x} = \hat{\theta}(x), which is an estimator of \mu = \theta(x), the population mean. Then a (1–\alpha) · 100% confidence interval for \hat{\theta}(x) is:

- Classical t-confidence interval:
  \hat{\theta}(x) \pm t\left(1 - \frac{\alpha}{2}\right) \cdot se(\hat{\theta}(x)), where df denotes the degrees of freedom.
- Asymptotic estimator of the standard error for inverse regression estimator \hat{X}_{new}:
  Let \hat{X}_{new} be the estimated value of the predictor X at a specified value of Y that is denoted Y_{new}. Let \hat{\beta}_0 be the estimated intercept and \hat{\beta}_1 the estimated slope of the regression line. The asymptotic estimate of the standard error is:

  s\hat{e}(\hat{X}_{new}) = \sqrt{\frac{MSE}{n \cdot \sum (X_i - \bar{X})^2}} with

  MSE = \frac{\sum_{i=1}^{n} (Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_i))^2}{n-2}

- \hat{X}_{new} and s\hat{e}(\hat{X}_{new}) are then used in the classical t-confidence interval equation.
- Bootstrap t-confidence interval:
  Select B independent samples of size n, with replacement, from the observed data. These samples may be denoted x^{*1}, x^{*2}, ..., x^{*B}, with each x^{*b} (b = 1, 2, ..., B) consisting of n data values. The bootstrap estimated standard error for \hat{\theta}(x) will be:

  s\hat{e}(\hat{\theta}(x)) = \sqrt{\frac{\sum_{b=1}^{B} (\hat{\theta}(x^{*b}) - \hat{\theta})^2}{B - 1}} with \hat{\theta} = \frac{\sum_{b=1}^{B} \hat{\theta}(x^{*b})}{B}

- \hat{\theta} and s\hat{e}(\hat{\theta}(x)) are then used in the classical t-confidence interval equation.
- Bootstrap percentile confidence interval:
  Order the B bootstrap estimates from smallest to largest. Identify the \left(\left(\frac{\alpha}{2} \cdot 100\right)\cdot B\right)^{th} and the \left(\left(1 - \frac{\alpha}{2} \cdot 100\right)\cdot B\right)^{th} values of the ordered values. These values represent the lower and upper limits for the (1–\alpha) · 100% confidence interval.