Section Outline

1. Prediction
2. Calibration

Prediction

Prediction is the use of a fitted regression model of the form

$$y = f(x_1, \ldots, x_k; \hat{\beta}_1, \ldots, \hat{\beta}_p) + N(0, s^2)$$

to estimate the values of the

1. regression function, $f(x_1^*, \ldots, x_k^*; \beta_1, \ldots, \beta_p)$, or
2. response variable, y^*,

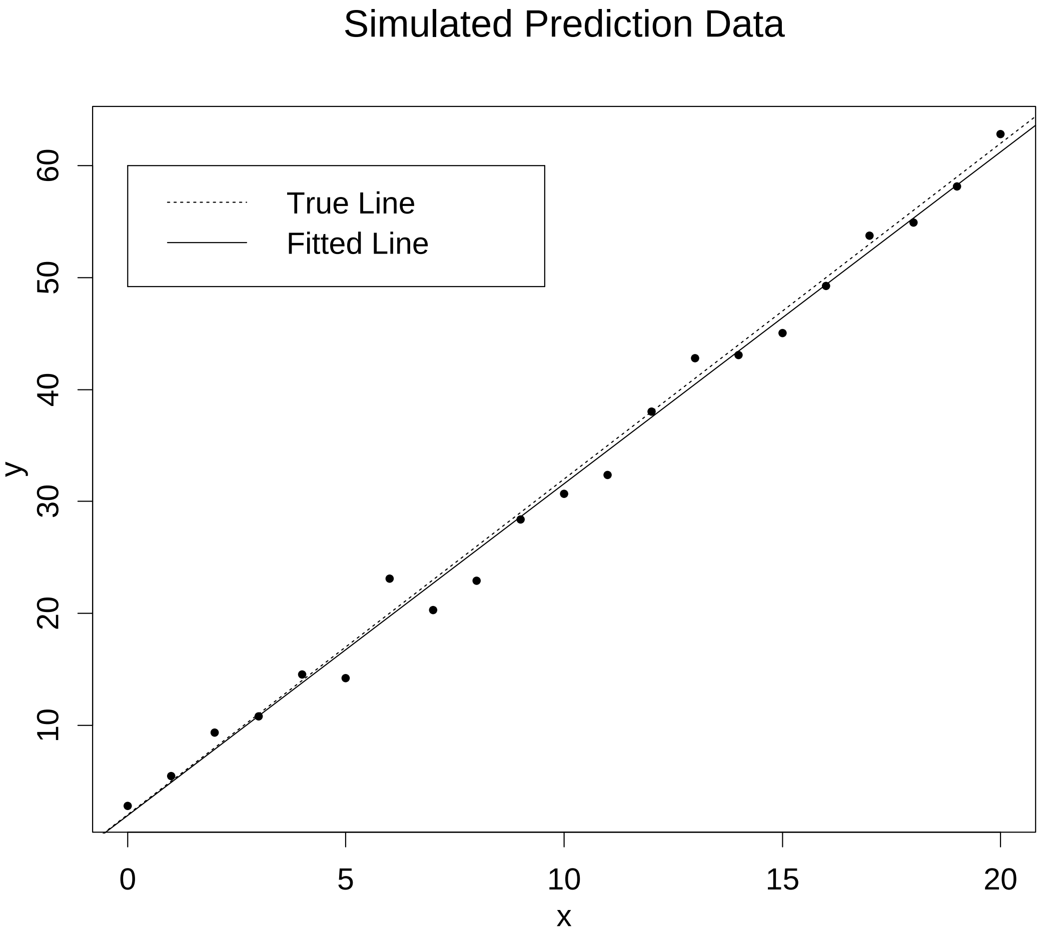
associated with specific values of the predictor variables $x_1^*, \ldots, x_k^*$.

In order to be able to compare the estimated values of $f(\vec{x}^*; \vec{\beta})$ or y^* to target values, or to be able to intercompare different estimates of $f(\vec{x}^*; \vec{\beta})$ or y^*, we also need to know the uncertainties of each estimate.

Two Types of Intervals Used in Prediction

confidence intervals - intervals that contain $f(x_1^*, \ldots, x_k^*; \beta_1, \ldots, \beta_p)$, the mean value of the population, for a specific set of values of $x_1^*, \ldots, x_k^*$,

prediction intervals - intervals that contain the value of a new measurement or observation from the process for a specific set of values of $x_1^*, \ldots, x_k^*$, and


Simulated Prediction Data
True Line
Fitted Line
y
10
20
30
40
50
60
x
0
5
10
15
20

Confidence Level of the Intervals

All of these intervals contain their correspondent quantities with a user-specified probability $100(1 - \alpha)\%$, called the confidence level of the interval.

α is a small number, usually between 0.01 and 0.2, chosen to give the confidence level the desired value.

Confidence Interval Formula for $f(\vec{x}^*; \vec{\beta})$ General Case

A confidence interval for

$$f(x_1^*, \ldots, x_k^*; \beta_1, \ldots, \beta_p)$$

is given by the following formula:

$$\hat{y}^* \pm t_{1-\alpha/2,n-p} \sqrt{\vec{d}^{*T} V \vec{d}^*}$$

where

- $\hat{y}^*$ is an estimate of $f(x_1^*, \ldots, x_k^*; \beta_1, \ldots, \beta_p)$
- $t_{1-\alpha/2,n-p}$ is a t distribution multiplier
- V is the variance-covariance matrix of the estimated parameters, and
- $\vec{d}^*$ is a column vector of partial derivatives of $f(x_1^*, \ldots, x_k^*; \beta_1, \ldots, \beta_p)$ with respect to $\beta_1, \ldots, \beta_p$ evaluated at $\hat{\beta}_1, \ldots, \hat{\beta}_p$ and $x_1^*, \ldots, x_k^*$

More Info on Confidence Interval Construction General Case

$\hat{y}^*$, the estimate of $f(x_1^*, \ldots, x_k^*; \beta_1, \ldots, \beta_p)$, can be found by plugging $x_1^*, \ldots, x_k^*$ into the estimated regression function.

$$\hat{y}^* = f(x_1^*, \ldots, x_k^*; \hat{\beta}_1, \ldots, \hat{\beta}_p)$$

V, or better yet, individual standard deviations for any specified values of $x_1^*, \ldots, x_k^*$, should be supplied by your regression software.

Usually, especially for statistically nonlinear models, $\vec{d}^*$ will have to be computed manually. Some software may provide generic tools for differentiation, however.

V and d^* for Statistically Linear Models

For any statistically linear model

$$d^{*T} = (1, x_1^*, \ldots, x_k^*)$$

Similarly, the variance-covariance matrix for any linear model is given by

$$V = s^2(X^T X)^{-1}$$

where

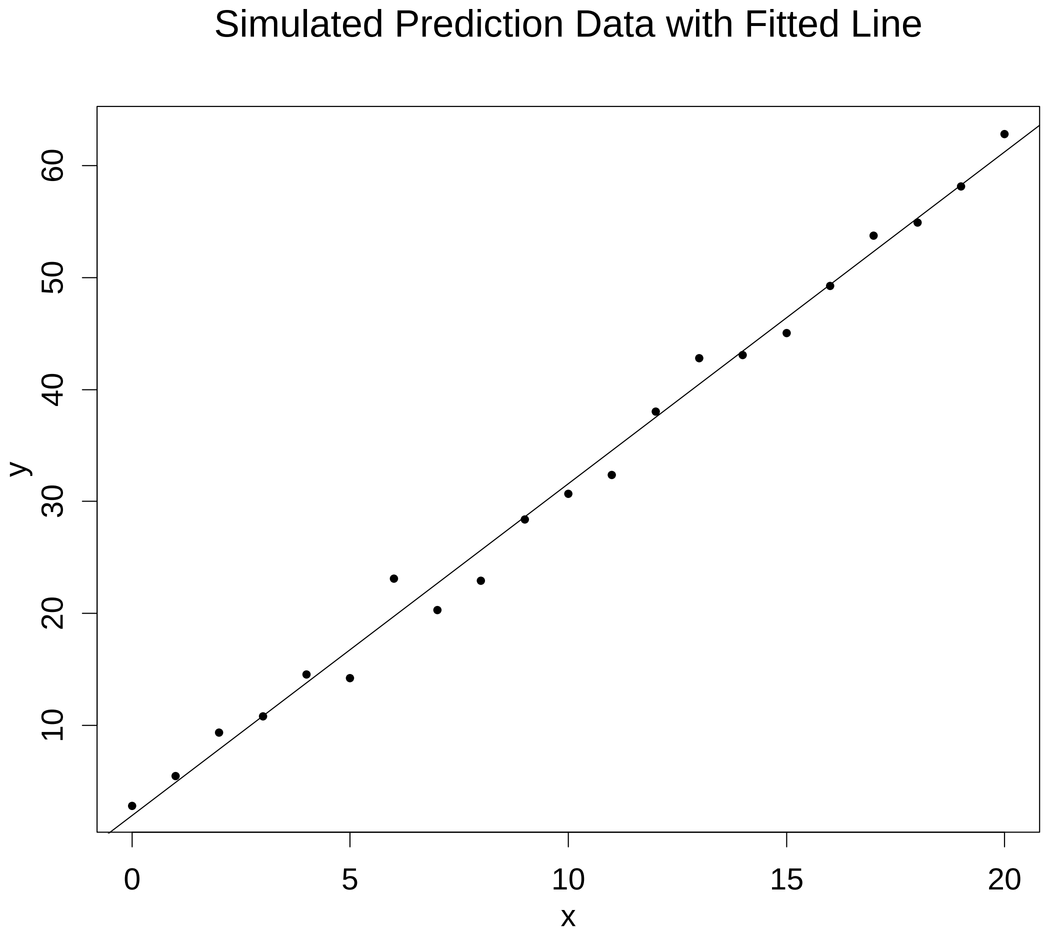
- s is the residual standard deviation, and
- X is the matrix with n rows and $p = k + 1$ columns, the first of which is a column of 1's for the intercept term, and each of the other columns given by the n values of one of the k variables used to fit the model,

$$X = [1|x_1|x_2|\ldots|x_k],$$

Confidence Interval for $f(\vec{x}^*; \vec{\beta})$ Straight Line Model

The formula for a $100(1-\alpha)\%$ confidence interval for $f(x^*; \beta_1, \beta_2)$ under the straight line model reduces to:

$$\hat{\beta}_1 + \hat{\beta}_2 x^* \pm t_{1-\alpha/2, n-2} s \sqrt{\frac{1}{n} + \frac{(x* - \bar{x})^2}{\Sigma_{i=1}^{n}(x_i - \bar{x})^2}}$$


Simulated Prediction Data with Fitted Line
y
10
20
30
40
50
60
x
0
5
10
15
20

Numeric Example of Confidence Interval Computations for Simulated Data

```
N = 21

Residual Standard Error = 1.665784

Multiple R-Square = 0.9922741

F-statistic = 2440.264 on 1 and 19 df, p-value = 0

               coef    std.err    t.stat    p.value
Intercept 1.941797 0.70178584  2.766936 0.01227545
    x     2.965458 0.06003068 49.399031 0.00000000

mean(x) = 10

20*var(x) = 770
```

Plugging in the numbers from the regression output yields:

$$\hat{y}^* = 1.941797 + 2.965458 \times 12.897 = 40.18731$$

$$\begin{aligned} U &= 2.093024(1.665784)\sqrt{\frac{1}{21} + \frac{(12.897 - 10)^2}{770}} \\ &= 0.8434117 \end{aligned}$$

$$\Downarrow$$

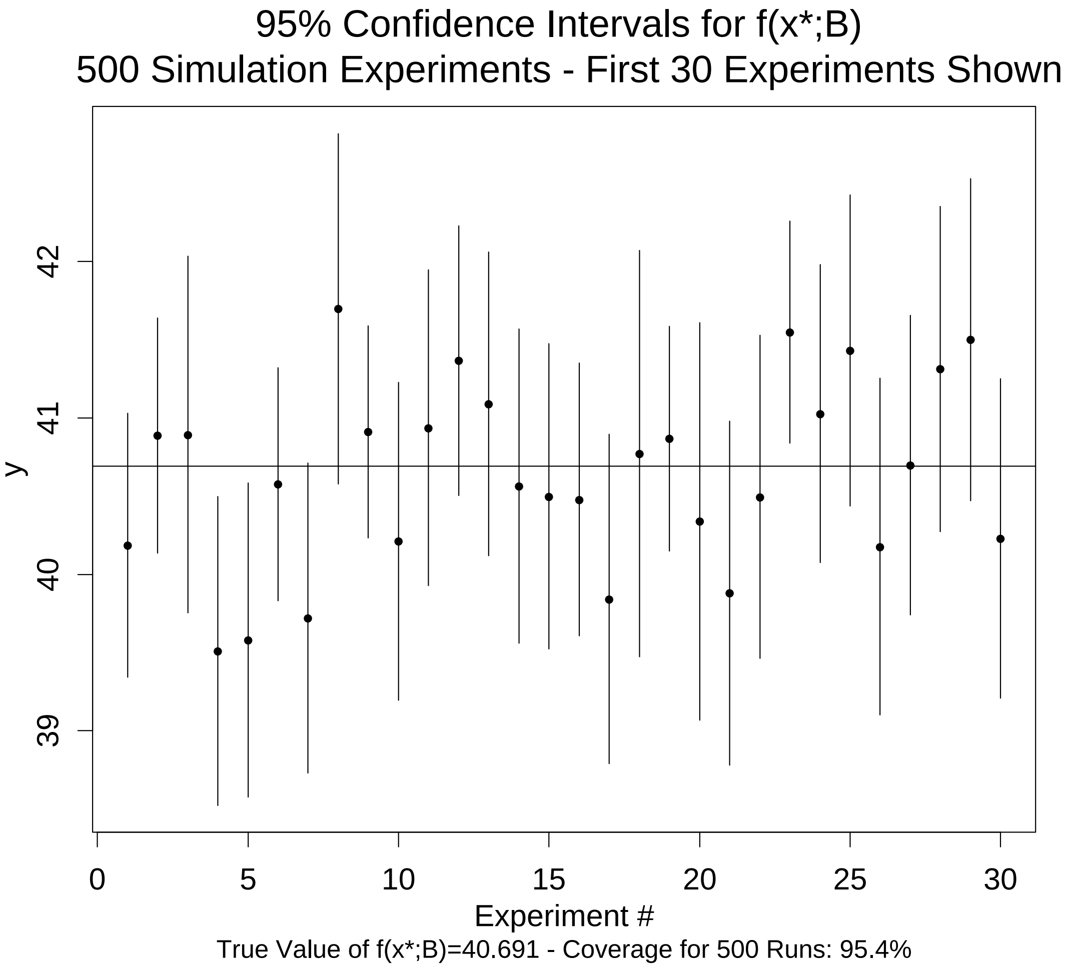
$$40.18731 \pm 0.8434117$$

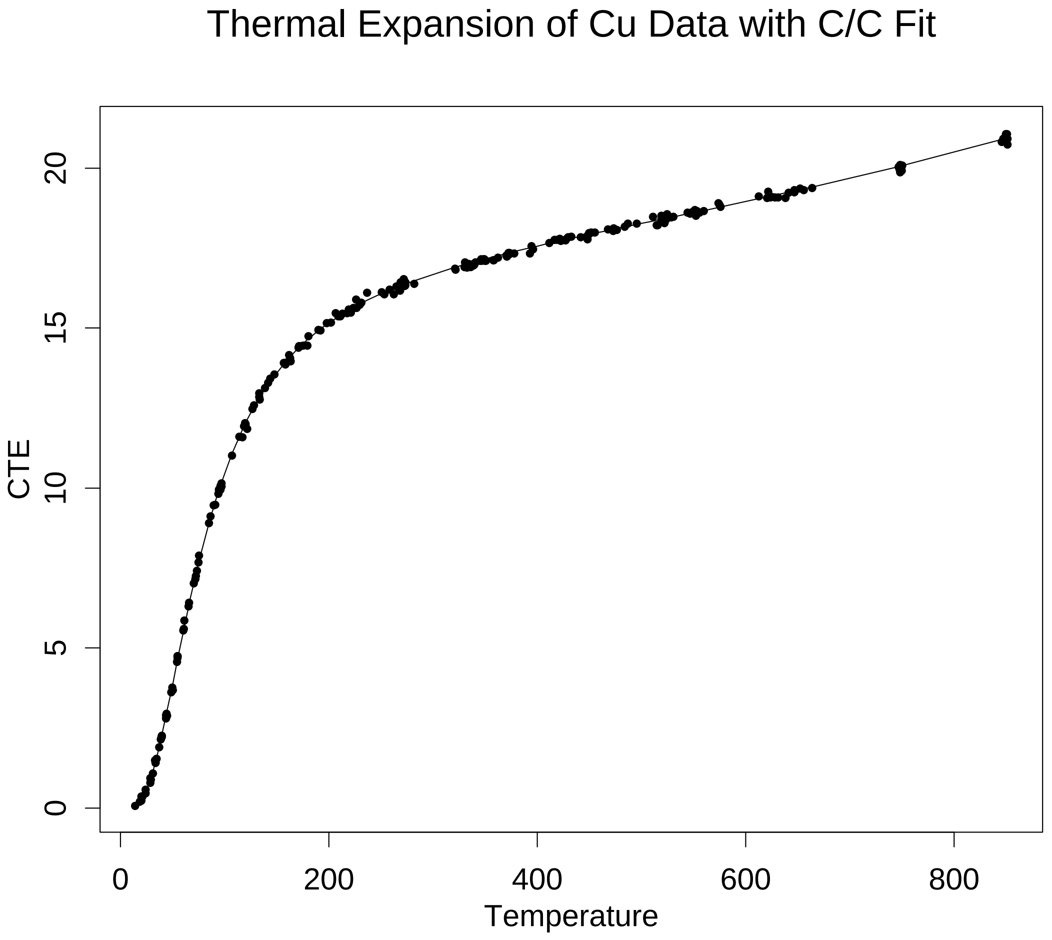
Interpretation of Confidence Intervals

Under long-term repetition of the steps

1. collect a set of regression data
2. fit the function f to the data, and
3. compute $\hat{y}^* \pm U$

$100(1-\alpha)\%$ of the intervals constructed will contain the true value of $f(x_1^*, \ldots, x_k^*; \beta_1, \ldots, \beta_p)$.


95% Confidence Intervals for f(x*;B)
500 Simulation Experiments - First 30 Experiments Shown
y
39
40
41
42
0
5
10
15
20
25
30
Experiment #
True Value of f(x*;B)=40.691 - Coverage for 500 Runs: 95.4%


Thermal Expansion of Cu Data with C/C Fit
CTE
0
5
10
15
20
0
200
400
600
800
Temperature

Output from C/C Fit to Cu Data

```
          Value  Std. Error   t value
b1  1.07766e+00 1.70702e-01   6.31312
b2 -1.22695e-01 1.20004e-02 -10.22430
b3  4.08642e-03 2.25085e-04  18.15500
b4 -1.42632e-06 2.75781e-07  -5.17192
b5 -5.76099e-03 2.47130e-04 -23.31160
b6  2.40539e-04 1.04494e-05  23.01930
b7 -1.23147e-07 1.30274e-08  -9.45294

Residual standard error: 0.0818039 on 229 degrees of freedom



Variance-Covariance Matrix (Unscaled):

              b1            b2            b3            b4
b1  4.354428e+00 -2.974657e-01  5.194511e-03 -4.944175e-06
b2 -2.974657e-01  2.151992e-02 -3.942427e-04  3.936012e-07
b3  5.194511e-03 -3.942427e-04  7.570837e-06 -8.073008e-09
b4 -4.944175e-06  3.936012e-07 -8.073008e-09  1.136534e-11
b5 -1.659381e-03  4.221995e-05  8.995716e-07 -4.089729e-09
b6  2.465934e-04 -1.848567e-05  3.488716e-07 -3.483624e-10
b7 -2.521217e-07  1.979725e-08 -3.989330e-10  5.341495e-13

              b5            b6            b7
b1 -1.659381e-03  2.465934e-04 -2.521217e-07
b2  4.221995e-05 -1.848567e-05  1.979725e-08
b3  8.995716e-07  3.488716e-07 -3.989330e-10
b4 -4.089729e-09 -3.483624e-10  5.341495e-13
b5  9.126447e-06  3.324414e-09 -1.599174e-10
b6  3.324414e-09  1.631693e-08 -1.749188e-11
b7 -1.599174e-10 -1.749188e-11  2.536104e-14
```

Confidence Interval for Cu Data

Using the output from p. 231 and the general formulas on pp. 222 through 224 yields the confidence interval

$$16.68267 \pm 0.01941922$$

for a temperature of 300 degrees.

In this case d^* is not $(1, x_1^*, \ldots, x_k^*)$. Instead, differentiation gives the values

```
        d1       d2       d3      d4        d5        d6        d7
0.05683349 17.05005 5115.014 1534504 -284.4404 -85332.11 -25599633
```

for d^* which are used with V to obtain an approximation to the uncertainty of $\hat{y}^*$.

Checking the approximation for this example by direct simulation gave an estimate of 94.2% for the true confidence level, based on 1000 replications.

Prediction Interval Formula for y^* General Case

A prediction interval for y^* is given by the following formula:

$$\hat{y}^* \pm t_{1-\alpha/2,n-p}\sqrt{s^2 + \vec{d}^{*T} V \vec{d}^*}$$

where

- $\hat{y}^*$ is an estimate of y^*,
- $t_{1-\alpha/2,n-p}$ is a t distribution multiplier
- s is the residual standard deviation
- V is the variance-covariance matrix of the estimated parameters, and
- $\vec{d}^*$ is a column vector of partial derivatives of $f(x_1^*, \ldots, x_k^*; \beta_1, \ldots, \beta_p)$ with respect to $\beta_1, \ldots, \beta_p$ evaluated at $\hat{\beta}_1, \ldots, \hat{\beta}_p$ and $x_1^*, \ldots, x_k^*$

More Info on Confidence Interval Construction General Case

Just as for confidence intervals:

$$\hat{y}^* = f(x_1^*, \ldots, x_k^*; \hat{\beta}_1, \ldots, \hat{\beta}_p)$$

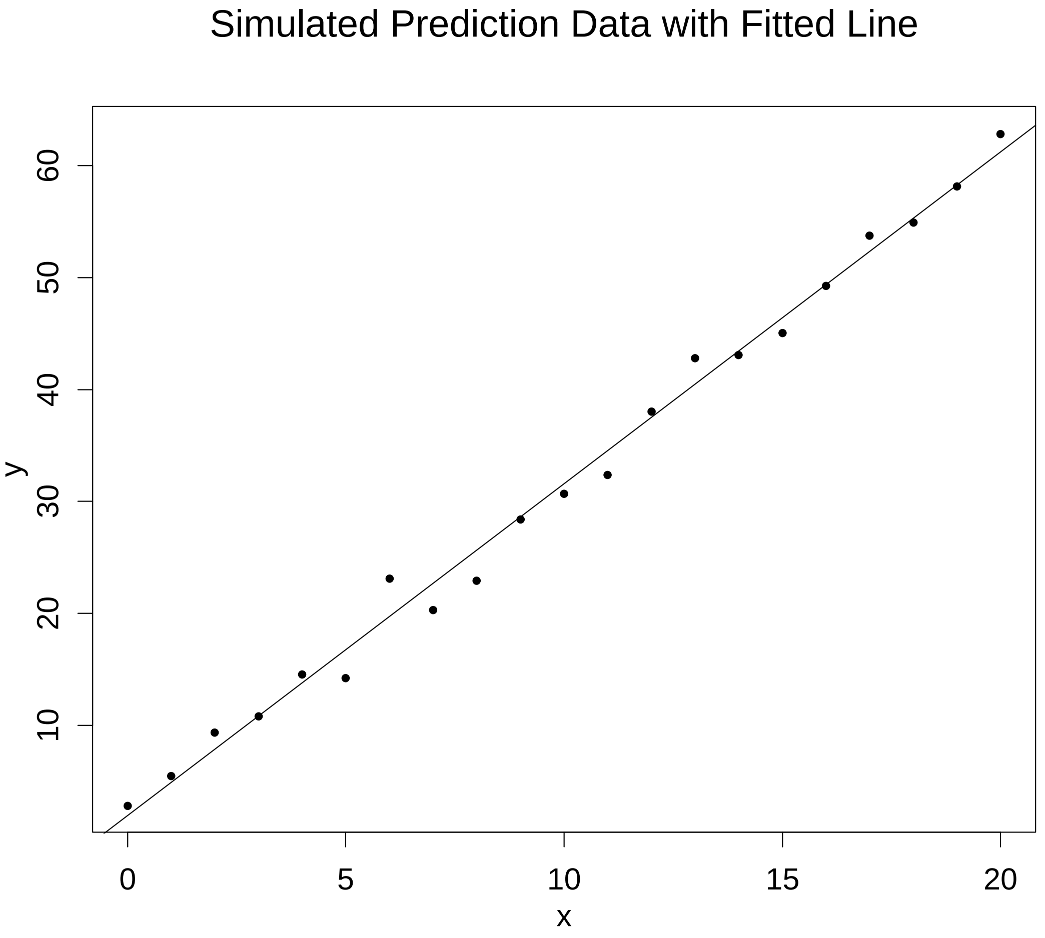
The general formulas for and comments made about V and d^* for confidence intervals also apply to prediction intervals.

Predictably :-), the simplifications of the formulas for V and d^* for statistically linear models are also unchanged.

Prediction Interval for y^* - Straight Line Model

The formula for a $100(1-\alpha)\%$ prediction interval for y^* under the straight line model reduces to:

$$\hat{\beta}_1+\hat{\beta}_2 x^* \pm t_{1-\alpha/2,n-2}\, s\sqrt{1+\frac{1}{n}+\frac{(x^*-\bar{x})^2}{\Sigma_{i=1}^{n}(x_i-\bar{x})^2}}$$


Simulated Prediction Data with Fitted Line
y
10
20
30
40
50
60
0
5
10
15
20
x

Numeric Example of Prediction Interval Computations for Simulated Data

```
N = 21

Residual Standard Error = 1.665784

Multiple R-Square = 0.9922741

F-statistic = 2440.264 on 1 and 19 df, p-value = 0

             coef    std.err    t.stat    p.value
Intercept 1.941797 0.70178584  2.766936 0.01227545
    x     2.965458 0.06003068 49.399031 0.00000000

mean(x) = 10

20*var(x) = 770
```

Plugging in the numbers from the regression output yields:

$$\hat{y}^* = 1.941797 + 2.965458 \times 12.897 = 40.18731$$

$$\begin{aligned} U &= 2.093024(1.665784)\sqrt{1 + \frac{1}{21} + \frac{(12.897 - 10)^2}{770}} \\ &= 3.587089 \end{aligned}$$

$$\Downarrow$$

$$40.18731 \pm 3.587089$$

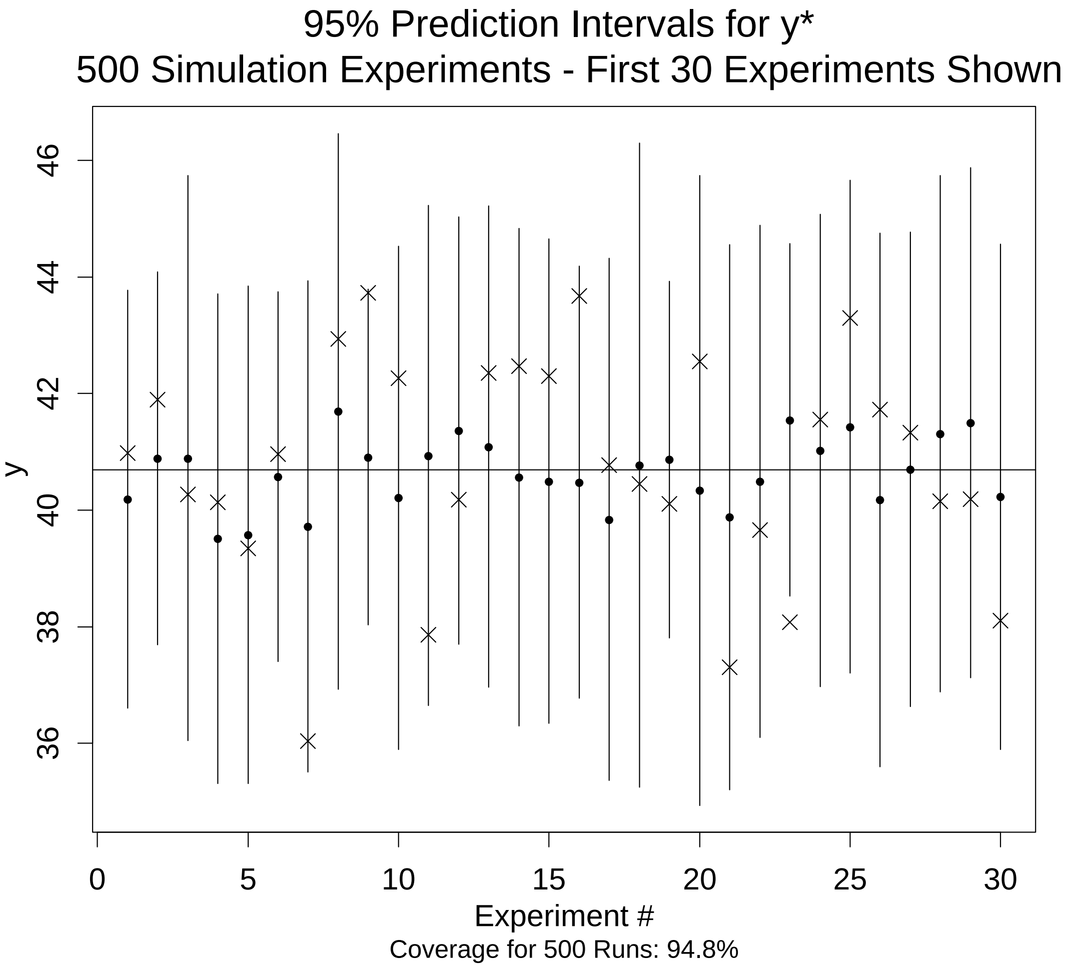
Interpretation of Prediction Intervals

Under long-term repetition of the steps

1. collect a set of regression data
2. fit the function f to the data, and
3. compute $\hat{y}^* \pm U$

$100(1-\alpha)\%$ of the intervals constructed will contain the value of a single new measurement, y^*, made at $x_1^*, \ldots, x_k^*$.

Unlike the case of the confidence interval, you may be able to directly see if a prediction interval worked by taking another measurement at $x_1^*, \ldots, x_k^*$ and checking to see if it falls inside the interval.


95% Prediction Intervals for y*
500 Simulation Experiments - First 30 Experiments Shown
y
36
38
40
42
44
46
0
5
10
15
20
25
30
Experiment #
Coverage for 500 Runs: 94.8%

Prediction Interval for Cu Data

Using the output from p. 231 again, with the general formulas for prediction intervals on pp. 233 and 234, yields the prediction interval

$$16.68267 \pm 0.16235$$

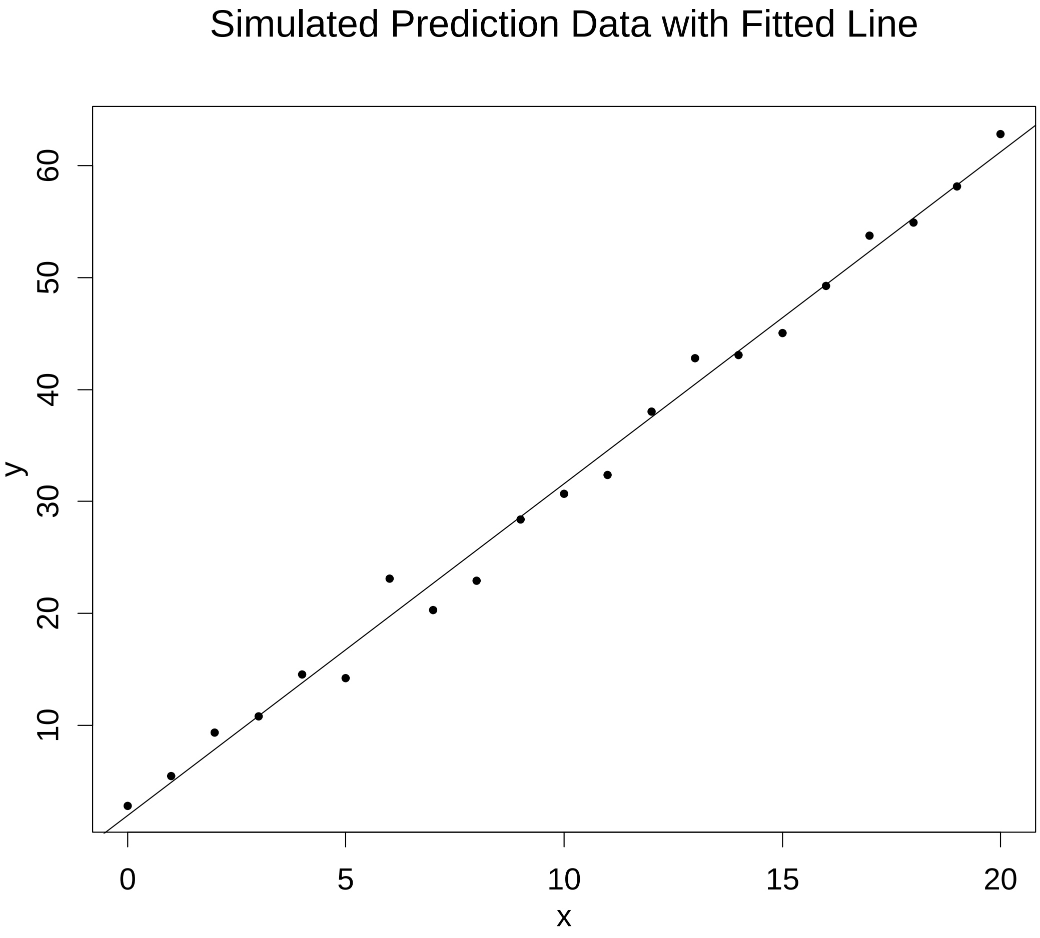
for a temperature of 300 degrees.

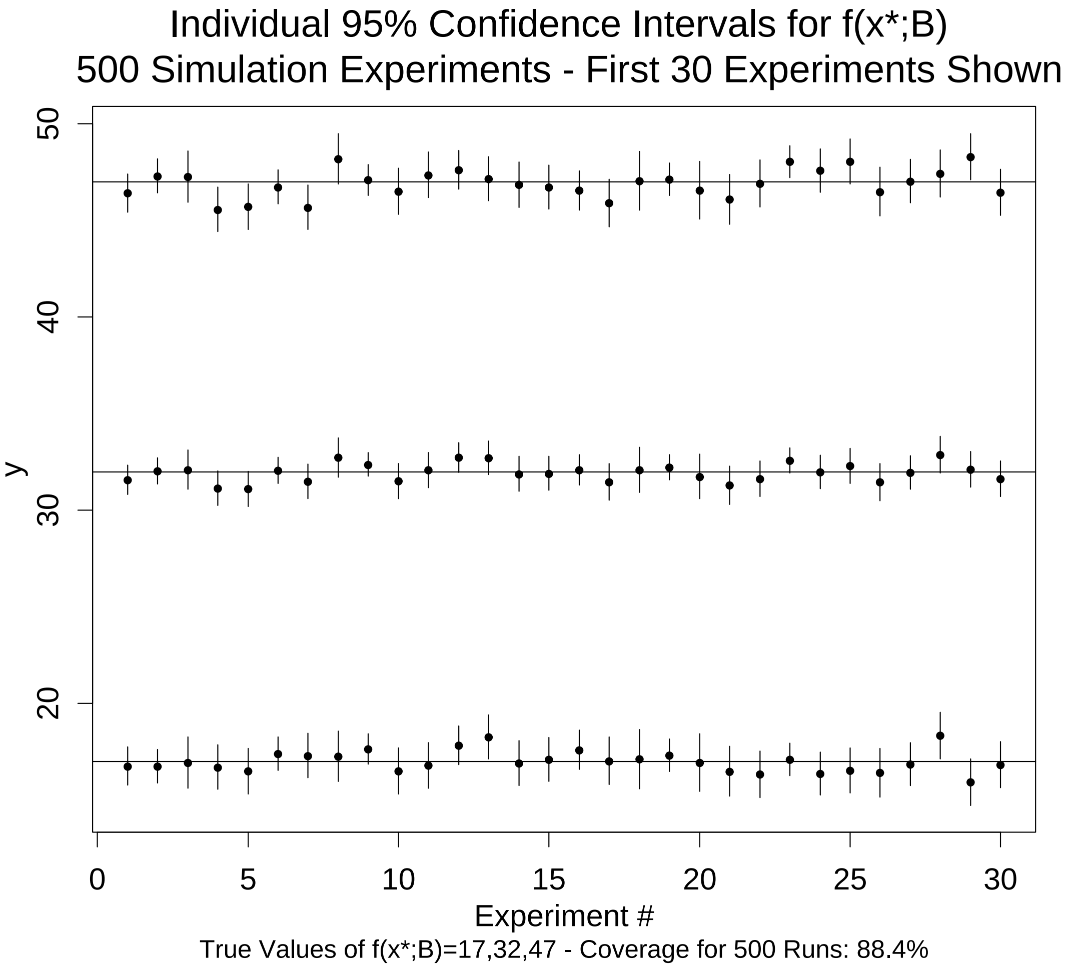
Checking the approximation to the uncertainty for this example by direct simulation gave an estimate of 95.6% for the true confidence level, based on 1000 replications.

Making Multiple Predictions

The interpretation of both confidence intervals and prediction intervals guarantee coverage for only one measurement for each set regression data.

If more predictions are made from each regression data set the stated coverage of the confidence or prediction intervals does not hold for all intervals simultaneously.


Simulated Prediction Data with Fitted Line
y
10
20
30
40
50
60
x
0
5
10
15
20


Individual 95% Confidence Intervals for f(x*;B)
500 Simulation Experiments - First 30 Experiments Shown
y
20
30
40
50
0
5
10
15
20
25
30
Experiment #
True Values of f(x*;B)=17,32,47 - Coverage for 500 Runs: 88.4%

m Simultaneous Confidence or Prediction Intervals

One method for computing a fixed number, m, of simultaneous confidence intervals is to replace the t distribution multipliers in the earlier formulas for an individual confidence or prediction interval with either $t_{1-\alpha/2m,n-p}$ or $\sqrt{mF_{1-\alpha,m,n-p}}$.

Using the smaller of these two multliers for any particular combination of confidence level and number of intervals yields intervals which cover all m values of the regression function or all m new measurements with confidence at least $100(1-\alpha)\%$.

Three 95% Simultaneous Confidence Intervals for the Simulated Prediction Data

To compute simultaneous 95% confidence intervals for three different values of the predictor variable, x_1^*, x_2^*, and x_3^*, first determine which multiplier to use.

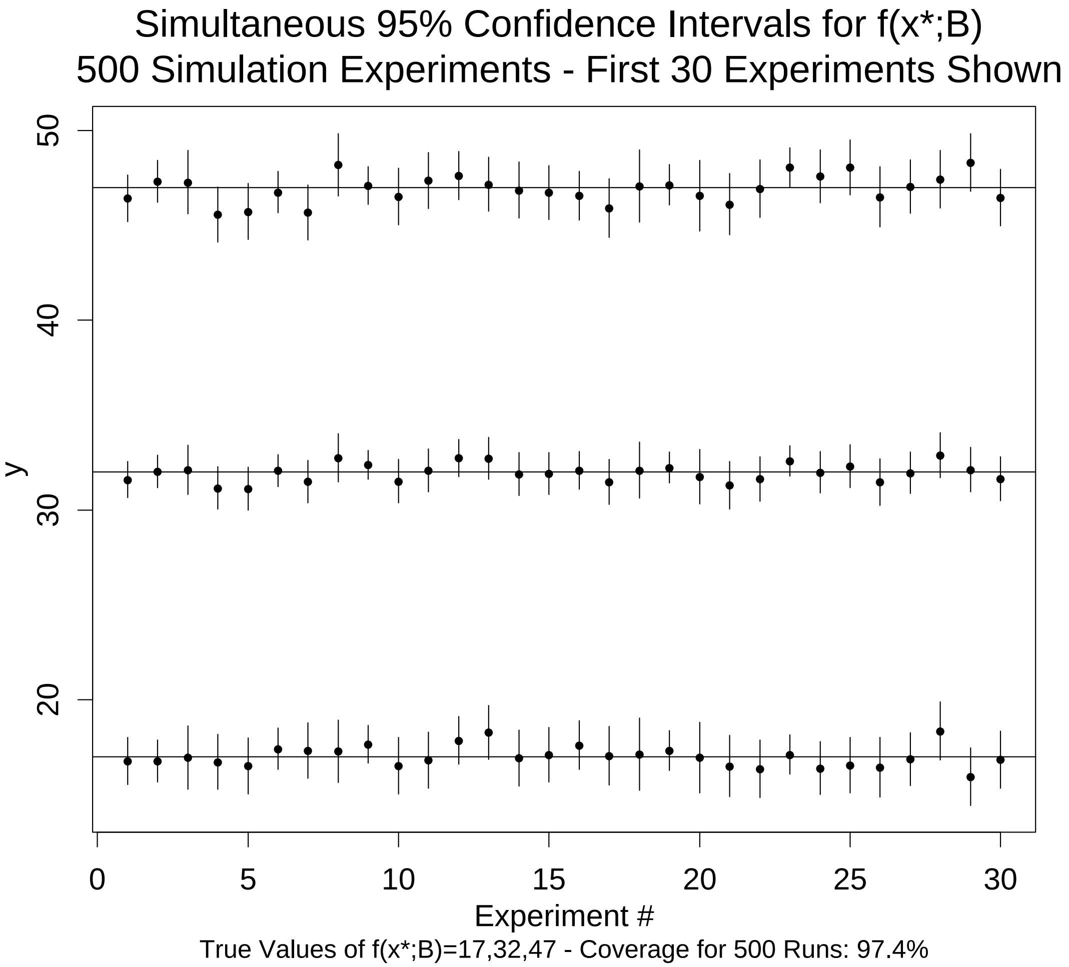
$$t_{0.99167,19} = 2.625293$$

and

$$\sqrt{3F_{0.95,3,19}} = 3.063013$$

so the t multiplier will provide shorter intervals in this case.

Now simply substitute the adjusted t value in the inidividual confidence interval formulas.


Simultaneous 95% Confidence Intervals for f(x*;B)
500 Simulation Experiments - First 30 Experiments Shown
y
20
30
40
50
0
5
10
15
20
25
30
Experiment #
True Values of f(x*;B)=17,32,47 - Coverage for 500 Runs: 97.4%

Interpretation of m Simultaneous Confidence or Prediction Intervals

Under long term repetition of the steps

1. collect a set of regression data,
2. fit the function f to the data, and
3. compute $\hat{y}_1^* \pm U_1, \ldots, \hat{y}_m^* \pm U_m$

at least $100(1-\alpha)\%$ of the sets of m intervals constructed will have the property that every interval in the set will contain its associated true value of $f(\vec{x}^*; \vec{\beta})$ or its associated new measurement.

Infinitely Many Simultaneous Confidence Intervals for Linear Models

If the t distribution multiplier in the formulas for individual confidence intervals is replaced with the factor $\sqrt{pF_{1-\alpha,p,n-p}}$ then infintely many simultaneous confidence intervals can be constructed.

Using this multiplier yields simultaneous confidence bands for the entire regression function so that an infinite number of confidence intervals can be computed without increasing the error rate.

These confidence bands are called Working-Hotelling confidence bands after the researchers who developed them.

Interpretation of Working-Hotelling Confidence Bands

Under long term repetition of the steps

1. collect a set of regression data,
2. fit the function f to the data, and
3. compute $\hat{y}_1^* \pm U_1, \hat{y}_2^* \pm U_2, \ldots$

every confidence interval constructed will contain its associated true value of $f(\vec{x}^*; \vec{\beta})$ for $100(1-\alpha)\%$ of the data sets, no matter how many intervals are computed.

Calibration

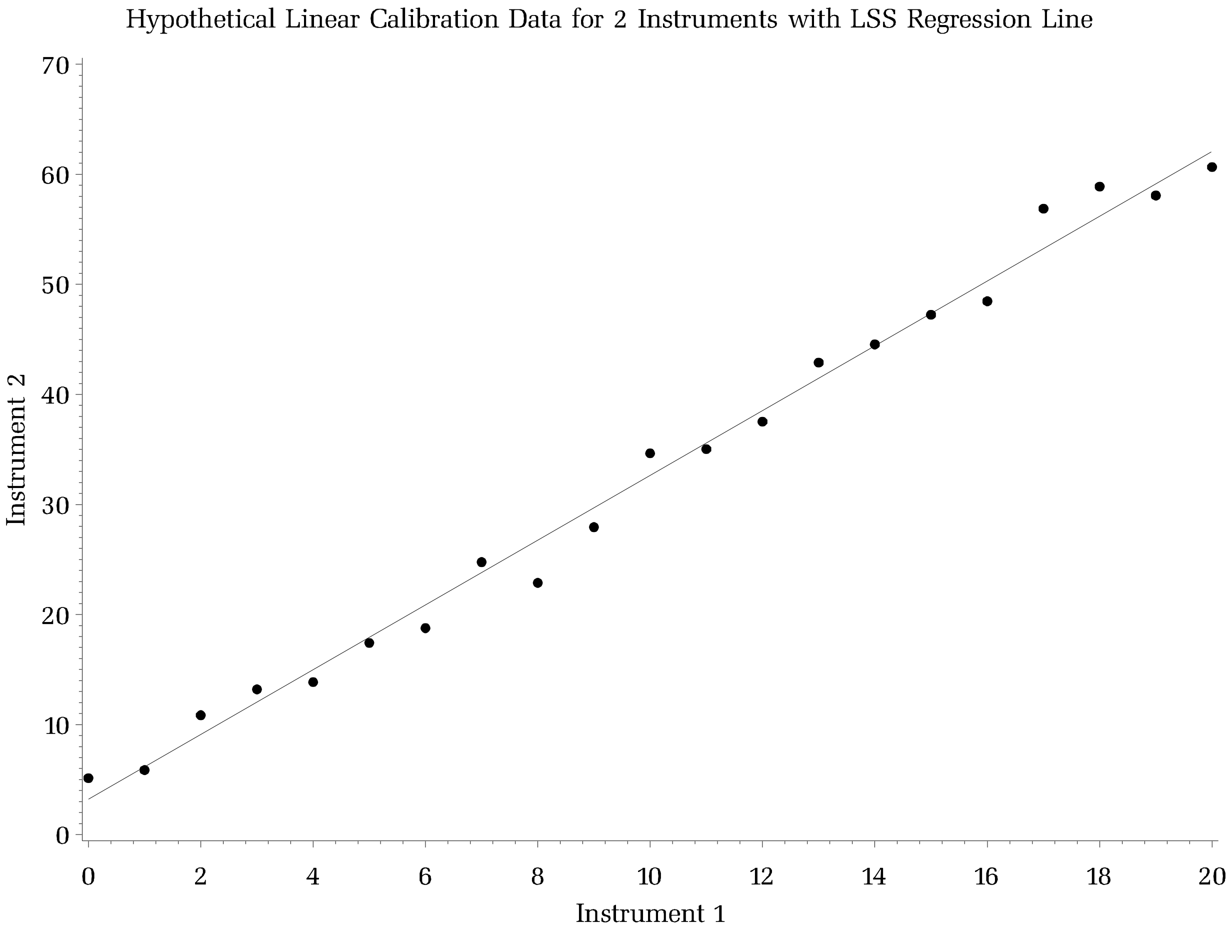
Calibration is the use of a fitted regression model of the form

$$y = f(x_1; \hat{\beta}_1, \ldots, \hat{\beta}_p) + N(0, s^2)$$

to determine the value of the predictor variable x^* associated with a newly observed value of the response variable y^*.

Any type of regression function with one predictor can be used for calibration, however the straight line is the most commonly used function.

One big advantage of the straight line model for calibration is the ease with which uncertainties can be computed. Computing uncertainties for more complicated calibration functions is usually relatively difficult.


Hypothetical Linear Calibration Data for 2 Instruments with LSS Regression Line
Instrument 2
0
10
20
30
40
50
60
70
Instrument 1
0
2
4
6
8
10
12
14
16
18
20

Types of Calibration Intervals

Unlike prediction problems, there is only one type of interval associated with calibration, the confidence interval.

This is because the predictor variable takes on only ‘true’ values, rather than taking on values from a population described by a probability distribution.

Like prediction problems, calibration confidence intervals contain the true value of the predictor variable with a user-specified confidence of $100(1-\alpha)\%$.

Point Estimate for x^*

A point estimate for x^*, the unknown predictor variable associated with an observed response variable, y^*, can be found by plugging y^* into the estimated regression function and solving for x^*.

For the straight line model this can be done anlytically

$$y^* = \hat{\beta}_1 + \hat{\beta}_2 x^*$$

$$\Downarrow$$

$$\hat{x}^* = \frac{(y^* - \hat{\beta}_1)}{\hat{\beta}_2}$$

For other functions it may only be practical to solve the equation numerically.

An Approximate Confidence Interval for x^* Straight Line Model

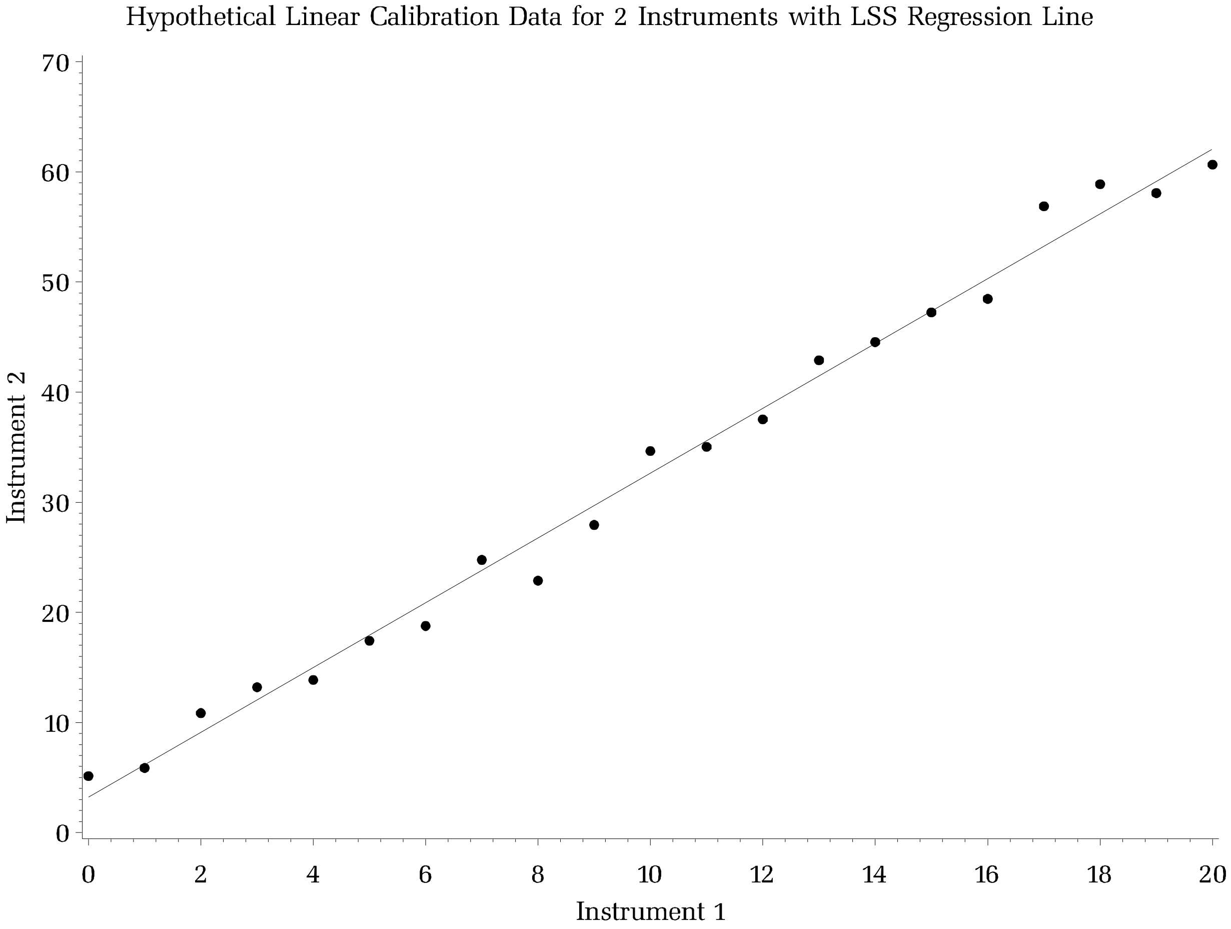
For the straight line model an approximate confidence interval for x^* is given by the following formula:

$$\hat{x^*} \pm t_{1-\alpha/2,n-2}\sqrt{\frac{s^2}{\hat{\beta}_2^2}(1+\frac{1}{n}+\frac{(\hat{x^*}-\bar{x})^2}{\Sigma(x_i-\bar{x})^2})}$$

The confidence level of this interval only holds if

$$c = \frac{s^2 t_{1-\alpha/2,n-2}^2}{\hat{\beta}_2^2\,\Sigma(x_i-\bar{x})^2} < 0.1$$

This criterion is usually met if the slope of the line is not near zero.


Hypothetical Linear Calibration Data for 2 Instruments with LSS Regression Line
Instrument 2
70
60
50
40
30
20
10
0
0
2
4
6
8
10
12
14
16
18
20
Instrument 1

Regression Output for Simulated Calibration Data

```
                ANOVA Table for Simulated Calibration Data

Dependent Variable: INS2

                      Analysis of Variance

                         Sum of         Mean
Source       DF        Squares        Square   F Value    Prob>F

Model         1     6667.26119    6667.26119  1873.995    0.0001
Error        19       67.59783       3.55778
C Total      20     6734.85902


    Root MSE        1.88621     R-square        0.9900
    Dep Mean       32.64167     Adj R-sq        0.9894


                Parameter Estimates

                Parameter     Standard    T for H0:
Variable  DF    Estimate         Error   Parameter=0   Prob > |T|

INTERCEP   1    3.215867   0.79464932          4.047       0.0007
INS1       1    2.942581   0.06797422         43.290       0.0001
```

Numeric Example of Confidence Interval Calculations

$$\hat{x}^* = \frac{41 - 3.215867}{2.942581} = 12.84047$$

The formula for an approximate 95% confidence interval for $\hat{x}^*$ is:

$$12.84047 \pm 2.093024\sqrt{\frac{3.55778}{2.942581^2}(1 + \frac{1}{21} + \frac{(12.84047 - 10)^2}{770})}$$

$$\Downarrow$$

$$12.84047 \pm 1.380061$$

$$c = \frac{2.093024^2 \times 3.55778}{2.942581^2 \times 770} \approx 0.00234$$

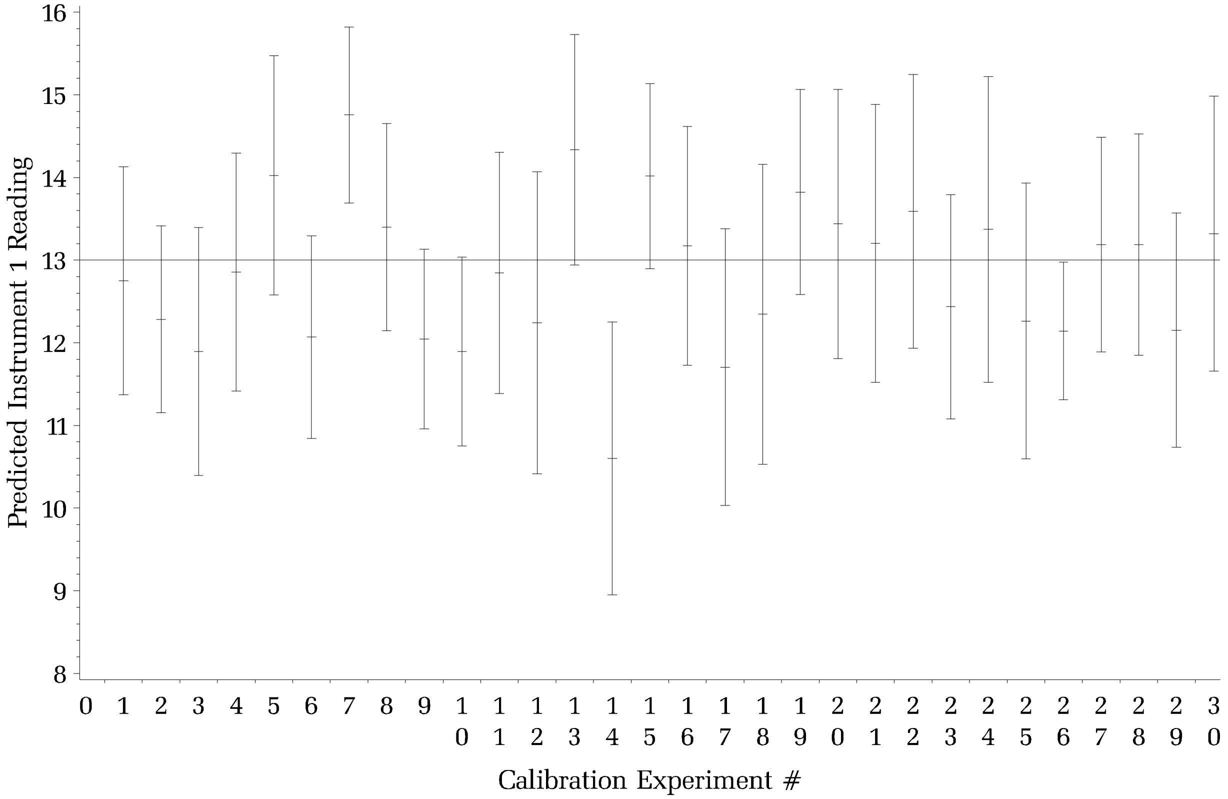
Interpretation of a Single Calibration Confidence Interval

Under long term repetition of the steps

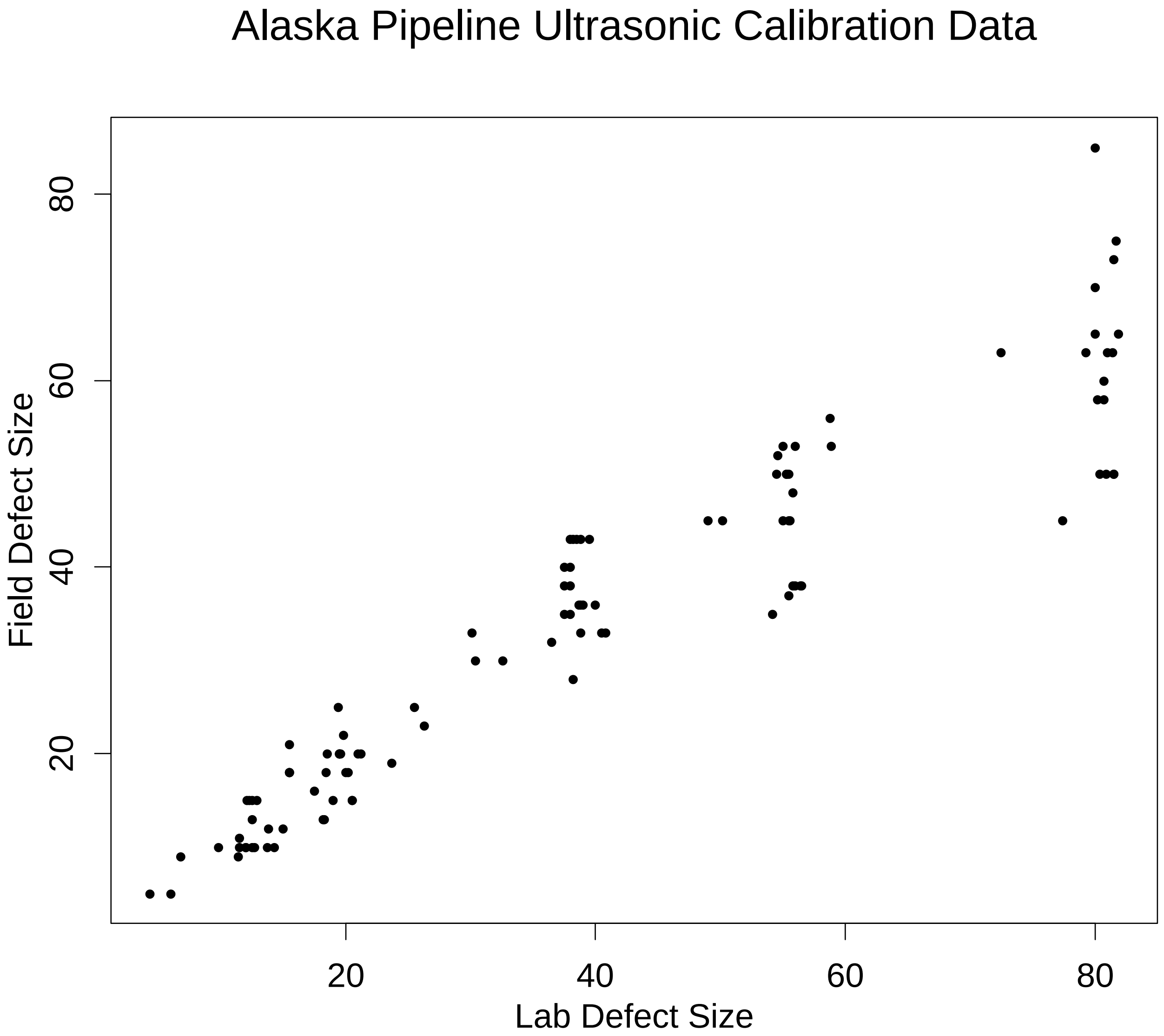
1. collect a set of regression data,
2. fit the function f to the data,
3. observe a new response, y^*, associated with some an unknown value of the predictor variable, x^*, and
4. compute $\hat{x}^* \pm U$,

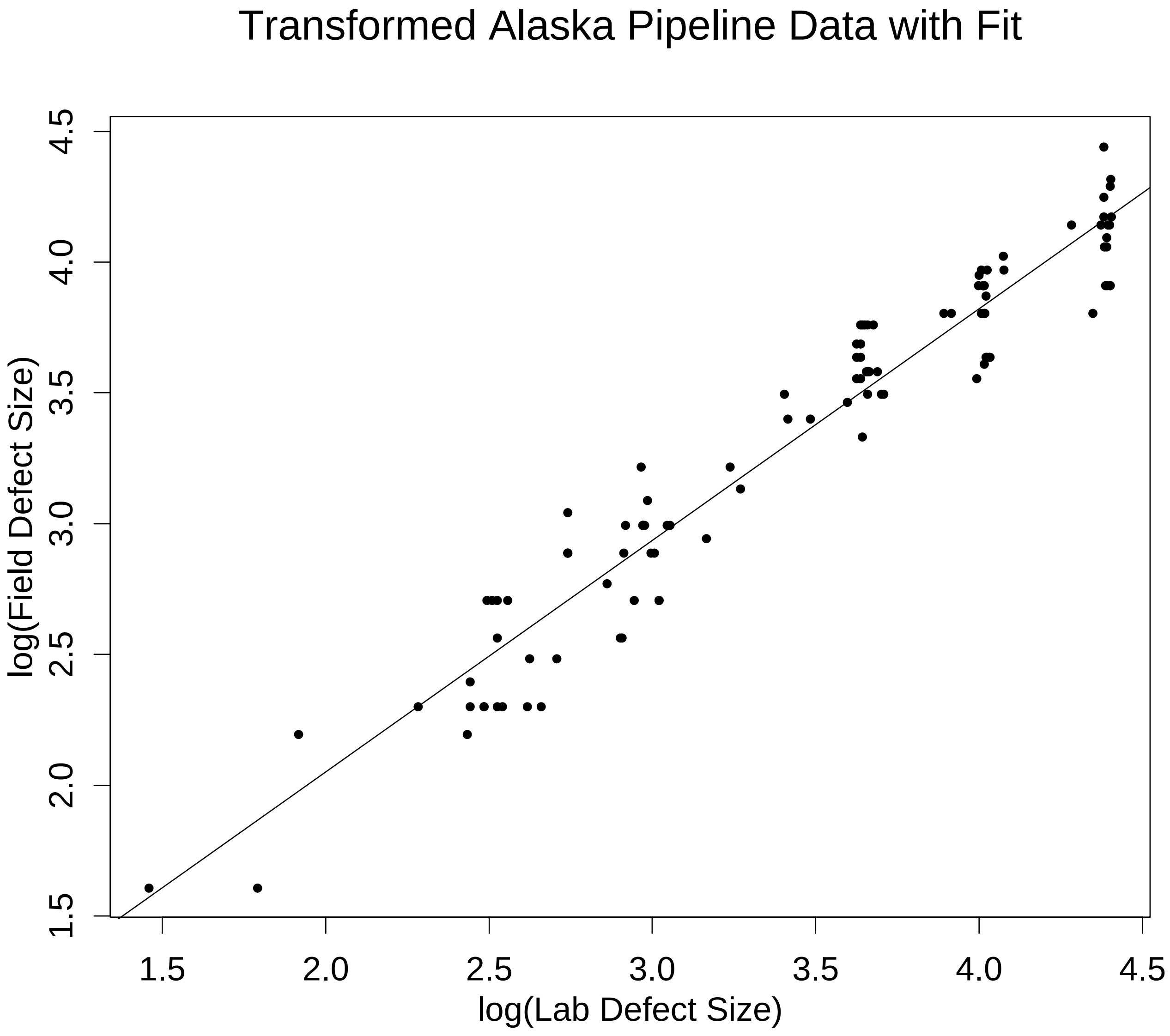
$100(1-\alpha)\%$ of the intervals constructed will contain the true value of x^* associated with the newly observed measurement, y^*.

95% Confidence Intervals for Predicted Instrument 1 Reading
Simulated Data for 500 Calibration Experiments – First 30 Experiments Shown



True Instrument 1 Reading Is 13 – Coverage for 500 Runs: 93.4%


Alaska Pipeline Ultrasonic Calibration Data
Field Defect Size
20
40
60
80
Lab Defect Size
20
40
60
80


Transformed Alaska Pipeline Data with Fit
log(Field Defect Size)
1.5
2.0
2.5
3.0
3.5
4.0
4.5
log(Lab Defect Size)
1.5
2.0
2.5
3.0
3.5
4.0
4.5

Regression Output for Transformed AK Pipeline Data

```
N = 107

Residual Standard Error = 0.1682604

Multiple R-Square = 0.9337104

F-statistic = 1478.958 on 1 and 105 df, p-value = 0

               coef    std.err    t.stat      p.value
Intercept 0.2813838 0.08092894  3.476924 0.0007390395
 log(LDS) 0.8851754 0.02301714 38.457221 0.0000000000

mean(log(lds)) = 3.444274

106*var(log(lds)) = 53.43934
```

Calibration Interval Computations for Transformed Data

$$FDS^* = 63.227 \Rightarrow \log(FDS^*) = 4.146731$$

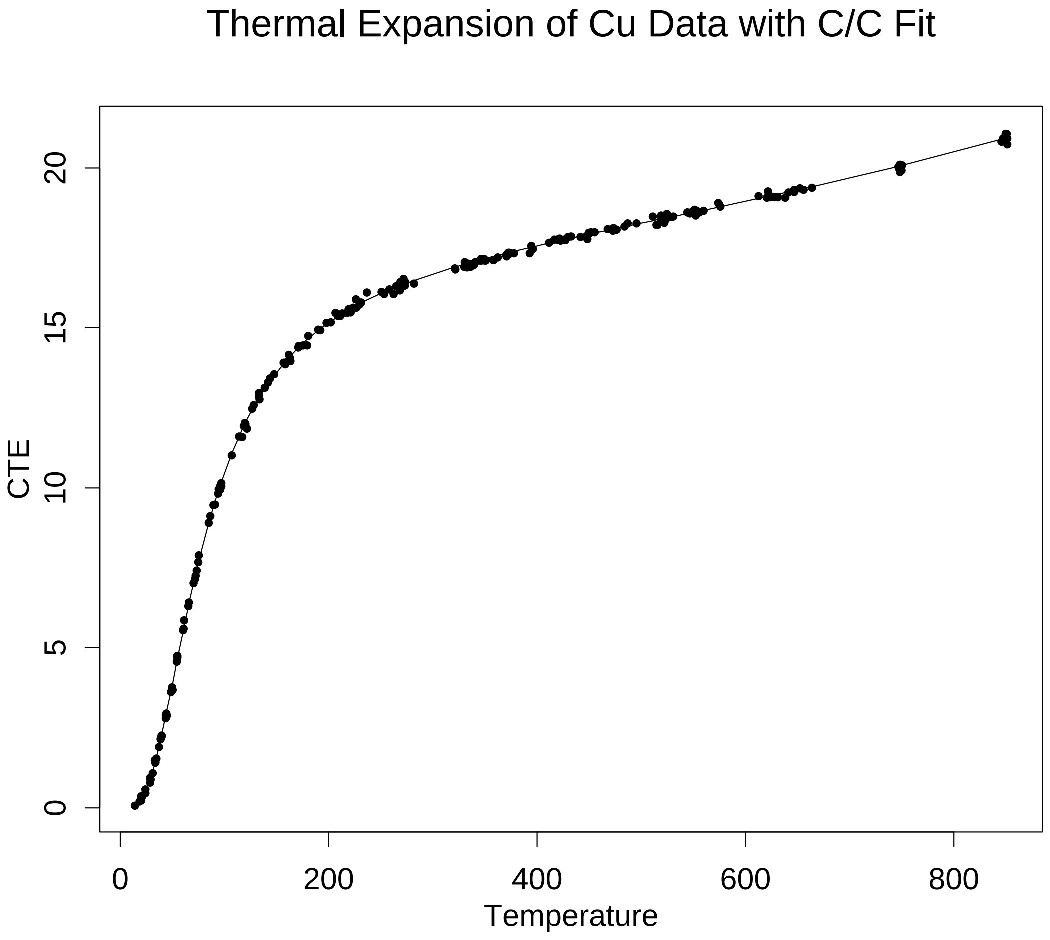
$$\log(L\hat{D}S^*) = \frac{4.146731 - 0.2813838}{0.8851754} = 4.366758$$

$$4.366758 \pm 1.982815\sqrt{\frac{0.1682604^2}{0.8851754^2}(1 + \frac{1}{107} + \frac{(4.366758 - 3.444274)^2}{53.43934})}$$

$$\Downarrow$$

$$4.366758 {\pm} 0.3816400 \mapsto [\exp(3.985118), \exp(4.748398)] \approx [53.79163, 115.3993]$$

$$c = \frac{0.1682604^2 \times 1.982815^2}{0.8851754^2 \times 53.43934} \approx 0.0027$$


Thermal Expansion of Cu Data with C/C Fit
CTE
0
5
10
15
20
Temperature
0
200
400
600
800

Approximate Calibration Intervals by Inversion of Prediction Intervals

The basic steps to this approach of getting a calibration interval are:

1. solve the equation
$$f(x^*; \hat{\beta}_1, \ldots, \hat{\beta}_p) - y^* = 0$$
to get an estimate of x^*,
2. solve the equation
$$f(x^*; \hat{\beta}_1, \ldots, \hat{\beta}_p) + U(x^*) - y^* = 0$$
to get the lower confidence bound for x^*,
3. solve the equation
$$f(x^*; \hat{\beta}_1, \ldots, \hat{\beta}_p) - U(x^*) - y^* = 0$$
to get the uppper confidence bound for x^*,

where $U(x^*) = t_{1-\alpha/2,n-p}\sqrt{s^2 + \vec{d}^{*T} V \vec{d}^*}$ as given on p. 233.

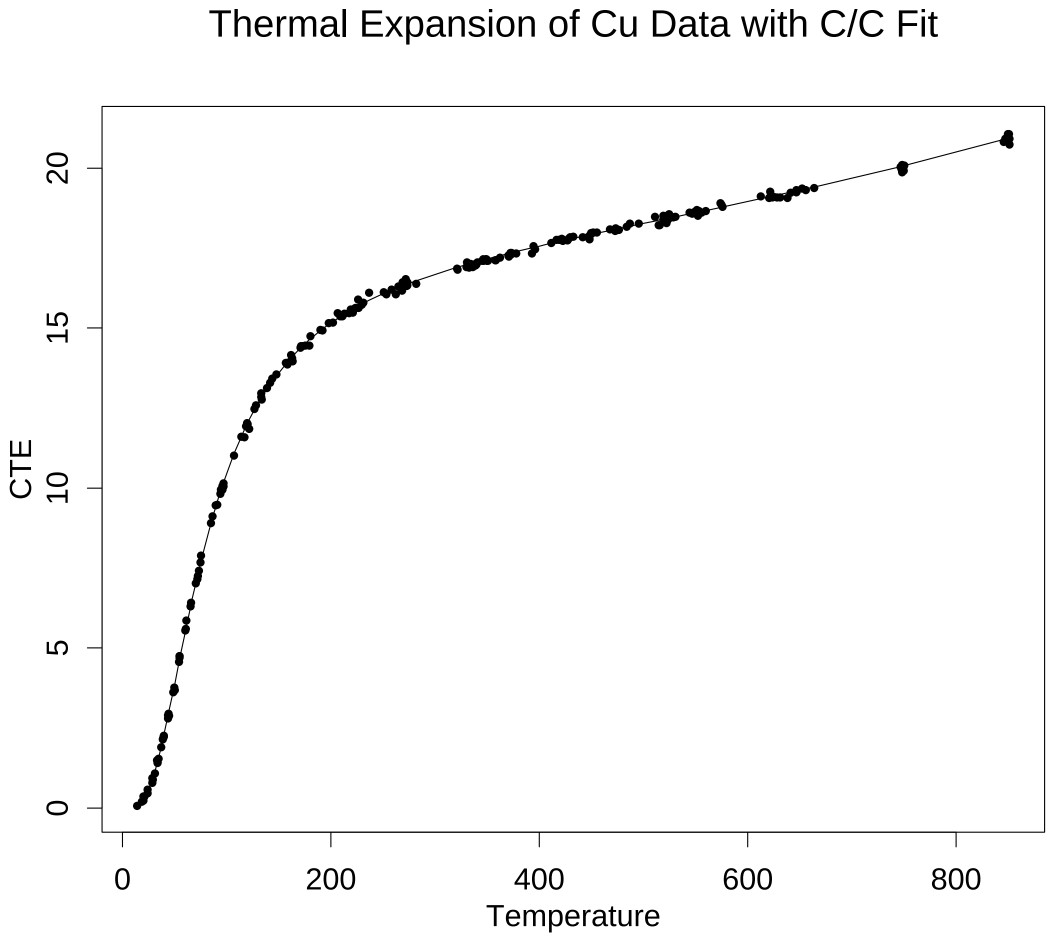
Example for the Cu Data

Assuming a new CTE of 10, a confidence level of 95%, and using general root finding software to solve the equations on the preceding page results in the calibration interval:

$$(93.77088, 95.40549, 97.08308)$$

This method does not provide symmetric intervals in general, but the amount of asymmetry is low.

Simulation of the performance of this interval gave an estimated true coverage of 95% in 1000 trials.

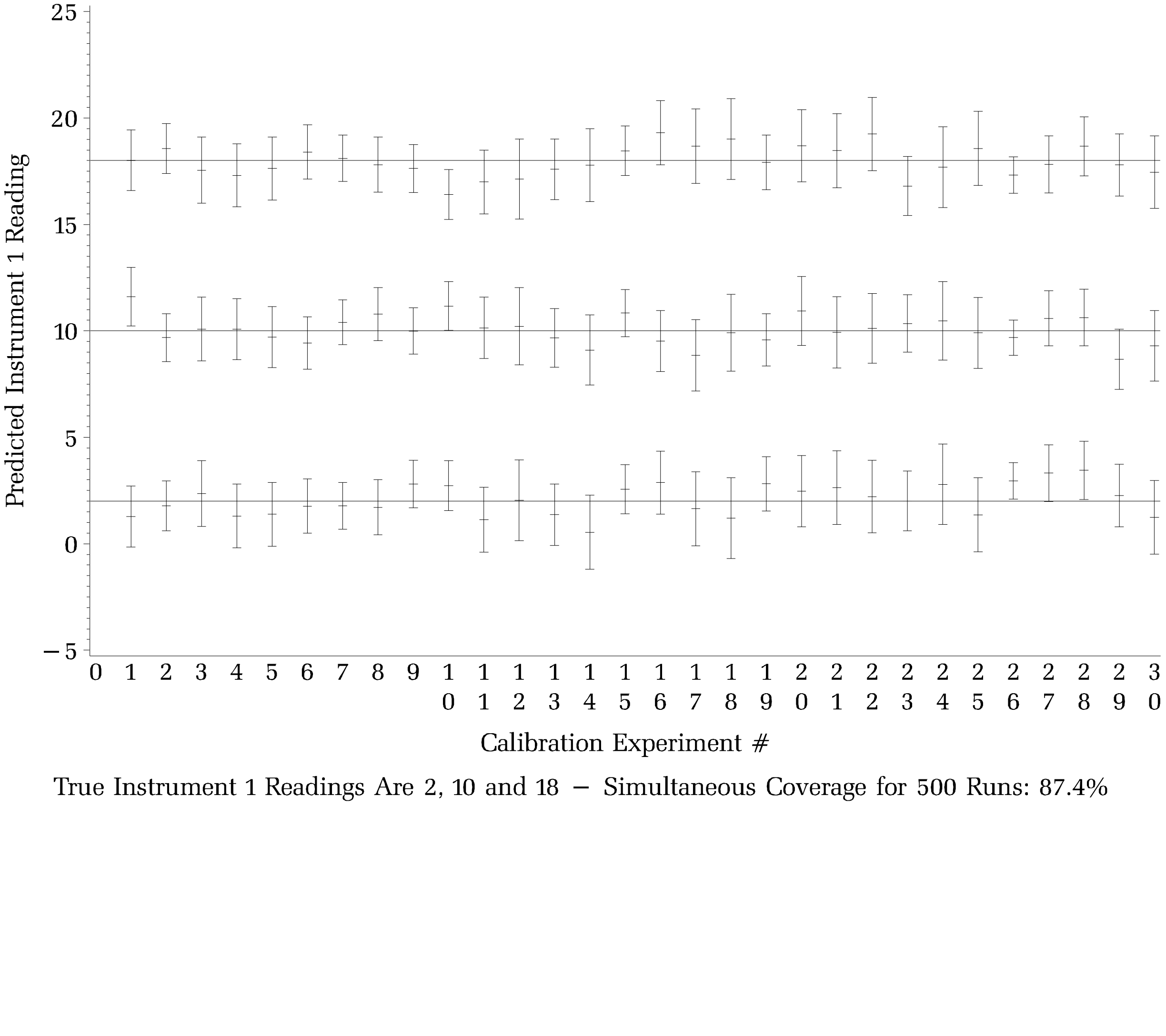

Thermal Expansion of Cu Data with C/C Fit
CTE
0
5
10
15
20
0
200
400
600
800
Temperature

Making Multiple Calibrations

Like the interpretation of both confidence intervals and prediction intervals, the interpretation of calibration intervals guarantee coverage for only one calibration for each set regression data.

If more calibrations are made from each regression data set the stated coverage of the intervals does not hold for all intervals simultaneously.

95% Individual Confidence Intervals for Predicted Instrument 1 Reading
Simulated Data for 500 Calibration Experiments – First 30 Experiments Shown



True Instrument 1 Readings Are 2, 10 and 18 – Simultaneous Coverage for 500 Runs: 87.4%

m Simultaneous Calibration Intervals

As for prediction, one method for computing m simultaneous confidence intervals is to replace the t distribution multiplier in the previous formula for an individual confidence interval with the the smaller of the two factors $t_{1-\alpha/(2m),n-2}$ or $\sqrt{mF_{1-\alpha,m,n-p}}$.

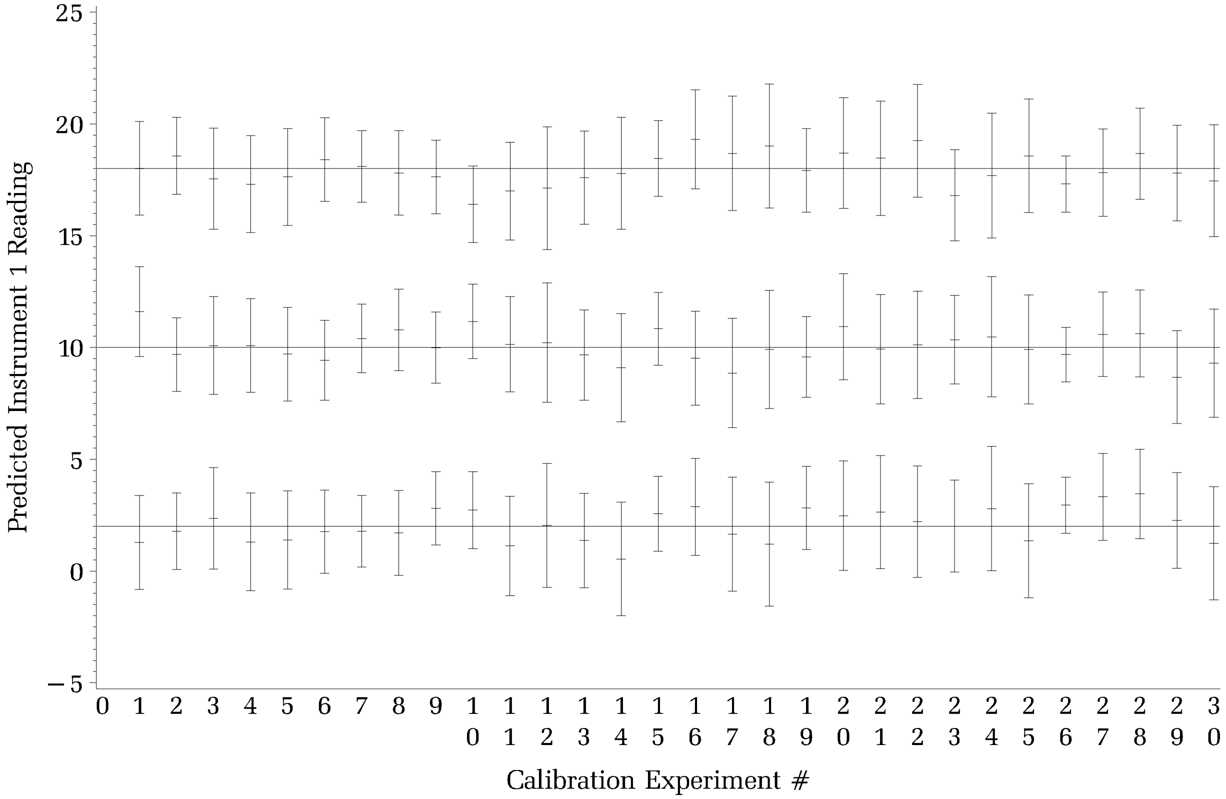
Interpretation of m Simultaneous Calibration Confidence Intervals

Under long term repetition of the steps

1. collect a set of regression data,
2. fit the function f to the data,
3. observe m new responses associated with m (unknown) values of the predictor variable, $x_1^*, \ldots, x_m^*$, and
4. compute $\hat{x}_1^* \pm U_1, \ldots, \hat{x}_m^* \pm U_m$

at least $100(1-\alpha)\%$ of the sets of m intervals constructed will have the property that every interval in the set will contain its corresponding true value of x_i^*.

95% Simultaneous Confidence Intervals for Predicted Instrument 1 Reading
Simulated Data for 500 Calibration Experiments – First 30 Experiments Shown



True Instrument 1 Readings Are 2, 10 and 18 – Simultaneous Coverage for 500 Runs: 98.6%

Section 3: Summary

Two of the main uses of regression models are:

1. prediction of a value of the regression function, f or of the value of a new measurement associated with a particular set of predictor variable values, and
2. calibration of a less-precise measurement technique to a higher precision technique by solving for the unobserved value of the high-precision technique associated with an observed measurement from the lower-precision technique.

For both prediction and calibration, repeated use of the same regression data may need to be accounted for if all prediction or calibration intervals must simultaneously contain their target values with a stated level of confidence.