STA 674

Regression Analysis And Design Of Experiments

Fitting Simple Linear Regression Models – Lecture 6

STA 674, RADOE:
Fitting Simple Linear Regression Models

- Last time: we nailed down the concepts of accuracy, precision and introduced (or recalled, for some) the term standard error
- Now we'll the apply the theoretic results about sampling distributions to data to make inferences in the forms of:
  - Confidence intervals
  - Significance, or hypothesis, tests

STA 674, RADOE:
Fitting Simple Linear Regression Models

Confidence intervals for LS Estimates

- **Definition:** A $(1-\alpha)100\%$ confidence interval for a parameter β is an interval (L, U) such that the probability that (L, U) covers the true value of the parameter is $(1-\alpha)$.

Most common case:

- A 95% confidence interval for the parameter is an interval (L, U) such that the probability that (L, U) covers the true value of β is 0.95.
- (Note that the 95% confidence interval corresponds to $\alpha = 0.05$.)

STA 674, RADOE:
Fitting Simple Linear Regression Models

Review: confidence intervals (CIs) for a population mean

Suppose $X_1, X_2, \ldots, X_n$ are normally distributed with common mean and standard deviation μ and σ.

A $(1-\alpha)100\%$ CI for μ is given by the endpoints:

$$L = \bar{x} - t_{\alpha/2,n-1} \frac{s_x}{\sqrt{n}}$$

$$U = \bar{x} + t_{\alpha/2,n-1} \frac{s_x}{\sqrt{n}}$$

n = number of data in sample
x bar = mean of sample; used to estimate
the population mean mu
df = n-1 (because we estimated mu with x bar, so
it can change?)
t distribution because we don't know sigma (we're using s of sample) and we're estimating mean with n-1 degrees of freedom
sx = standard deviation of sample
sx/sqrt (n) = standard error of sample x bar...sx is used to approximate

STA 674, RADOE:
Fitting Simple Linear Regression Models

Example: confidence interval for a population mean

A quality control inspector is concerned with the average amount of weight that can be held by a type of steel beam. A random sample of five beams is tested with the following amounts of weight added before the beams begin to show stress (in thousands of pounds):

9, 11, 10, 10, 8

Assuming that the population of weights is normally distributed, construct a 95% confidence interval estimate of the population mean weight that can be held.

x bar = 9.6; sx = 1.14; t of 0.025 = 2.7764 (used to multiply sx/sqrt(n) for standard error...used for CI)
LCI = 8.18; UCI = 11.02

STA 674, RADOE:
Fitting Simple Linear Regression Models

Example: confidence interval for a population mean

A quality control inspector is concerned with the average amount of weight that can be held by a type of steel beam. A random sample of five beams is tested with the following amounts of weight added before the beams begin to show stress (in thousands of pounds):

9, 11, 10, 10, 8

Assuming that the population of weights is normally distributed, construct a 95% confidence interval estimate of the population mean weight that can be held.

$\bar{x} = 9.6$; $s_x = 1.14$; and $t_{0.025,4} = 2.7764$

STA 674, RADOE:

Fitting Simple Linear Regression Models

Example: confidence interval for a population mean

A quality control inspector is concerned with the average amount of weight that can be held by a type of steel beam. A random sample of five beams is tested with the following amounts of weight added before the beams begin to show stress (in thousands of pounds):

9, 11, 10, 10, 8

Assuming that the population of weights is normally distributed, construct a 95% confidence interval estimate of the population mean weight that can be held.

$\bar{x} = 9.6$; $s_x = 1.14$; and $t_{0.025,4} = 2.7764$

$L = 8.18$ and $U = 11.02$

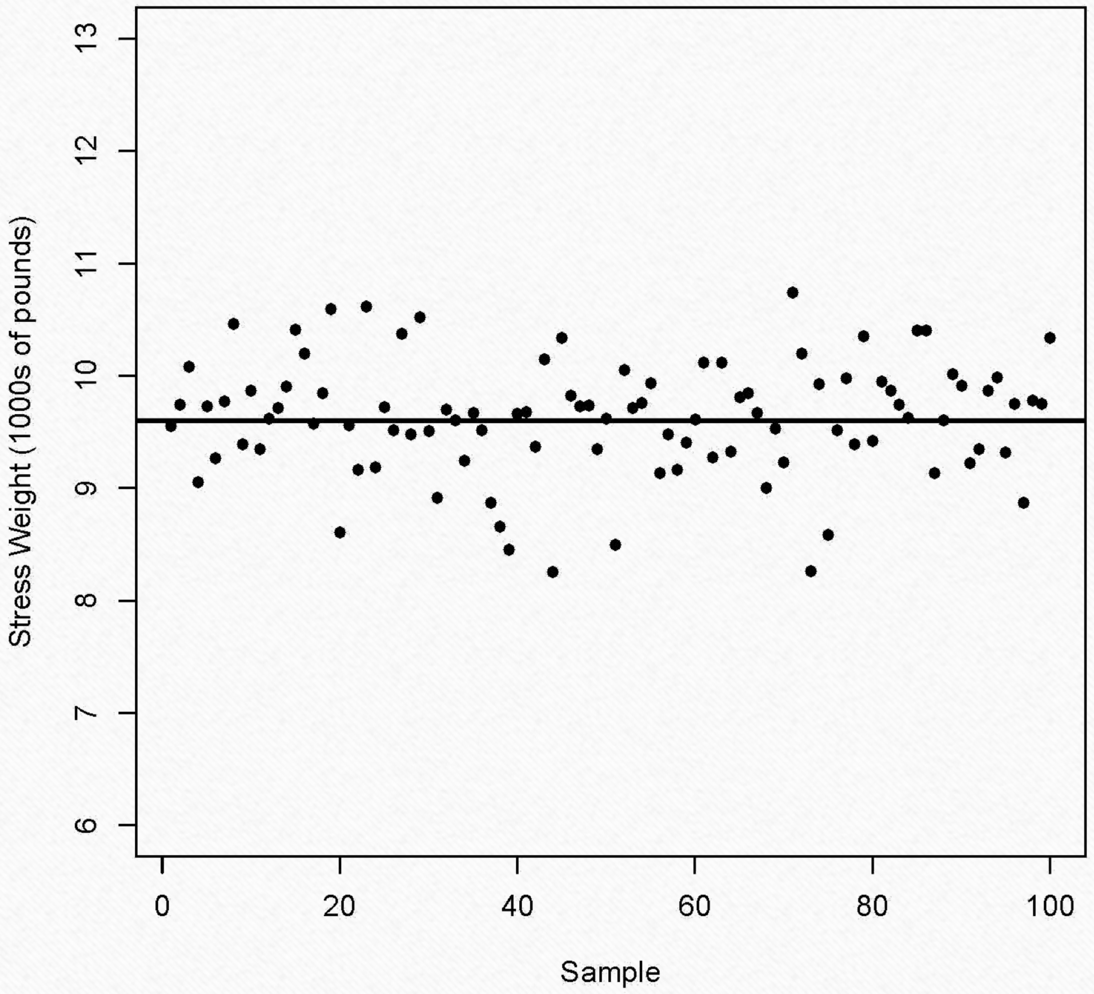
Basic Confidence Limits Assuming Normality			
Parameter	Estimate	95% Confidence Limits	
Mean	9.60000	8.18429	11.01571

STA 674, RADOE:
Fitting Simple Linear Regression Models

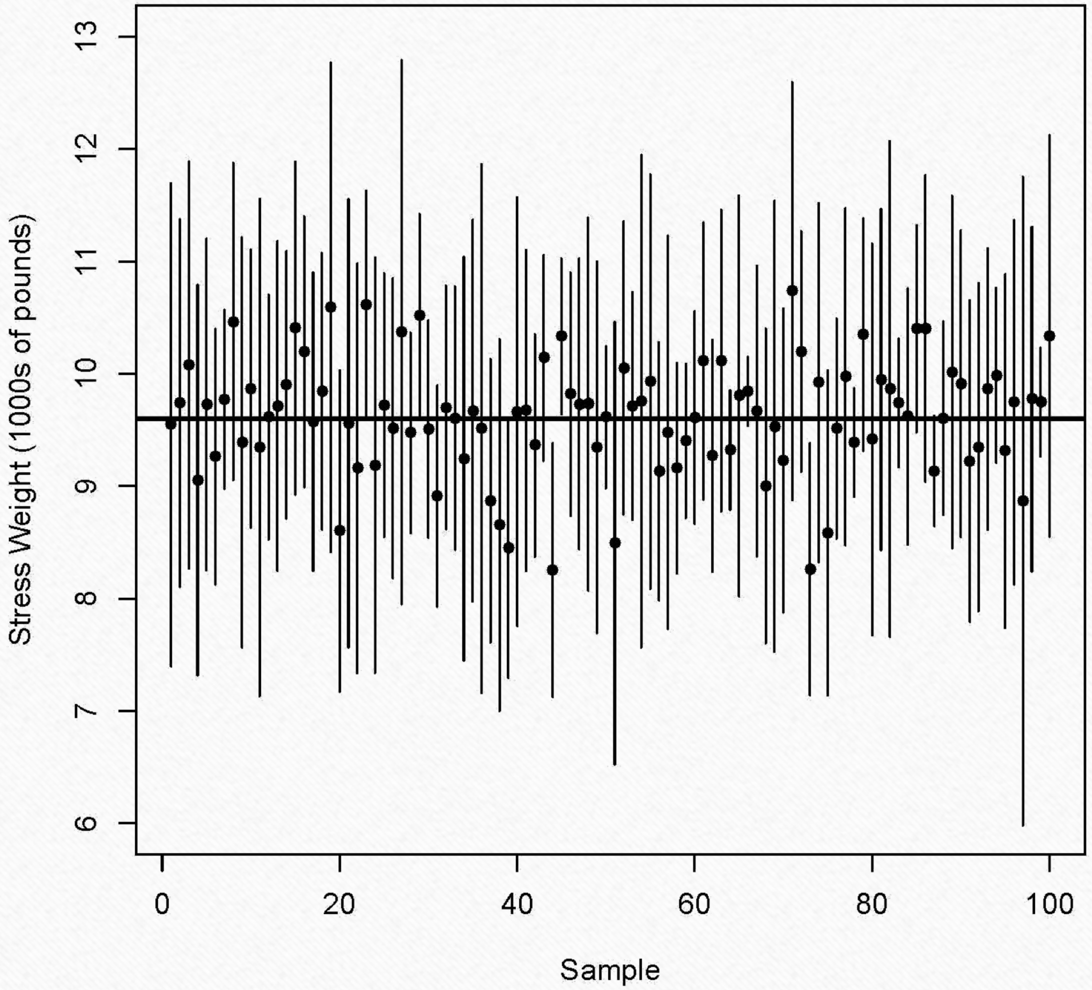
Interpretation of confidence intervals (for the population mean)

- A probability interpretation: if you were to repeat the experiment many, many times and compute the $(1-\alpha)100\%$ interval for each data set then $(1-\alpha)100\%$ of the intervals would cover the population mean.

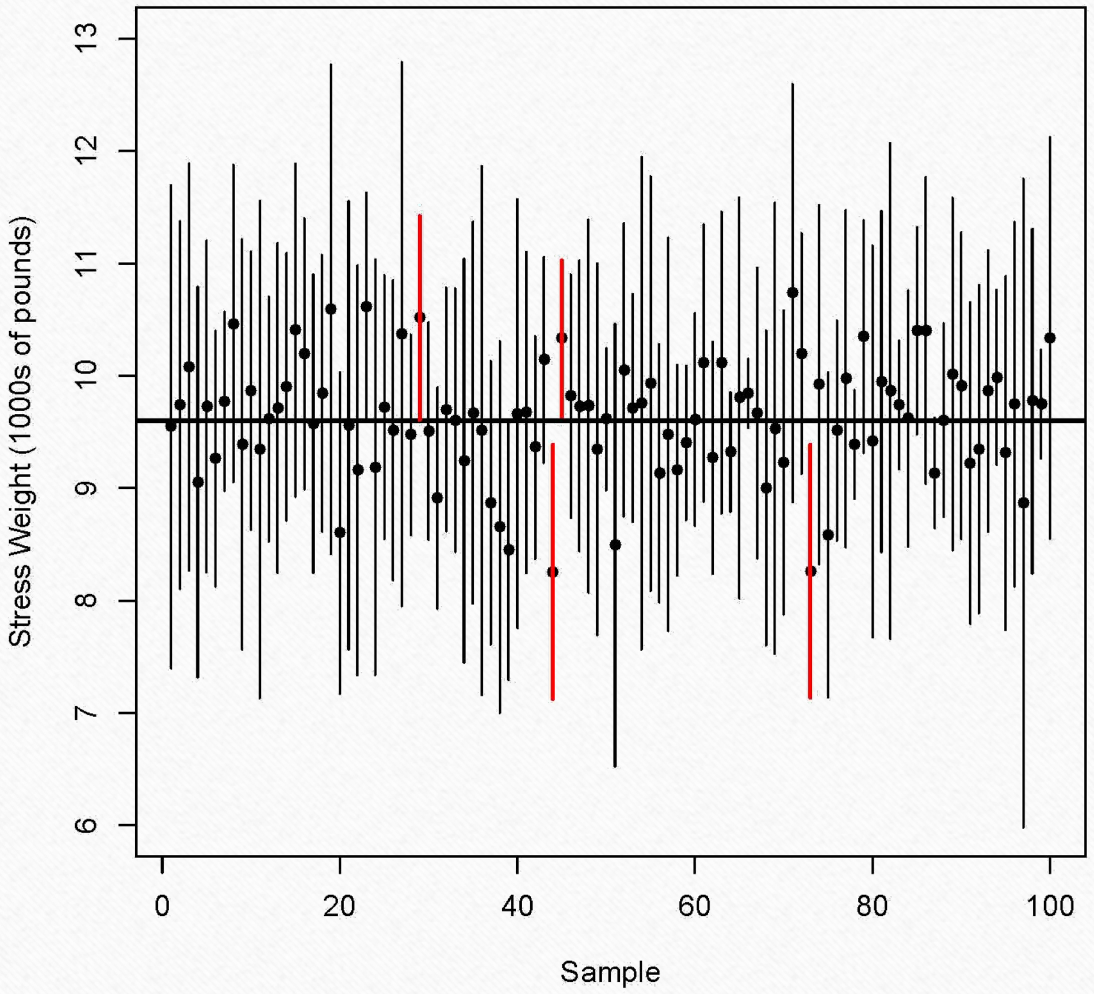
Probability interpretation



Probability interpretation

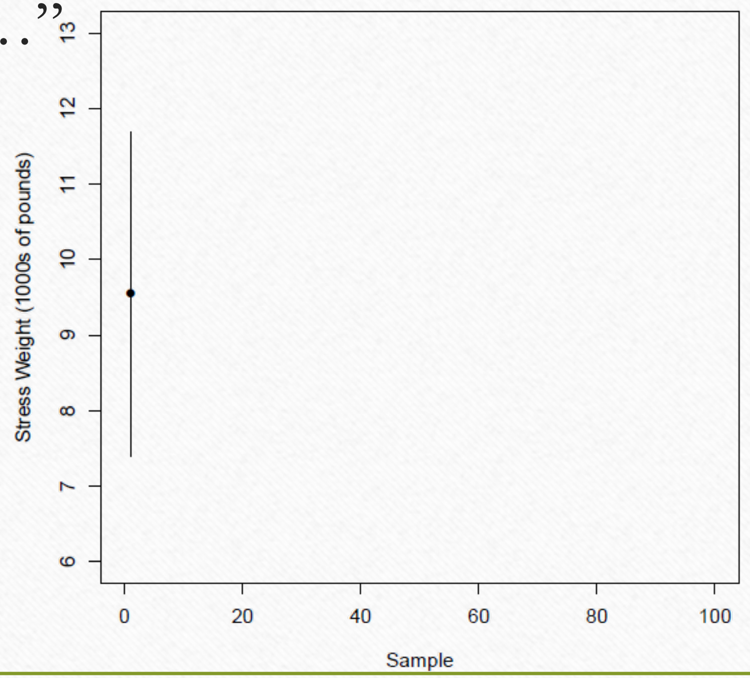


Probability interpretation



STA 674, RADOE:
Fitting Simple Linear Regression Models

- Problem with a probability interpretation: it started with “if you were to repeat the experiment many, many times …”



STA 674, RADOE:
Fitting Simple Linear Regression Models

Interpretation of confidence intervals (for the population mean)

- A ***probability*** interpretation: if you were to repeat the experiment many, many times and compute the $(1-\alpha)100\%$ interval for each data set then $(1-\alpha)100\%$ of the intervals would cover the population mean.
- A ***heuristic*** Interpretation: we cannot say whether or not the confidence interval for a single data set covers the population mean. However, since 95% of intervals like this cover the mean it seems reasonable to assume that the interval we computed covers the mean. We say that we are 95% confident that the interval covers the population mean.
- The values inside the 95% confidence interval represent a *reasonable* range of guesses for the *population* mean.

Confidence is in the process of constructing the interval