

# Assumptions and More Inference for Simple Linear Regression (Ch. 9.1)

Dason Kurkiewicz

Iowa State University

July 2, 2013

# Outline

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

Assumptions for SLR

SLR: Inference for the Mean Response at some  $x$

Simultaneous Confidence Intervals for  $\mu_{y|x}$

- Recall our model:

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

$$\epsilon_1, \dots, \epsilon_n \sim \text{iid } N(0, \sigma^2)$$

- Under the model, the true mean response at some observed covariate value  $x_i$  is:

$$\mu_{y|x_i} = \beta_0 + \beta_1 x_i$$

# Assumptions in SLR

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

- ▶ Correct Model (sometimes called 'Linearity')
- ▶ Constant Variance
- ▶ Normally distributed error terms
- ▶ Independence

# Correct Model

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

- ▶ We assume that we have specified a correct functional form relating the predictor to the response.
- ▶ To check this we can use:
  - ▶ Scatterplot
  - ▶ Residual by Predicted plot
  - ▶ Residual by Predictor ( $x$ ) plot

# Correct Model

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

- ▶ The Residual by Predicted plot is one of the best ways to check these assumptions.
- ▶ Things we want to see
  - ▶ A horizontal band of points with no discernable pattern left over
- ▶ Things we don't want to see
  - ▶ Left over pattern!
  - ▶ Any curvature left over - implies we should fit a higher order polynomial
  - ▶ A sinusoidal pattern
  - ▶ ANY pattern

# Correct Model

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

# Constant Variance

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

- ▶ It is assumed that  $\epsilon_i \sim N(0, \sigma^2)$
- ▶ Notice that the error terms all have the same variance
- ▶ Once again we can use the residual by predicted plot to check this assumption
- ▶ The residuals themselves actually have a slightly different variance so it is typical to 'standardize' them.

$$e_i \sim N\left(0, \sigma^2\left(1 - \frac{1}{n} - \frac{(x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2}\right)\right)$$



# Constant Variance

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

# Normally distributed errors

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

- ▶ It is assumed that  $e_i \sim N(0, \sigma^2)$
- ▶ We should check if the errors are normally distributed
- ▶ We will use the residuals as estimates for the errors and check for approximate normality.
- ▶ Make a normal qq-plot for the residuals (or standardized residuals)

# Independence

- ▶ The most important assumption
- ▶ It is also the hardest to check in practice
- ▶ Typically have to rely on what you know about the data
- ▶ Things to watch for:
  - ▶ Data collected through time - consecutive points might be more correlated than points separated by larger periods of time.
  - ▶ Units that have a natural grouping.
    - ▶ Maybe you have multiple measurements of multiple experimental units. These measurements shouldn't be considered independent.

# Outline

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

Assumptions for SLR

SLR: Inference for the Mean Response at some  $x$

Simultaneous Confidence Intervals for  $\mu_{y|x}$

# SLR: mean response at $x$

- ▶ Recall our model:

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

$$\varepsilon_1, \dots, \varepsilon_n \sim \text{iid } N(0, \sigma^2)$$

- ▶ Under the model, the true mean response at some observed covariate value  $x_i$  is:

$$\mu_{y|x_i} = \beta_0 + \beta_1 x_i$$

- ▶ Now, if some new covariate value  $x$  is within the range of the  $x_i$ 's, we can estimate the true mean response at this new  $x$ :

$$\hat{\mu}_{y|x} = b_0 + b_1 x$$

# SLR: mean response at $x$

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

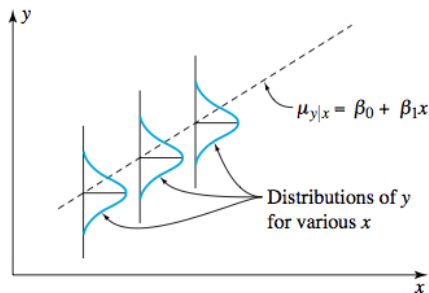
Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

- ▶ But how good is the estimate?



- ▶ That's why we do inference.

## SLR: mean response at $x$

- Under the model,  $\hat{\mu}_{y|x}$  is normally distributed with:

$$E(\hat{\mu}_{y|x}) = \mu_{y|x} = \beta_0 + \beta_1 x$$
$$Var(\hat{\mu}_{y|x}) = \sigma^2 \left( \frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_i (x_i - \bar{x})^2} \right)$$

- We can construct a  $N(0, 1)$  random variable by standardizing:

$$Z = \frac{\hat{\mu}_{y|x} - \mu_{y|x}}{\sigma \sqrt{\frac{1}{n} \frac{(x - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}}} \sim N(0, 1)$$

- Replacing  $\sigma$  with  $s_{LF} = \sqrt{\frac{1}{n-2} \sum_i (y_i - \hat{y}_i)^2}$ :

$$T = \frac{\hat{\mu}_{y|x} - \mu_{y|x}}{s_{LF} \sqrt{\frac{1}{n} \frac{(x - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}}} \sim t_{n-2}$$

## SLR: mean response at $x$

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

- ▶ To test  $H_0 : \mu_{y|x} = \#$ , we can use the test statistic:

$$K = \frac{\hat{\mu}_{y|x} - \#}{s_{LF} \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}}}$$

which has a  $t_{n-2}$  distribution if  $H_0$  is true and the model is correct.

- ▶ A 2-sided  $1 - \alpha$  confidence interval for  $\mu_{y|x}$  is:

$$\left( \hat{\mu}_{y|x} - t_{n-2, 1-\alpha/2} \cdot s_{LF} \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}}, \right. \\ \left. \hat{\mu}_{y|x} + t_{n-2, 1-\alpha/2} \cdot s_{LF} \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}} \right)$$

and the one-sided intervals are analogous.



## Pressing pressures and specimen densities for a ceramic compound

A mixture of  $\text{Al}_2\text{O}_3$ , polyvinyl alcohol, and water was prepared, dried overnight, crushed, and sieved to obtain 100 mesh size grains. These were pressed into cylinders at pressures from 2,000 psi to 10,000 psi, and cylinder densities were calculated.

x (pressure in psi)	y (density in g/cc)
2000.00	2.49
2000.00	2.48
2000.00	2.47
4000.00	2.56
4000.00	2.57
4000.00	2.58
6000.00	2.65
6000.00	2.66
6000.00	2.65
8000.00	2.72
8000.00	2.77
8000.00	2.81
10000.00	2.86
10000.00	2.88
10000.00	2.86

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
x

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

## Example: ceramics

- First, I'll make a 2-sided 95% confidence interval for the true mean density of the ceramics at 4000 psi.

$$\hat{\mu}_{y|x} = 2.375 + 0.0000487(4000) = 2.5697 \text{ g/cc}$$

With  $t_{n-2, 1-\alpha/2} = t_{13, 0.975} = 2.160$ , the margin of error in the confidence interval is:

$$\begin{aligned} t_{n-2, 1-\alpha/2} \cdot s_{LF} \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}} \\ = 2.160(0.0199) \sqrt{\frac{1}{15} + \frac{(4000 - 6000)^2}{1.2 \times 10^8}} = 0.0136 \text{ g/cc} \end{aligned}$$

Hence, the 95% CI is:

$$(2.5697 - 0.0136, 2.5697 + 0.0136) = (2.5561, 2.5833)$$

- We're 95% confident that the true mean density of the ceramics at 4000 psi is between 2.5561 g/cc and 2.5833 g/cc.

# Your turn: ceramics

- ▶ Calculate and interpret a 2-sided 95% confidence interval for the true mean density at 5000 psi, given:
  - ▶  $\hat{\mu}_{y|x} = 2.375 + 0.0000487x$
  - ▶ The margin of error is  $t_{n-2, 1-\alpha/2} s_{LF} \sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{\sum_i (x_i - \bar{x})^2}}$
  - ▶  $\sum_i (x_i - \bar{x})^2 = 1.2 \times 10^8$
  - ▶  $n = 15, \bar{x} = 6000$ .
  - ▶  $s_{LF} = 0.0199$
  - ▶  $t_{13, 0.975} = 2.16$
- ▶ Test  $H_0 : \beta_0 = 0$  vs.  $H_a : \beta_0 \neq 0$  at significance level  $\alpha = 0.05$  using the method of p-values.

## Answers: ceramics

- ▶ Make a 2-sided 95% confidence interval for the true mean density of the ceramics at 5000 psi:

$$\hat{\mu}_{y|x} = 2.375 + 0.0000487(5000) = 2.6183 \text{ g/cc}$$

With  $t_{n-2, 1-\alpha/2} = t_{13, 0.975} = 2.160$ , the margin of error in the confidence interval is:

$$\begin{aligned} t_{n-2, 1-\alpha/2} \cdot s_{LF} \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}} \\ = 2.160(0.0199) \sqrt{\frac{1}{15} + \frac{(5000 - 6000)^2}{1.2 \times 10^8}} = 0.0118 \text{ g/cc} \end{aligned}$$

Hence, the 95% CI is:

$$(2.6183 - 0.0118, 2.6183 + 0.0118) = (2.6065, 2.6301)$$

- ▶ We're 95% confident that the true mean density of the ceramics at 5000 psi is between 2.6065 g/cc and 2.6301 g/cc.

# Answers: ceramics

Now for the hypothesis test:

1.  $H_0 : \beta_0 = 0, H_a : \beta_0 \neq 0$
2.  $\alpha = 0.05$
3.  $\beta_0$  is just  $\mu_{y|x=0}$ . The test statistic is:

$$K = \frac{b_0 - 0}{s_{LF} \sqrt{\frac{1}{n} + \frac{(0 - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}}} = \frac{b_0}{s_{LF} \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum_i (x_i - \bar{x})^2}}}$$

- ▶  $K \sim t_{n-2}$  assuming:
  - ▶  $H_0$  is true.
  - ▶ The model  $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$  is correct, with  $\varepsilon_1, \dots, \varepsilon_n \sim \text{iid } N(0, 1)$ .

## Answers: ceramics

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

4. The moment of truth:

$$b_0 = 2.375$$

$$K = \frac{2.375}{0.0199 \sqrt{\frac{1}{15} + \frac{6000^2}{1.2 \times 10^8}}} = 197.09$$

$$\text{p-value} = P(|t_{13}| > 197.09) \ll 0.0001$$

5. With a p-value  $\ll 0.0001 < \alpha$ , we reject  $H_0$  and conclude  $H_a$ .
6. There is overwhelming evidence that the intercept of the “true” line is different from 0.

# Ceramics: back to the JMP output

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

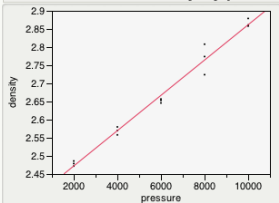
Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

## ▼ Bivariate Fit of density By pressure



Linear Fit

## ▼ Linear Fit

density = 2.375 + 4.8667e-5\*pressure

## ▼ Summary of Fit

RSquare	0.982193
RSquare Adj	0.980824
Root Mean Square Error	0.019909
Mean of Response	2.667
Observations (or Sum Wgts)	15

## ▼ Lack Of Fit

## ▼ Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	0.28421333	0.284213	717.0604
Error	13	0.00515267	0.000396	Prob > F
C. Total	14	0.28936600		<.0001*

## ▼ Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	2.375	0.012055	197.01	<.0001*
pressure	4.8667e-5	1.817e-6	26.78	<.0001*

# Ceramics: back to the JMP output

## ▼ Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	2.375	0.012055	197.01	<.0001*
pressure	4.8667e-5	1.817e-6	26.78	<.0001*

- ▶ The test statistic  $K$  is under “t Ratio” for the intercept.
- ▶ “Prob> |t|” for the intercept is the p-value for the significance test you just did.
- ▶ “Estimate” for the intercept is  $b_0$ .
- ▶ “Std Error” for the intercept is:

$$\widehat{SD}(b_0) = s_{LF} \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{\sum_i (x_i - \bar{x})^2}}$$



## Be careful with Inference on $\beta_0$

- ▶ In this case and many others,  $\beta_0 = \mu_{y|x=0}$  is beyond the range of our data.
- ▶ Estimating beyond the range of our covariate values is called **extrapolation**, which is dangerous for linear regression.
- ▶ Only extrapolate when:
  - ▶ You know your process or system well, and can describe it with the right differential equations.
  - ▶ You estimate the parameters of the resulting model using *nonlinear* regression:
    - ▶ Example: special case of the Michaelis-Menten model for enzyme kinetics with reaction speed  $y$  and substrate concentration  $x$ :

$$Y_i = \frac{\theta_1 x_i}{\theta_2 + x_i} + \varepsilon_i$$

- ▶ See Nonlinear Regression Analysis and Its Applications by Bates and Watts for more information on nonlinear regression.

# Outline

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

Assumptions for SLR

SLR: Inference for the Mean Response at some  $x$

Simultaneous Confidence Intervals for  $\mu_{y|x}$

# Simultaneous confidence intervals

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

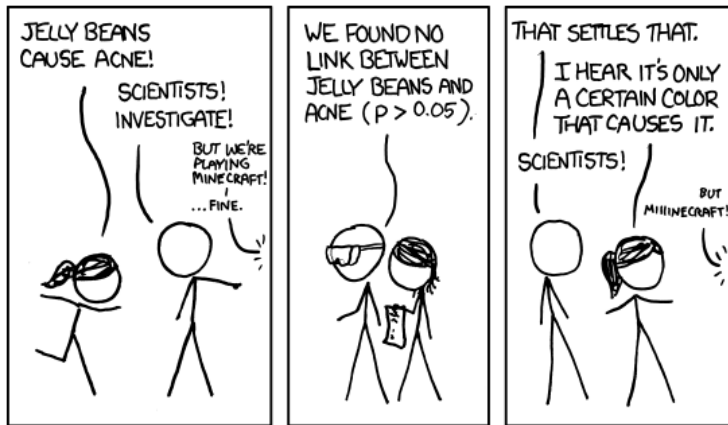
Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

- ▶ Situations will arise when you'll want to do inference on  $\mu_{y|x=2000}, \mu_{y|x=4000}, \mu_{y|x=6000}, \dots$ , all at once.
- ▶ When you compute several confidence intervals at once or do multiple tests at once, you need to account for the simultaneity.
- ▶ On average, for every 20 tests you do at  $\alpha = 0.05$ , we expect 1 of those tests to conclude  $H_a$  by chance alone.
  - ▶ Remember:  $\alpha = P(\text{reject } H_0 \text{ assuming } H_0 \text{ is true})$ .

Example: <http://xkcd.com/882/>



Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

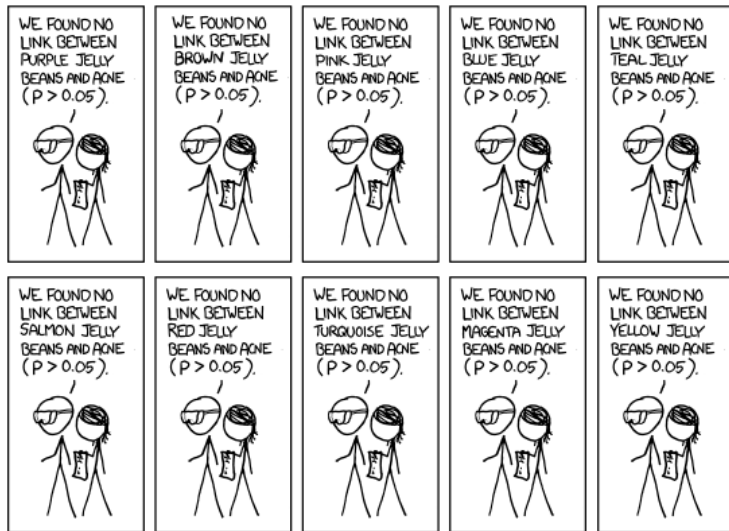
Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

Example: <http://xkcd.com/882/>



Assumptions and More Inference for Simple Linear Regression (Ch. 9.1)

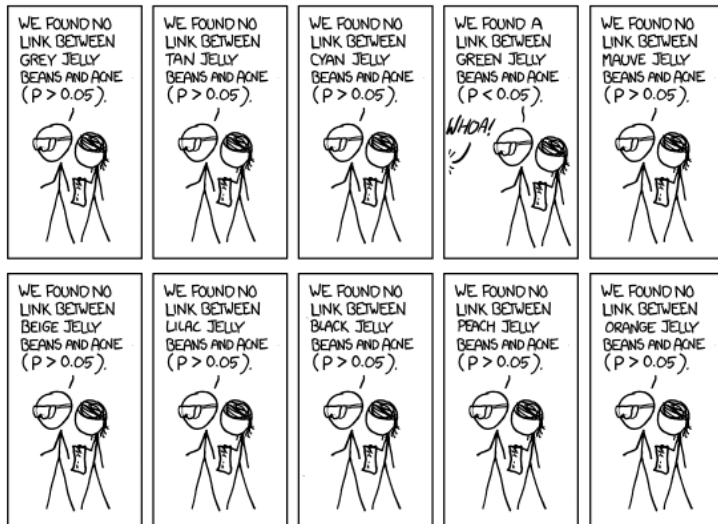
Dason Kurkiewicz

Assumptions for SLR

SLR: Inference for the Mean Response at some  $x$

Simultaneous Confidence Intervals for  $\mu_{y|x}$

Example: <http://xkcd.com/882/>



Assumptions and More Inference for Simple Linear Regression (Ch. 9.1)

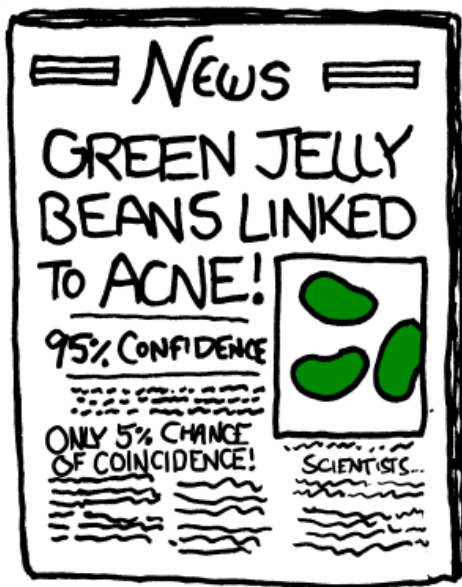
Dason Kurkiewicz

Assumptions for SLR

SLR: Inference for the Mean Response at some  $x$

Simultaneous Confidence Intervals for  $\mu_{y|x}$

Example: <http://xkcd.com/882/>



Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

# Simultaneous confidence intervals for $\mu_{y|x}$

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

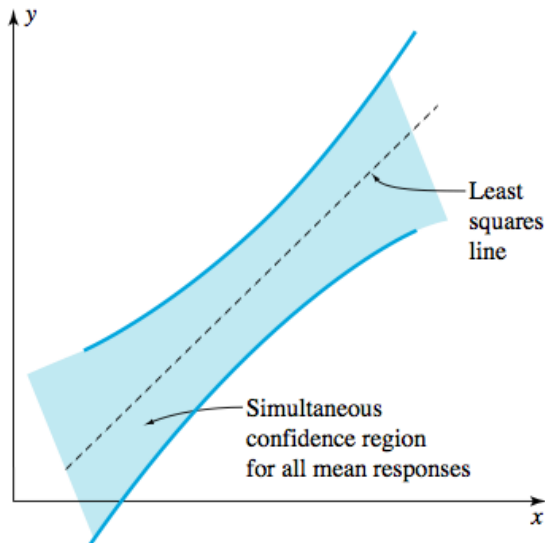
- ▶ For simultaneous confidence intervals for  $\mu_{y|x}$  for multiple values of  $x$ , use:

$$b_0 + b_1x \pm \sqrt{2F_{2,n-2,1-\alpha}} \cdot s_{LF} \cdot \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}}$$

- ▶ This formula accounts for the fact that we're computing  $k$  confidence intervals at the same time.



# Simultaneous confidence intervals for $\mu_{y|x}$



Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

## Example: ceramics

► Given:

- $n = 15$
- $\bar{x} = 6000$
- $\sum_i (x_i - \bar{x})^2 = 1.2 \times 10^8$
- $\hat{y} = 2.375 + 4.87 \times 10^{-5}x$ ,  $s_{LF} = 0.0199$ .
- The simultaneous confidence interval formula is:

$$b_0 + b_1x \pm \sqrt{2F_{2,n-2,1-\alpha/2} \cdot s_{LF} \cdot \sqrt{\frac{1}{n} + \frac{(x - \bar{x})^2}{\sum_i (x_i - \bar{x})^2}}}$$

- I will calculate simultaneous 95% confidence intervals for the mean responses  $\mu_{y|x}$  at  $x = 2000, 4000, 6000, 8000$ , and  $10000$ .

## Example: ceramics

- Using  $F_{2,n-2,1-\alpha} = F_{2,13,0.95} = 3.81$ , the intervals are of the form:

$$2.375 + 4.87 \times 10^{-5}x \pm \sqrt{2 \cdot 3.81 \cdot 0.0199} \cdot \sqrt{\frac{1}{15} + \frac{(x - 6000)^2}{1.2 \times 10^8}}$$
$$= 2.375 + 4.87 \times 10^{-5}x \pm 0.0549 \sqrt{0.066 + 8.33 \times 10^{-9}(x - 6000)^2}$$

x, pressure	CI, compact form	CI
2000	$2.4723 \pm 0.0246$	(2.4477, 2.4969)
4000	$2.5697 \pm 0.0174$	(2.5523, 2.5871)
6000	$2.6670 \pm 0.0142$	(2.6528, 2.6812)
8000	$2.7643 \pm 0.0174$	(2.7469, 2.7817)
10000	$2.8617 \pm 0.0246$	(2.8371, 2.8863)

# Ceramics; plotting simultaneous confidence regions

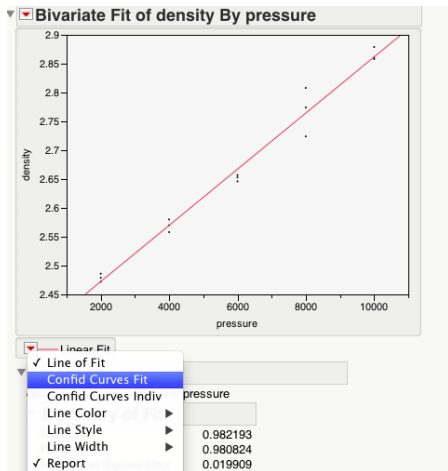
Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$



# Ceramics; plotting simultaneous confidence regions

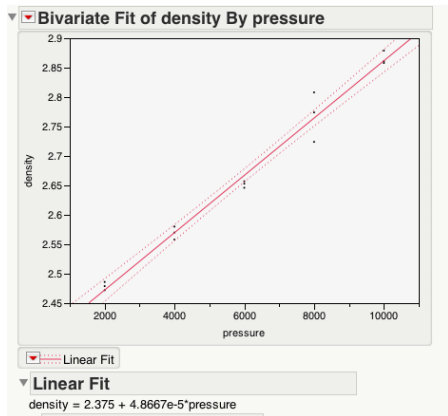
Assumptions and More Inference for Simple Linear Regression (Ch. 9.1)

Dason Kurkiewicz

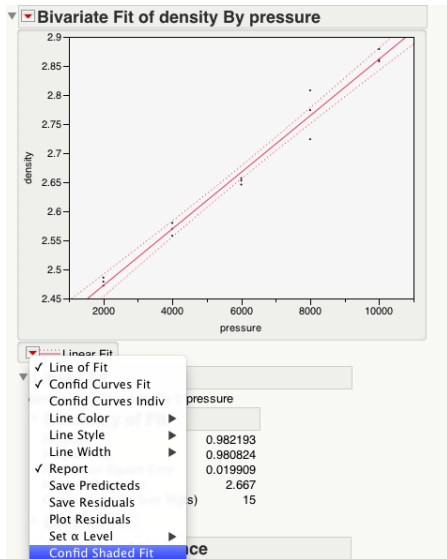
Assumptions for SLR

SLR: Inference for the Mean Response at some  $x$

Simultaneous Confidence Intervals for  $\mu_{y|x}$



# Ceramics; plotting simultaneous confidence regions



Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

# Ceramics; plotting simultaneous confidence regions

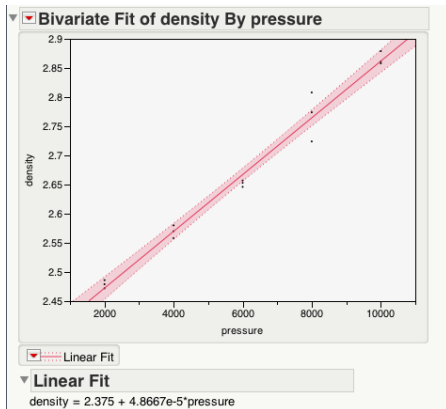
Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$



# Ceramics: calculating the margins of error in JMP

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

The screenshot shows the JMP software interface. On the left, a dialog box titled 'Simultaneous conf. margin err' is open. It has fields for 'Column Name' (Simultaneous conf. margin err), 'Data Type' (Numeric), 'Modeling Type' (Continuous), and 'Format' (Best). There are also checkboxes for 'Lock', 'Use thousands separator', and 'Initialize Data' (Missing/Empty). At the bottom, there is a 'Formula' section with a text area containing 'optional item' and buttons for 'Edit Formula', 'Suppress Eval', and 'Ignore Errors'. On the right, the 'ceramics.jmp' data table is visible. It has columns for 'pressure', 'density', and 'Simultaneous conf. margin'. The data is organized into rows, with a summary section at the bottom showing counts for 'All rows', 'Selected', 'Excluded', 'Hidden', and 'Labelled'.

	pressure	density	Simultaneous conf. margin
1	2000	2.486	*
2	2000	2.479	*
3	2000	2.472	*
4	4000	2.558	*
5	4000	2.57	*
6	4000	2.58	*
7	6000	2.646	*
8	6000	2.657	*
9	6000	2.653	*
10	8000	2.724	*
11	8000	2.774	*
12	8000	2.808	*
13	10000	2.861	*
14	10000	2.879	*
15	10000	2.858	*



# Ceramics: calculating the margins of error in JMP

Simultaneous conf. margin err

Table Columns ▼

- pressure
- density
- Simultaneous confidence intervals

Functions (grouped) ▼

- Row
- Numeric
- Transcendental
- Trigonometric
- Character
- Comparison
- Conditional
- Probability
- Discrete Probability

OK

Cancel

Apply

Clear

Help

0.0549 \*  $\sqrt{0.066 + 0.00000000888 * (pressure - 6000)^2}$

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

# Ceramics: calculating the margins of error in JMP

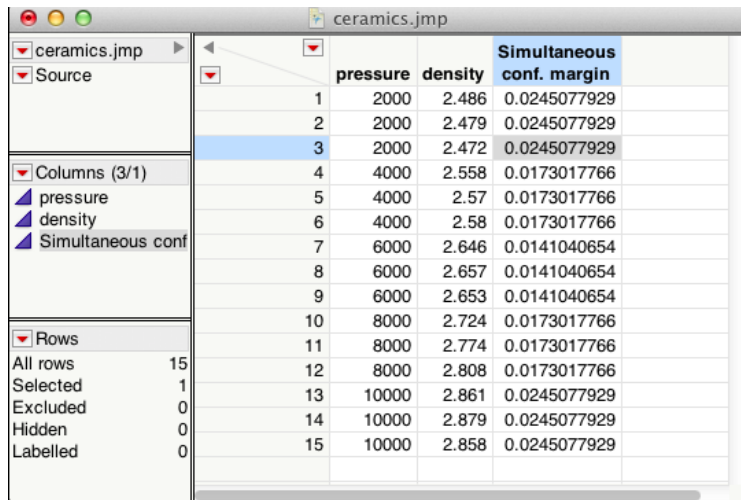
Assumptions and More Inference for Simple Linear Regression (Ch. 9.1)

Dason Kurkiewicz

Assumptions for SLR

SLR: Inference for the Mean Response at some  $x$

Simultaneous Confidence Intervals for  $\mu_{y|x}$



	pressure	density	Simultaneous conf. margin
1	2000	2.486	0.0245077929
2	2000	2.479	0.0245077929
3	2000	2.472	0.0245077929
4	4000	2.558	0.0173017766
5	4000	2.57	0.0173017766
6	4000	2.58	0.0173017766
7	6000	2.646	0.0141040654
8	6000	2.657	0.0141040654
9	6000	2.653	0.0141040654
10	8000	2.724	0.0173017766
11	8000	2.774	0.0173017766
12	8000	2.808	0.0173017766
13	10000	2.861	0.0245077929
14	10000	2.879	0.0245077929
15	10000	2.858	0.0245077929

# Ceramics: calculating the margins of error in JMP

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

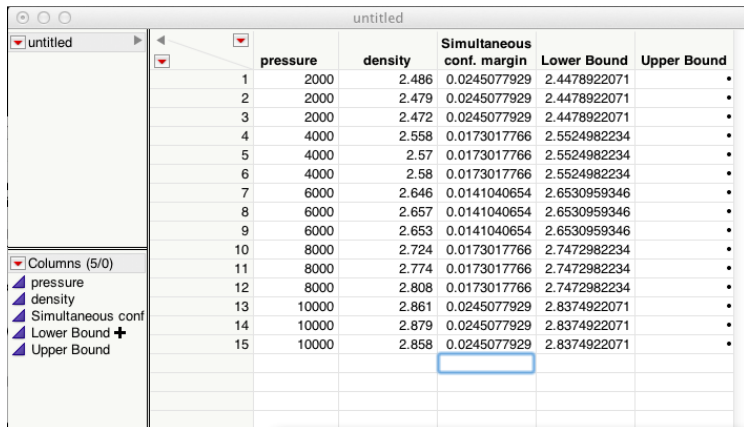
Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

The screenshot shows the JMP software interface. On the left, the 'Column Properties' dialog box for 'Lower Bound' is open, showing settings for Data Type (Numeric), Modeling Type (Continuous), and Format (Best). The 'Formula' tab is selected, showing the formula  $2.375 + 0.0000487 * \text{pressure}$ . The 'Lower Bound' dialog box is also open, showing the 'Table Columns' list with 'Simultaneous confidence intervals' selected. The 'Functions (grouped)' list is open, showing various mathematical functions. The 'Formula' editor at the bottom shows the formula  $2.375 + 0.0000487 * \text{pressure} - \text{Simultaneous conf. margin}$ .

# Ceramics: calculating the margins of error in JMP

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz



untitled

	pressure	density	Simultaneous conf. margin	Lower Bound	Upper Bound
1	2000	2.486	0.0245077929	2.4478922071	•
2	2000	2.479	0.0245077929	2.4478922071	•
3	2000	2.472	0.0245077929	2.4478922071	•
4	4000	2.558	0.0173017766	2.5524982234	•
5	4000	2.57	0.0173017766	2.5524982234	•
6	4000	2.58	0.0173017766	2.5524982234	•
7	6000	2.646	0.0141040654	2.6530959346	•
8	6000	2.657	0.0141040654	2.6530959346	•
9	6000	2.653	0.0141040654	2.6530959346	•
10	8000	2.724	0.0173017766	2.7472982234	•
11	8000	2.774	0.0173017766	2.7472982234	•
12	8000	2.808	0.0173017766	2.7472982234	•
13	10000	2.861	0.0245077929	2.8374922071	•
14	10000	2.879	0.0245077929	2.8374922071	•
15	10000	2.858	0.0245077929	2.8374922071	•

Columns (5/0)

- pressure
- density
- Simultaneous conf.
- Lower Bound +
- Upper Bound

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

# Ceramics: calculating the margins of error in JMP

Assumptions and  
More Inference for  
Simple Linear  
Regression (Ch.  
9.1)

Dason Kurkiewicz

Assumptions for  
SLR

SLR: Inference for  
the Mean  
Response at some  
 $x$

Simultaneous  
Confidence  
Intervals for  $\mu_{y|x}$

The screenshot shows two overlapping windows in the JMP software. The background window is titled "'Upper Bound' in Table 'untitled'" and contains settings for a column named "Upper Bound". The column is set to be "Locked", "Numeric", "Continuous", and in "Best" format. The "Initialize Data" is set to "Missing/Empty". The "Formula" tab is selected, showing a formula editor with the text "optional item" and a "Remove" button. The foreground window is titled "Upper Bound" and contains a "Table Columns" list with "pressure", "density", "Simultaneous confidence interval", "Lower Bound", and "Upper Bound". The "Functions (grouped)" list includes "Row", "Numeric", "Transcendental", "Trigonometric", "Character", "Comparison", "Conditional", "Probability", and "Discrete Probability". The "OK", "Cancel", "Apply", "Clear", and "Help" buttons are on the right. The main area of the "Upper Bound" window displays the formula:  $2.375 + 0.0000487 * \text{pressure} + \text{Simultaneous conf. margin}$ .

# Ceramics: calculating the margins of error in JMP

Assumptions and More Inference for Simple Linear Regression (Ch. 9.1)

Dason Kurkiewicz

▼ untitled						
		pressure	density	Simultaneous conf. margin	Lower Bound	Upper Bound
	1	2000	2.486	0.0245077929	2.4478922071	2.4969077929
	2	2000	2.479	0.0245077929	2.4478922071	2.4969077929
	3	2000	2.472	0.0245077929	2.4478922071	2.4969077929
	4	4000	2.558	0.0173017766	2.5524982234	2.5871017766
	5	4000	2.57	0.0173017766	2.5524982234	2.5871017766
	6	4000	2.58	0.0173017766	2.5524982234	2.5871017766
	7	6000	2.646	0.0141040654	2.6530959346	2.6813040654
	8	6000	2.657	0.0141040654	2.6530959346	2.6813040654
	9	6000	2.653	0.0141040654	2.6530959346	2.6813040654
	10	8000	2.724	0.0173017766	2.7472982234	2.7819017766
	11	8000	2.774	0.0173017766	2.7472982234	2.7819017766
	12	8000	2.808	0.0173017766	2.7472982234	2.7819017766
	13	10000	2.861	0.0245077929	2.8374922071	2.8865077929
	14	10000	2.879	0.0245077929	2.8374922071	2.8865077929
	15	10000	2.858	0.0245077929	2.8374922071	2.8865077929

Columns (5/1)  
pressure  
density  
Simultaneous conf  
Lower Bound +  
Upper Bound +

Assumptions for SLR

SLR: Inference for the Mean Response at some  $x$

Simultaneous Confidence Intervals for  $\mu_{y|x}$