SIMPLE LINEAR REGRESSION

In linear regression, we consider the frequency distribution of one variable (Y) at each of several levels of a second variable (X).

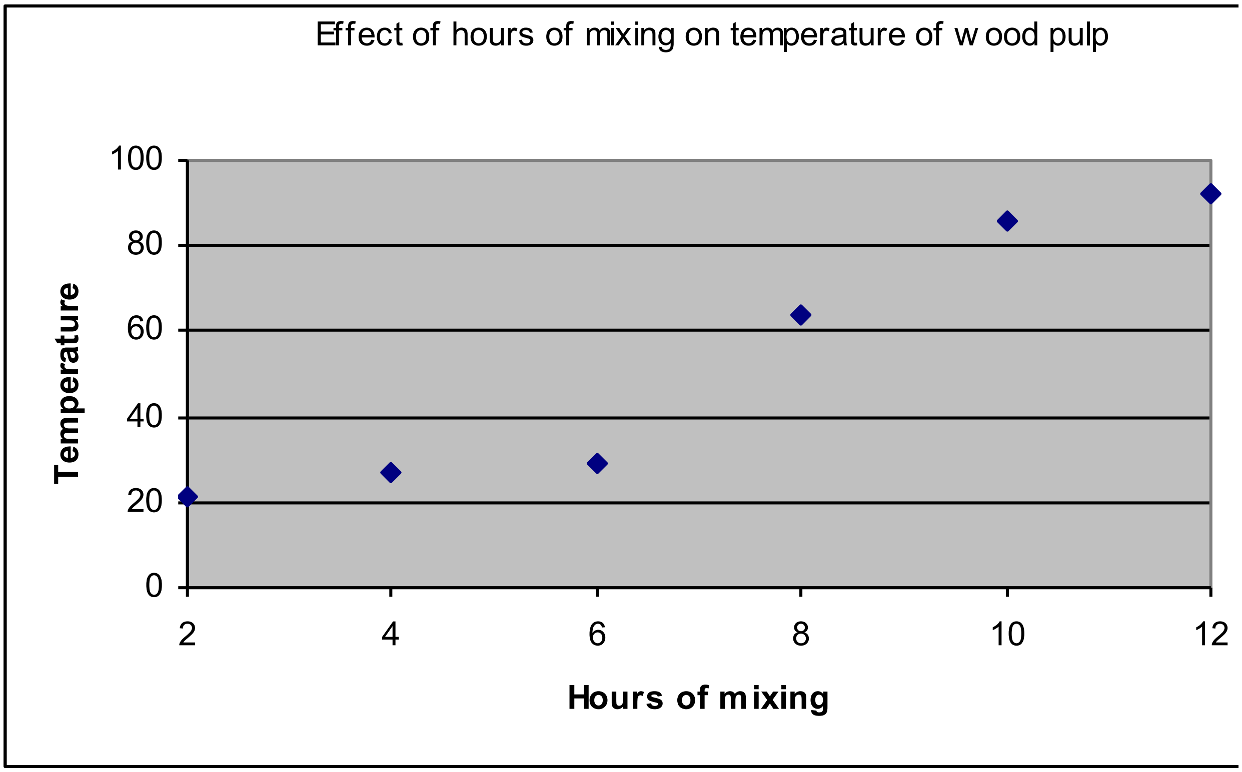
Y is known as the dependent variable. The variable for which you collect data.
X is known as the independent variable. The variable for the treatments.

Determining the Regression Equation

One goal of regression is to draw the "best" line through the data points.
The best line usually is obtained using means instead of individual observations.

Example
Effect of hours of mixing on temperature of wood pulp

Hours of mixing (X)	Temperature of wood pulp (Y)	XY
2	21	42
4	27	108
6	29	174
8	64	512
10	86	860
12	92	1104
$\sum X=42$	$\sum Y=319$	$\sum XY=2800$
$\sum X^2=364$	$\sum Y^2=21,967$	$n=6$



The equation for any straight line can be written as: $\hat{Y} = b_0 + b_1 X$

where: b_o = Y intercept, and
b_1 = regression coefficient = slope of the line

The linear model can be written as: $Y_i = \beta_0 + \beta_1 X + \varepsilon_i$

where: e_i=residual = $Y_i - \hat{Y}_i$

With the data provided, our first goal is to determine the regression equation

Step 1. Solve for b_1

$$b_1 = \frac{\sum (X-\overline{X})(Y-\overline{Y})}{\sum (X-\overline{X})^2} = \frac{\sum XY - \frac{(\sum X \sum Y)}{n}}{\sum X^2 - \frac{(\sum X)^2}{n}} = \frac{\text{SS Cross Products}}{\text{SS X}} = \frac{\text{SSCP}}{\text{SS X}}$$

for the data in this example

$\sum X = 42 \qquad \sum Y = 319 \qquad \sum XY = 2{,}800 \quad \sum X^2 = 364 \qquad \sum Y^2 = 21{,}967$

$$b_1 = \frac{\sum XY - \frac{(\sum X \sum Y)}{n}}{\sum X^2 - \frac{(\sum X)^2}{n}} = \frac{2800 - \frac{(42x319)}{6}}{364 - \frac{42^2}{6}} = \frac{567}{70} = 8.1$$

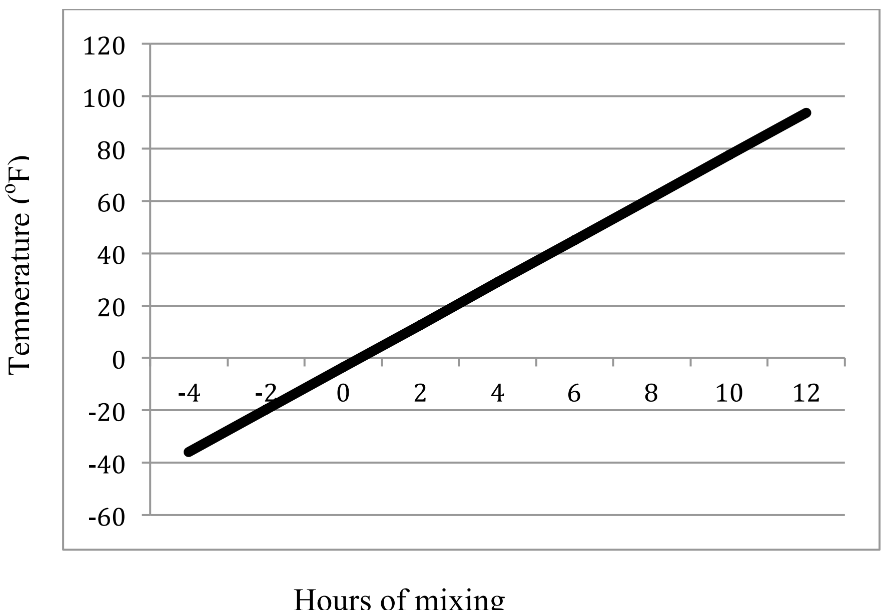
The number calculated for b_1, the regression coefficient, indicates that for each unit increase in X (i.e., hours of mixing), Y (i.e., wood pulp temperature) will increase 8.1 units (i.e., degrees).

The regression coefficient can be a positive or negative number.

To complete the regression equation, we need to calculate b_o.

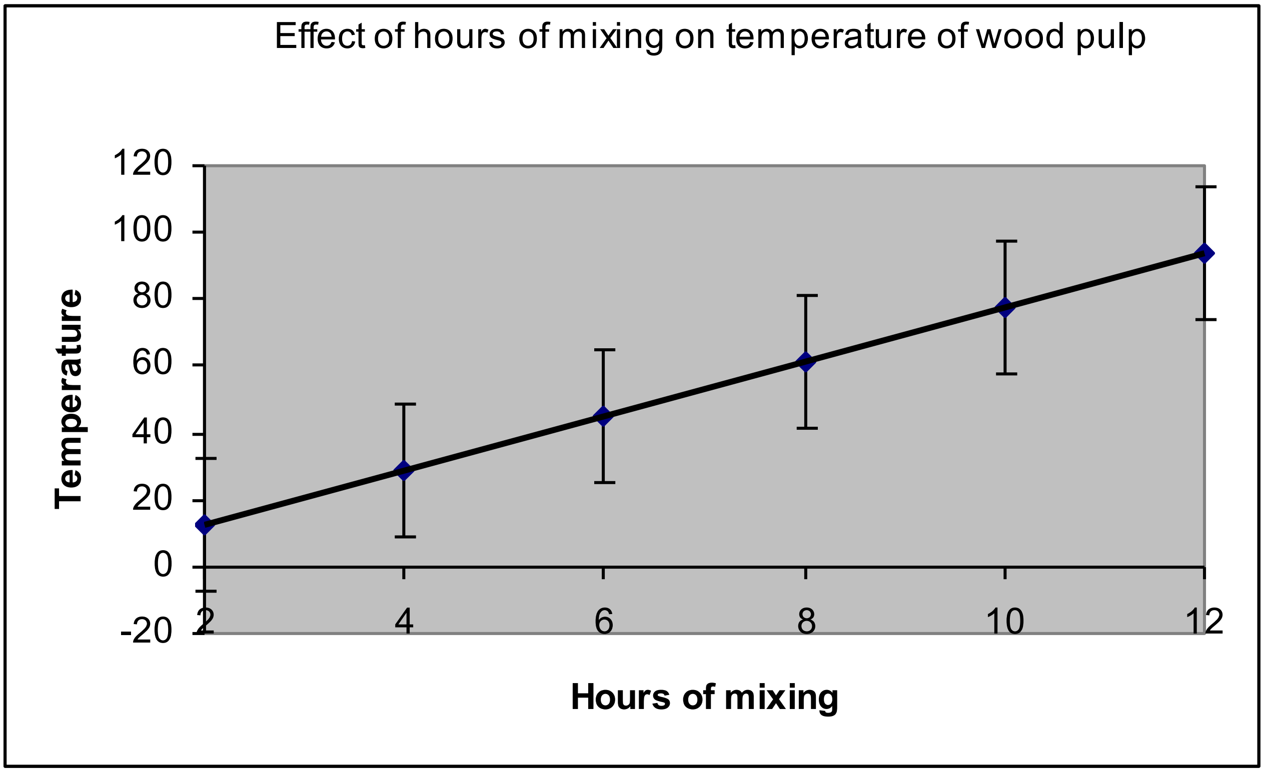
$$b_0 = \overline{Y} - b_1\overline{X} = \frac{319}{6} - 8.1\left(\frac{42}{6}\right) = -3.533$$

Therefore, the regression equation is: $\hat{Y}_i = -3.533 + 8.1X$



Assumptions of Regression

1. There is a linear relationship between X and Y
2. The values of X are known constants and presumably are measure without error.
3. For each value of X, Y is independent and normally distributed: Y~N(0, $\boldsymbol{\sigma}^2$).
4. Sum of deviations from the regression line equals zero: $\sum\left(Y_i - \hat{Y}_i\right) = 0$.
5. Sum of squares for error are a minimum.



If you square the deviations and sum across all observations, you obtain the definition formulas for the following sums of squares:

$\sum\left(\hat{Y}_i - \overline{Y}\right)^2$ = Sum Squares Due to Regression

$\sum\left(Y_i - \hat{Y}_i\right)^2$ = Sum Squares Due to Deviation from Regression (Residual)

$\sum\left(Y_i - \overline{Y}\right)^2$ = Sum Squares Total

Testing the hypothesis that a linear relationship between X and Y exists

The hypotheses to test that a linear relationship between X and Y exists are:

H_o: $ß_1 = 0$
H_A: $ß_1 \neq 0$

These hypotheses can be tested using three different methods:

1. *F*-test
2. *t*-test
3. Confidence interval

Method 1. F-test

The ANOVA to test H_o: $\$_1 = 0$ can be done using the following sources of variation, degrees of freedom, and sums of squares:

SOV	df	Sum of Square
Due to regression	1	$\dfrac{\left(\sum XY - \dfrac{(\sum X \sum Y)}{n}\right)^2}{\sum X^2 - \dfrac{\left(\sum X\right)^2}{n}} = \dfrac{SSCP^2}{SS\,X}$
Residual	n-2	Determined by subtraction
Total	n-1	$\sum Y^2 - \dfrac{\left(\sum Y\right)^2}{n} = SS\,Y$

Using data from the example:

$\sum X = 42$ $\sum Y = 319$ $\sum XY = 2,800$ $\sum X^2 = 364$ $\sum Y^2 = 21,967$

Step 1. Calculate Total SS =

$$\sum Y^2 - \frac{\left(\sum Y\right)^2}{n} = 21,967 - \frac{319^2}{6} = 5,006.833$$

Step 2. Calculate SS Due to Regression =

$$\frac{\left(\sum XY - \frac{(\sum X \sum Y)}{n}\right)^2}{\sum X^2 - \frac{(\sum X)^2}{n}} = \frac{\left(2800 - \frac{(42x319)}{6}\right)^2}{364 - \frac{42^2}{6}} = \frac{321{,}489}{70} = 4{,}592.7$$

Step 3. Calculate Residual SS = SS Deviation from Regression

Total SS - SS Due to Regression

5006.833 - 4592.7 = 414.133

Step 4. Complete ANOVA

SOV	df	SS	MS	*F*
Due to Regression	1	4592.7	4592.7	Due to Reg. MS/Residual MS = 44.36**
Residual	4	414.133	103.533	
Total	5	5006.833		

The residual mean square is an estimate of $\sigma^2_{Y|X}$, read as variance of Y given X.
This parameter estimates the statistic $\sigma^2_{Y|X}$.

Step 5. Because the F-test on the Due to Regression SOV is significant, we reject H_o: $\beta_1 = 0$ at the 99% level of confidence and can conclude that there is a linear relationship between X and Y.

Coefficient of Determination - *r^2*

From the ANOVA table, the coefficient of variation can be calculated using the formula

r^2 = SS Due to Regression / SS Total

This value always will be positive and range from 0 to 1.0.

As r^2 approaches 1.0, the association between X and Y improves.

r^2 x 100 is the percentage of the variation in Y that can be explained by having X in the model.

For our example: r^2 = 4592.7 / 5006.833 = 0.917.

We can conclude that 91.7% (i.e. 0.917 x 100) of the variation in wood pulp temperature can be explained by hours of mixing.

Method 2. *t*-test

The formula for the t-test to test the hypothesis H_o: $\beta_1=0$ is:

$$t = \frac{b_1}{s_{b_1}}$$

where: b_1 the regression coefficient, and

$$s_{b_1} = \sqrt{\frac{s^2_{Y|X}}{SS\,X}}$$

Remember that $s^2_{Y|X}$ = Residual MS = [SS Y - (SSCP^2 / SS X)] / (n-2)

For our example:

Step 1. Calculate $s^2_{b_1}$

We know from previous parts of this example:

SS Y = 5006.833
SSCP = 567.0
SS X = 70.0

Therefore, $s^2_{b_1} = (s^2_{Y|X} / SS\,X)$

$$\left. \frac{SS\,Y - \dfrac{SSCP^2}{SS\,X}}{n-2} \right/ SS\,X$$

$$= \left. \frac{5006.833 - \dfrac{567^2}{70}}{6-2} \right/ 70$$

$$= 1.479$$

Step 2. Calculate t statistic

$$t = \frac{b_1}{s_{b_1}}$$

$$= \frac{8.1}{\sqrt{1.479}}$$

$$= 6.66$$

Step 3. Look up table *t* value

Table $t_{\alpha/2,\,(n-2)\,df} = t_{.05/2,\,4df} = 2.776$

Step 4. Draw conclusions

Since the table *t* value (2.776) is less that the calculated *t*-value (6.66), we reject H_o: $\beta_1=0$ at the 95% level of confidence. Thus, we can conclude that there is a linear relationship between hours of mixing and wood pulp temperature at the 95% level of confidence.

Method 3. Confidence Interval

The hypothesis H_o: $\beta_1=0$ can be tested using the confidence interval:

$$CI = b_1 \pm t_{\alpha/2,(n-2)df}(s_{b_1})$$

For this example:

$$CI = b_1 \pm t_{\alpha/2,(n-2)df}(s_{b_1})$$

$$= 8.1 \pm 2.776\sqrt{1.479}$$

$$= 4.724 \le \beta_1 \le 11.476$$

We reject Ho: $\beta_1=0$ at the 95% level of confidence since the CI does not include 0.

Predicting Y Given X

Regression analysis also can be used to predict a value for Y given X.

Using the example, we can predict the temperature of **one batch** of wood pulp after mixing X hours.

In this case, we predict an individual outcome of Y_X drawn from the distribution of Y.

This estimate is distinct from estimating mean or average of a distribution of Y.

The value of an individual Y at a given X will take on the form of the confidence interval:

$$CI = \hat{Y} \pm t_{\alpha/2,(n-2)df}(s_{Y|X=X_0})$$

where $s_{Y|X=X_0} = \sqrt{s^2_{Y|X=0}}$, and

$$s^2_{Y|X=X_0} = s^2_{Y|X}\left[1+\frac{1}{n}+\frac{(X-\overline{X})^2}{SS\,X}\right]$$

Remember $s^2_{Y|X}$ is the Residual Mean Square

Example 1

We wish to determine the temperature of the one batch of wood pulp after mixing two hours (i.e., $Y_{X=2}$).

Step 1. Using the regression equation, solve for $\hat{Y}$ when X=2.

Remember $\hat{Y}$ =-3.533 + 8.1X

$\hat{Y}$ = -3.533 + 8.1(2) = 12.667

Step 2. Solve for $s^2_{Y|X=2}$

$$s^2_{Y|X=X_0} = s^2_{Y|X}\left[1+\frac{1}{n}+\frac{(X-\overline{X})^2}{SS\,X}\right]$$

$$= 103.533\left[1+\frac{1}{6}+\frac{(2-7)^2}{70}\right]$$

$$= 157.765$$

Step 3. Calculate the confidence interval

$$CI = \hat{Y} \pm t_{\alpha/2,(n-2)df}(s_{Y|X=X_0})$$

$$= 12.667 \pm 2.776\sqrt{157.65}$$

$$= 12.667 \pm 34.868$$

Therefore: LCI = -22.201 and UCI = 47.535

Note: This CI is not used to test a hypothesis

This CI states that if we mix the wood pulp for two hours, we would expect the temperature to fall within the range of -22.201 and 47.535 degrees 95% of the time.

We would expect the temperature to fall outside of this range 5% of the time due to random chance.

Example 2

We wish to determine the temperature of the one batch of wood pulp after mixing seven hours (i.e., $Y_{X=7}$).

Step 1. Using the regression equation, solve for $\hat{Y}$ when X=7.

Remember $\hat{Y}$ =-3.533 + 8.1X

$\hat{Y}$ = -3.533 + 8.1(7) = **53.167**

Step 2. Solve for $s^2_{Y|X=7}$

$$s^2_{Y|X=X_0} = s^2_{Y|X}\left[1 + \frac{1}{n} + \frac{(X - \overline{X})^2}{SS\,X}\right]$$

$$= 103.533\left[1 + \frac{1}{6} + \frac{(7-7)^2}{70}\right]$$

$$= 120.789$$

Step 3. Calculate the confidence interval

$$\text{CI} = \hat{Y} \pm t_{\alpha/2,(n-2)df}(s_{Y|X=X_0})$$

$$= 53.167 \pm 2.776\sqrt{120.789}$$

$$= 53.167 \pm 30.509$$

$$\text{Therefore: LCI} = 22.658 \text{ and UCI} = 83.676$$

Note: For X=7 (i.e., at the mean of X), the variance $s^2_{Y|X=X_0}$ is at a minimum.

This CI states that if we mix the wood pulp for seven hours, we would expect the temperature to fall within the range of 22.658 and 83.676 degrees 95% of the time.

We would expect the temperature to fall outside of this range 5% of the time due to random chance.

Predicting $\overline{Y}$ Given X

Regression analysis also can be used to predict a value for $\overline{Y}$ given X.

Using the example, we can predict the **average** temperature of wood pulp after mixing X hours.

In this case, we predict an individual outcome of $\overline{Y}_X$ drawn from the distribution of Y.

This estimate is distinct from distribution of Y for a X.

The value of an individual Y at a given X will take on the form of the confidence interval:

$$\text{CI} = \hat{Y} \pm t_{\alpha/2,(n-2)df}(s_{\overline{Y}|X=X_0})$$

where $s_{\overline{Y}|X=X_0} = \sqrt{s^2_{\overline{Y}|X=0}}$, and

$$s^2_{\overline{Y}|X=X_0} = s^2_{Y|X}\left[\frac{1}{n} + \frac{(X-\overline{X})^2}{\text{SS X}}\right]$$

Example 1

We wish to determine the **average** temperature of the wood pulp after mixing two hours (i.e., $Y_{X=2}$).

Step 1. Using the regression equation, solve for $\hat{Y}$ when X=2.

Remember $\hat{Y} = -3.533 + 8.1X$

$\hat{Y} = -3.533 + 8.1(2) \quad = \quad$ **12.667**

Step 2. Solve for $s^2_{\bar{Y}|X=2}$

$$s^2_{\bar{Y}|X=2} = s^2_{Y|X}\left[\frac{1}{n} + \frac{(X-\bar{X})^2}{SS\,X}\right]$$

$$= 103.533\left[\frac{1}{6} + \frac{(2-7)^2}{70}\right]$$

$$= 54.232$$

Step 3. Calculate the confidence interval

$$CI = \hat{Y} \pm t_{\alpha/2,(n-2)df}(s_{Y|X=X_0})$$

$$= 12.667 \pm 2.776\sqrt{54.232}$$

$$= 12.667 \pm 20.443$$

Therefore: LCI = -7.776 and UCI = 33.110

Note: This CI is not used to test a hypothesis

This CI states that if we mix the wood pulp for two hours any number of times, we would expect the **average** temperature to fall within the range of -7.776 and 33.110 degrees 95% of the time.

We would expect the temperature to fall outside of this range 5% of the time due to random chance.

Example 2

We wish to determine the average temperature of wood pulp after mixing seven hours.

Step 1. Using the regression equation, solve for $\hat{Y}$ when X=7.

Remember $\hat{Y}$ =-3.533 + 8.1X

$$\hat{Y} = -3.533 + 8.1(7) \quad = \quad 53.167$$

Step 2. Solve for $s^2_{\bar{Y}|X=7}$

$$s^2_{\bar{Y}|X=7} = s^2_{Y|X}\left[\frac{1}{n} + \frac{(X-\bar{X})^2}{SS\,X}\right]$$

$$= 103.533\left[\frac{1}{6} + \frac{(7-7)^2}{70}\right]$$

$$= 17.256$$

Step 3. Calculate the confidence interval

$$CI = \hat{Y} \pm t_{\alpha/2,(n-2)df}(s_{Y|X=X_0})$$

$$= 53.167 \pm 2.776\sqrt{17.256}$$

$$= 53.167 \pm 11.532$$

Therefore: LCI = 41.635 and UCI = 64.669

Note: For X=7 (i.e., at the mean of X), the variance $s^2_{\bar{Y}|X=X_0}$ is at a minimum.

Comparing $s^2_{Y|X=X_0}$ and $s^2_{\bar{Y}|X=X_0}$

$s^2_{Y|X=X_0}$ is always greater than $s^2_{\bar{Y}|X=X_0}$.

Comparing the formulas:

$$s^2_{Y|X=X_0} = s^2_{Y|X}\left[1+\frac{1}{n}+\frac{(X-\bar{X})^2}{SS\,X}\right] \text{ and}$$

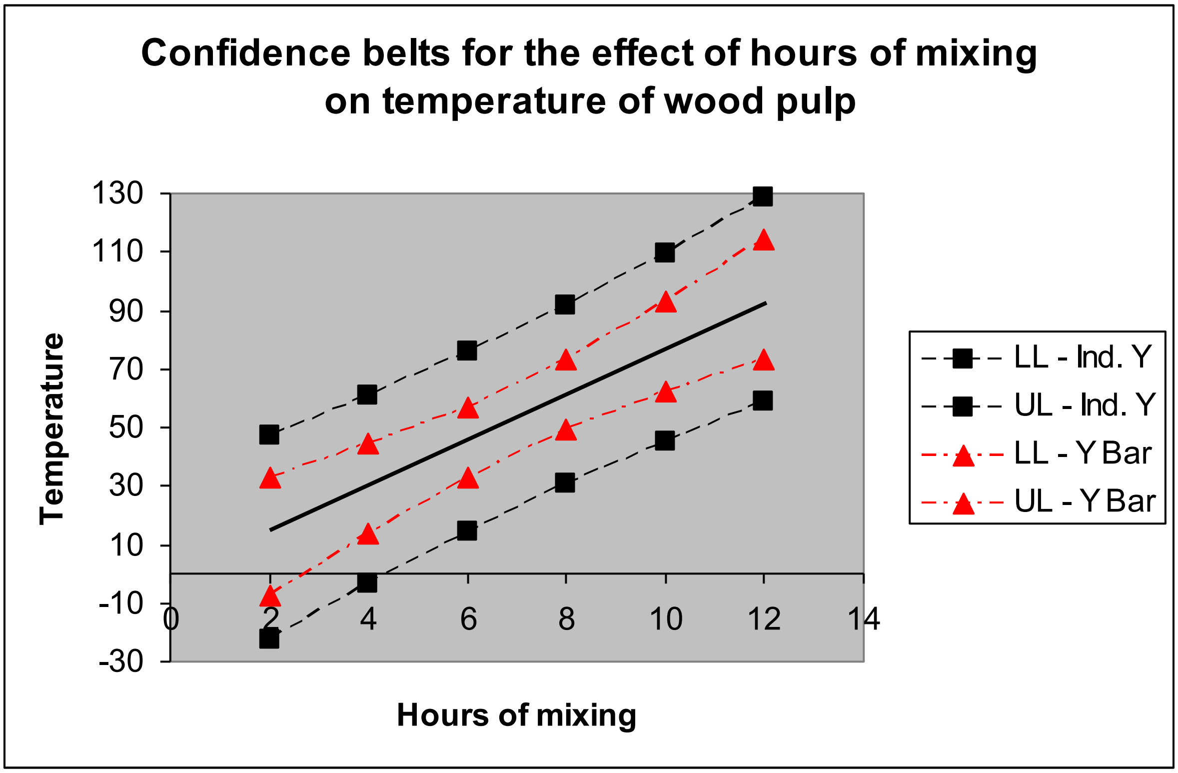
$$s^2_{\bar{Y}|X=X_0} = s^2_{Y|X}\left[\frac{1}{n}+\frac{(X-\bar{X})^2}{SS\,X}\right].$$

Notice that in the formula for $s^2_{Y|X=X_0}$ you add one while in the formula for $s^2_{\bar{Y}|X=X_0}$ you do not.

Comparison of $s^2_{Y|X=X_0}$ and $s^2_{\bar{Y}|X=X_0}$.

X	$s^2_{Y|X=X_0}$	$s^2_{\bar{Y}|X=X_0}$
2	157.767	54.232
7	120.789	17.256

We can draw the two confidence intervals as "confidence belts" about the regression line.



Notice that:

1. The confidence belts are symmetrical about the regression line
2. The confidence belts are narrowest at the mean of X, and
3. The confidence belts for the distribution based on means are narrower than the distribution based on an individual observation.

Determining if Two Independent Regression Coefficients are Different

It may be desirable to test the homogeneity of two b_1*'s* to determine if they are estimates of the same β_1.

This can be done using a t-test to test the hypotheses:

H_o: $\beta_1 = \beta_{1'}$

H_A: $\beta_1 \neq \beta_{1'}$

$$\text{Where } t = \frac{b_1 - b_{1'}}{\sqrt{(\text{Residual 1 MS} + \text{Residual 2 MS})\left(\frac{1}{X_1SS} + \frac{1}{X_2SS}\right)}}$$

The Table *t*-value has $(n_1 - 2) + (n_2 - 2)$ df.

Example

X	Y_1	Y_2
1	11	22
2	16	29
3	14	34
4	18	45
5	23	58
$\sum X = 15$	$\sum Y_1 = 82$	$\sum Y_2 = 187$
$\sum X^2 = 55$	$\sum Y_1^2 = 1426$	$\sum Y_2^2 = 7827$

Step 1. Determine regression coefficient for each Y

for Y_1 $\sum XY = 272$

Thus $b_1 = [272 - (15x82)/5] / [55 - 15^2/5] = 2.6$

for Y_2 $\sum XY = 651$

Thus $b_{1'} = [651 - (15x187)/5] / [55 - 15/5] = 9.0$

Step 2. Calculate Residual MS for each Y

$$\text{Remember Residual MS} = \frac{\text{SS Y} - \left(\frac{\text{SSCP}^2}{\text{SS X}}\right)}{n - 2}$$

$$\text{Residual 1 MS} = \frac{\left(1426 - \frac{82^2}{5}\right) - \left(\frac{26^2}{10}\right)}{5 - 2}$$

$$= 4.5$$

$$\text{Residual 2 MS} = \frac{\left(7827 - \frac{187^2}{5}\right) - \left(\frac{90^2}{10}\right)}{5 - 2}$$

$$= 7.7$$

Step 3. Solve for *t*

$$\frac{2.6 - 9.0}{\sqrt{(4.5 + 7.7) \text{x} \left(\frac{1}{10} + \frac{1}{10}\right)}}$$

$$= \frac{-6.4}{\sqrt{12.2(0.2)}}$$

$$= -4.10$$

Step 4. Look up table *t*-value with $(n_1 - 2) + (n_2 - 2)$ df

$t_{0.05/2,\, 6\, df} = -2.447$

Step 4. Make conclusions

Because the absolute value of the calculated *t*-value (-4.10) is greater than the absolute value of the tabular *t*-value (2.776), we can conclude at the 95% level of confidence that the two regression coefficients are not estimating the same $ß_1$.

Summary - Some Uses of Regression

1. Determine if there is a linear relationship between an independent and dependent variable.

2. Predict values of Y at a given X

   Most accurate near the mean of X.

   Should avoid predicting values of Y outside the range of the independent variables that were used.

3. Can adjust Y to a common base by removing the effect of the independent variables (Analysis of Covariance).

4. ANOVA (CRD, RCBD, and LS) can be done using regression

5. Compare homogeneity of two regression coefficients.

SAS Commands

```
options pagneo=1;
data reg;
input x y;
datalines;
2      21
4      27
6      29
8      64
10     86
12     92
;
proc reg;
model y=x/cli clm;
title 'SAS Output for Linear Regression Example in Class';
run;
```

Printout of Data

Obs	x	y	xy
1	2	21	42
2	4	27	108
3	6	29	174
4	8	64	512
5	10	86	860
6	12	92	1104

Descriptive statistics

The MEANS Procedure

Variable	Sum	Mean	Uncorrected SS	Corrected SS
x	42.00	7.00	364.00	70.00
y	319.00	53.17	21967.00	5006.83
xy	2800.00	466.67	2264264.00	957597.33

- This summary table includes the values that were calculated for the data on page 1 of the regression notes.
- The uncorrected SS gives you the values when you square the value for each observations and then sum the squared values.
- Disregard the values for Mean, Uncorrected SS, and Corrected SS for xy.

Simple Linear Regression, Including Confidence Intervals

The REG Procedure
Model: MODEL1
Dependent Variable:
y

Number of Observations Read	6
Number of Observations Used	6

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	4592.70000	4592.70000	44.36	0.0026
Error	4	414.13333	103.53333		
Corrected Total	5	5006.83333			

Root MSE	10.17513	**R-Square**	0.9173
Dependent Mean	53.16667	**Adj R-Sq**	0.8966
Coeff Var	19.13818		

Parameter Estimates					
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	1	-3.53333	9.47253	-0.37	0.7281
x	1	8.10000	1.21616	6.66	0.0026

Results from the *F*-test of the hypothesis H_o: $ß_1 = 0$.

r^2 value – the coefficient of determination

- Values used to generate the regression equation.
- Intercept = Y-intercept
- x = regression coefficient
- $\hat{Y}_i = -3.533 + 8.1X$

Results from the *t*-test of the hypothesis H_o: $ß_1 = 0$

Simple Linear Regression, Including Confidence Intervals

The REG Procedure
Model: MODEL1
Dependent Variable:
y

Output Statistics								
Obs	**Dependent Variable**	**Predicted Value**	**Std Error Mean Predict**	**95% CL Mean**		**95% CL Predict**		**Residual**
1	21.0000	12.6667	7.3642	-7.7797	33.1130	-22.2068	47.5401	8.3333
2	27.0000	28.8667	5.5287	13.5164	44.2169	-3.2850	61.0184	-1.8667
3	29.0000	45.0667	4.3283	33.0492	57.0841	14.3662	75.7672	-16.0667
4	64.0000	61.2667	4.3283	49.2492	73.2841	30.5662	91.9672	2.7333
5	86.0000	77.4667	5.5287	62.1164	92.8169	45.3150	109.6184	8.5333
6	92.0000	93.6667	7.3642	73.2203	114.1130	58.7932	128.5401	-1.6667

Sum of Residuals	0
Sum of Squared Residuals	414.13333
Predicted Residual SS (PRESS)	868.05699

- Dependent variable is Y_i.
- Predicted value is $\hat{Y}_i$ based on the regression equation.
- 95% CL Mean is the confidence limits for a predicted mean value of Y at a given level of X.
- 95% CL for the confidence limits for the predicted value of an individual value of Y at a given level of X.

SAS Commands for Data with Multiple Replicates

```
options pageno=1;
data reg;
input x rep y;
xy=x*y;
datalines;
10 1 432
20 1 334
30 1 224
40 1 199
50 1 134
60 1 125
10 2 410
20 2 403
30 2 300
40 2 280
50 2 247
60 2 225
;;
ods graphics off;
ods rtf file='repreg.rtf';
proc print;
title 'Printout of Data -  All Data';
run;
proc reg;
model y=x;
title 'Regression on All Data - Reps Kept Separate';
run;
proc sort;
by x;
proc means mean noprint;
by x;
var y;
output out=new mean=meany;
*Comment:  The previous statement takes the average of Y1 and Y2 to create
a new variable Y.  Performing regression on means provides a better
coefficient of determination';
run;
proc reg;
model meany=x;
title 'Regression on the Mean of Y';
run;
ods rtf close;
```

Printout of Data - All Data

Obs	x	rep	y	xy
1	10	1	432	4320
2	20	1	334	6680
3	30	1	224	6720
4	40	1	199	7960
5	50	1	134	6700
6	60	1	125	7500
7	10	2	410	4100
8	20	2	403	8060
9	30	2	300	9000
10	40	2	280	11200
11	50	2	247	12350
12	60	2	225	13500

Regression on All Data - Reps Kept Separate

The REG Procedure
Model: MODEL1
Dependent Variable: y

Number of Observations Read	12
Number of Observations Used	12

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	91188	91188	34.77	0.0002
Error	10	26229	2622.88524		
Corrected Total	11	117417			

Root MSE	51.21411	**R-Square**	0.7766
Dependent Mean	276.08333	**Adj R-Sq**	0.7543
Coeff Var	18.55024		

Parameter Estimates					
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	1	454.73333	33.71326	13.49	<.0001
x	1	-5.10429	0.86568	-5.90	0.0002

Regression on the Mean of Y

The REG Procedure
Model: MODEL1
Dependent Variable: meany

Number of Observations Read	6
Number of Observations Used	6

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	45594	45594	60.26	0.0015
Error	4	3026.67619	756.66905		
Corrected Total	5	48621			

Root MSE	27.50762	**R-Square**	0.9377
Dependent Mean	276.08333	**Adj R-Sq**	0.9222
Coeff Var	9.96352		

Parameter Estimates					
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	1	454.73333	25.60820	17.76	<.0001
x	1	-5.10429	0.65756	-7.76	0.0015