

## StatCrunch and Regression

These instructions use the data from table 11.1, page 466 in Rosner. One you load that data into Statcrunch (provided in file rosner\_11.1.xls), it should look as shown below. You want to use regression to determine if there is a relationship between estriol level in pregnant women and the infant birth weight (examples 11.1 and 11.3). You can use StatCrunch to produce a graphic similar to figure 11.1, page 465. Choose STAT / REGRESSION / SIMPLE LINEAR.

*NOTE: some of the StatCrunch results discussed here go beyond the first three sections of chapter 11, the only sections that are REQUIRED reading.*

rosner\_11.1.xls

StatCrunch	Data	Stat	Graphics
Row	estriol	birth weight	
1	7	25	
2	9	25	
3	9	25	
4	12	27	
5	14	27	
6	16	27	
7	16	24	
8	14	30	
9	16	30	
10	16	31	
11	17	30	
12	19	31	
13	21	30	
14	24	28	
15	15	32	

Stat	Graphics	Help
Summary Stats		
Tables		
Z statistics		
Proportions		
T statistics		
Variance		
Regression	Simple Linear	
ANOVA	Multiple Linear	
Nonparametrics	Logistic	

You will then see the screen shown on the right. Choose ESESTRIOL as the X-variable (or, the independent variable) and BIRTH WEIGHT as the Y-VARIABLE (or, the dependent variable). You want to find out if changes in estriol have an effect on birth weight. Thus, estriol is the independent variable (a variable that you do not vary in a controlled study, rather that you measure at varying levels in a study population). You want to see if birth weight "depends" on estriol level, making birth weight the dependent or Y-variable. Click on NEXT.

On the next screen, click on the box to the left of "Predict Y for X=" and enter 15. If you look at example 11.9, you will see that you are asked to predict the value of birth weight when the estriol level is 15. Entering 15 on this screen will provide that predicted value in the StatCrunch output. Click on NEXT.

Simple Linear Regression

X-Variable: estriol

Y-Variable: birth weight

Where: --optional--

Group by: --optional--

? Snapshot Cancel < Back Next > Calculate

Simple Linear Regression

Options:

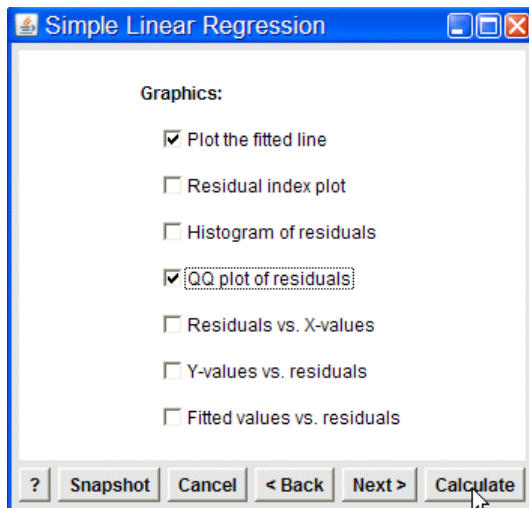
☒ Predict Y for X = 15

☐ Save residuals

☐ Save fitted values

? Snapshot Cancel < Back Next > Calculate

On the next screen, click on the boxes "Plot the fitted line" and "QQ plot of residuals" (no QQ plot is shown in Posner, but you can use the QQ plot as one measure of the success on fitting a regression equation to the data in table 11.1). Then click on CALCULATE.



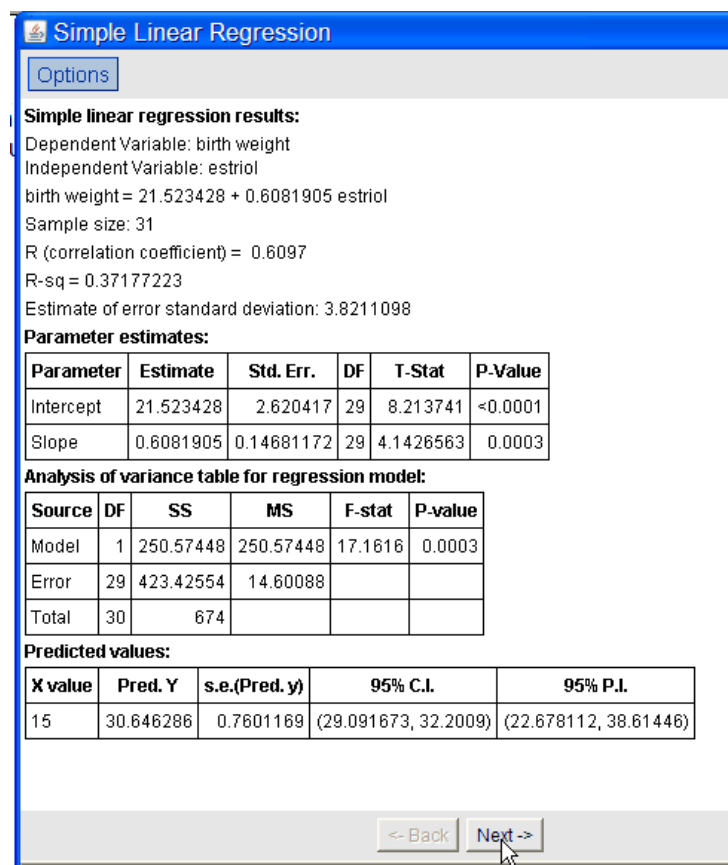
**Simple Linear Regression**

**Graphics:**

- ☒ Plot the fitted line
- ☐ Residual index plot
- ☐ Histogram of residuals
- ☒ QQ plot of residuals
- ☐ Residuals vs. X-values
- ☐ Y-values vs. residuals
- ☐ Fitted values vs. residuals

? Snapshot Cancel < Back Next > Calculate

You should see the screen shown to the right. On the fourth line, you will see an equation that is similar to that shown in figure 11.1 in Rosner. That equation shows the relationship between changes in estriol and birth weight. There is also a correlation coefficient (the same value that you will find in example 11.24, page 493) and R-sq (the same value found in example 11.13, page 479).



**Simple Linear Regression**

**Options**

**Simple linear regression results:**

Dependent Variable: birth weight  
 Independent Variable: estriol  
 birth weight = 21.523428 + 0.6081905 estriol  
 Sample size: 31  
 R (correlation coefficient) = 0.6097  
 R-sq = 0.37177223  
 Estimate of error standard deviation: 3.8211098

**Parameter estimates:**

Parameter	Estimate	Std. Err.	DF	T-Stat	P-Value
Intercept	21.523428	2.620417	29	8.213741	<0.0001
Slope	0.6081905	0.14681172	29	4.1426563	0.0003

**Analysis of variance table for regression model:**

Source	DF	SS	MS	F-stat	P-value
Model	1	250.57448	250.57448	17.1616	0.0003
Error	29	423.42554	14.60088		
Total	30	674			

**Predicted values:**

X value	Pred. Y	s.e.(Pred. y)	95% C.I.	95% P.I.
15	30.646286	0.7601169	(29.091673, 32.2009)	(22.678112, 38.61446)

<- Back Next ->

Look box "Slope" in box labeled "Parameter Estimates" ... you will see a T-Statistic that is a measure of the relationship between estriol and birth weight. It matches the value in example 11.14, page 482 (there is a significant relationship given that the p-value is < 0.05)

The next box labeled "Analysis of variance ..." matches table 11.3, page 478, and is yet another measure how well the regression equation allows you to predict birth weight from various values of estriol. Last in the first section of output is "Predicted values" that matches the value in example 11.9, page 472.

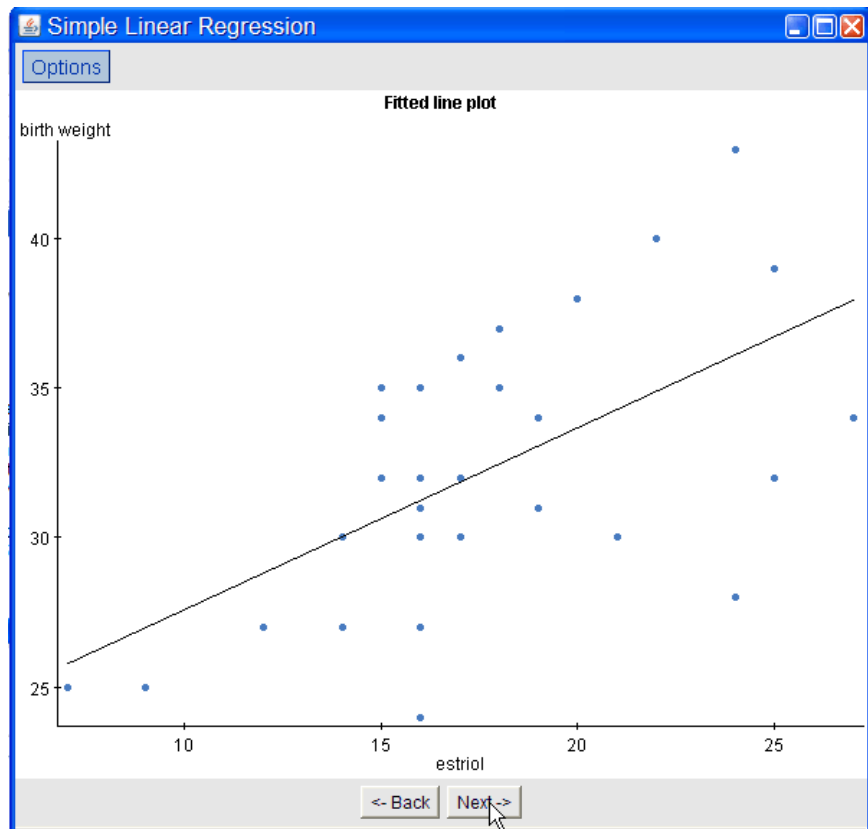
Click on NEXT.

Finally, the StatCrunch results on the right are similar to figure 11.1, page 465. They look a tiny bit different in that when two dots occupy the same space in the plot in Rosner, he shows two dots right next to each other rather than a single dot (there are two observations with an estriol value of 9 and a birth weight of 25, and two observations with an estriol value of 15 and a birth weight value of 34). The StatCrunch plot shows a single dot.

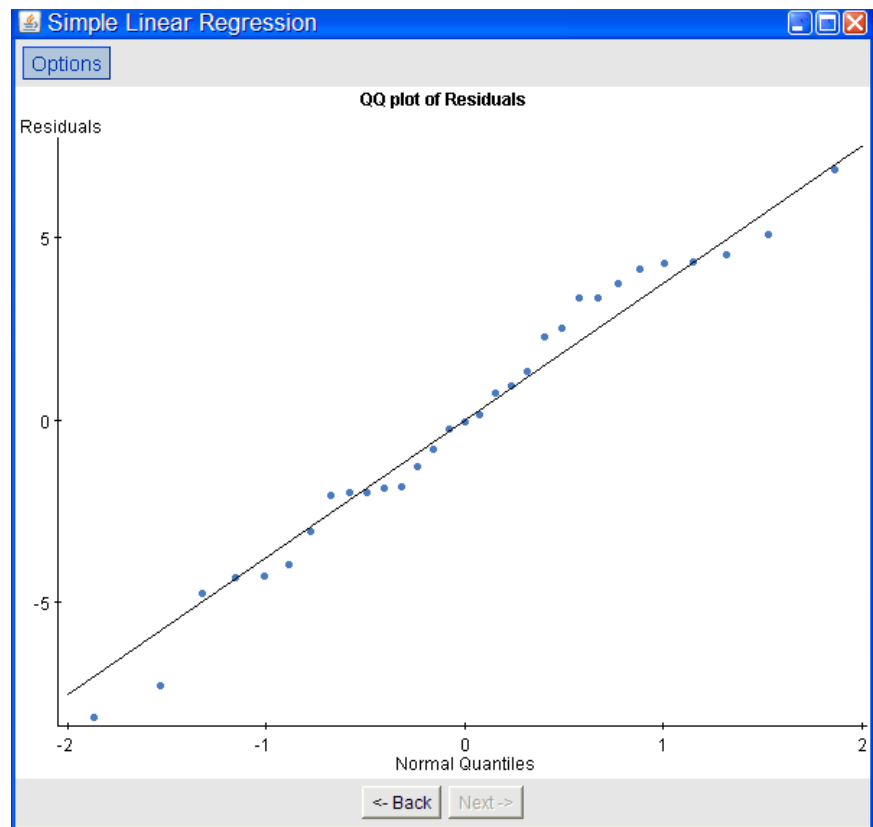
The "eyeball" test for this plot is whether the line appears to follow the pattern of the observed values ... or the predicted values (the line) are close to the observed data.

The statistical tests are the t-statistic and the analysis of variance.

Click on NEXT.



As discussed in some of the notes in the Nonparametric module, a QQ plot is an "eyeball" test for normality. Another test for "success" in fitting a regression line is that the residuals (a residual is the difference between an observed and predicted value) are normally distributed. Though assumption #2 in equation 11.13 uses the term "for any given value of  $x$ " it is customary to just examine all the residuals at one time for normality (as shown on the right).



As stated at the beginning, some of the output discussed here is beyond the REQUIRED reading. If you want to try another set of data that is used in Rosner, you are also provided with the data from table 11.4, page 479 in file rosner11.4.xls. You should be able to generate the results shown in the various examples in Rosner associated with that data.