

Homework 10, due November 29 in lecture **Solutions**

(assignment for these solutions can be found on the last page)

1. *Show that considered by itself, weight of mother at final prenatal visit is not a significant predictor of child's birth weight.*

Regression Analysis: Birth_Weight versus Weight_Final_Mom

The regression equation is
 $\text{Birth_Weight} = 2791 + 10.4 \text{ Weight_Final_Mom}$

Predictor	Coef	SE Coef	T	P
Constant	2790.9	601.7	4.64	0.000
Weight_Final_Mom	10.385	7.756	1.34	0.187

2. *Show that if gestation period and sex of the child are also entered into the model, then weight of mother at final prenatal visit is a significant predictor of child's birth weight.*

Regression Analysis: Birth_Weight versus Weight_Final, Gestation_Pe, ...

The regression equation is
 $\text{Birth_Weight} = -3602 + 14.3 \text{ Weight_Final_Mom} + 157 \text{ Gestation_Period} - 442 \text{ Sex_Child_0M_1F}$

Predictor	Coef	SE Coef	T	P
Constant	-3602	2374	-1.52	0.136
Weight_Final_Mom	14.280	6.935	2.06	0.045
Gestation_Period	156.54	57.53	2.72	0.009
Sex_Child_0M_1F	-441.8	142.0	-3.11	0.003

3. *Interpret the coefficients fit in part 2). Think carefully about how to interpret the coefficient for sex of child.*

The interpretation of the regression coefficients is the expected change in the response variable per unit increase in a particular predictor variable with the other predictor variables held constant.

Birth weight is expected to increase 14.3 grams per kilogram increase in mother's weight at final prenatal visit with gestation period and child's sex held constant.

Birth weight is expected to increase 157 grams per week increase in gestation period with mother's weight at final prenatal visit and child's sex held constant.

Birth weight is expected to be 442 grams less for girls (1) than for boys (0) (that is like a unit increase in sex) with mother's weight at final prenatal visit and gestation period held constant.

4. Does the model fit in part 2) appear to give good predictions? Explain.

First, good predictions. The $R^2=0.286$, which is not large, so the model may not give great predictions.

S = 470.745 **R-Sq = 28.6%** R-Sq(adj) = 23.7%

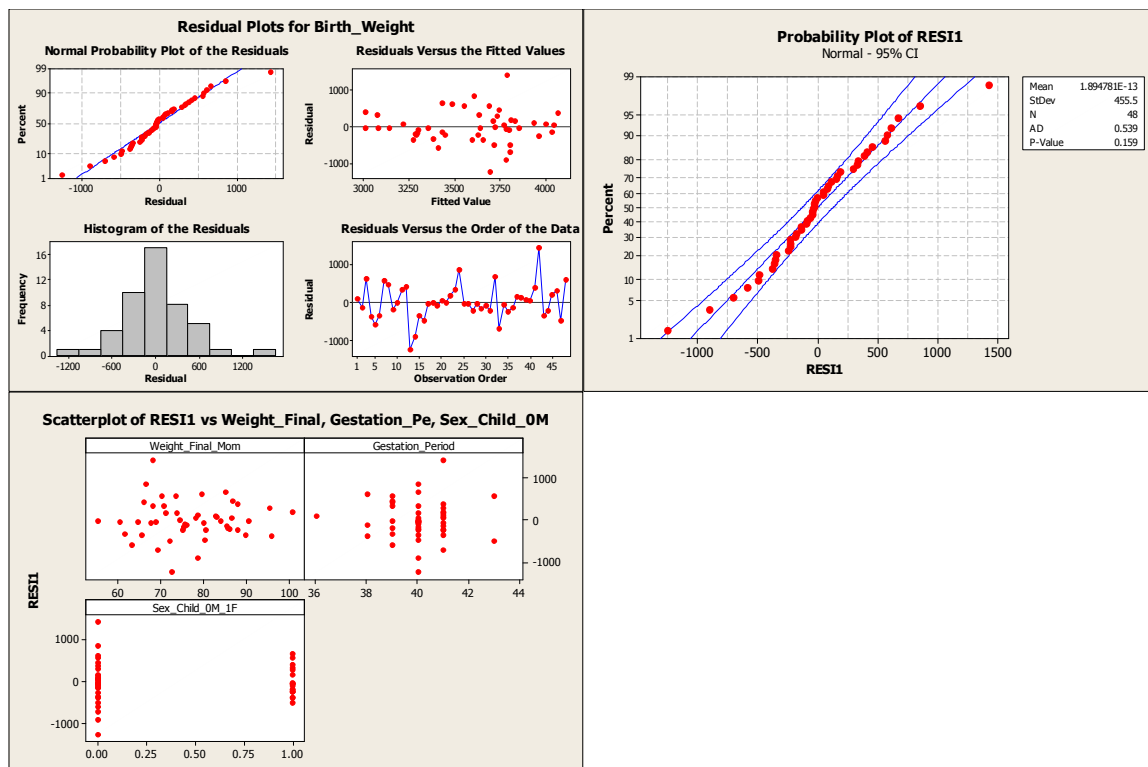
Now, model fit.

We assess model fit by checking the residuals for normality, constant variance, and no patterns.

In the residual plots on the left, the two left plots give us the normal probability plot and histogram, which both appear reasonably normal. On the right, a closer look at the residuals looks good, confirmed by the Anderson-Darling p-value=0.159.

In the upper right plot in the residual plots, there appears to be more variation near the center of the observations than on the ends, but that may be due to having more observations in that range.

The scatterplots at the bottom plot the residuals on the y-axis, and the three predictor variables on the x-axis. Residuals should not exhibit any pattern when plotted against any of these. We see there is roughly constant variance for Weight_Final_Mom, and there may be an outlier at top. For gestation period, it is roughly constant and one possible outlier at top. For sex, boys have more variation than girls.



5. *Minitab reports some unusual observations. Are they having a serious impact on the model you fit in part 2)? Explain.*

I created a column Unusual_groups that are 0 if not unusual, 1 if R, and 2 if X as unusual observation table.

Group 1 (large residual) are consistently large in all scatterplots. However, they are near the center of the data and balance each other. They are not likely effecting the regression line. (We can refit the model to check, if desired.)

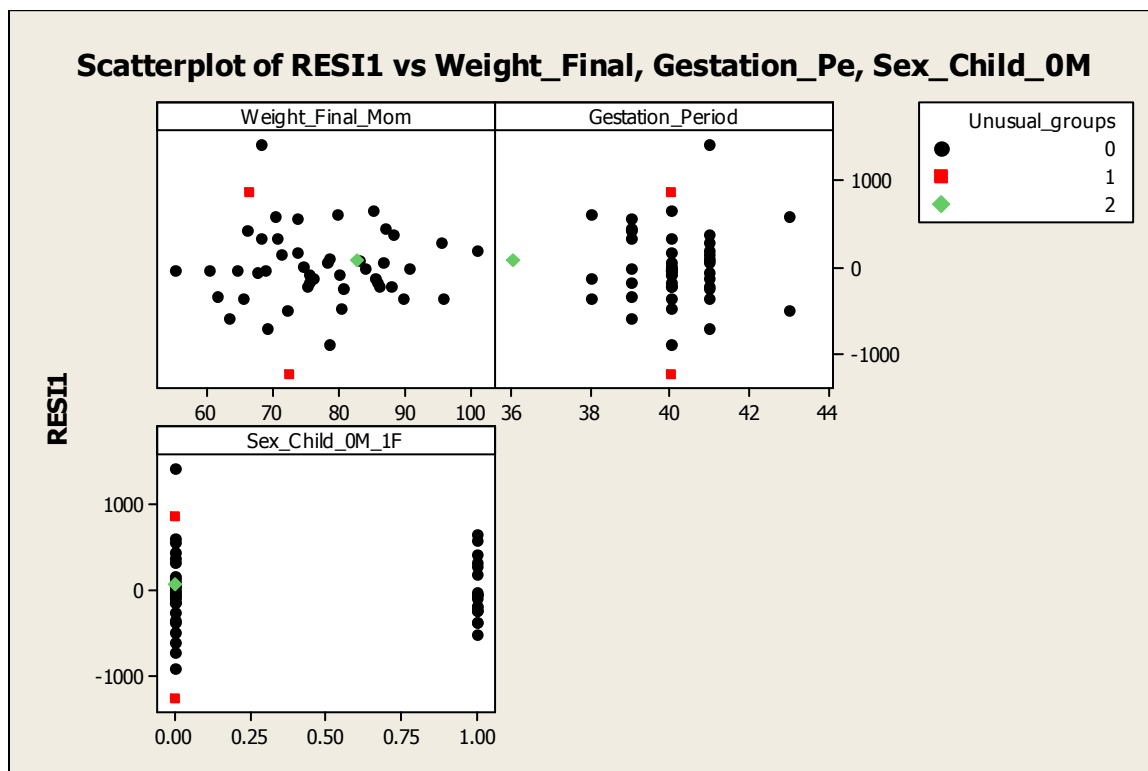
Group 2 (large influence) has just the one observation which is clearly an outlier having a short Gestation_Period. This observation probably is effecting the regression.

What follows the scatterplot is the regression with these three observations removed.

Unusual Observations

Obs	Weight_Final_Mom	Birth_Weight	Fit	SE Fit	Residual	St Resid
1	83	3300.0	3214.4	246.0	85.6	0.21 X
13	73	2450.0	3694.9	90.0	-1244.9	-2.69R
42	68	5220.0	3790.0	117.7	1430.0	3.14R

R denotes an observation with a large standardized residual.
 X denotes an observation whose X value gives it large influence.



The fit below is comparable to our first analysis. First, note that there are two unusual observations now, that were not unusual under the previous model. The three predictor variables have roughly the same coefficients, so the regression line has not changed much. The proportion of variation has gone up a little. However, unless there is a good reason for having removed those observations, we should not remove them and stand by the previous analysis.

Regression Analysis: Birth_Weight versus Gestation_Pe, Sex_Child_0M, ...

The regression equation is

$$\text{Birth_Weight}_1 = -3891 + 163 \text{ Gestation_Period}_1 - 457 \text{ Sex_Child_0M_1F}_1 + 15.0 \text{ Weight_Final_Mom}_1$$

Predictor	Coef	SE Coef	T	P
Constant	-3891	2432	-1.60	0.117
Gestation_Period_1	162.65	59.70	2.72	0.009
Sex_Child_0M_1F_1	-456.6	130.5	-3.50	0.001
Weight_Final_Mom_1	15.019	6.379	2.35	0.023

S = 426.289 R-Sq = 35.2% R-Sq(adj) = 30.4%

Analysis of Variance

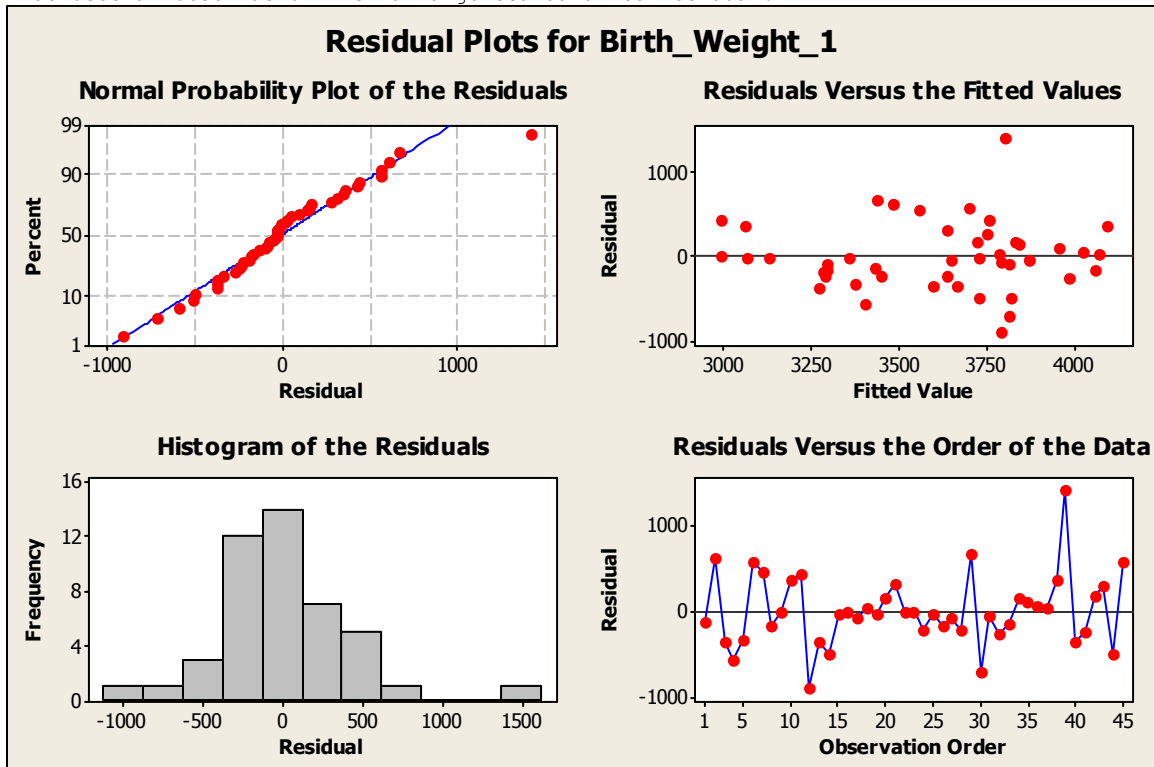
Source	DF	SS	MS	F	P
Regression	3	4045023	1348341	7.42	0.000
Residual Error	41	7450607	181722		
Total	44	11495630			

Source	DF	Seq SS
Gestation_Period_1	1	1214629
Sex_Child_0M_1F_1	1	1823221
Weight_Final_Mom_1	1	1007173

Unusual Observations

Obs	Gestation_Period_1	Birth_Weight_1	Fit	SE Fit	Residual	St Resid
12	40.0	2885.0	3795.2	85.7	-910.2	-2.18R
39	41.0	5220.0	3801.7	111.7	1418.3	3.45R

R denotes an observation with a large standardized residual.



Stat 538 – Biostatistics I
Homework 10 solutions

5/5

Erik Barry Erhardt, UNM

Stat 538, 2005, Homework 10, due November 29 in lecture

Multiple linear regression.

This uses the same data set as homework 9. You should support your answers with appropriate output from Minitab.

- 1) Show that considered by itself, weight of mother at final prenatal visit is not a significant predictor of child's birth weight.
- 2) Show that if gestation period and sex of the child are also entered into the model, then weight of mother at final prenatal visit is a significant predictor of child's birth weight.
- 3) Interpret the coefficients fit in part 2). Think carefully about how to interpret the coefficient for sex of child.
- 4) Does the model fit in part 2) appear to give good predictions? Explain.
- 5) Minitab reports some unusual observations. Are they having a serious impact on the model you fit in part 2)? Explain.