Stat 346 Homework 6

Nathan Jarus

Apr. 3, 2014

1 Problem 1

1.1 Part a

Math Salaries

Correlation Matrix

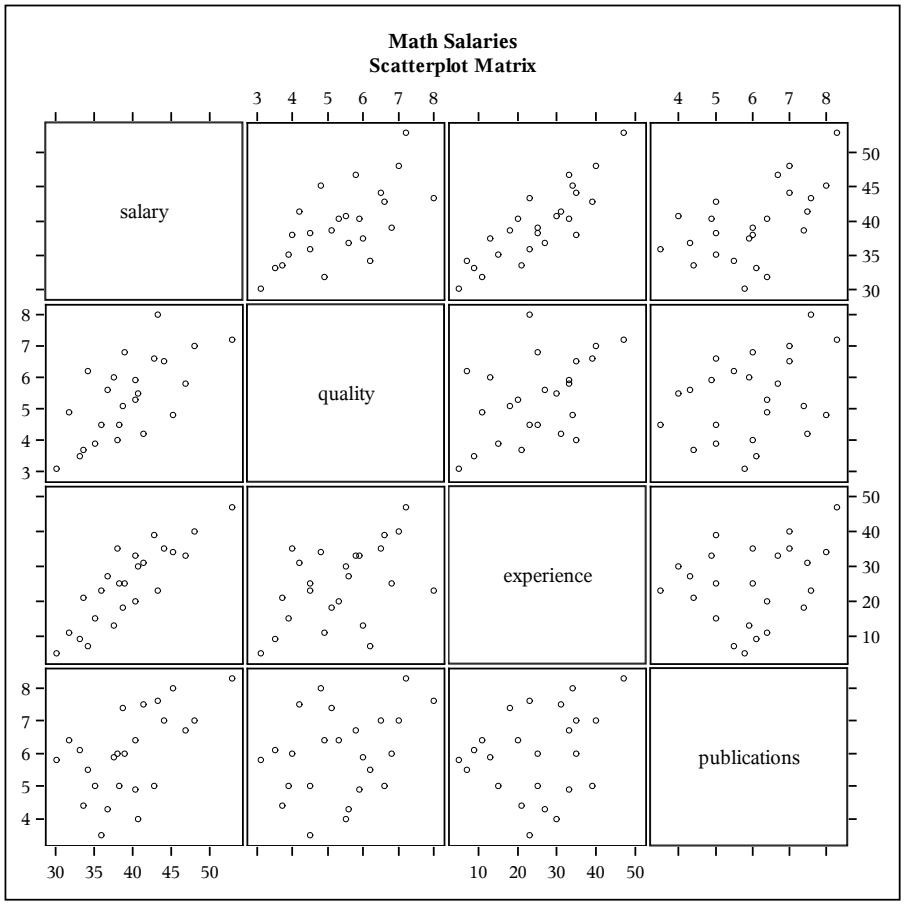
The CORR Procedure

4 Variables:	salary quality experience publications

Simple Statistics						
Variable	**N**	**Mean**	**Std Dev**	**Sum**	**Minimum**	**Maximum**
salary	24	39.50000	5.47429	948.00000	30.10000	52.90000
quality	24	5.35833	1.29142	128.60000	3.10000	8.00000
experience	24	24.95833	11.22295	599.00000	5.00000	47.00000
publications	24	5.98750	1.30361	143.70000	3.50000	8.30000

Pearson Correlation Coefficients, N = 24				
	salary	**quality**	**experience**	**publications**
salary	1.00000	0.66710	0.85856	0.55820
quality	0.66710	1.00000	0.46695	0.32276
experience	0.85856	0.46695	1.00000	0.25375
publications	0.55820	0.32276	0.25375	1.00000

17:57 Wednesday, April 2, 2014 **1**



1.2 Part b

Math Salaries

Simple Regression

The REG Procedure

Model: MODEL1

Dependent Variable: salary

Number of Observations Read	24
Number of Observations Used	24

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	306.73233	306.73233	17.64	0.0004
Error	22	382.52767	17.38762		
Corrected Total	23	689.26000			

Root MSE	4.16985	**R-Square**	0.4450
Dependent Mean	39.50000	**Adj R-Sq**	0.4198
Coeff Var	10.55657		

Parameter Estimates					
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	1	24.34766	3.70666	6.57	<.0001
quality	1	2.82781	0.67327	4.20	0.0004

Math Salaries

Simple Regression

The REG Procedure

Model: MODEL2

Dependent Variable: salary

Number of Observations Read	24
Number of Observations Used	24

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	508.06883	508.06883	61.69	<.0001
Error	22	181.19117	8.23596		
Corrected Total	23	689.26000			

Root MSE	2.86984	R-Square	0.7371
Dependent Mean	39.50000	Adj R-Sq	0.7252
Coeff Var	7.26541		

Parameter Estimates					
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	1	29.04785	1.45400	19.98	<.0001
experience	1	0.41878	0.05332	7.85	<.0001

Math Salaries

Simple Regression

The REG Procedure

Model: MODEL3

Dependent Variable: salary

Number of Observations Read	24
Number of Observations Used	24

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	214.76157	214.76157	9.96	0.0046
Error	22	474.49843	21.56811		

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Corrected Total	23	689.26000			

Root MSE	4.64415	**R-Square**	0.3116
Dependent Mean	39.50000	**Adj R-Sq**	0.2803
Coeff Var	11.75734		

Parameter Estimates					
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**
Intercept	1	25.46502	4.54765	5.60	<.0001
publications	1	2.34405	0.74284	3.16	0.0046

$$salary = 24.34766 + 2.82781 quality \tag{1}$$

$$salary = 29.04785 + 0.41878 experience \tag{2}$$

$$salary = 25.46502 + 2.34405 publications \tag{3}$$

1.3 Part c

Math Salaries

Multiple Regression

The REG Procedure

Model: MODEL1

Dependent Variable: salary

Number of Observations Read	24
Number of Observations Used	24

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	3	627.81700	209.27233	68.12	<.0001
Error	20	61.44300	3.07215		

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Corrected Total	23	689.26000			

Root MSE	1.75276	**R-Square**	0.9109
Dependent Mean	39.50000	**Adj R-Sq**	0.8975
Coeff Var	4.43735		

Parameter Estimates							
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**98.334% Confidence Limits**	
Intercept	1	17.84693	2.00188	8.92	<.0001	12.61649	23.07737
quality	1	1.10313	0.32957	3.35	0.0032	0.24203	1.96423
experience	1	0.32152	0.03711	8.66	<.0001	0.22456	0.41848
publications	1	1.28894	0.29848	4.32	0.0003	0.50908	2.06880

$$salary = 17.84693 + 1.10313quality + 0.32152experience + 1.28894publications \tag{4}$$

1.4 Part d

The regression coefficients are all smaller in MLR than they were in SLR. This is due to each variable having to explain less of the variation in salary as they are all taken into account.

1.5 Part e

	Quality	Experience	Publications
H_0	$b_1 = 0$	$b_2 = 0$	$b_3 = 0$
H_a	$b_1 \neq 0$	$b_2 \neq 0$	$b_3 \neq 0$
Test Statistic	$\frac{b_1}{sb_1}$	$\frac{b_2}{sb_2}$	$\frac{b_3}{sb_3}$
DoF	23	23	23
p-value	0.0032	< 0.0001	0.0003

Each variable has $> 95\%$ likelyhood that it has a significant linear relationship with the result variable.

1.6 Part f

The Bonferonni procedure for a family of 3 variables and a 95% family confidence interval gives us an individual confidence interval of 98.333%, meaning $\alpha = 0.01666$.

	Estimate	Confidence	Limits
Intercept	17.84693	12.61649	23.07737
Quality	1.10313	0.24203	1.96423
Experience	0.32152	0.22456	0.41848
Publications	1.28894	0.50908	2.06880

1.7 Part g

For a mathematician with a quality index of 5.4, a publication index of 6.0, and 17 years of experience, we would predict a salary of \$37003. There is a 95% probability that the salary is between \$33219 and \$40787.

1.8 Part h

We would use the narrower of the two intervals generated by the Bonferonni and Scheffe methods.

1.9 Part i

On average, such a mathematician would make \$37003, with 95% of such mathematicians' salaries falling in the interval $(\$36027, \$37979)$.

1.10 Part j

	Adjusted R^2
Quality	0.4198
Experience	0.7252
Publications	0.2803
Multiple	0.8975

Since the multiple linear regression has the highest adjusted R^2 value, it best explains the variation in salary.

1.11 Part k

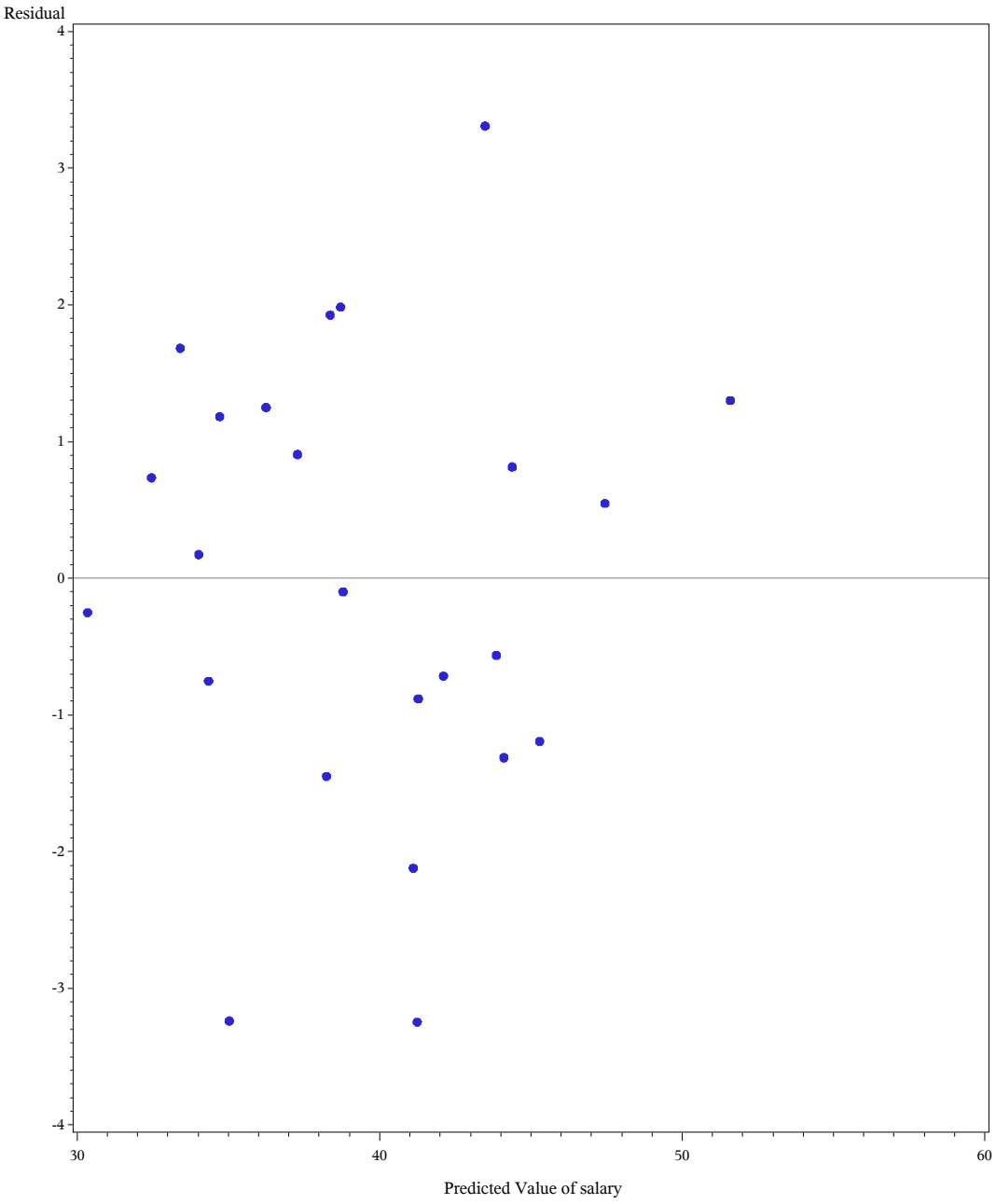
Figure 1.11 shows no trend between residuals and predicted values, indicating that the constant variance assumption holds.

Figures 1.11, 1.11, and 1.11 show no overall trends between the predictor variables and salary, indicating that the constant variance and linearity assumptions hold.

Figure 1.11 shows a relatively straight normal probability plot, indicating that the normal distribution of errors holds.

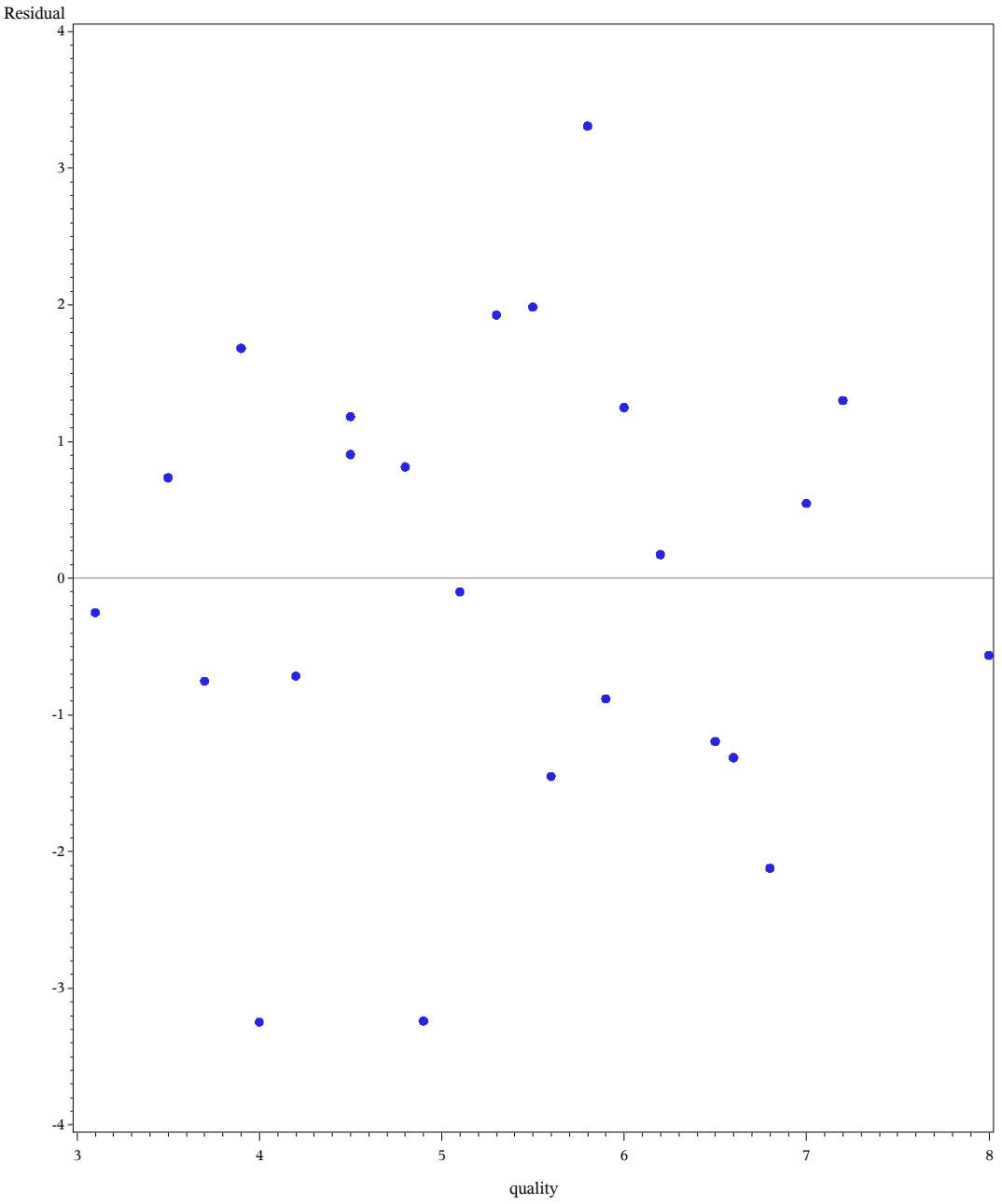
17:57 Wednesday, April 2, 2014 **1**

Math Salaries
Residuals vs. Predicted Salary



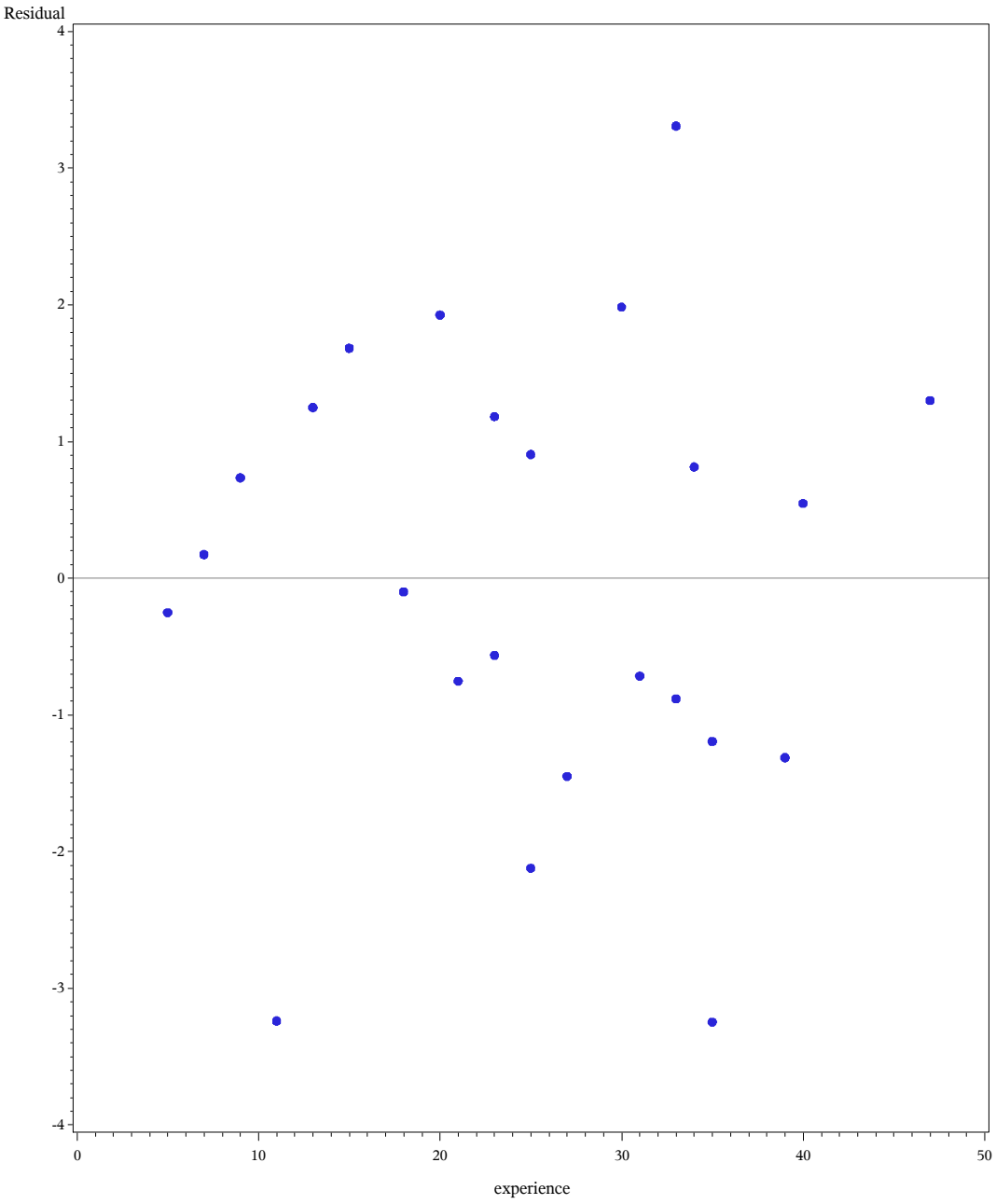
17:57 Wednesday, April 2, 2014 **1**

Math Salaries
Residuals vs. Quality



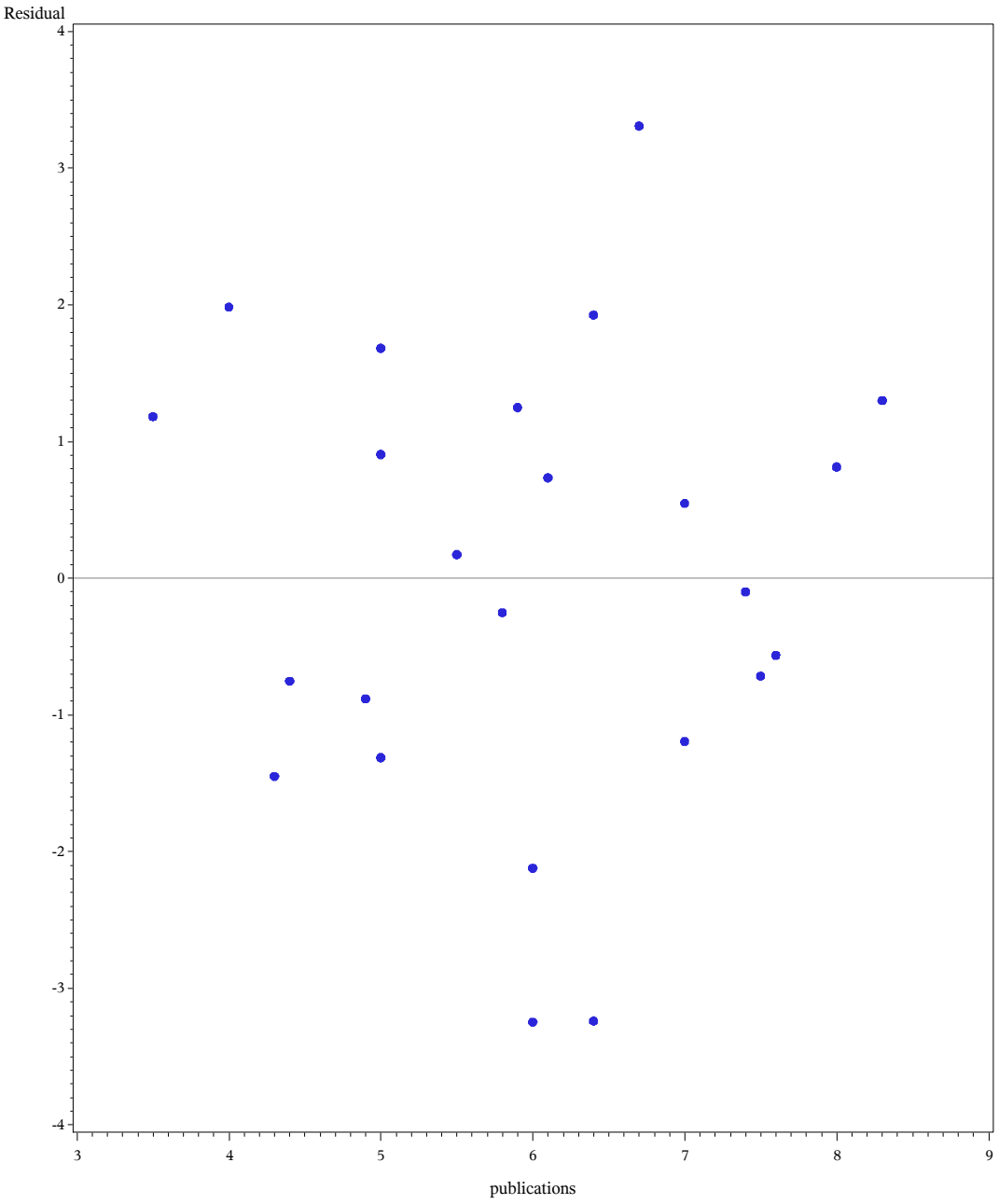
17:57 Wednesday, April 2, 2014 **1**

Math Salaries
Residuals vs. Experience



17:57 Wednesday, April 2, 2014 **1**

Math Salaries
Residuals vs. Publications



Math Salaries
Normal Probability

17:57 Wednesday, April 2, 2014 **1**