Math 3080 § 1. **Second Midterm Exam** Name: Example

Treibergs March 10, 2010

From my Math 3080 Midterms given Feb. 9 and March 30, 2005.

[1.] Complete the ANOVA table, given that there are five samples from each of the three populations.

Source of Variation	df	SS	MS	F	f-crit
Treatment			9.8		
Error					
Total		31.6			

No cheating! Solution below.

[2.] *In a recent study, researchers tried to predict the height of California redwood trees y (ft) in terms of the diameter at breast height x (in.) They measured 21 trees and recorded the following summary statistics. Test at the $\alpha = 0.01$ level whether there is a linear relationship between these two variables. State your assumptions.*

$$\sum_{i=1}^n x_i = 615 \qquad \sum_{i=1}^n y_i = 3298.8 \qquad \sum_{i=1}^n x_i^2 = 22387 \qquad \sum_{i=1}^n x_i y_i = 108310 \qquad \sum_{i=1}^n y_i^2 = 561090$$

Assume that the pairs (x_i, y_i) are a random sample taken from a fixed bivariate normal distribution. Compute the correlation coefficient.

$$r^2 = \frac{S_{xy}^2}{S_{xx}S_{xy}} = \frac{\left(\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}\right)^2}{\left(\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right)\left(\sum y_i^2 - \frac{(\sum y_i)^2}{n}\right)} = \frac{\left(108310 - \frac{615 \cdot 3298.8}{21}\right)^2}{\left(561094 - \frac{3298.8^2}{21}\right)\left(22387 - \frac{615^2}{21}\right)} = 0.729427313$$

Then the test for correlation is $\mathcal{H}_0 : \rho = 0$ vs $\mathcal{H}_a : \rho \neq 0$. Then the U statistic satisfies the t distribution with $n - 2 = 19$ df. The null hypothesis is rejected provided that $|u| \geq t_{\alpha/2, n-2} = t_{.005,19} = 2.861$. The statistic is given by

$$u = \pm\sqrt{\frac{(n-2)r^2}{1-r^2}} = \sqrt{\frac{19 \cdot 0.729427313}{1 - 0.729427313}} = 7.15691458,$$

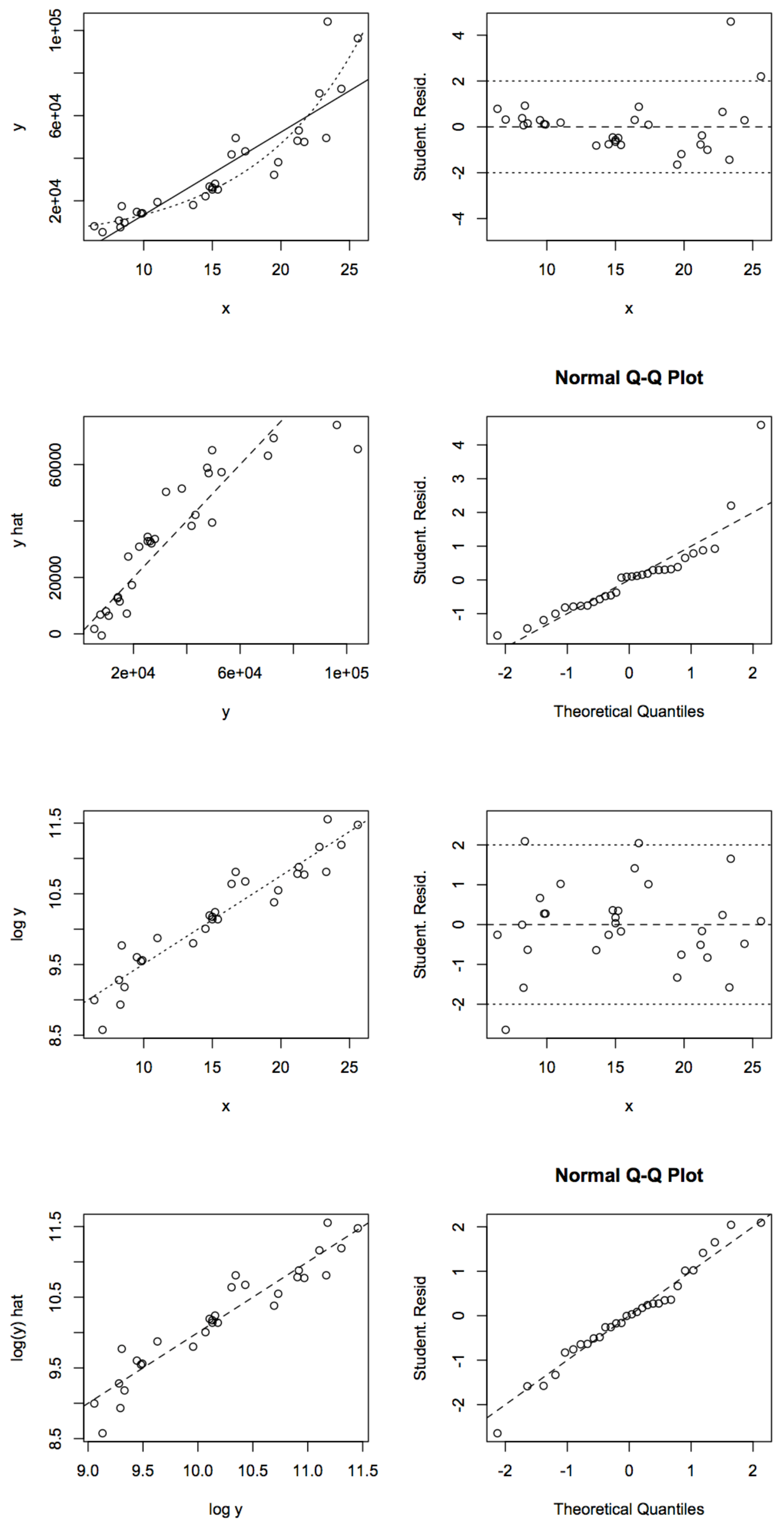
thus we reject the null hypothesis. With 99% confidence, we conclude that there is evidence that there is a linear relationship between height and diameter.

[3.] *In his masters thesis "Certain mechanical properties of wood-foam composites," Connors (2002) studied the relationship between density and stiffness of particleboards. x is density in lb./ft^3 and y is stiffness in lb./in^2. The first model is a simple linear regression $y = \beta_0 + \beta_1 x$, and the second is $\ln(y) = \beta_0 + \beta_1 x$. Discuss the two models with regard to quality of fit and whether model assumptions are satisfied. Which is the better model? What do both models predict to be the mean Y when density is 14 lb/ft^3? (The attached figure shows the data list, ANOVA printouts from MacAnova, (x, y) scatterplot, x vs studentized residuals, $\hat{y}$ vs y and a normal PP-plot of residuals. The top four panels show the first model. The bottom four show the second model. The dotted curve in the first panel is the fitted regression line of the second model. What is its equation?*

x	y	x	y
9.5	14814	8.4	17502
9.8	14007	11.0	19433
8.3	7573	9.9	14191
8.6	9714	6.4	8076
7.0	5304	8.2	10728
17.4	43243	15.0	25319
15.2	28028	16.4	41792
16.7	49499	15.4	25312
15.0	26222	14.5	22148
14.8	26751	13.6	18036
25.6	96305	23.4	104170
24.4	72594	23.3	49512
19.5	32207	21.2	48218
22.8	70453	21.7	47661
19.8	38138	21.3	53045

```
Model used is y=x
                 Coef       StdErr           t
CONSTANT       -25435       6104.7     -4.1664
x                3885       370.01        10.5
.
N: 30, MSE: 1.3508e+08, DF: 28, R-sq: 0.79746
Regression F(1,28): 110.25, Durbin-Watson: 1.1866
                DF          SS          MS            F      P-value
CONSTANT         1  3.6053e+10  3.6053e+10    266.90186     < 1e-08
x                1  1.4892e+10  1.4892e+10    110.24678     < 1e-08
ERROR1          28  3.7822e+09  1.3508e+08
```

```
Model used is {log(y)} = x
                 Coef       StdErr           t
CONSTANT       8.2574      0.12864      64.189
x             0.12493     0.007797      16.022
.
N: 30, MSE: 0.059982, DF: 28, R-sq: 0.90166
Regression F(1,28): 256.71, Durbin-Watson: 1.5526
                DF          SS          MS            F      P-value
CONSTANT         1      3115.1      3115.1  51933.01915     < 1e-08
x                1      15.398      15.398    256.71220     < 1e-08
ERROR1          28      1.6795    0.059982
```


y
x
Student. Resid.
x
Normal Q-Q Plot
y hat
y
Student. Resid.
Theoretical Quantiles
log y
x
Student. Resid.
x
Normal Q-Q Plot
log(y) hat
log y
Student. Resid
Theoretical Quantiles

For $x^* = 14$, the predicted value from the first model is $y^* = \hat{\beta}_0 + \hat{\beta}_1 x^* = -25436 + 3885 \cdot 14 = 28954$. The predicted value from model two is $(\ln(y))^* = \hat{\beta}_0 + \hat{\beta}_1 x = 8.2574 + 0.12493 \cdot 14 = 10.006$ so that exponentiating we get the y prediction $\exp((\ln(y))^*) = \exp(10.006) = 22168$. Thus, the exponential of the regression line from the second model gives the curved line in panel 1. Its equation is $y = \exp(8.2574 + 0.12493x)$.

Looking at the scatterplots, the data points curve upward in the first panel but are straighter in the fifth. Thus there is some nonlinear nature in the original data. The R^2 values are 0.79746 for the first model and 0.90166 for the second, so the linear model fits the transformed points better. Looking at the standardized residual plots, the first model exhibits heteroscedasticity (nonconstancy of variance.) The vertical spread of residuals increases with x, indicating that the variance is not constant as it shoud be. On the other hand, model two displays a residual plot of points spread in a horizontal band centered about zero, as it should be for the assumed normal $N(0, \sigma)$ errors, independent of x. The second model does much better. In the y vs $\hat{y}$ plots, the cloud of points is curved in panel 3 but aligns with $\hat{y} = y$ in panel 7. Again, the second model shows no violation of assumptions, whereas the curvaure in the first model indicates a nonlinear behavior. Finally, the PP-plot in panel 4 shows skewing, also indicating that there are nonlinear effects. The points line up nicely with the 45° line, as they should do for normal residuals. Again, the second model does much better than the first.

In conclusion, the second model does a better job of representing the data. If fits the points better and the regression assumptions are better met.

[1.] The completed ANOVA table is

Source of Variation	df	SS	MS	F	f-crit
Treatment	2	19.6	9.8	9.8	3.89
Error	12	12.0	1.0		
Total	14	31.6			

More Problems.

Problems 1-3 taken from Prof. Roberts' Math 3080 Exams given Spring 2004

(1.) *A bicycle manufacturing company is considering switching from their current brand of tier, X, to one of three other brands A, B, C. Random samples of five tires from each brand we tested and the time required to reduce the tire tread to a specified level was recorded for each tire. We have the following data on the sample means and the partially complete ANOVA table*

	X	A	B	C	Overall mean
Sample mean	207	223	197	209	$\bar{X}_{..} = 209$

ANOVA

Source of Variation	df	SS	MS	F	f-crit
Treatment					
Error					
Total		2082			

i) *Complete the ANOVA table and state the assumptions you need to make about the population sampled to do an ANOVA test. [Hint: Although you don't have variances for each sample, you can use SST to find SSE.]*

ii) *Do the sample results indicate that there is a difference (at $\alpha = .05$) in mean time among the four tire brands? Justify your answer by carrying out the ANOVA test.*

This is a one factor fixed effects model. Let X_{ij} be the tire life of the in hours for brand i and sample number j. We are assuming that the tire life is predicted by the model $X_{ij} = \mu + \alpha_i + \epsilon_{ij}$ such that $E(X_{ij}) = \mu + \alpha_i$ for each i, j where α_i is the constant deviation from the grand mean μ so $\sum_i \alpha_i = 0$, and the ϵ_{ij} are IID $N(0, \sigma^2)$ normal variables. The number of levels (brands) is $I = 4$. The number of replications id $J = 5$. The treatment means are $\bar{X}_{i\cdot}$ (e.g., $\bar{X}_{1\cdot} = 207$.) Hence the treatment SS is $SSTr = J[\sum_i (\bar{X}_{i\cdot} - \bar{X}_{\cdot\cdot})^2 = 5[(207-209)^2 + (223-209)^2 + (197-209)^2 + (209-209)^2] = 1770$. The Fundamental identity implies $SSE = SST - SSTr = 2082 - 344 = 1738$. Treatment $df = I - 1 = 4 - 1 = 3$. Error $df = I(J-1) = 4 \cdot 4 = 16$. Total $df = IJ - 1 = 4 \cdot 5 - 1 = 19$. $MSTr = SSTR/(I-1) = 344/3 = 114.667$. $MSE = SSE/((I(J-1)) = 1738/16 = 91.474$. $F = MSTr/MSE = 114.667/91.474 = 5.278$. Summarizing

Source of Variation	df	SS	MS	F
Treatment	3	1720	573.333	25.341
Error	16	362	22.625	
Total	19	2082		

We are testing the hypothesis $\mathcal{H}_0 : \alpha_1 = \cdots = \alpha_I = 0$ vs. the alternative $\mathcal{H}_1 : \alpha_i \neq 0$ for some i. The rejection region at the $\alpha = .05$ level is $F > f_{\alpha, I-1, I(J-1)} = f_{.05,3,16} = 3.24$ by the table. Since our computed $F = 25.341$ exceeds this, we reject the null hypothesis: there is strong evidence that not all mean times $\mu_i = \mu + \alpha_i$ are equal.

(2.) *Suppose y = amount of residual chlorine in the pool (ppm) and x = hours of cleaning it. It was decided that the model $y = Ae^{Bx}$ would best explain the relationship between x and y. A simple linear regression of $\ln y$ on x gave the following results*

```
    Regression equation:   ln(chlorine) = 0.558 ? 0.0239(hours)
                           S = 0.05635 R-sq = .759 R-sq(adjusted) = 0.699
 Summary Statistics for:
               X (hours):  n= 6 Mean=7.00 StDev = 1.53
```

i) *What would you use for A, B in the model $Y = Ae^{Bx}$?*

Taking logs we obtain the intrinsically linear model $\ln Y = \ln A + Bx$ so that $A = \exp(.558) = 1.747$ and $B = -0.0239$.

ii) *Based on the model, give an interval estimate (95%) for the mean amount of chlorine (ppm) in the pool 15 hours after the next cleaning. Show your work.*

The estimate for the mean amount of chlorine after $x^* = 15$ hours is $\ln(y)^* = 0.558 - 0.0239x^* = 0.558 - 0.0239(15) = .1995$ so that $y^* = \exp(.1995) = 1.2208$. We need to find S_{xx} which is given by

$$S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2 = (n-1)(\text{Std. Dev.})^2 = 5 \cdot (1.53)^2 = 11.70.$$

The 95% or $\alpha = .05$ confidence interval for $E(\log(Y)|X = 15)$ is given by $\ln(y)^* \pm t_{\alpha/2, n-2} s_{\ln(y)^*}$ where $t_{.025,4} = 2.776$ and

$$s_{\ln(y)^*} = s\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}} = 0.05635\sqrt{\frac{1}{6} + \frac{(15-7)^2}{11.70}} = .1338$$

It follows that the CI for $E(\ln(Y)^*)$ is $1.2208 \pm 2.776 \cdot 0.1338$ or $(.849, 1.592)$. Exponentiating, the corresponding interval for Y is $(2.237, 4.914)$.

(3.) *Data on the gain of reading speed (y words/min.) and the number of weeks in a speed reading program (x) was recorded for 10 students selected at random from the program. A statistical software program indicated that the simple linear regression model for y was statistically significant, and provided the following information. Use this to answer the following questions.*

```
Regression equation:   sp.gain = 2.64 + 11.8 weeks
                       S2 = MSE = 115.6
Summary Statistics for
      No of Wks (x):   mean = 6.1; standard deviation = 2.807; min = 2; max = 11
  Gain in speed (y):   mean = 69.1; standard deviation = 33.3; min = 21; max = 130
```

i) *Give a 95% confidence interval estimate for the mean gain in speed for students who have been in the program for 9 weeks.*

The expected value at $x^* = 9$ weeks is $y^* = 2.64 + 11.8 \cdot 9 = 108.84$. We need to find S_{xx} which is given by

$$S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2 = (n-1)(\text{Std. Dev.})^2 = 9 \cdot (2.807)^2 = 70.91.$$

The 95% or $\alpha = .05$ confidence interval for $E(Y|X = 9)$ is given by $y^* \pm t_{\alpha/2, n-2} s_{y^*}$ where $t_{.025,8} = 2.306$ and

$$s_{y^*} = s\sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{S_{xx}}} = \sqrt{115.6}\sqrt{\frac{1}{10} + \frac{(9-6.1)^2}{70.91}} = 5.027$$

It follows that the CI for $E(Y^*)$ is $108.84 \pm 2.306 \cdot 5.027$ or $(97.25, 120.43)$.

ii) *What was the R-sq value for this model and what does it measure in the practical context of the problem?*

Using the fact that $\hat{\beta}_1 = S_{xy}/S_{xx}$, $S_{yy} = (n-1)(\text{St.Dev.}y)^2 = 9(33.3)^2 = 9980.01$ and S_{xx} from above, we find that

$$R = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} = \hat{\beta}_1 \sqrt{\frac{S_{xx}}{S_{yy}}} = 11.8\sqrt{\frac{70.91}{9980.01}} = .995.$$

$R^2 = .990$ is the coefficient of determination, which says under the hypotheses of the linear regression model, the proportion of the observed speed gain in reading accounted for by the number of weeks in the program through the simple linear model is 99.0%.

---

(4) *The following table are measurements of tensile strength x (in ksi) and Brinell hardness y for 10 specimens of cold drawn copper. Find a 95% confidence interval for ρ, the population correlation coefficient. Can you conclude that $\rho < 0.3$? Can you conclude that $\rho \neq 0$? What assumptions are you making about the data?*

x:	106.2	106.3	105.3	106.1	105.4	106.3	104.7	105.4	105.5	105.1
y:	35.0	37.2	39.8	35.8	41.3	40.7	38.7	40.2	38.1	41.6

We assume that the data points are a random sample from a bivariate normal distribution. In computing the correlation coefficiient, we find $S_{xx} = \sum x_i^2 - (\sum x_i)^2/n = 111579.79 -$

$(1056.3)^2/10 = 2.821$, $S_{xy} = \sum x_i y_i - (\sum x_i)(\sum y_i)/n = 41020.79 - (388.4)(1056.3)/10 = -5.902$ so $R < 0$ and $S_{yy} = \sum y_i^2 - (\sum y_i)^2/n = 15132.6 - (388.4)^2/1047.144 = 47.144$. Hence

$$R^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}} = \frac{(-5.902)^2}{2.821 \cdot 47.144} = .26192$$

Let us test whether $\mathcal{H}_0 : \ \rho = 0$ vs $\rho \neq 0$ using the fact that under $\mathcal{H}_0$, U has a t-distribution with $n-2$ degrees of freedom. We reject $\mathcal{H}_0$ if $|U| \geq t_{\alpha/2,n-2} = t_{.025,8} = 2.306$. We compute

$$U = R\sqrt{\frac{(n-2)}{1-R^2}} = -\sqrt{\frac{8 \cdot 0.26192}{1 - 0.26192}} = -1.685.$$

Thus, at the $\alpha = .05$ level, there is no strong evidence to indicate that $\rho \neq 0$.

Assume that we only conduct one of the tests (otherwise we need to use the Bonferroni or other estimate for simultaneaous tests.) Now let us test whether $\mathcal{H}_0 : \rho = .3$ vs $\mathcal{H}_a : \rho < .3$. Although there are not enough data points to use the Fisher Z-transform with a lot of confidence (we should have at least 20 points, using Sen & Srivastava's rule of thumb,) we shall use it anyway. We accept $\mathcal{H}_a$ provided that $z \leq -z_\alpha = -z_{.05} = -1.645$. Computing, $r = -.5118$ and

$$Z = \frac{\ln\left(\frac{1+R}{1-R}\right) - \ln\left(\frac{1+\rho_0}{1-\rho_0}\right)}{\frac{2}{\sqrt{n-3}}} = \frac{\ln\left(\frac{1+.5118}{1-.5118}\right) - \ln\left(\frac{1+.3}{1-.3}\right)}{\frac{2}{\sqrt{7}}} = -2.314$$

so we accept $\mathcal{H}_a$.

Using the Z statistic, we may find a level α, two sided CI for ρ. Computing $z_{\alpha/2} = z_{.025} = 1.960$,

$$v = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right) = \frac{1}{2}\ln\left(\frac{1-.5118}{1.5118}\right) = -.5652.$$

The CI for Z is $v \pm z_{\alpha/2}/\sqrt{n-3} = -.5652 \pm 1.960/\sqrt{7}$ or $(-1.306, .1756)$. Transforming back to ρ we have the CI for ρ is $(\tanh(v_1), \tanh(v_2)) = (\tanh(-1.306), \tanh(.1756) = (-.8632, .1739)$.

(7) *A windmill is used to generate direct current. Data are collected on 45 different days to determine the relationship between wind speed x mi/h and current y kA. Compute the least squares line for predicting Model I: y from x and the least squares line from predicting Moldel II: y from $\ln x$. Which of these two models fits best? Use ANOVA tables, scatterplots including the fitted lines, at least one residual plot, a plot of $\hat{y}$ vs y and the normal PP-plot. Are the model assumptions satisfied?*

Wind Speed	Curr.	Wind Speed	Curr.	Wind Speed	Curr.	Wind Speed	Curr.	Wind Speed	Curr.	Wind Speed	Curr.	Wind Speed	Curr.
4.2	1.9	1.8	0.3	1.6	1.1	10.7	3.2	9.2	2.9	2.6	1.4	2.3	1.2
1.4	0.7	5.8	2.3	2.3	1.5	5.3	2.3	4.4	1.8	7.7	2.8	11.9	3.0
6.6	2.2	7.3	2.6	4.2	1.5	5.1	1.9	8.0	2.6	6.1	2.4	8.6	2.5
4.7	2.0	7.1	2.7	3.7	2.1	4.9	2.3	10.5	3.0	5.5	2.2	5.6	2.1
2.6	1.1	6.4	2.4	5.9	2.2	8.3	3.1	5.1	2.1	4.7	2.3	4.2	1.7
5.8	2.6	4.6	2.2	6.0	2.6	7.1	2.3	5.8	2.5	4.0	2.0	6.2	2.3
7.7	2.6	6.6	2.9	6.9	2.6								

Model used is Current=WindSp

	Coef	StdErr	t
CONSTANT	0.83325	0.11355	7.338
WindSp	0.23542	0.018375	12.812

N: 45, MSE: 0.084662, DF: 43, R-sq: 0.79242
Regression F(1,43): 164.15, Durbin-Watson: 1.8832

	DF	SS	MS	F	P-value
CONSTANT	1	213.42	213.42	2520.86932	< 1e-08
WindSp	1	13.897	13.897	164.15015	< 1e-08
ERROR1	43	3.6405	0.084662		

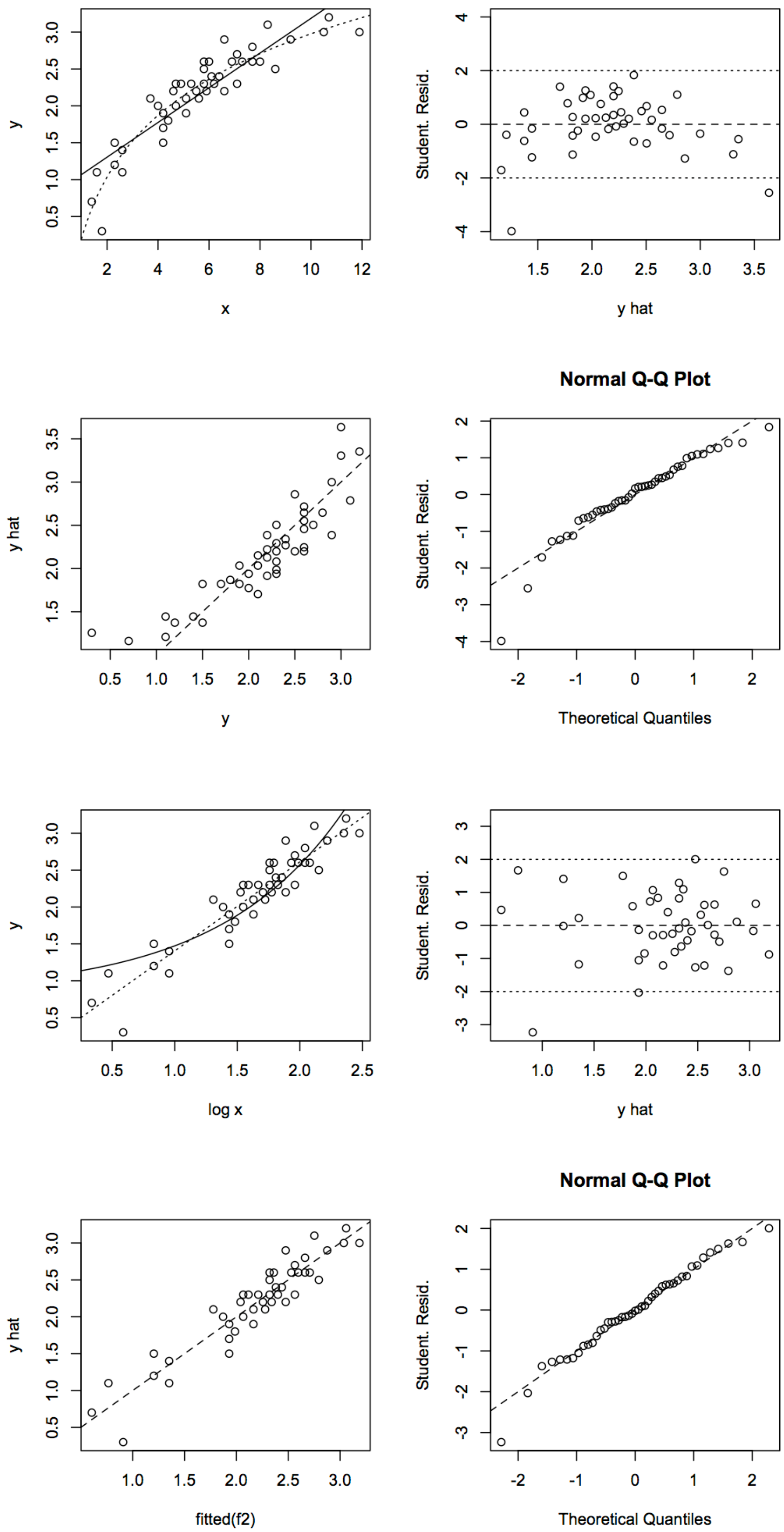
\medskip

Model used is Current = {log(WindSp)}

	Coef	StdErr	t
CONSTANT	0.19878	0.11677	1.7023
log(WindSp)	1.2066	0.068272	17.673

N: 45, MSE: 0.049354, DF: 43, R-sq: 0.87899
Regression F(1,43): 312.34, Durbin-Watson: 1.9299

	DF	SS	MS	F	P-value
CONSTANT	1	213.42	213.42	4324.27902	< 1e-08
log(WindSp)	1	15.416	15.416	312.34371	< 1e-08
ERROR1	43	2.1222	0.049354		



y
x
Student. Resid.
y hat
Normal Q-Q Plot
y hat
y
Student. Resid.
Theoretical Quantiles
y
log x
Student. Resid.
y hat
Normal Q-Q Plot
y hat
fitted(f2)
Student. Resid.
Theoretical Quantiles

For Model I, the regression line is $y = 0.833 + 0.235x$. For Model II, the regression line is $y = 0.199 + 1.207 \ln(x)$.

The top of the figure is Model I, The bottom is Model II. The upper left panel denotes the scatterplots (x_i, y_i) and $(\ln(x_i), y_i)$. The straight lines are the least squares regression lines. The curved line in the first panel is the fitted line from Model II, $(x, 0.199 + 1.207 \ln(x))$. the curved line in the fifth panel is the fitted line from Model I: $(\ln(x), 0.833 + 0.235x)$. Note that the Model II seems to capture the curvature in panel 1. The transformation of variables seems to straighten out the cloud of points.

The second and sixth panels are plots of $(\hat{y}_i, res_i)$ for each model. Note that the residual cloud seems to be bowed downward in the second panel. This indicates heteroscedasticity of the data: The mean of ε_i seems to depend on $\hat{y}$ whereas it should be constant. After transformation, the $\hat{y}_i$ vs residual plot seems to fall in a horizontal band centered on zero. Its residual plot shows the lesser pattern.

The third and seventh panels show the QQ-plot of residuals vs normal scores. The third panel is more skewed than the seventh. The seventh does not show appreciable deviation from normality. Thus the transformed Model II's residuals shows no gross violation of normality.

Finally, the fourth and eighth panels show $(y_i, \hat{y}_i)$. Panel four shows an upward curvature, indicating a systematic dependence, which is indicative of the straight line model not fitting the curve very well. But the eighth panel cloud of points follows the $y = \hat{y}$ line nicely, as it should do if the model is doing a good job.

Looking at the ANOVA tables from MacAnova, we see that $r^2 = 0.79242$ for Model I and $r^2 = 0.87899$ for Model II, which means that the transformed points are more linear in Model II. Model II does the better job fitting the line. The assumptions seem to be satisfied for Model II.

(8) *The article "The influence of temperature and sunshine on the alpha acid contents of hops," (Agricultural Meteorology, 1974) reports the following data on yield (y), mean temperature over the period between date of coming into hops and date of picking (x_1), and mean percentage of sunshine during the period (x_2), for the fuggle variety of hop.*

x_1 :	16.7	17.4	18.4	16.8	18.9	17.1	17.3	18.2	21.3	21.2	20.7	18.5
x_2 :	30	42	47	47	43	41	48	44	43	50	56	60
y :	210	110	103	103	91	76	73	70	68	53	45	31

Here are partial outputs from **R** *for the linear model and the quadratic model. Use the partial F-test to decide at the $\alpha = 0.05$ level whether the quadratic terms can be dropped from the quadratic model.*

```
lm(formula = y ~ x1 + x2)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  415.113     82.517   5.031 0.000709
x1            -6.593      4.859  -1.357 0.207913
x2            -4.504      1.071  -4.204 0.002292
Residual standard error: 24.45 on 9 degrees of freedom
Multiple R-squared: 0.768,Adjusted R-squared: 0.7164
F-statistic:  14.9 on 2 and 9 DF,  p-value: 0.001395 104

lm(formula = y ~ x1 + x2 + I(x1^2) + I(x1 * x2) + I(x2^2))
Coefficients:
```

```
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 1201.6196  1426.0798   0.843    0.432
x1           -43.6759   150.0072  -0.291    0.781
x2           -24.5398    12.7796  -1.920    0.103
I(x1^2)        0.3155     3.9908   0.079    0.940
I(x1 * x2)     0.5572     0.9400   0.593    0.575
I(x2^2)        0.1085     0.1128   0.962    0.373
Residual standard error: 23.22 on 6 degrees of freedom
Multiple R-squared: 0.8606,Adjusted R-squared: 0.7444
F-statistic: 7.408 on 5 and 6 DF,  p-value: 0.01507
```

We need SSE and DF for the full and reduced models. There are $n = 12$ data points, $k = 5$ variables in the full quadratic model and $j = 2$ in the reduced linear submodel not counting the constant. We are given only residual standard errors, which are $s = \sqrt{MSE} = \sqrt{SSE/(n-k-1)}$. Thus for the full model, $SSE(f) = (n-k-1)s(f)^2 = (23.22)^2 = 3235.0104$ and $SSE(r) = (n-j-1)s(r)^2 = 9(24.45)^2 = 5380.2225$. The partial F test statistic is

$$F = \frac{(SSE(r) - SSE(f))/(DF(r) - DF(f))}{SSE(f)/DF(f)} = \frac{(5380.2225 - 3235.0104)/(9-6)}{3235.0104/6} = 1.3262$$

The critical f-value is $f_{DF(r)-DF(f),DF(f)}(\alpha) = f_{3,6}(0.05) = 4.76$. In fact, the P-value is .3505. Thus we are unable to reject the null hypothesis: the quadratic terms may be dropped.