Yu-Hon Lin
Stat 463
Chunming Li
Homework 5

8.4 Elasticity of moissanite. Moissanite is a popular abrasive material because of its extreme hardness. Another important property of moissanite is elasticity. The elastic properties of the material were investigated in the *Journal of Applied Physics* (September 1993). A diamond anvil cell was used to compress a mixture of moissanite, sodium chloride, and gold in a ratio of 33:99:1 by volume. The compressed volume, y, of the mixture (relative to the zero-pressure volume) was measured at each of 11 different pressures (GPa). The results are displayed in the table (p. 397). A MINITAB printout for the straight-line regression model $E(y) = \beta_0 + \beta_1 x$ and a MINITAB residual plot are displayed at left.

MINITAB Output for Exercise 8.4

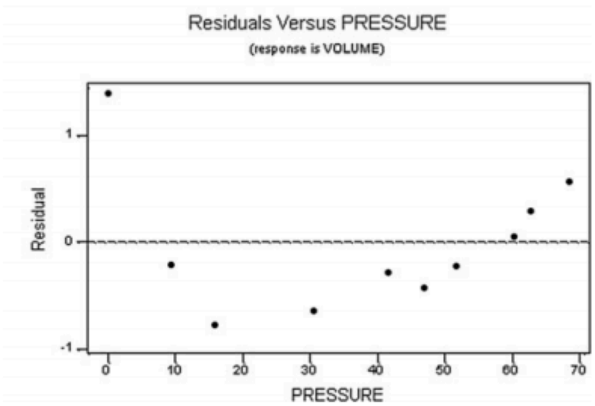
```
The regression equation is
VOLUME = 98.6 - 0.256 PRESSURE

Predictor        Coef     SE Coef        T       P
Constant      98.6149      0.4037   244.26   0.000
PRESSURE    -0.255594    0.008646   -29.56   0.000

S = 0.6484      R-Sq = 99.0%     R-Sq(adj) = 98.9%

Analysis of Variance

Source          DF        SS        MS        F        P
Regression       1    367.34    367.34   873.87    0.000
Residual Error   9      3.78      0.42
Total           10    371.12
```



MOISSANITE

COMPRESSED VOLUME y, %	PRESSURE x, GPa
100	0
96	9.4
93.8	15.8
90.2	30.4
87.7	41.6
86.2	46.9
85.2	51.6
83.3	60.1
82.9	62.6
82.9	62.6
81.7	68.4

Source: Bassett, W. A., Weathers, M. S., and Wu, T. C. "Compressibility of SiC up to 68.4 GPa," *Journal of Applied Physics*, Vol. 74. No. 6, Sept. 15, 1993, p. 3825 (Table 1). Reprinted with permission from Journal of Applied Physics. Copyright © 1993, American Institute of Physics.

(a) Calculate the regression residuals.

$$\beta_0 = 98.6149, \beta_1 = -0.255594, \hat{\varepsilon} = y - \hat{y} = y - 98.6149 + 0.255594x$$

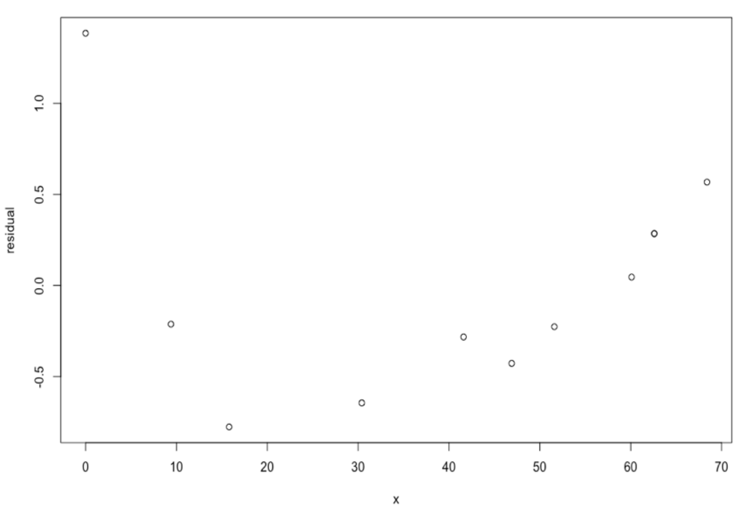
x	y	$\hat{y}$	$\hat{\varepsilon}$
0	100	98.6149	**1.3851**
9.4	96	96.2123	**–0.2123**
15.8	93.8	94.5765	**–0.7765**
30.4	90.2	90.8448	**–0.6448**
41.6	87.7	87.9822	**–0.2822**
46.9	86.2	86.6275	**–0.4275**
51.6	85.2	85.4262	**–0.2262**
60.1	83.3	83.2537	**0.0463**
62.6	82.9	82.6147	**0.2853**
62.6	82.9	82.6147	**0.2853**
68.4	81.7	81.1323	**0.5677**

(b) Plot the residuals against x. Do you detect a trend?

Console Terminal

~/

```
> x<-c(0,9.4,15.8,30.4,41.6,46.9,51.6,60.1,62.6,62.6,68.4)
> residual<-c(1.3851,-0.2123,-0.7765,-0.6448,-0.2822,-0.4275,-0.2262,0.0463,0.2853,0.2853,0.5677)
> plot(residual~x)
>
```



The plot reveals a clear parabolic trend, implying a lack of fit.

Ans: Lack of linear fit

(c) Propose an alternative model based on the plot, part (b).

$E(y) = \beta_0 + \beta_1 x + \beta_2 x^2$

Ans: $\boldsymbol{E(y) = \beta_0 + \beta_1 x + \beta_2 x^2}$

(d) Fit and analyze the model you proposed in part (c).

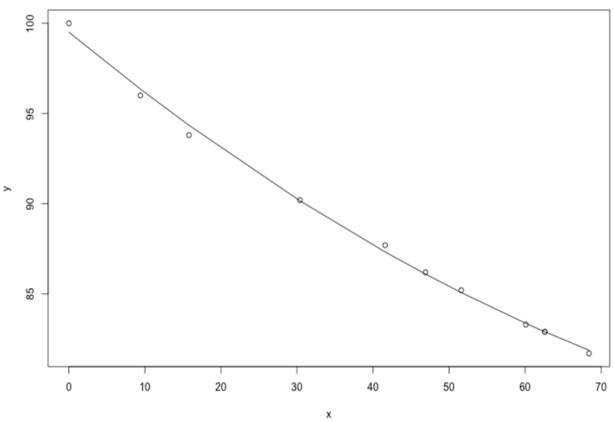
```
Console  Terminal ×
~/
> x<-c(0,9.4,15.8,30.4,41.6,46.9,51.6,60.1,62.6,62.6,68.4)
> y<-c(100,96,93.8,90.2,87.7,86.2,85.2,83.3,82.9,82.9,81.7)
> fit<-lm(y~x+I(x^2))
> plot(x,y)
> points(x, predict(fit), type="l")
> summary(fit)

Call:
lm(formula = y ~ x + I(x^2))

Residuals:
     Min       1Q   Median       3Q      Max 
-0.54312 -0.12355  0.00034  0.11405  0.49717 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 99.502832   0.268286  370.88  < 2e-16 ***
x           -0.347272   0.018306  -18.97 6.17e-08 ***
I(x^2)       0.001311   0.000254    5.16 0.000864 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3305 on 8 degrees of freedom
Multiple R-squared:  0.9976,	Adjusted R-squared:  0.9971 
F-statistic:  1694 on 2 and 8 DF,  p-value: 3.077e-11
```



$\beta_0 = 99.502832, \beta_1 = -0.347272, \beta_2 = 0.001311,$

$\hat{\varepsilon} = y - \hat{y} = y - 99.502832 + 0.347272x - 0.001311x^2$

x	y	$\hat{y}$	$\hat{\varepsilon}$
0	100	99.502832	**0.497168**
9.4	96	96.3543152	**–0.3543152**
15.8	93.8	94.3432124	**–0.5432124**
30.4	90.2	90.157337	**0.04266304**
41.6	87.7	87.325081	**0.37491904**
46.9	86.2	86.0994639	**0.10053609**
51.6	85.2	85.0741215	**0.12578704**
60.1	83.3	83.3671299	**–0.0671299**
62.6	82.9	82.9010992	**–0.0010992**
62.6	82.9	82.9010992	**–0.0010992**
68.4	81.7	81.8830194	**–0.1830194**

Conclusion: $t_{\beta_2} = 5.16$ ($p-$value: 0.000864), the model adequacy is improved.

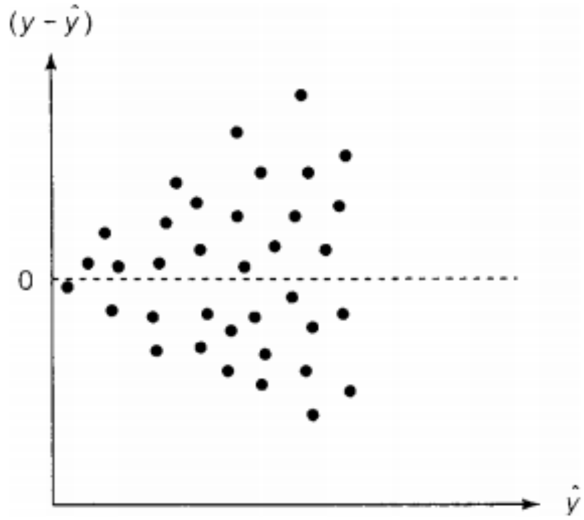
8.13 Assembly line breakdowns. Breakdowns of machines that produce steel cans are very costly. The more breakdown, the fewer cans produced, and the smaller the company's profits. To help anticipate profit loss, the owners of a can company would like to find a model that will predict the number of breakdowns on the assembly line. The model proposed by the company's statisticians is the following:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon$$

where y is the number of breakdowns per 8-hour shift,

$$x_1 = \begin{cases} 1 & \text{if afternoon shift} \\ 0 & \text{otherwise} \end{cases} \quad x_2 = \begin{cases} 1 & \text{if midnight shift} \\ 0 & \text{otherwise} \end{cases}$$

x_3 is the temperature of the planet (°F), and x_4 is the number of inexperienced personnel working on the assembly line. After the model is fit using the least squares procedure, the residuals are plotted against $\hat{y}$, as shown in the accompanying figure.



(a) Do you detect a pattern in the residual plot? What does this suggest about the least squares assumptions?

Ans: Yes; assumption of equal variances violated.

(b) Given the nature of the response variable y and the pattern detected in part (a), what model adjustments would you recommend?

Ans: Use transformation $y^* = \sqrt{y}$