12.6.6

We can write table as

Source	Df	Sum of squares	Mean squares	F-statistic	p-value
Regression	1	87.5889	87.5889	1.2454	0.2905
Error	10	703.3278	70.3328		
Total	11	790.9167			

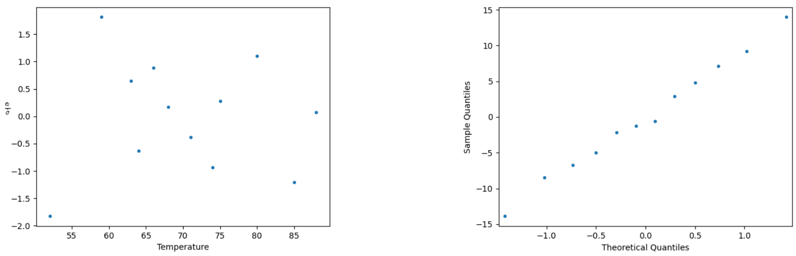
The coefficient of determination is

$$R^2 = \frac{SSR}{SST} = 0.1107$$

The t-statistic for testing $H_0 : \beta_1 = 0$ is

$$t = \frac{\hat{\beta}_1}{\hat{\sigma}/\sqrt{S_{xx}}} = 1.1160 \Rightarrow t^2 = 1.2454 = F$$

The large p-value implies that there is no sufficient evidence to conclude that unloading time depends upon the temperature.



(a) Standardized residual

(b) Normal probability plot

12.7.2

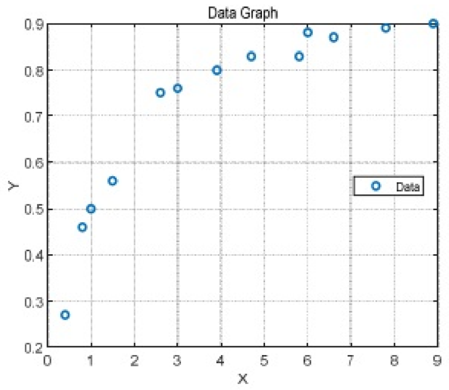
There are no points that might be considered to be outliers since absolute value of standardized residual is smaller than 3.
The residual plot has no patterns that suggest that the fitted regression model is not appropriate.
The normal probability plot is approximately on a straight line, then the error terms are normally distributed.

12.7.6

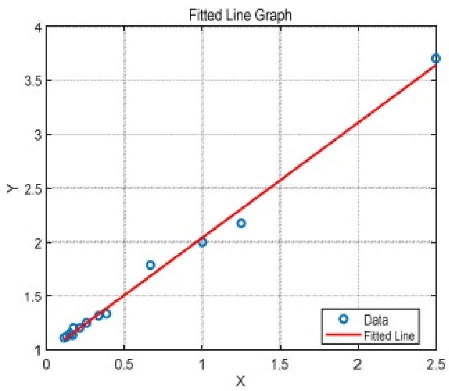
The plot shows that the residuals against the temperature increase

- You get 10 points for plot whether you arrive at any conclusion.

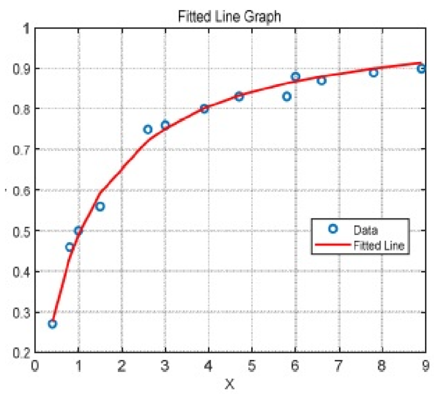
12.8.2 Plot of the data:



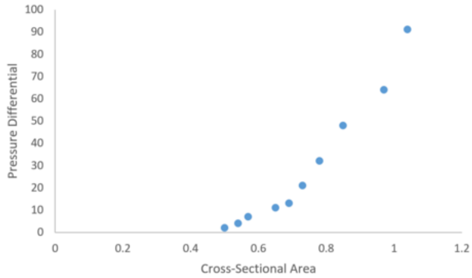
Use model $1/y = \gamma_1 + \gamma_0 \times 1/x$.**(+2 points)** By replacing $1/x$ by $x^{'}$ and $1/y$ by $y^{'}$, we get $\gamma_1 = \hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = 1.067$ and $\gamma_0 = \hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = 0.974$. We have straight line $y^{'} = 0.974 + 1.067x^{'}$.**(+5 points)**



The fitted model back in terms of the original variables is $y = \frac{x}{1.067+0.974x}$. For $x = 2$, predicted value is $y = \frac{2}{1.067+0.974\times 2} = 0.663$.**(+3 points)**

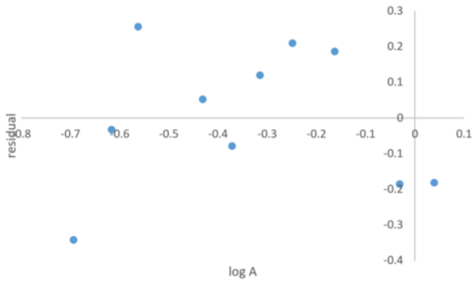


12.8.4 (a) A plot of given data is given by

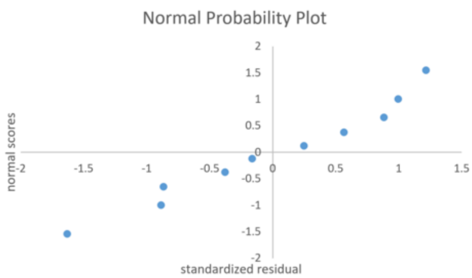


Thus, the model $P = \gamma_0 A^{\gamma_1}$ looks appropriate.

Apply the residual anaysis. After calculate all the values of (b), residual plot is given by



And using $\hat{\sigma}^2 = 0.0442$, the residual probability plot is given by



It looks like a single line so the model $P = \gamma_0 A^{\gamma_1}$ might be appropriate.

12.8.4

(b) We have $\log P = \gamma_1 \log A + \log \gamma_0$. By using simple linear regression model $\log A$ as input and $\log P$ as output. Lwr $x_i = \log A_i$ and $y_i = \log P_i$ where (A_i, P_i) is i-th data. Then with some computation, we can conclude that $\sum_{i=1}^{10} x_i = -3.390$ $\sum_{i=1}^{10} y_i = 28.039$, $\sum_{i=1}^{10} x_i^2 = 1.689$, $\sum_{i=1}^{10} y_i^2 = 92.428$, and $\sum_{i=1}^{10} x_i y_i = -6.811$. By using it, we can conclude that $\hat{\gamma}_1 = 4.993$, $\log \hat{\gamma}_0 = 4.497$. Therefore, we can conclude that $\hat{\gamma}_0 = 89.722$, $\hat{\gamma}_1 = 4.993$.

(c) Using values in (b), we can compute $\hat{\sigma}^2 = 0.0442$, $S_{xx} = 0.540$, and s.e.$(\hat{\gamma}_1) = 0.286$. Therefore, 95% confidence interval of γ_1 is $(\hat{\beta}_1 - t_{\alpha/2,n-2} \times \text{s.e.}(\hat{\beta}_1), \hat{\beta}_1 + t_{\alpha/2,n-2} \times \text{s.e.}(\hat{\beta}_1)) = (4.3335, 5.6527)$. Similarly, 95% confidence interval of $\log \gamma_0$ is $(\log \hat{\gamma}_0 - t_{\alpha/2,n-2}\hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{x}}{S_{xx}}}, \log \hat{\gamma}_0 + t_{\alpha/2,n-2}\hat{\sigma}\sqrt{\frac{1}{n} + \frac{\bar{x}}{S_{xx}}}) = (4.2256, 4.7678)$. By taking exponential, we can conclude that 95% confidence interval of γ_0 is $(68.4185, 117.6589)$.

12.8.6 Taking ln to both side of $e^{y/\gamma_0} = \gamma_1/x^2$, we have

$$y = \gamma_0 \ln(\gamma_1) - 2\gamma_0 \ln(x).$$

Therefore, the simple linear regression model with $\ln(x)$ as the input variable is given by

$$y = \hat{\beta}_0 + \hat{\beta}_1 \ln(x) = \hat{\gamma}_0 \ln(\hat{\gamma}_1) - 2\hat{\gamma}_0 \ln(x).$$

That is, we have $\hat{\beta}_0 = \hat{\gamma}_0 \ln(\hat{\gamma}_1)$ and $\hat{\beta}_1 = -2\hat{\gamma}_0$. Thus, we have

$$\hat{\gamma}_0 = -\frac{\hat{\beta}_1}{2} \quad \text{and} \quad \hat{\gamma}_1 = e^{-\frac{2\hat{\beta}_0}{\hat{\beta}_1}}.$$

12.9.4 The data set of the times taken to unload a truck at a warehouse given in DS 12.2.2., the simple linear regression model is applied here.

i	y_i	x_i	y_i^2	$x_i y_i$	x_i^2
1	64	52	4096	3328	2704
2	53	68	2809	3604	4624
3	58	64	3364	3712	4096
4	59	88	3481	5192	7744
5	49	80	2401	3920	6400
6	54	75	2916	4050	5625
7	38	59	1444	2242	3481
8	48	63	2304	3024	3969
9	68	85	4624	5780	7225
10	63	74	3969	4662	5476
11	58	71	3364	4118	5041
12	47	66	2209	3102	4356
$\sum$	659	845	36981	46734	60741

There are "n" = 12 points in this data set. Hand calculations would be started by finding the following five sums:

$$S_y = \sum y_i = 659, \qquad S_x = \sum x_i = 845,$$

$$S_{yy} + \frac{S_y^2}{n} = \sum y_i^2 = 36981, \quad S_{xx} + \frac{S_x^2}{n} = \sum x_i^2 = 60741$$

$$S_{xy} + \frac{S_x S_y}{n} = \sum x_i y_i = 46734$$

These quantities would be used to calculate the estimates of the regression coefficients, and their standard errors.

$$\widehat{\beta_1} = \frac{nS_{xy}}{nS_{xx}} = \frac{12 \cdot 46734 - 659 \cdot 845}{12 \cdot 60741 - 845^2} \approx 0.266$$

$$\widehat{\beta_0} = \frac{1}{n} S_y - \widehat{\beta_1} \frac{1}{n} S_x \approx 36.194$$

$$\widehat{\sigma}^2 = \frac{1}{n(n-2)} \left[nS_{yy} - \widehat{\beta_1}^2 nS_{xx} \right] \approx 70.333$$

$$\text{s.e}(\widehat{\beta_1}) = \sqrt{\frac{n\widehat{\sigma}^2}{nS_{xx}}} \approx 0.238$$

12.9.4 The normality assumption allows us to construct a t-value

$$t = \frac{\widehat{\beta_1}}{\text{s.e}(\widehat{\beta_1})} \approx 1.116 \sim\ t_{n-2}.$$

The product-moment correlation coefficient might also be calculated:

$$r = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} \approx 0.333$$

And the $t = r\sqrt{n-2}/(\sqrt{1-r^2})$ is approximately 1.116. So, we guess that the two t-statistics calculating using different methods have the same values. **(+10 points)** One can prove it. First, observe that

$$r^2 = \frac{S_{xy}^2}{S_{xx}S_{yy}} = \widehat{\beta_1}^2 \frac{S_{xx}}{S_{yy}}.$$

Therefore,

$$(n-2)\widehat{\sigma}^2 = S_{yy} - \widehat{\beta_1}^2 S_{xx} = S_{yy} - S_{yy}r^2 = S_{yy}(1-r^2).$$

So, we have:

$$r\sqrt{n-2}/(\sqrt{1-r^2}) = \frac{\sqrt{n-2}}{\sqrt{1-r^2}\sqrt{S_{yy}}} \frac{S_{xy}}{\sqrt{S_{xx}}} = \frac{\sqrt{n-2}}{\sqrt{(n-2)\widehat{\sigma}^2}} \frac{S_{xy}}{\sqrt{S_{xx}}} = \frac{S_{xy}}{\widehat{\sigma}\sqrt{S_{xx}}}$$

And the above is equals to $\frac{\widehat{\beta_1}}{\text{s.e.}(\widehat{\beta_1})}$, because

$$\frac{\widehat{\beta_1}}{\text{s.e.}(\widehat{\beta_1})} = \frac{S_{xy}}{s_{xx}} \Bigg/ \sqrt{\frac{\widehat{\sigma}^2}{S_{xx}}} = \frac{S_{xy}}{\widehat{\sigma}\sqrt{S_{xx}}}.$$

12.12.19 Figure suggest use of exponential model :

$$y = \gamma_0 e^{\gamma_1 x}.$$

Therefore we transform by taking ln both side to obtain

$$\ln y = \ln \gamma_0 + \gamma_1 \ln x.$$

Define $\ln y_i = z_i$ and $\ln x_i = w_i$. Then we have :

i	**Diameter**	**Strength**	w_i	z_i	w_i^2	$w_i z_i$	z_i^2
1	2.5	81	0.91629	4.39445	0.83959	4.02659	19.3112
2	3	167	1.09861	5.11799	1.20695	5.62269	26.1939
3	4	244	1.38629	5.49717	1.92181	7.62069	30.2189
4	5	484	1.60944	6.18209	2.59029	9.94968	38.2182
5	6	623	1.79176	6.43455	3.21040	11.5292	41.4034
6	8	1140	2.07944	7.03878	4.32408	14.6367	49.5445
7	9	1455	2.19723	7.28276	4.82780	16.0019	53.0386
8	11	2457	2.39790	7.80670	5.74990	18.7196	60.9445
9	13	3140	2.56495	8.05198	6.57897	20.6529	64.8344
10	16	6170	2.77259	8.72745	7.68725	24.1976	76.1685
$\sum$	77.5	15961	18.8145	66.5339	38.9370	132.958	459.876

There are "n" = 10 points in this data set. Hand calculations would be started by finding the following five sums:

$$S_w = \sum w_i = 18.814, \qquad S_z = \sum z_i = 66.554,$$

$$S_{ww} + \frac{S_w^2}{n} = \sum w_i^2 = 38.947, \quad S_{zz} + \frac{S_z^2}{n} = \sum z_i^2 = 459.876$$

$$S_{wz} + \frac{S_w S_z}{n} = \sum w_i z_i = 132.958$$

These quantities would be used to calculate the estimates of the regression coefficients (you can see the formulas in the solution of 12.9.4)

$$\widehat{\gamma_1} = 2.1979 \quad \text{and} \quad \widehat{\ln \gamma_0} = 2.5181.$$

And the R^2 is $\widehat{\gamma_1}^2 \frac{S_{ww}}{S_{zz}} = 0.994$. The form is

$$y = 12.405 e^{21979x} \qquad \because 12.405 = e^{\widehat{\ln \gamma_0}} \quad \textbf{(+10 points)}$$