Project Analysis: Ball Bearings

Anjani Chaudhary and Kyle Garza

12/3/2017

I. Abstract

Given the data published by Lieblein and Zelen in 1956 and the analysis of Caroni, we will attempt to recreate a snippet of their analysis and create our own model from the same data. The data contains 210 observations of the rating life of ball bearings, the load during the test, the number of balls each bearing has, the diameter of those balls, and more. After recreating Caroni's analysis and using a weighted least squares regression along with a backward variable selection algorithm on the regressors, we found our conclusions were similar to that of Caroni's with one exception. Caroni found that the value for *p* did not vary from one company to another, but we noticed a small variation in *p* for company B.

II. Introduction

In the article by Lieblein and Zelen[2], a particular model for the life span of ball bearings is theoretically derived and its coefficients are tested to see if the model is applicable across the whole industry or to specific manufacturers, or by bearing type.

Our goal is to check the hypothesized model by replicating the analysis presented in the article by Caroni[1] and completing our own exploratory analysis of the given data.

We found our data in the Journal for Statistics Education at the American Statistical Association Website. The data consists of 210 tests, each test involved running a batch of bearings of the same type under the same conditions, measuring L_{10} and L_{50} for each test. The data came with 11 total columns, two of which were the response variables L_{10} and L_{50}. There were three regressors, *P*, *D*, and *Z*, that were the main focus of the analysis. An additional interaction analysis was done on Company and Bearing Type (for Company B only). The Test Number and Year were largely ignored, and Number of bearings and Weibull Slope were used to fix the linear regression assumptions.

III. Article Summary and Replicated analysis

The main concern of the article is to check the reliability of ball bearings using a relationship between the load placed on the bearing during use, the diameter of the balls in the bearing, the number of balls in the bearing, and the type of bearing. Bearing life is a measure of the number of millions of revolutions where 90% (for L_{10}) of the ball bearings can be expected to survive. The following model for the rating of the life span for ball bearings, *L*, is the industry standard.

$$L = \left[\frac{f Z^{a} D^{b}}{P} \wedge \right]^{p} \tag{1}$$

where *Z* is the number of balls in the tested bearing, *D* is the diameter of the balls, *P* is the operational load, and *f*, *a*, *b*, and *p* are all constants to be determined. In particular, we will focus on the parameter *p* as there was some concern in the industry mentioned in the Lieblein and Zelen article that *p* could be either 3 or 4.

If we take the natural logarithm of both sides of equation (1) our model can be expressed as a linear combination of the natural logarithms of our variables *Z*, *D*, and *P*.

$$ln(L) = p\, ln(f) + ap\, ln(Z) + bp\, ln(D) - p\, ln(P) \tag{2}$$

We will do our analyses using the following parameters:

$$\beta_0 = p\, ln(f),\ \beta_1 = ap,\ \beta_2 = bp,\ and\, \beta_3 =- p$$

Several hypotheses were tested.

(a) All the parameters of the equation are the same for each one of the three companies.
(b) The parameter -b_3 (hence, *p*) is the same for each company.
(c) All the parameters of the equation are the same for each one of the three types of bearing produced by Company *B*
(d) The parameter -b_3 is the same for each type of bearing produced by Company *B*.

However, for this paper we did our analyses only on hypotheses (a) and (b).

To test hypothesis (a) we will want a model that allows the parameters to vary freely between the companies. To do this we can add in the indicator variables *B* and *C* for companies B and C respectively, as well as six interaction variables that will take the form *B*ln(*Z*), one for each combination of company and the other three regressor variables.

$$\begin{aligned} ln(L) = \beta_0 &+ \beta_1 ln(Z) + \beta_2 ln(D) + \beta_3 ln(P) + \beta_4 B + \beta_5 C \\ &+ \beta_6 Bln(Z) + \beta_7 Bln(D) + \beta_8 Bln(P) \\ &+ \beta_9 Cln(Z) + \beta_{10} Cln(D) + \beta_{11} Cln(P) \end{aligned} \tag{3}$$

To test (b) we will remove the interaction terms from equation (3) that contain ln(*P*).

In order to do a least-squares regression we will need to test that our assumptions are not violated. According to Montgomery et al. these assumptions are:

1. *The relationship between the response y and the regressors is linear, at least approximately.*
2. *The error term ε has zero mean.*
3. *The error term ε has constant variance σ^2.*
4. *The errors are uncorrelated.*
5. *The errors are normally distributed.*

We will test a few of these here for all three models that we will be analyzing.

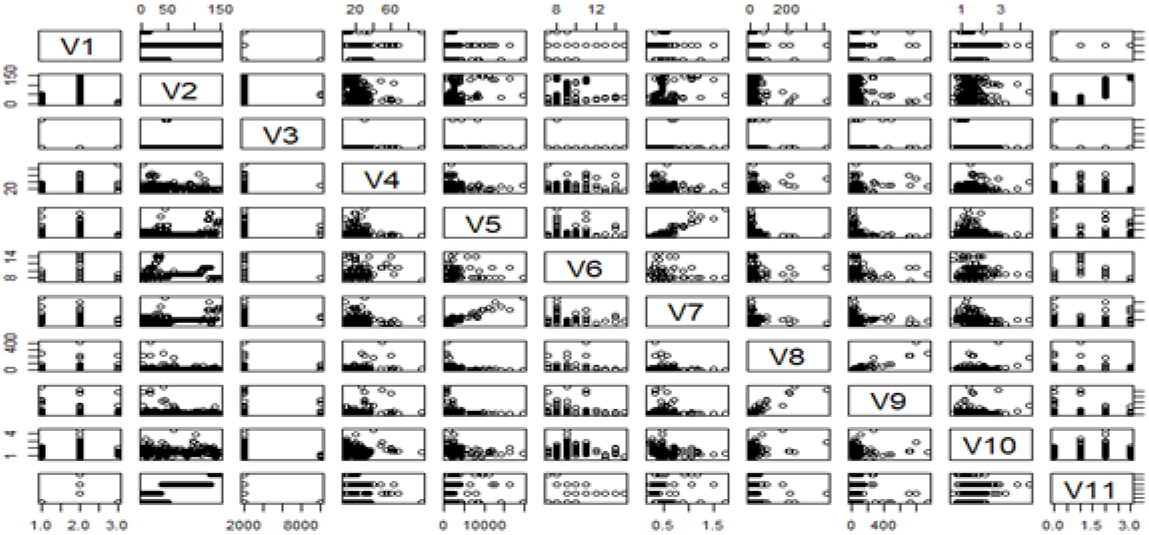
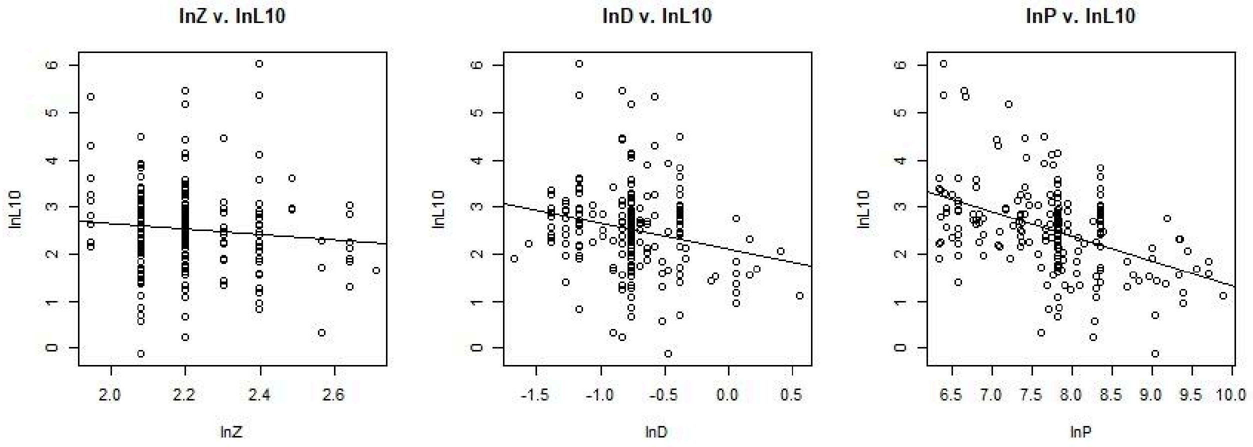


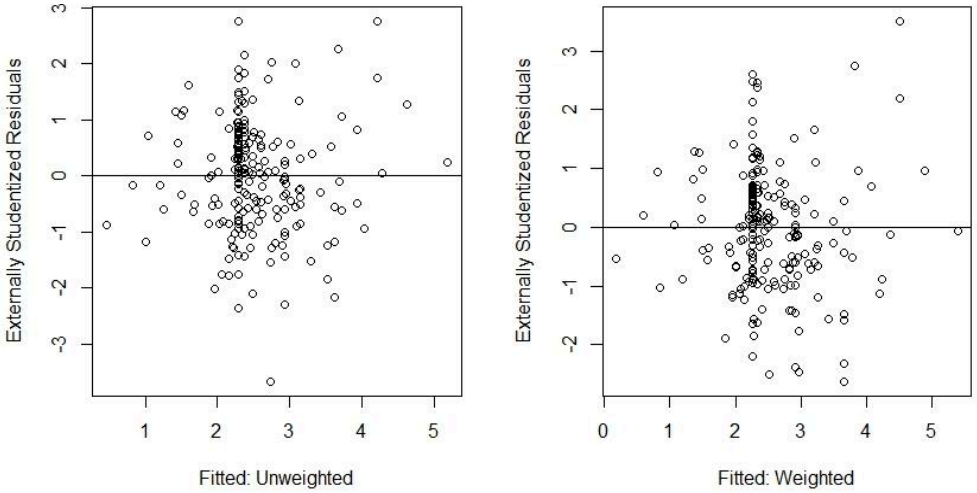
Figure: scatter plot of regressors

First, looking at the graphs of the regressors versus the response variables in the scatter plots below, there is a weak linear correlation between each of the regressors ln$Z$, ln$D$, and ln$P$, and the response lnL_{10}.

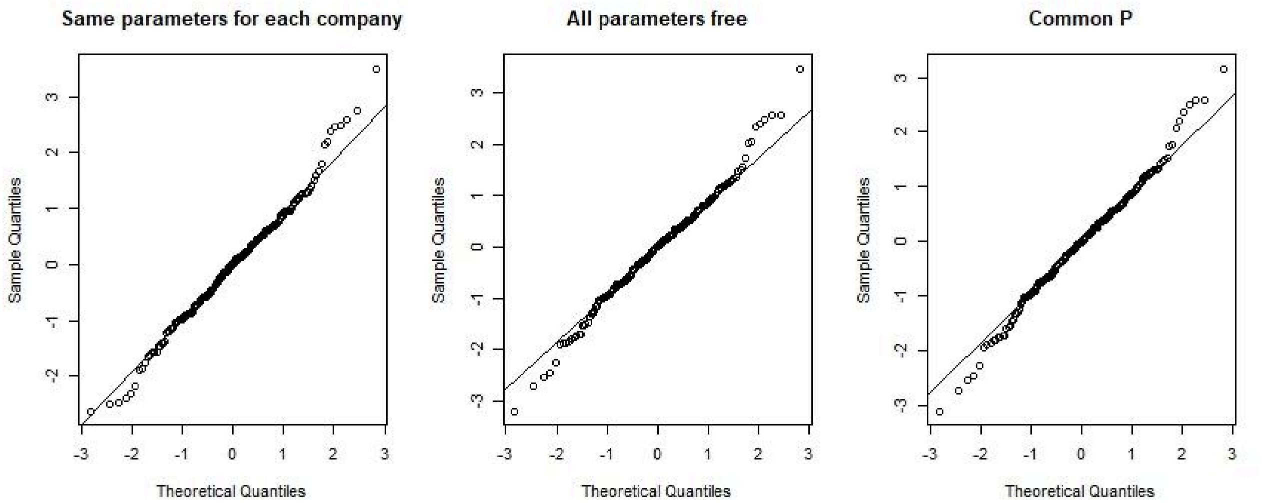


The assumption that the error term ε has zero mean is taken care of by the method of least-squares.

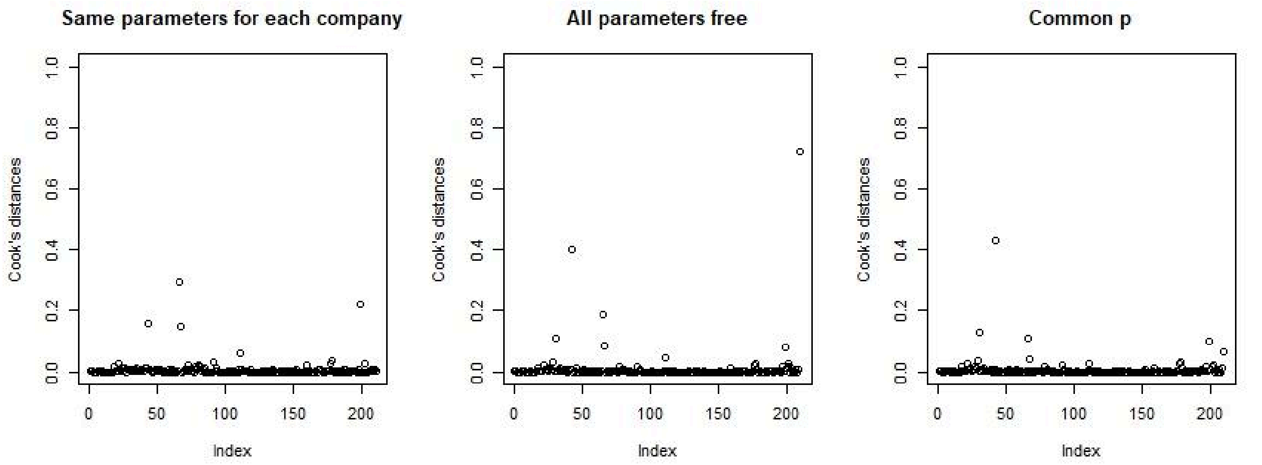
Caroni fit their regression models using weighted least squares. The author notes that "if unweighted analyses are used and the residuals are examined in relation to the number of bearings in the test, there does not seem to be any sign of the expected heteroscedasticity." They go on to say, "similar results should be obtained using ordinary least squares." The authors point can be clearly seen in the residual plots below. Regardless, we chose to use weighted least-squares to handle the non-constant variance due to the difference in the number (N) of bearings tested ($\sigma^2 \sim 1/N$), and so that our results would be consistent with that of Caroni.



From the normal probability plots below we may conclude that the externally standardized residuals are close to a normal distribution for all three models.



Given that there are 210 observations there are no excessively large values. No outliers were found using Cook's distance statistic. Looking at the model that assumed a common p, one observation, $D_{43} = 0.43$, stood out, but had negligible effect on *p* when removed.



To test the hypothesis that all the parameters of the equation are the same for each one of the three companies we ran an extra sum of squares test on our two nested models (Model 2 - Model 1 in the table below) using the anova() function in R. From the analysis, whose results are shown in the table, there is a 0.09674% chance that the observed differences in the two models are due to random variation. Using the standard significance level $\alpha = 0.05$ that we will be using throughout this paper, this is sufficient evidence to reject the hypothesis that all the parameters of the equation are the same for each one of the three companies.

To test the second hypothesis that the parameter $-\beta_3$ (hence, *p*) is the same for each company we repeated the extra sum of squares test, but this time on Model 3 - Model 1. The results are in the table. Since the probability here is 75.27%, there is not sufficient evidence to reject the hypothesis, and we may conclude that the parameter -β_3 is the same for each company.

<table>
<tr><td colspan="7">Analysis of Variance Table</td></tr>
<tr><td>Model 1</td><td colspan="6">$ln(L_{10}) \sim ln(Z) + ln(D) + ln(P) + B + C + Bln(Z) + Bln(D) + Bln(P) + Cln(Z) + Cln(D)$</td></tr>
<tr><td>Model 2</td><td colspan="6">$ln((L_{10}) \sim ln(Z) + ln(D) + ln(P)$</td></tr>
<tr><td>Model 3</td><td colspan="6">$ln(L_{10}) \sim ln(Z) + ln(D) + ln(P) + B + C + Bln(Z) + Bln(D) + Cln(Z) + Cln(D)$</td></tr>
<tr><td></td><td>Res. DF</td><td>RSS</td><td>DF</td><td>SS</td><td>F</td><td>P(>F)</td></tr>
<tr><td>1</td><td>198</td><td>1944.5</td><td></td><td></td><td></td><td></td></tr>
<tr><td>2 (2-1)</td><td>206</td><td>2215.4</td><td>8</td><td>270.88</td><td>3.4478</td><td>0.0009674</td></tr>
<tr><td>3 (3-1)</td><td>200</td><td>1950.1</td><td>2</td><td>5.5869</td><td>0.2844</td><td>0.7527</td></tr>
</table>

The article by Caroni went on to show the analyses for hypotheses (c) and (d). The results were similar to (a) and (b). Caroni also suggested that analyses could be done for other models such as: (1) assume $p = 3$ and use the new response variable $ln(L_{10}) + 3ln(P)$, (2) we could similarly assume $a = ⅔$, or $b = 1.8$, or (3) fit equation (3) to each company separately, then add the three residual sums of squares from the three analyses (what Lieblein and Zelen actually did).

IV. Our Models and Our Analysis Results

In this next section we will be using the Lieblein-Zelen data to build a model from the ground up. Taking a queue from the article, though, we will not assume all the variables given can be possible regressors, and we will only look at L_{10} and not L_{50} as a response. Our possible regressors will be limited to *P*, *Z*, *D*, and we will include indicator variables for the values in the "Company" column of the data. However, we will only add in the company data after we have properly analyzed the regressors *P*, *Z* and *D*.

Model Equation (i) : $L_{10} = \beta_0 + \beta_1 Z + \beta_2 D + \beta_3 P$

Running a quick lm(), summary(), and anova() in R for Model Equation (i) we get the following results:

$$L_{10} = -39.266424 + 3.618477Z + 121.429771D - 0.011051P$$

$$R^2 = 0.08589,\ R^2_{Adj} = 0.07258$$

It is acceptable to use a test for lack-of-fit since we have true replicate observations on L_{10} for multiple levels of each of the regressors, the straight-line character was doubtful, based on the article analysis, and the assumptions normality, independence, and constant variance are met. The results of the test for lack-of-fit in R give a p-value of $p < 2.2 \times 10^{-16}$ with $F_0 = \frac{2905.0}{272.4} = 10.6658$on 101 over 105 degrees of freedom. Thus, our data does not adequately fit this linear model, and we must look for an appropriate transformation.

1. Error has mean 0

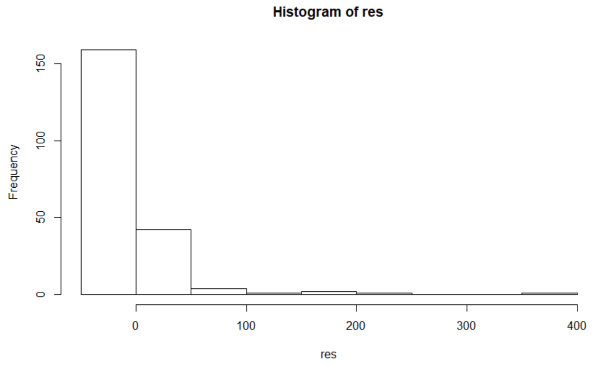


Figure: Histogram of residuals

2. Error has constant variance

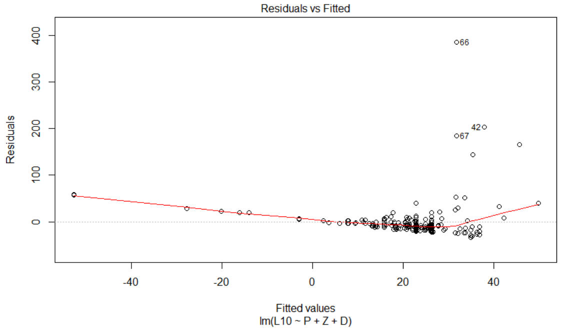


Figure: Fitted values versus residuals

3. Errors are normally distributed

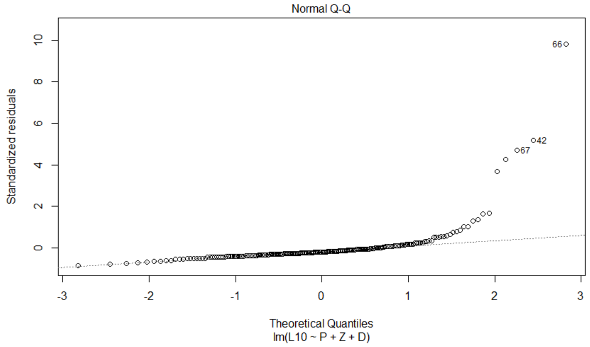
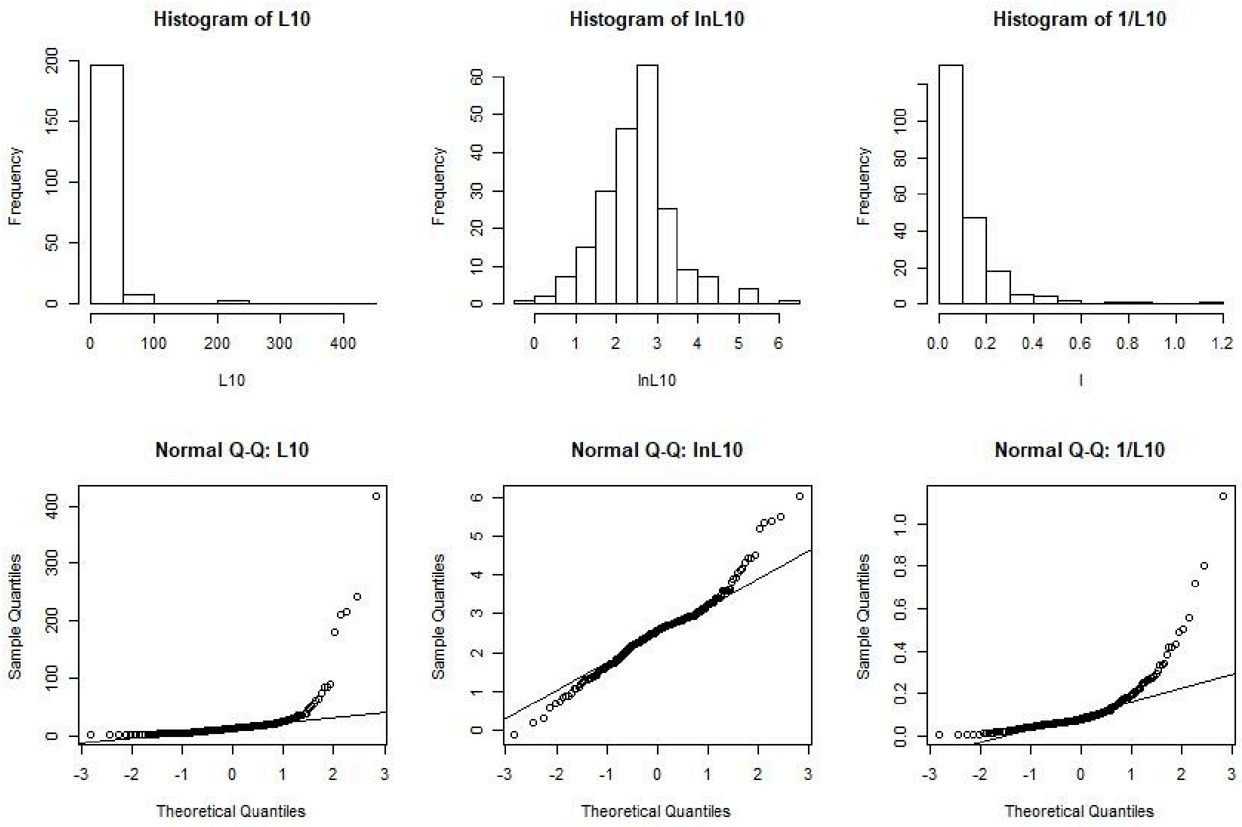


Figure: Q-Q normal plot

To determine which transformations should be used on our data we made histograms and Normal Q-Q plots of all the variables in question as well as the natural logarithmic and inverse transformations of those variables. From the histograms of L_{10}, $\ln(L_{10})$, and $1/L_{10}$, we can safely assume a logarithmic transformation on L_{10} is appropriate. While certainly not as pretty as L_{10}, similar checks on the regressors suggest logarithmic transformations as well.



Applying the natural logarithmic transformations gives the following model. As above in the article summary section, while using a weighted least-squares will yield similar results to not, for consistency's sake we will use a weighted least-squares with weights given by the number (*N*) of bearings tested in each observation.

Model Equation (ii): $lnL_{10} = \beta_0 + \beta_1 lnZ + \beta_2 lnD + \beta_3 lnP$with $Weight = N$, where *N*= number of bearings.

R produces the following results:

$$ln(L_{10}) = 20.4329 + 2.1611ln(Z) + 4.3322ln(D) - 2.5115ln(P)$$

$$R^2 = 0.5508, \; R^2_{Adj} = 0.5442$$

Since our R^2 has increased greatly from 8.589% to 55.08%, this model is a vast improvement over the first. However, using a test for lack-of-fit in R we get a *p*-value of p = 0.0001795 with $F_0 = \frac{14.52}{7.13} = 2.0350$on 101 over 105 degrees of freedom. From the lack-of-fit test there is sufficient evidence to suggest that this model is still not accurate, and we should consider adding additional regressors to our model.

1. Error has mean 0

2. Error has constant variance

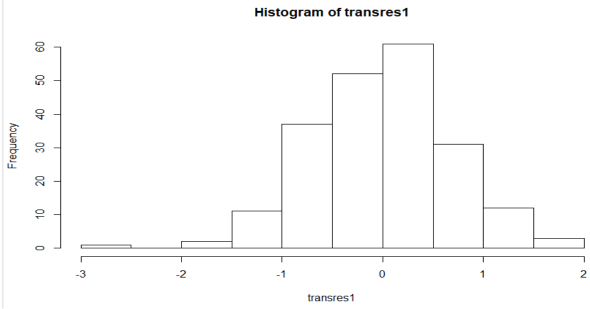


Figure: Histogram of residuals

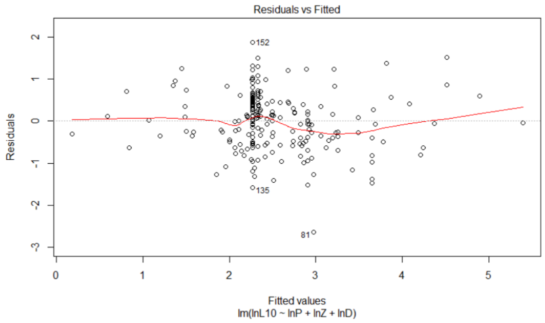


Figure: Fitted values versus residuals

3. Errors are normally distributed

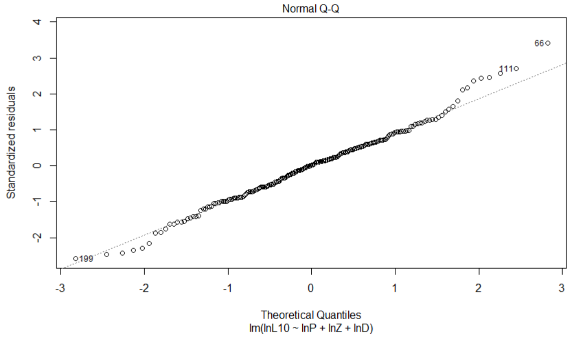
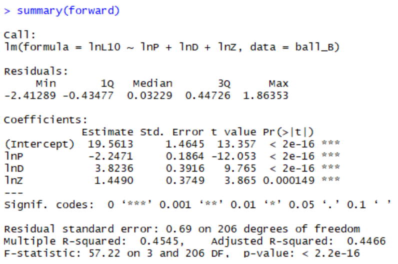


Figure: Normal Q-Q Plot

We ran three variable selection algorithms, forward, backward, and stepwise, to determine if all of the regressors were necessary. The output of those algorithms are below. In both cases all three variables were selected. Our next step will be to add in additional possible regressors.

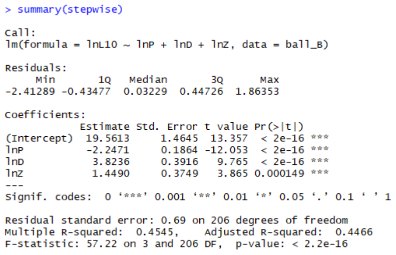
Variable selection:

a) Forward selection:



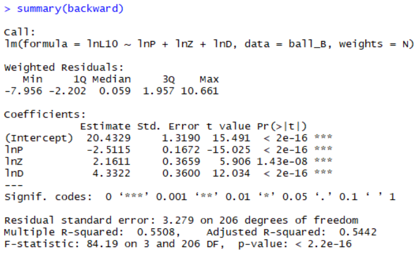
```
> summary(forward)

Call:
lm(formula = lnL10 ~ lnP + lnD + lnZ, data = ball_B)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.41289 -0.43477  0.03229  0.44726  1.86353 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  19.5613     1.4645  13.357  < 2e-16 ***
lnP          -2.2471     0.1864 -12.053  < 2e-16 ***
lnD           3.8236     0.3916   9.765  < 2e-16 ***
lnZ           1.4490     0.3749   3.865 0.000149 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.69 on 206 degrees of freedom
Multiple R-squared:  0.4545,	Adjusted R-squared:  0.4466 
F-statistic: 57.22 on 3 and 206 DF,  p-value: < 2.2e-16
```

b) Stepwise Selection

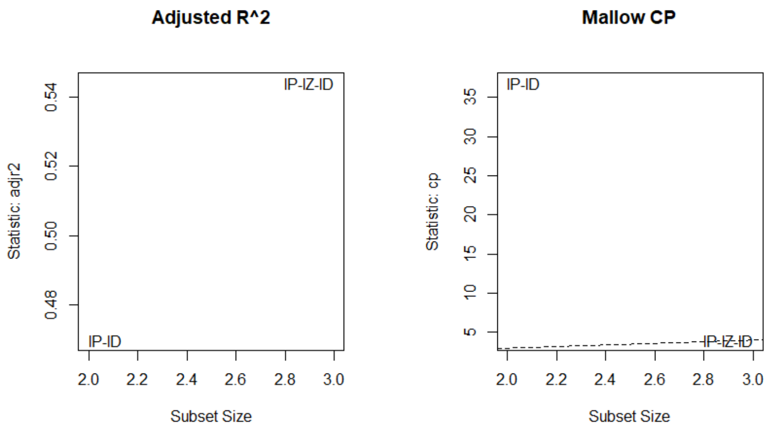


```
> summary(stepwise)

Call:
lm(formula = lnL10 ~ lnP + lnD + lnZ, data = ball_B)

Residuals:
     Min       1Q   Median       3Q      Max 
-2.41289 -0.43477  0.03229  0.44726  1.86353 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  19.5613     1.4645  13.357  < 2e-16 ***
lnP          -2.2471     0.1864 -12.053  < 2e-16 ***
lnD           3.8236     0.3916   9.765  < 2e-16 ***
lnZ           1.4490     0.3749   3.865 0.000149 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.69 on 206 degrees of freedom
Multiple R-squared:  0.4545,	Adjusted R-squared:  0.4466 
F-statistic: 57.22 on 3 and 206 DF,  p-value: < 2.2e-16
```

c)Backward Selection



```
> summary(backward)

Call:
lm(formula = lnL10 ~ lnP + lnZ + lnD, data = ball_B, weights = N)

Weighted Residuals:
   Min     1Q Median     3Q    Max 
-7.956 -2.202  0.059  1.957 10.661 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  20.4329     1.3190  15.491  < 2e-16 ***
lnP          -2.5115     0.1672 -15.025  < 2e-16 ***
lnZ           2.1611     0.3659   5.906 1.43e-08 ***
lnD           4.3322     0.3600  12.034  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.279 on 206 degrees of freedom
Multiple R-squared:  0.5508,	Adjusted R-squared:  0.5442 
F-statistic: 84.19 on 3 and 206 DF,  p-value: < 2.2e-16
```



From the adjusted R^2 plot, we can see the highest R^2 comes from the model with ln$P$+ln$D$+ln$Z$. This gives the best subset of the regressors from the three variable model.

Model Equation (iii) :

$$\begin{aligned} lnL_{10} = \beta_0 + \beta_1 ln(Z) \ + \beta_2 ln(D) + \beta_3 ln(P) \ + \beta_4 B \ + \beta_5 C \\ + \beta_6 Bln(Z) \ + \beta_7 Bln(D) \ + \beta_8 Bln(P) \\ + \beta_9 Cln(Z) \ + \beta_{10} Cln(D) \ + \beta_{11} Cln(P) \end{aligned}$$

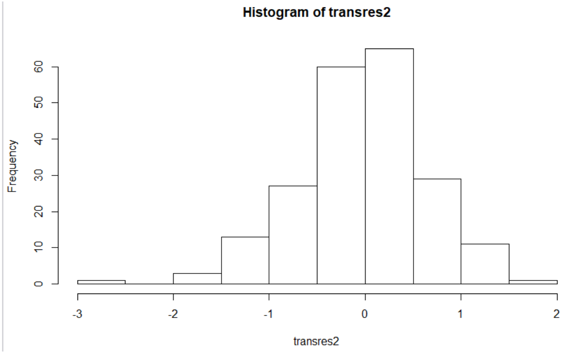
The above Model Equation (iii) adds two indicator variables and six interaction variables. The indicator variables, *B* and *C*, represent companies B and C, respectively with company A being the base case when both *B* and *C* are equal to zero. It is possible that the lifespan of the tested bearings could vary greatly from one company to another, and the contributions of each of the regressors could be different from one company to another. These additions to our model should help us identify any correlations between the companies and the original regressors (ln$Z$, ln$D$, and ln$P$).

The parameters and the values of R^2 and adjusted R^2, once again given by R, are as follows.

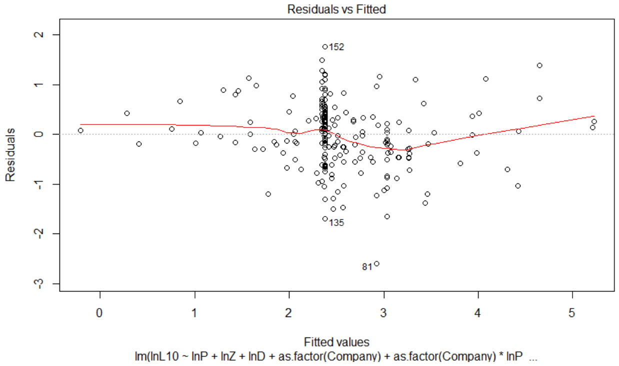
$$\begin{aligned} ln(L_{10}) \ = \ 26.8505 \ + \ 1.1577 ln(Z) \ + \ 5.1646 ln(D) \ - \ 2.9908 ln(P) \ - \ 3.1904 B \ + \ 1.7921 C \\ + \ 0.6804 Bln(Z) \ - \ 0.1973 Bln(D) \ + \ 0.2331 Bln(P) \\ - \ 1.1759 Cln(Z) \ - \ 0.8365 Cln(D) \ - \ 0.1430 Cln(P) \\ R^2 \ = \ 0.6057, \ \ R^2_{Adj} \ = \ 0.5838 \end{aligned}$$

Again, we have a nice improvement to our model as the adjusted R^2 goes from 54.42% up to 58.38%. However, the test for lack-of-fit in R gives a p-value of $p = 0.001768$ with $F_0 = \frac{1195.44}{749.08} = 1.5959$on 93 over 105 degrees of freedom. This result is better, but there is still some work to be done.

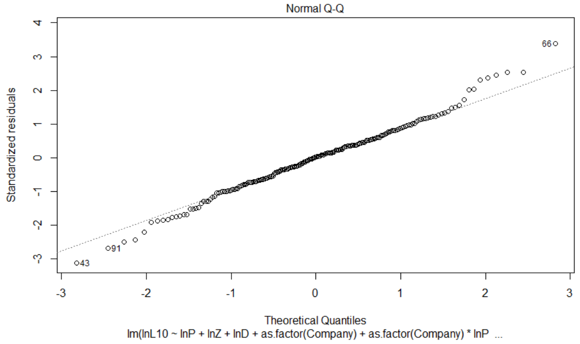
1. Error has mean 0



2. Error has constant variance

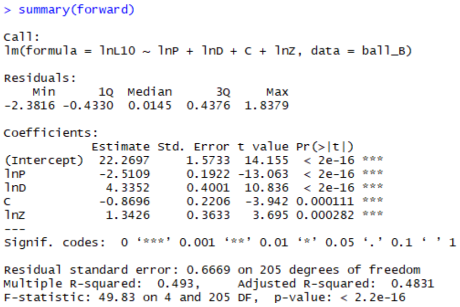


3. Errors are normally distributed



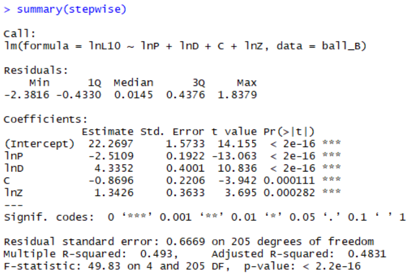
After doing variable selection procedure, forward and stepwise method gives same model.

a) Forward Selection



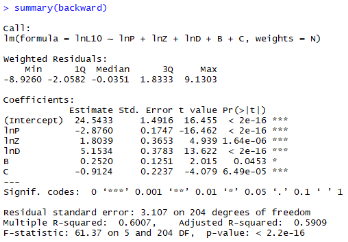
```
> summary(forward)

Call:
lm(formula = lnL10 ~ lnP + lnD + C + lnZ, data = ball_B)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.3816 -0.4330  0.0145  0.4376  1.8379 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  22.2697     1.5733  14.155  < 2e-16 ***
lnP          -2.5109     0.1922 -13.063  < 2e-16 ***
lnD           4.3352     0.4001  10.836  < 2e-16 ***
C            -0.8696     0.2206  -3.942 0.000111 ***
lnZ           1.3426     0.3633   3.695 0.000282 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6669 on 205 degrees of freedom
Multiple R-squared:  0.493,	Adjusted R-squared:  0.4831 
F-statistic: 49.83 on 4 and 205 DF,  p-value: < 2.2e-16
```

b) Stepwise Selection

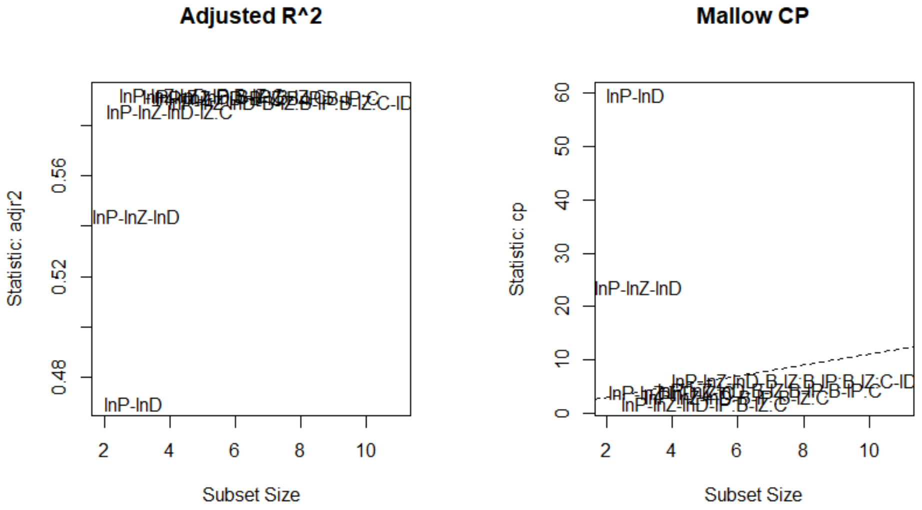


```
> summary(stepwise)

Call:
lm(formula = lnL10 ~ lnP + lnD + C + lnZ, data = ball_B)

Residuals:
    Min      1Q  Median      3Q     Max 
-2.3816 -0.4330  0.0145  0.4376  1.8379 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  22.2697     1.5733  14.155  < 2e-16 ***
lnP          -2.5109     0.1922 -13.063  < 2e-16 ***
lnD           4.3352     0.4001  10.836  < 2e-16 ***
C            -0.8696     0.2206  -3.942 0.000111 ***
lnZ           1.3426     0.3633   3.695 0.000282 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6669 on 205 degrees of freedom
Multiple R-squared:  0.493,	Adjusted R-squared:  0.4831 
F-statistic: 49.83 on 4 and 205 DF,  p-value: < 2.2e-16
```

c) Backward Selection



```
> summary(backward)

Call:
lm(formula = lnL10 ~ lnP + lnZ + lnD + B + C, weights = N)

Weighted Residuals:
    Min      1Q  Median      3Q     Max 
-8.9260 -2.0582 -0.0351  1.8333  9.1303 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  24.5433     1.4916  16.455  < 2e-16 ***
lnP          -2.8760     0.1747 -16.462  < 2e-16 ***
lnZ           1.8039     0.3653   4.939 1.64e-06 ***
lnD           5.1534     0.3783  13.622  < 2e-16 ***
B             0.2520     0.1251   2.015   0.0453 *  
C            -0.9124     0.2237  -4.079 6.49e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.107 on 204 degrees of freedom
Multiple R-squared:  0.6007,	Adjusted R-squared:  0.5909 
F-statistic: 61.37 on 5 and 204 DF,  p-value: < 2.2e-16
```



The adjusted R^2 and Mallow CP graph which is not so clear. Fitting best model using "regsubset" method we got the best model with five variables with below equation.

$$ln(L_{10}) = 24.69388 + 1.79801ln(Z) + 5.14191ln(D) - 2.89425ln(P) + 0.03236Bln(P) - 0.46229Cln(Z)$$

$$R^2 = 0.6021, \ R^2_{Adj} = 0.5923$$

V. Conclusion and Discussion

With only about 60% of the variance in the data explained by our model, it is fair to say that there is more to be done with this analysis. We did not consider the different types of bearings used by company B for example, nor did we consider the year each test took place. Both of these things should be considered for additional study.

It should be noted that the results of the Caroni article suggests that *p* is the same for all companies, but the results of our analysis shows that *p* is slightly smaller for company B. They did find that not all parameters would be the same for all companies, so it is not surprising to see that the term *C*ln(*Z*) was included in our final model.

VI. References

[1] Caroni, Chrys. “Modeling the Reliability of Ball Bearings.” *Journal of Statistics Education*, Volume 10, Number 3, 2002.

[2] Lieblein, J. and Zelen, M. “Statistical Investigation of the fatigue Life of Deep-Groove Ball Bearings.” *Journal of Research of the National Bureau of Standards*, Volume 57, Number 5, November 1956, Research Paper 2719, 273-316.

[3] Montgomery, Douglas C. et al. *Introduction to Linear Regression Analysis*. 5th ed., Wiley, 2012.