UNIT 11 HW

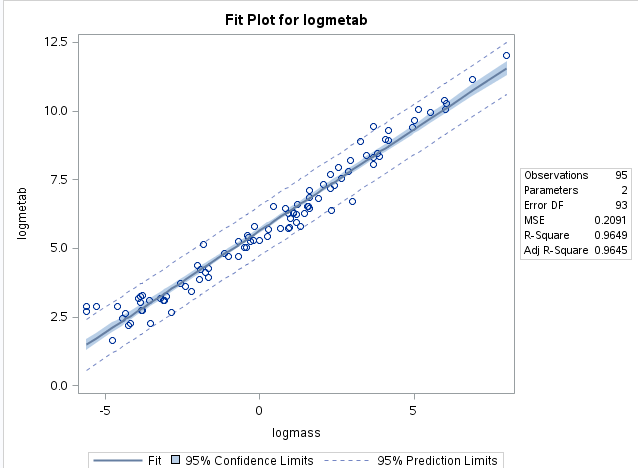
1. Problem 26 Chapter 8

The Metabolic data on Canvas has the average mass, metabolic rate and average lifespan of 95 different species of mammals. Kleiber’s law states that the metabolic rate of an animal species, on average, is proportional to its mass raised to the power of ¾. Judge the adequacy of this theory with these data.

Be sure and provide:

* + 1. A Scatterplot with confidence intervals of the regression line and prediction intervals of the regression line. In SAS and R!

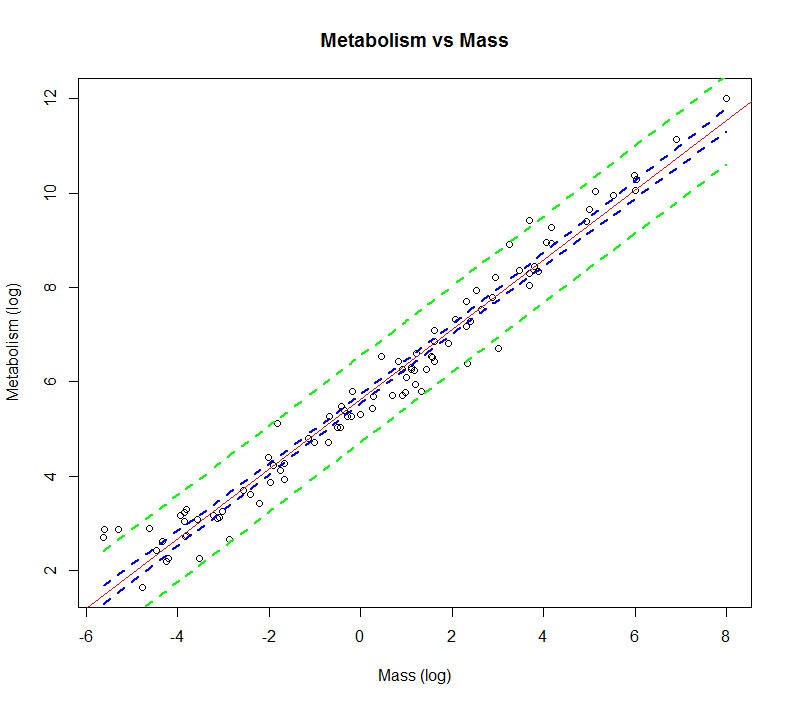
**SAS Output**



**SAS Code Used**

|  |
| --- |
| data work.meta\_trans;  set work.meta;  logmass = log(Mass);  logmetab = log(Metab);  run;  title "Regression of Log Metabolism vs Log Mass";  proc reg data=work.meta\_trans;  model logmetab = logmass / cli clm r;  run; |

**R Output**

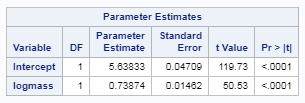


**R Code Used**

|  |
| --- |
| dfmeta["logmetab"] = log(dfmeta$Metab)  dfmeta["logmass"] = log(dfmeta$Mass)  meta\_fit = lm(logmetab~logmass, data=dfmeta)  plot(dfmeta$logmass, dfmeta$logmetab, xlab = "Mass (log)", ylab="Metabolism (log)",  main = "Metabolism vs Mass")  abline(meta\_fit, col="red")  newx = dfmeta$logmass  newx = sort(newx)  pred\_conf = predict(meta\_fit, newdata = data.frame(logmass = newx), interval = c("confidence"), type = c("response"), level = 0.95)  pred\_p = predict(meta\_fit, newdata = data.frame(logmass = newx), interval = c("predict"), type = c("response"), level = 0.95)  lines(newx, pred\_conf[,2], col="blue", lty =2, lwd =2)  lines(newx, pred\_conf[,3], col="blue", lty = 2, lwd = 2)  lines(newx, pred\_p[,2], col="green", lty =2, lwd =2)  lines(newx, pred\_p[,3], col="green", lty = 2, lwd = 2) |

* + 1. A table showing the t-statistics and pvalues for the significance of the regression parameters: . In SAS and R!

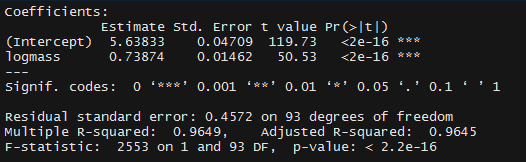
**SAS Output**



**SAS Code Used**

|  |
| --- |
| title "Regression of Log Metabolism vs Log Mass";  proc reg data=work.meta\_trans;  model logmetab = logmass / cli clm r;  run; |

**R Output**



|  |
| --- |
| summary(meta\_fit) |

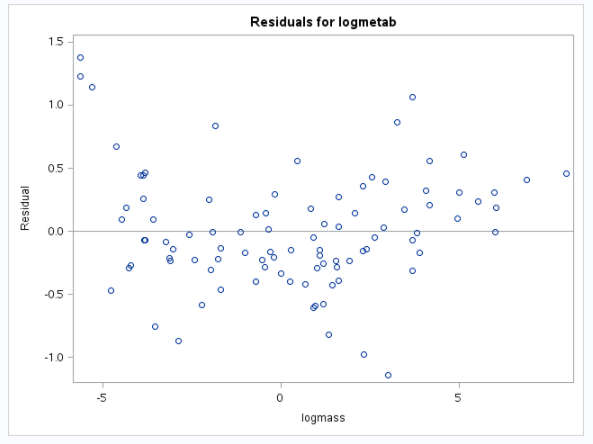
* + 1. The regression equation.

Log(Metabolism) = -5.63833 + 0.73874\*log(mass)

* + 1. Interpretation of the model given any transformation you may use.

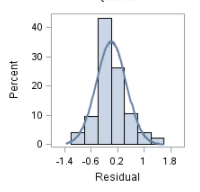
Since both of the variables have been log transformed the interpretation of this regression is a bit difficult. For each doubling of X (mass) there is a 2 ^ 0.7387 change in the median value of Y (metabolic rate). In other words, for each doubling of mass there is a 1.69 change in the value of metabolic rate. Not quite ¾ but getting close.

* + 1. A scatterplot of residuals. In SAS and R!



|  |
| --- |
| #resdiuals of model  metaresid = resid(meta\_fit)  plot(metaresid, main = "Residual Plot of Metabolic Regression Model", xlab = "logmass", ylab = "residuals")  abline(0,0, col="red", lwd=2) |

* + 1. A histogram of residuals with normal distribution superimposed. (from SAS).



* + 1. Provide a measure of the amount of variation in the response that is accounted for by the explanatory variable. Interpret this measure clearly indicating the units of the response and the explanatory variables.

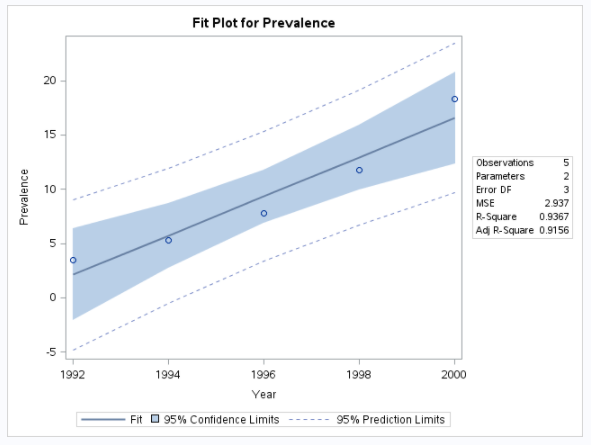
The amount of variation in the response variable that is explained by the explanatory variable will be the R2 of the regression model. The R2 of regression model is 0.9649 indicating that 96.49 percent of the variation in metabolic rate is explained by the value of animal mass.

1. Problem 29 Chapter 8

The Autism data on Canvas item shows the prevalence of autism per 10,000 ten-year old children in the United States in each of five years. Analyze the data to describe the change in the distribution of autism prevalence per year in this time period.

* + 1. A Scatterplot with confidence intervals of the regression line and prediction intervals of the regression line. In SAS or R!

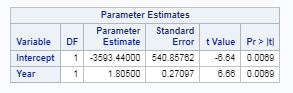
**SAS Output**



**SAS Code Use**

|  |
| --- |
| title "Autism Rate vs Year";  proc reg data=work.autism;  model prevalence = year / cli clm r;  run; |

* + 1. A table showing the t-statistics and pvalues for the significance of the regression parameters: . In SAS or R!



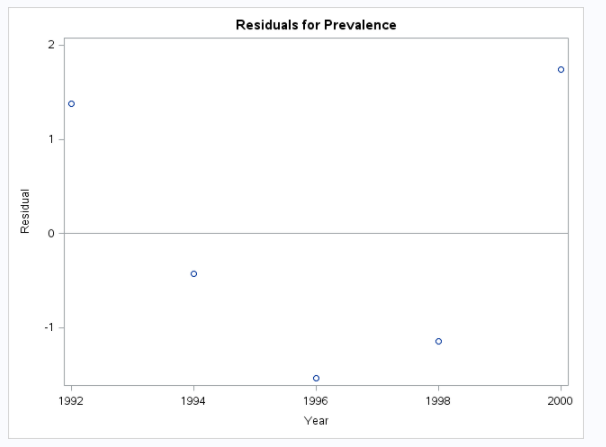
* + 1. The regression equation.

Autism Prevalence = -3593.4 + 1.805\*Year

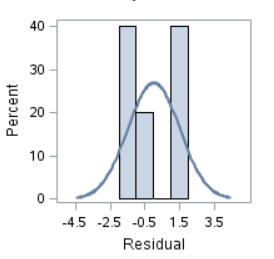
* + 1. Interpretation of the model given any transformation you may use.

Based on analysis of the data and residuals of the data I see no reason to transform the data. Thus, the model can be interpreted as each increase in Year (a five year period) there is a 1.8 increase in the prevalence of autism per 10,000 children.

* + 1. A scatterplot of residuals. In SAS or R!



* + 1. A histogram of residuals with normal distribution superimposed. (from SAS).



* + 1. Provide a measure of the amount of variation in the response that is accounted for by the explanatory variable. Interpret this measure clearly indicating the units of the response and the explanatory variables.

The amount of variation in the response variable that is explained by the explanatory variable will be the R2 of the regression model. The R2 of regression model is 0.9367 indicating that 93.67 percent of the variation in Autism Rate is explained by the value of Year.