

STA108 Project I

Jasper Dong

2023-05-01

Analysis of Variance Table

Response: length

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
infection	1	72.227	72.227	47.141	0.0000000004233 ***
Residuals	109	167.004	1.532		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Analysis of Variance Table

Response: length

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
facility	1	32.373	32.373	17.058	0.00007128 ***
Residuals	109	206.859	1.898		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Analysis of Variance Table

Response: length

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Xray	1	33.881	33.881	17.984	0.00004687 ***
Residuals	109	205.351	1.884		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

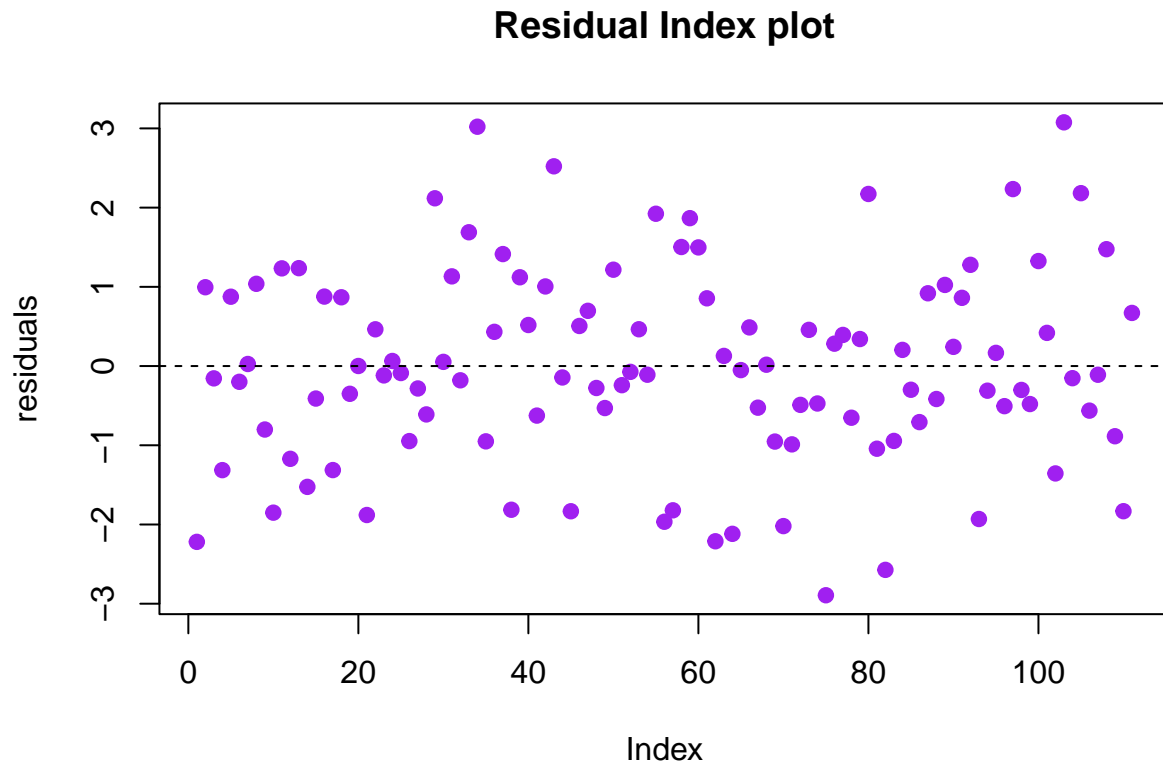
Analysis of Variance Table

Model 1: length ~ 1

Model 2: length ~ infection

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	110	239.23				
2	109	167.00	1	72.227	47.141	0.0000000004233 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1



	2.5 %	97.5 %
(Intercept)	6.0537164	7.6447306
infection	0.4337345	0.7857618

Shapiro-Wilk normality test

data: ei
W = 0.9908, p-value = 0.662

Fligner-Killeen test of homogeneity of variances

data: SENIC\$ei1 and SENIC\$Group
Fligner-Killeen:med chi-squared = 0.0067169, df = 1, p-value = 0.9347

R Appendix

```
knitr::opts_chunk$set(echo = FALSE, comment = NA)
options(scipen = 999) #Remove the scientific notation
library(readr)
library(MASS)
```

```

library(ggplot2)
SENIC <- read_csv("SENIC.csv")
# removing outliers
outliers = c(47, 112)
SENIC = SENIC[-outliers,]
##### MODELS #####
infect.model = lm(length ~ infection, data = SENIC)
facility.model = lm(length ~ facility, data = SENIC)
xray.model = lm(length ~ Xray, data = SENIC)
SENIC$ei1 = infect.model$residuals
SENIC$yhat1 = infect.model$fitted.values
# ANOVA tables for each model
anova(infect.model)
anova(facility.model)
anova(xray.model)
# Assessing Infection Linearity
options(scipen = 8)
reduced.model = lm(length ~ 1, data = SENIC)
anova.table = anova(reduced.model, infect.model)
anova.table
# Assessing Infection Independence
plot(infect.model$residuals, main = "Residual Index plot", xlab = "Index", ylab = "residuals", pch = 19)
abline(h = 0, lty = 2)
# Confidence Interval for Parameters
alpha = 0.05
infect.CIs = confint(infect.model, level = 1-alpha)
infect.CIs
# Shapiro-Wilks Test for Infection Normality
ei = infect.model$residuals
infect.SWtest = shapiro.test(ei)
infect.SWtest
# Fligner-Killeen test for Infection Homoscedasticity
SENIC$ei1 = infect.model$residuals
Group = rep("Lower", nrow(SENIC))
Group[SENIC$length < median(SENIC$length)] = "Upper"
Group = as.factor(Group)
SENIC$Group = Group
infect.FKtest = fligner.test(SENIC$ei1, SENIC$Group)
infect.FKtest

```