P3

November 6, 2021

[1]:
```
#3
```

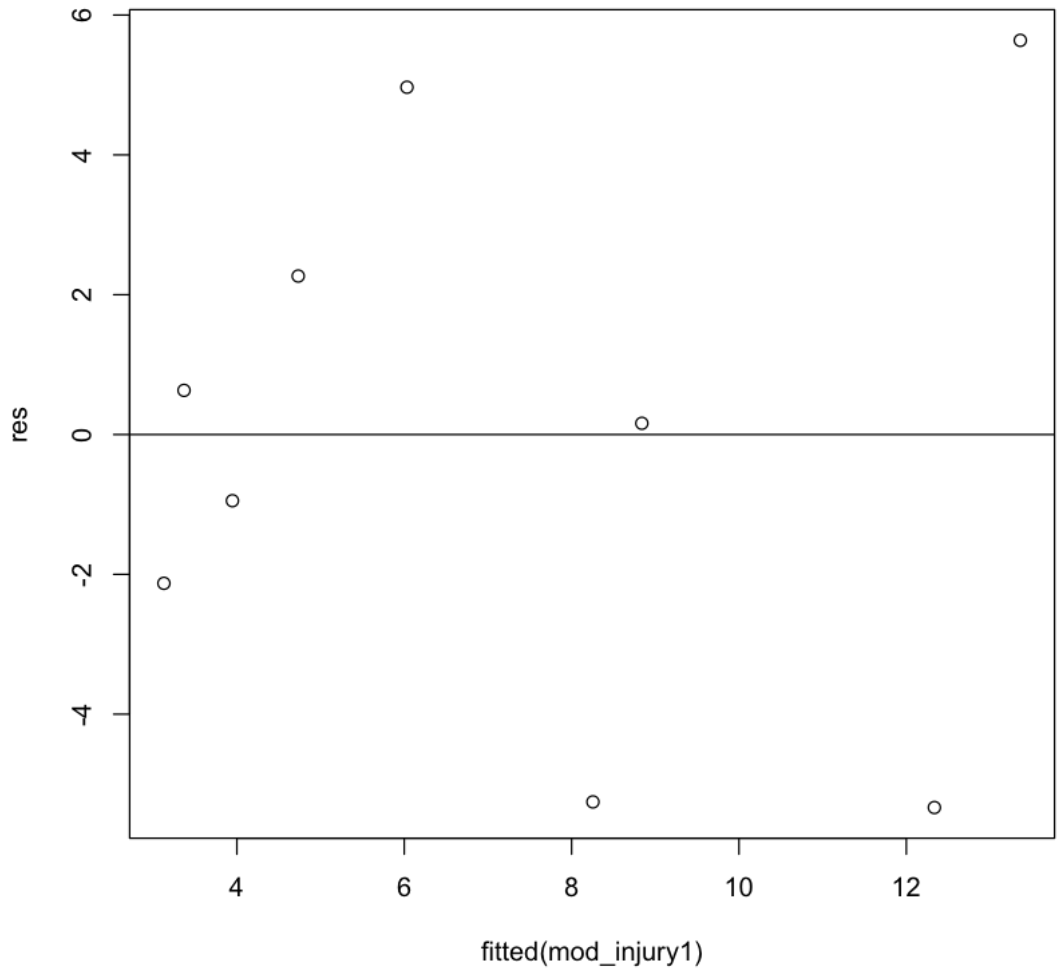
(a)

[4]:
```
data_injury <- read.table("Table6.6.txt",
                          head = TRUE,
                          sep = "\t")
#data_injury
```

[6]:
```
mod_injury1 <- lm(Y ~ N, data_injury)
print(summary(mod_injury1))
```

```
Call:
lm(formula = Y ~ N, data = data_injury)

Residuals:
    Min      1Q  Median      3Q     Max
-5.3351 -2.1281  0.1605  2.2670  5.6382

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.1402     3.1412  -0.045   0.9657
N            64.9755    25.1959   2.579   0.0365 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.201 on 7 degrees of freedom
Multiple R-squared:  0.4872,    Adjusted R-squared:  0.4139
F-statistic:  6.65 on 1 and 7 DF,  p-value: 0.03654
```

[7]:
```
res <- resid(mod_injury1)
plot(fitted(mod_injury1), res)
abline(0,0)
```



AS shown above the error is not heteroscedastic, and the "E" assumption is violated.

(b)

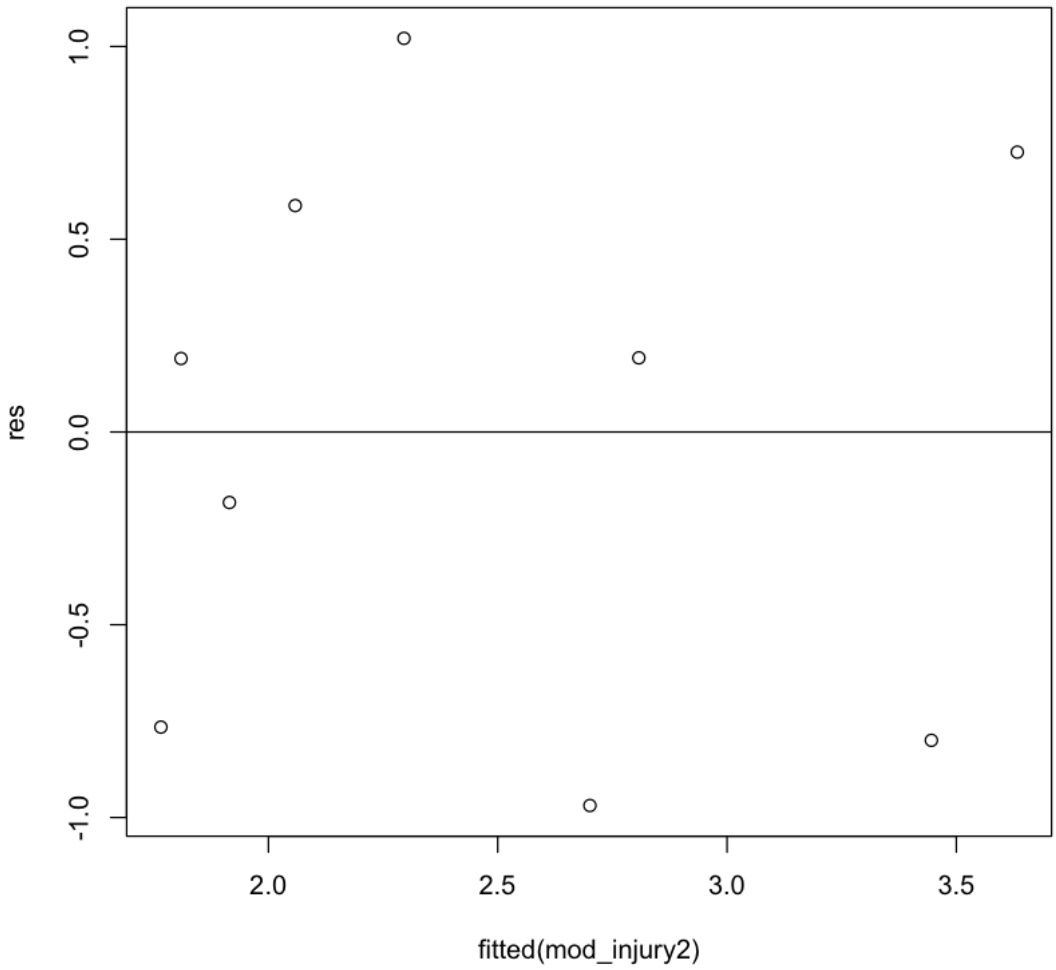
[9]:
```
mod_injury2 <- lm(sqrt(Y) ~ N, data_injury)
print(summary(mod_injury2))
```

```
Call:
lm(formula = sqrt(Y) ~ N, data = data_injury)

Residuals:
    Min      1Q  Median      3Q     Max
-0.9690 -0.7655  0.1906  0.5874  1.0211
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.1692     0.5783   2.022   0.0829 .
N            11.8564     4.6382   2.556   0.0378 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7733 on 7 degrees of freedom
Multiple R-squared:  0.4828,    Adjusted R-squared:  0.4089
F-statistic: 6.535 on 1 and 7 DF,  p-value: 0.03776
```

Regression line equation: $\sqrt{Y} = \beta_0 + \beta_1 N + \varepsilon$ That is $\sqrt{Y} = 1.17 + 11.86N + \varepsilon$

[10]:
```
res <- resid(mod_injury2)
plot(fitted(mod_injury2), res)
abline(0,0)
```



Now there is no heteroscedasticity.