

Midterm Q5

Bhagyarathi Raman

3/9/2022

```
library(AER)
```

```
## Loading required package: car
```

```
## Loading required package: carData
```

```
## Loading required package: lmtest
```

```
## Loading required package: zoo
```

```
##
```

```
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      as.Date, as.Date.numeric
```

```
## Loading required package: sandwich
```

```
## Loading required package: survival
```

```
data("Affairs")
```

5. Goodness-of-fit

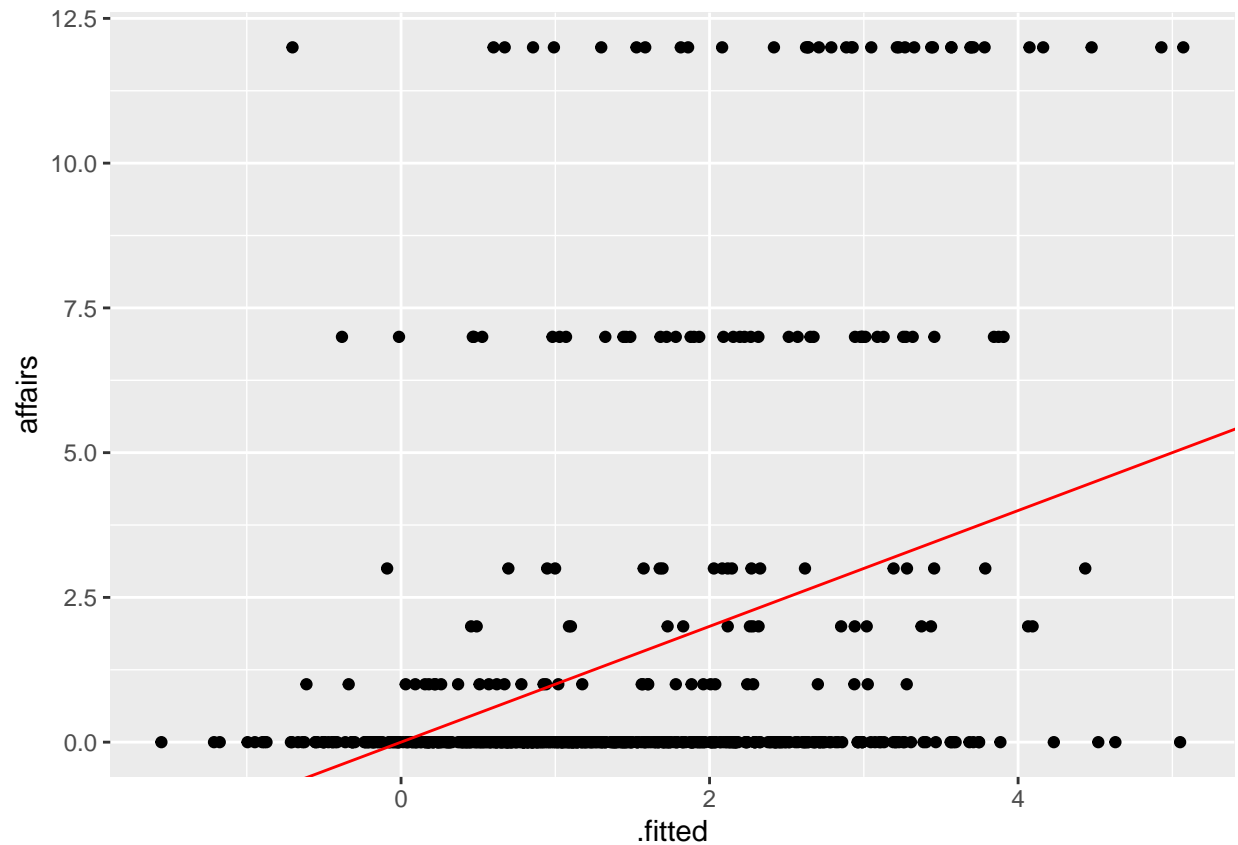
- Compute the fit plot for the g-model, does the model do a good or bad job of predicting affairs? Explain.

```
library(ggplot2)
```

```
g=lm (affairs~gender+age+yearsmarried+children+education+religiousness+occupation+rating, data = Affairs)
```

```
mod.g<-fortify(g)
```

```
qplot(.fitted, affairs, data = mod.g) + geom_abline(intercept = 0, slope = 1,  
  color = "red")
```



The fit plot for the G-model does not do a good job predicting affairs, as the majority of the data l

- b. What is the Pearson correlation between the g-model-fitted-values and actual values, is it good or bad? Explain.

```
cor(mod.g$.fitted, mod.g$affairs)
```

```
## [1] 0.3629574
```

- c. What is the R2 for the g-model? What does R2 mean for the fitted model? Is it good or bad? Explain.

```
summary(g)$r.sq
```

```
## [1] 0.1317381
```

About 0.1317381 Indicates that 13.17% of the outcome-variable-deviation, the rest of the outcome-vari

- d. Theoretically, what is the relation between the R2 and the Pearson correlation (actual-values vs fitted-values)? Does the relation hold for the computed values of the g-model? If not, why not?

```
c(cor(g$fitted,Affairs$affairs)^2, summary(g)$r.sq)
```

```
## [1] 0.1317381 0.1317381
```

```
#Theoretically,  $r^2$  is equal to  $R^2$ . Yes, it holds for this case.
```