STAT 021 – HW 5

Hunter Luber

10/28/2019

Question 1

Sketch (by hand) residual plots (with $\hat{y}_i$, predicted response values, on the horizontal axis) that show each of the following: (5 points) (1) Constant variance and linearity (2) non-constant variance and linearity (3) constant variance and non-linearity (4) non-constant variance and non-linearity

SEE ATTACHED HARD COPY FOR MY SOLUTIONS TO QUESTIONS 1.1-1.4

Question 2

Suppose we have two random variables X and Y. What are the differences among the following assumptions regarding X and Y: uncorrelated, independent, have the same variance, and have the same distribution? (5 points)

Uncorrelated and independent, although somewhat similar, are not the same. The definition of being independent is that the probability of X and Y occurring together is the same as their individual probabilities when separated. In other words, the outcome of X would have no influence on the outcome of Y. Uncorrelated is a different term. For two random variables to be uncorrelated, it means that their covariance is equal to zero.

Having the same variance and having the same distribution are also similar, although not the same as one another. Two random variables may very well have the same variance. However, they could have the same variance and still have very different distributions. One of the easiest ways to tell if this is the case is to look at the modality of the resulting graph. Data on both of the random variables could have the same variance, but one may have many more peaks than the other, and actually appear quite different when presented as a plot.

Question 3

Read the Wikipedia page for Simpson's Paradox: https://en.wikipedia.org/wiki/Simpson%27s_paradox. Then, import the "Stand your ground" data set uploaded on Moodle. This data (from 2015) is related to the Stand Your Ground law in Florida. Each observational unit consists of a case where the Stand Your Ground law was a part of the defense strategy, the defendent's race (white or non-white), the victim's race (white or non-white), and a binary variable indicating whether or not the defendent was convicted. With this categorical data we are not going to fit a regression model, but we are going to examine this data and look out for Simpson's paradox. (10 points)

```
## Parsed with column specification:
## cols(
##   Convicted = col_character(),
##   Accused = col_character(),
##   WhiteVictim = col_double(),
##   MinVictim = col_double()
## )

## Observations: 220
## Variables: 4
```

```
## $ Convicted   <chr> "Yes", "Yes", "Yes", "Yes", "Yes", "Yes", "Yes", "...
## $ Accused     <chr> "White", "White", "White", "White", "White", "Whit...
## $ WhiteVictim <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,...
## $ MinVictim   <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,...

## [1] "Convicted"   "Accused"     "WhiteVictim" "MinVictim"
```

a.) Create and print the following tables to summarize the data: 1. Defendant's race vs. convicted for all observational units;

```
##
##            No Yes
##   Minority 60  29
##   White    86  45
```

2. Defendant's race vs. convicted for cases with minority victims only;

```
## , ,  = 0
##
##
##            No Yes
##   Minority 15  10
##   White    67  40
##
## , ,  = 1
##
##
##            No Yes
##   Minority 45  19
##   White    19   5
```

3. Defendant's race vs. convicted for cases with white victims only;

```
## , ,  = 0
##
##
##            No Yes
##   Minority 45  19
##   White    19   5
##
## , ,  = 1
##
##
##            No Yes
##   Minority 15  10
##   White    67  40
```

4. The table created by adding Tables 2 and 3 together.

```
##
##            No Yes
##   Minority 60  29
##   White    86  45

## , ,  = 0
##
##
##            No Yes
##   Minority 60  29
```

```
##    White     86  45
##
## , ,  = 1
##
##
##            No Yes
##    Minority 60  29
##    White     86  45
```

Table 1 and Table 2 + Table 3 are the same output. They simply display slightly differently here, due to the way that I have coded them.

b.) What are the overall conviction rates for minority and white defendants, respectively? What are the csyonviction rates for minority and white defendants among the cases with white victims? What are the conviction rates for minority and white defendants among the cases with minority victims?

```
##   [1] "White"    "White"    "White"    "White"    "White"    "White"
##   [7] "White"    "White"    "White"    "White"    "White"    "White"
##  [13] "White"    "White"    "White"    "White"    "White"    "White"
##  [19] "White"    "White"    "White"    "White"    "White"    "White"
##  [25] "White"    "White"    "White"    "White"    "White"    "White"
##  [31] "White"    "White"    "White"    "White"    "White"    "White"
##  [37] "White"    "White"    "White"    "White"    "White"    "White"
##  [43] "White"    "White"    "White"    "White"    "White"    "White"
##  [49] "White"    "White"    "White"    "White"    "White"    "White"
##  [55] "White"    "White"    "White"    "White"    "White"    "White"
##  [61] "White"    "White"    "White"    "White"    "White"    "White"
##  [67] "White"    "White"    "White"    "White"    "White"    "White"
##  [73] "White"    "White"    "White"    "White"    "White"    "White"
##  [79] "White"    "White"    "White"    "White"    "White"    "White"
##  [85] "White"    "White"    "White"    "White"    "White"    "White"
##  [91] "White"    "White"    "White"    "White"    "White"    "White"
##  [97] "White"    "White"    "White"    "White"    "White"    "White"
## [103] "White"    "White"    "White"    "White"    "White"    "White"
## [109] "White"    "White"    "White"    "White"    "White"    "White"
## [115] "White"    "White"    "White"    "White"    "Minority" "Minority"
## [121] "Minority" "Minority" "Minority" "Minority" "Minority" "Minority"
## [127] "Minority" "Minority" "Minority" "Minority" "Minority" "Minority"
## [133] "Minority" "Minority" "Minority" "Minority" "Minority" "Minority"
## [139] "Minority" "Minority" "Minority" "Minority" "Minority" "Minority"
## [145] "Minority" "Minority" "Minority" "Minority" "Minority" "Minority"
## [151] "Minority" "Minority" "Minority" "Minority" "Minority" "Minority"
## [157] "Minority" "Minority" "Minority" "Minority" "Minority" "Minority"
## [163] "Minority" "Minority" "Minority" "Minority" "Minority" "Minority"
## [169] "Minority" "Minority" "Minority" "Minority" "Minority" "Minority"
## [175] "Minority" "Minority" "Minority" "Minority" "Minority" "Minority"
## [181] "Minority" "Minority" "Minority" "Minority" "Minority" "Minority"
## [187] "Minority" "Minority" "Minority" "Minority" "Minority" "Minority"
## [193] "Minority" "Minority" "Minority" "Minority" "Minority" "White"
## [199] "White"    "White"    "White"    "White"    "White"    "White"
## [205] "White"    "White"    "White"    "White"    "White"    "White"
## [211] "Minority" "Minority" "Minority" "Minority" "Minority" "Minority"
## [217] "Minority" "Minority" "Minority" "Minority"

## # A tibble: 74 x 4
##    Convicted Accused WhiteVictim MinVictim
```

```
##      <chr>      <chr>          <dbl>        <dbl>
##   1 Yes        White              1            0
##   2 Yes        White              1            0
##   3 Yes        White              1            0
##   4 Yes        White              1            0
##   5 Yes        White              1            0
##   6 Yes        White              1            0
##   7 Yes        White              1            0
##   8 Yes        White              1            0
##   9 Yes        White              1            0
##  10 Yes        White              1            0
## # ... with 64 more rows

## # A tibble: 131 x 4
##      Convicted Accused WhiteVictim MinVictim
##      <chr>      <chr>          <dbl>        <dbl>
##   1 Yes        White              1            0
##   2 Yes        White              1            0
##   3 Yes        White              1            0
##   4 Yes        White              1            0
##   5 Yes        White              1            0
##   6 Yes        White              1            0
##   7 Yes        White              1            0
##   8 Yes        White              1            0
##   9 Yes        White              1            0
##  10 Yes        White              1            0
## # ... with 121 more rows

## # A tibble: 89 x 4
##      Convicted Accused   WhiteVictim MinVictim
##      <chr>      <chr>          <dbl>        <dbl>
##   1 Yes        Minority           1            0
##   2 Yes        Minority           1            0
##   3 Yes        Minority           1            0
##   4 Yes        Minority           1            0
##   5 Yes        Minority           1            0
##   6 Yes        Minority           1            0
##   7 Yes        Minority           1            0
##   8 Yes        Minority           1            0
##   9 No         Minority           1            0
##  10 No         Minority           1            0
## # ... with 79 more rows

## # A tibble: 45 x 4
##      Convicted Accused WhiteVictim MinVictim
##      <chr>      <chr>          <dbl>        <dbl>
##   1 Yes        White              1            0
##   2 Yes        White              1            0
##   3 Yes        White              1            0
##   4 Yes        White              1            0
##   5 Yes        White              1            0
##   6 Yes        White              1            0
##   7 Yes        White              1            0
##   8 Yes        White              1            0
##   9 Yes        White              1            0
```

```
## 10 Yes        White              1           0
## # ... with 35 more rows

## # A tibble: 29 x 4
##    Convicted Accused  WhiteVictim MinVictim
##    <chr>     <chr>          <dbl>     <dbl>
##  1 Yes       Minority           1         0
##  2 Yes       Minority           1         0
##  3 Yes       Minority           1         0
##  4 Yes       Minority           1         0
##  5 Yes       Minority           1         0
##  6 Yes       Minority           1         0
##  7 Yes       Minority           1         0
##  8 Yes       Minority           1         0
##  9 Yes       Minority           0         1
## 10 Yes       Minority           0         1
## # ... with 19 more rows
```

Out of the 220 total defendants accused, 74 were convicted. There were 131 white defendants and 89 minority defendants. 45 white defendants were convicted, and 29 minority defendants were convicted. Therefore, the overall conviction rate for white defendants is $45/131 = 34.35\%$, and the overall conviction rate for minority defendants is $29/89 = 32.58\%$. Together, this would show that the overall conviction rate for white defendants is slightly higher, by about 2%.

```
## # A tibble: 132 x 4
##    Convicted Accused WhiteVictim MinVictim
##    <chr>     <chr>         <dbl>     <dbl>
##  1 Yes       White             1         0
##  2 Yes       White             1         0
##  3 Yes       White             1         0
##  4 Yes       White             1         0
##  5 Yes       White             1         0
##  6 Yes       White             1         0
##  7 Yes       White             1         0
##  8 Yes       White             1         0
##  9 Yes       White             1         0
## 10 Yes       White             1         0
## # ... with 122 more rows

## # A tibble: 107 x 4
##    Convicted Accused WhiteVictim MinVictim
##    <chr>     <chr>         <dbl>     <dbl>
##  1 Yes       White             1         0
##  2 Yes       White             1         0
##  3 Yes       White             1         0
##  4 Yes       White             1         0
##  5 Yes       White             1         0
##  6 Yes       White             1         0
##  7 Yes       White             1         0
##  8 Yes       White             1         0
##  9 Yes       White             1         0
## 10 Yes       White             1         0
## # ... with 97 more rows

## # A tibble: 25 x 4
##    Convicted Accused  WhiteVictim MinVictim
```

```
##     <chr>     <chr>          <dbl>     <dbl>
##  1 Yes       Minority           1         0
##  2 Yes       Minority           1         0
##  3 Yes       Minority           1         0
##  4 Yes       Minority           1         0
##  5 Yes       Minority           1         0
##  6 Yes       Minority           1         0
##  7 Yes       Minority           1         0
##  8 Yes       Minority           1         0
##  9 No        Minority           1         0
## 10 No        Minority           1         0
## # ... with 15 more rows

## # A tibble: 40 x 4
##     Convicted Accused WhiteVictim MinVictim
##     <chr>     <chr>         <dbl>     <dbl>
##  1 Yes       White             1         0
##  2 Yes       White             1         0
##  3 Yes       White             1         0
##  4 Yes       White             1         0
##  5 Yes       White             1         0
##  6 Yes       White             1         0
##  7 Yes       White             1         0
##  8 Yes       White             1         0
##  9 Yes       White             1         0
## 10 Yes       White             1         0
## # ... with 30 more rows

## # A tibble: 10 x 4
##     Convicted Accused  WhiteVictim MinVictim
##     <chr>     <chr>          <dbl>     <dbl>
##  1 Yes       Minority           1         0
##  2 Yes       Minority           1         0
##  3 Yes       Minority           1         0
##  4 Yes       Minority           1         0
##  5 Yes       Minority           1         0
##  6 Yes       Minority           1         0
##  7 Yes       Minority           1         0
##  8 Yes       Minority           1         0
##  9 Yes       Minority           1         0
## 10 Yes       Minority           1         0
```

There were 132 cases with a white victim. When filtered to solely cases that had white victims, the conviction rates for white defendants were $40/107 = 37.38\%$. The conviction rates for minority defendants were $10/25 = 40\%$. This would show that the conviction rates for minority defendants is higher than for white defendants by almost 3%, when white victims were involved.

```
## # A tibble: 88 x 4
##     Convicted Accused WhiteVictim MinVictim
##     <chr>     <chr>         <dbl>     <dbl>
##  1 Yes       White             0         1
##  2 Yes       White             0         1
##  3 Yes       White             0         1
##  4 Yes       White             0         1
##  5 Yes       White             0         1
##  6 No        White             0         1
```

```
##  7 No        White           0         1
##  8 No        White           0         1
##  9 No        White           0         1
## 10 No        White           0         1
## # ... with 78 more rows

## # A tibble: 24 x 4
##    Convicted Accused WhiteVictim MinVictim
##    <chr>     <chr>         <dbl>     <dbl>
##  1 Yes       White           0         1
##  2 Yes       White           0         1
##  3 Yes       White           0         1
##  4 Yes       White           0         1
##  5 Yes       White           0         1
##  6 No        White           0         1
##  7 No        White           0         1
##  8 No        White           0         1
##  9 No        White           0         1
## 10 No        White           0         1
## # ... with 14 more rows

## # A tibble: 64 x 4
##    Convicted Accused  WhiteVictim MinVictim
##    <chr>     <chr>          <dbl>     <dbl>
##  1 Yes       Minority         0         1
##  2 Yes       Minority         0         1
##  3 Yes       Minority         0         1
##  4 Yes       Minority         0         1
##  5 Yes       Minority         0         1
##  6 Yes       Minority         0         1
##  7 Yes       Minority         0         1
##  8 Yes       Minority         0         1
##  9 Yes       Minority         0         1
## 10 Yes       Minority         0         1
## # ... with 54 more rows

## # A tibble: 5 x 4
##   Convicted Accused WhiteVictim MinVictim
##   <chr>     <chr>         <dbl>     <dbl>
## 1 Yes       White           0         1
## 2 Yes       White           0         1
## 3 Yes       White           0         1
## 4 Yes       White           0         1
## 5 Yes       White           0         1

## # A tibble: 10 x 4
##    Convicted Accused  WhiteVictim MinVictim
##    <chr>     <chr>          <dbl>     <dbl>
##  1 Yes       Minority         1         0
##  2 Yes       Minority         1         0
##  3 Yes       Minority         1         0
##  4 Yes       Minority         1         0
##  5 Yes       Minority         1         0
##  6 Yes       Minority         1         0
##  7 Yes       Minority         1         0
##  8 Yes       Minority         1         0
```

```
##  9 Yes        Minority           1         0
## 10 Yes        Minority           1         0
```

There were 88 total cases that involved a minority victim. When filtered to solely cases that had minority victims, there were 24 cases that involved a white defendant, and 64 cases that involved a minority defendant. The conviction rate for white defendants against minority victims was $5/24 = 20.83\%$. The conviction rate for minority defendants against minority victims was $10/64 = 15.625\%$. Therefore, with respect to solely minority victims, white defendants seem to be almsot 5% more likely to be convicted than do minority defendants.

c.) Explain what is going on here in terms of Simpson's paradox and interpret what this means with respect to racial bias in the criminal justice system.

There is a Simpson's paradox that is evident here. Overall, it would appear that white defendants have about a 2% greater conviction rate than do minority defendants, out of the cases studied. When white victims are considered, convictions for minority defendants are about 3% higher than for white defendants. However, when minority victims are considered, convictions for white defendants are about 5% larger than are conviction rates for minority defendants.

This presents an interesting dilemma to parse out. However, there were a large number of cases with white victim (132) and a much smaller number of cases with a minority victim (88). This discrepancy could have some influence on the data. Interestingly, there were 131 cases with a white defendant, and 89 cases with a minority defendant. These numbers seem to match up eerily similarly. These numbers would lead us to believe that there could be a confounding variable present that is not readily visible in the study.

Overall, this would lead me to believe that there might be a racial bias towards white victims getting their cases to trial. Often, the common conception is that there is more minority crime than there is white crime. However, this study seems to debunk that theory, at least within the data set provided on instances of the Stand Your Ground Law. For each group, it appears that minority defendants are more likely to be convicted against white victims, and white defendants are more likely to be convicted against minority victims. This is an interesting dilemma. The numbers do match up, however, with the difference between those two groups being 2%, which explains the overall combined conviction rate for white defendants being slightly larger than the conviction rate for minority defendants.