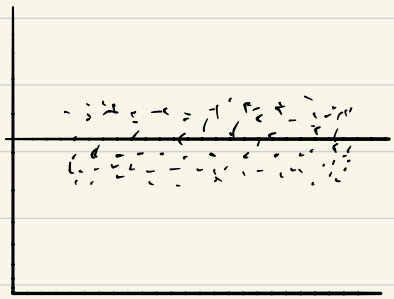
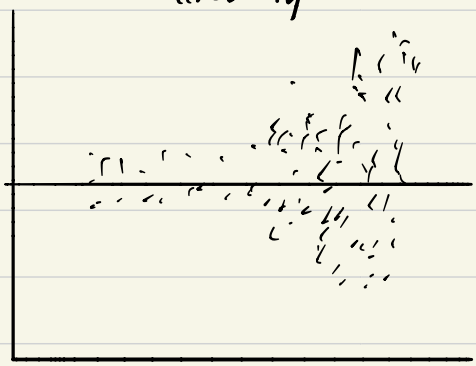
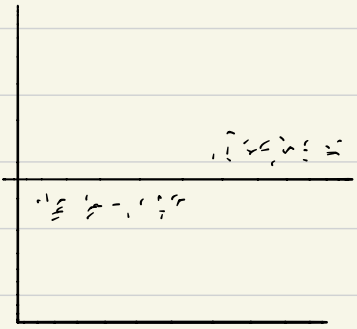
i) Constant variance and linearity



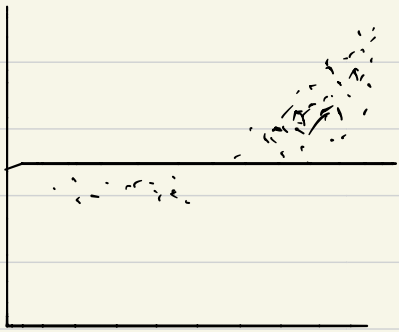
ii) non-constant variance and linearity



iii) constant variance and non-linearity



iv) non-constant variance and non-linearity



Stat 021 Homework 5

Suzanne Thornton

Due: Friday, Nov. 1, 12:00pm

Q2) Suppose we have two random variables X and Y. What are the differences among the following assumptions regarding X and Y:

- X and Y are uncorrelated

When two random variables are uncorrelated, it means that there is no linear relationship between these two.

- X and Y are independent

Two random variables are called "dependent" if the probability of events regarding one variable influence the probability of the other variable. Thus, if they are independent, one variable doesn't influence the other's distribution of probability. Mathematically, if random variables X and Y are independent, Pr(A and B) = P(A)P(B).

- X and Y have the same variance **When two random variables have the same variance, it means that their random errors are spread out in the same manner.**
- X and Y have the same distribution? **If two random variables have the same distribution, it means they can be characterized with the same type of distribution, which includes Poisson, all beta, all t-distributed, and etc.**

Q3) a) Create and print the following tables to summarize the data:

```
1. Defendant's race vs convicted for all observational units;

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

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

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

```
stand_you_ground <- read.csv("~/stand_your_ground.csv")
table1 <- stand_you_ground %>%
  select(Convicted,Accused) %>%
  filter(Convicted == 'Yes')

head(table1)
```

```
##   Convicted Accused
## 1       Yes   White
## 2       Yes   White
## 3       Yes   White
## 4       Yes   White
## 5       Yes   White
## 6       Yes   White
```

```
table2 <- stand_you_ground %>%
  filter(MinVictim == 1,Convicted=='Yes') %>%
  select(Convicted,Accused)
```

```
head(table2)
```

```
##   Convicted  Accused
## 1       Yes    White
## 2       Yes    White
## 3       Yes    White
## 4       Yes    White
## 5       Yes    White
## 6       Yes Minority
```

```
table3<- stand_you_ground %>%
  filter(WhiteVictim == 1,Convicted == 'Yes') %>%
  select(Convicted,Accused)


head(table3)
```

```
##   Convicted Accused
## 1       Yes   White
## 2       Yes   White
## 3       Yes   White
## 4       Yes   White
## 5       Yes   White
## 6       Yes   White
```

```
table4 <- rbind(table2,table3)

head(table4)
```

```
##   Convicted  Accused
## 1       Yes    White
## 2       Yes    White
## 3       Yes    White
## 4       Yes    White
## 5       Yes    White
## 6       Yes Minority
```

b) What are the overall conviction rates for minority and white defendants, respectively? What are the conviction rates for minority and white defendants among the cases with minority victims? What are the conviction rates for minority and white defendants among the cases with white victims? **1. Conviction rate for white defendants: 45/(86+45)=.34// Conviction rate for minority defendants: 29/(60+29)=.32//**

2. Conviction rate for white defendants among cases with minority victims: 5/(19+5)=.20// Conviction rate for minority defendants: 19/(19+45)=.30//

3. Conviction rate for white defendants among cases with white victims: 40/(40+67)=.37 // Conviction rate for minority defendants: 10/(10+15)=.4//

```
stand_you_ground %>%
  filter(Accused == 'White') %>%
  count(Convicted)
```

```
## # A tibble: 2 x 2
##   Convicted     n
##   <fct>     <int>
## 1 No           86
```

```
## 2 Yes          45
```

```
stand_you_ground %>%
  filter(Accused == 'Minority') %>%
  count(Convicted)
```

```
## # A tibble: 2 x 2
##   Convicted     n
##   <fct>     <int>
## 1 No           60
## 2 Yes          29
```

```
stand_you_ground %>%
  filter(Accused == 'White',MinVictim == '1') %>%
  count(Convicted)
```

```
## # A tibble: 2 x 2
##   Convicted     n
##   <fct>     <int>
## 1 No           19
## 2 Yes           5
```

```
stand_you_ground %>%
  filter(Accused == 'Minority', MinVictim == '1') %>%
  count(Convicted)
```

```
## # A tibble: 2 x 2
##   Convicted     n
##   <fct>     <int>
## 1 No           45
## 2 Yes          19
```

```
stand_you_ground %>%
  filter(Accused == 'White',WhiteVictim == '1') %>%
  count(Convicted)
```

```
## # A tibble: 2 x 2
##   Convicted     n
##   <fct>     <int>
## 1 No           67
## 2 Yes          40
```

```
stand_you_ground %>%
  filter(Accused == 'Minority', WhiteVictim == '1') %>%
  count(Convicted)
```

```
## # A tibble: 2 x 2
##   Convicted     n
##   <fct>     <int>
## 1 No           15
## 2 Yes          10
```

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. **Overall, the conviction rates are rather similar; it is even slightly higher for white people. Thus, we would think that the criminal justice system is biased against white people. However, if we look into the data a little more closely, we can see that minority defendants are convited at a higher rate in each cases of white victims and minority victims, letting us know tha the criminal justice system is a little more biased against minorities**