Chapter 3 Exercise 5

Allen Church

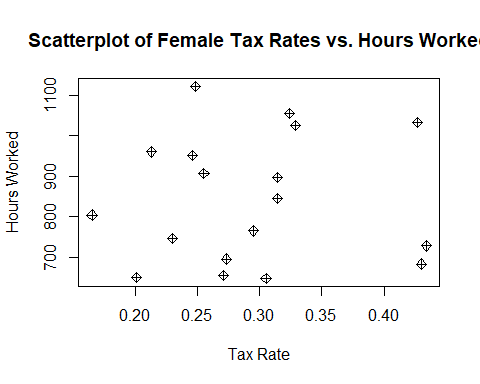
September 24, 2019

#Load datasets and view  
load("C:/Users/WB537822/Desktop/stats1/Ch3\_Exercise4\_Divorce\_rates\_Women.RData")  
load("C:/Users/WB537822/Desktop/stats1/Ch3\_Exercise4\_Divorce\_rates\_Men.RData")

5a. For each data set (women and men), create scatterplot of hours worked on y axis and tax rates on x axis.

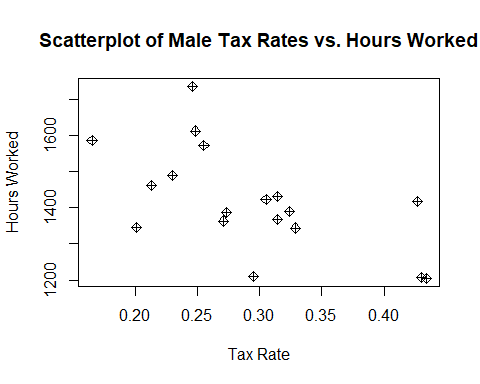
Scatterplot for women dataset

plot(Wdata$taxrate, Wdata$hours, main="Scatterplot of Female Tax Rates vs. Hours Worked",  
 xlab="Tax Rate", ylab = "Hours Worked", pch=9)



Scatterplot for men dataset

plot(Mdata$taxrate, Mdata$hours, main="Scatterplot of Male Tax Rates vs. Hours Worked",  
 xlab="Tax Rate", ylab = "Hours Worked", pch=9)



5b. For each dataset, set estimate an OLS regression in which hours worked is regressed on tax rates. Report the estimated regression equation and interpret coefficients. Explain any differences in coefficients.

Regression for women dataset. The formula for this regression is the dependent variable = hours worked and the independent variable = tax rate.

ols\_tax\_female <- lm(Wdata$hours ~ Wdata$taxrate)  
summary(ols\_tax\_female)

##   
## Call:  
## lm(formula = Wdata$hours ~ Wdata$taxrate)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -195.2 -139.8 -15.0 119.0 281.0   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 827.05 151.97 5.442 5.43e-05 \*\*\*  
## Wdata$taxrate 53.46 502.42 0.106 0.917   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 160.2 on 16 degrees of freedom  
## Multiple R-squared: 0.0007071, Adjusted R-squared: -0.06175   
## F-statistic: 0.01132 on 1 and 16 DF, p-value: 0.9166

Regression for men dataset. The formula for this regression is the dependent variable = hours worked and the independent variable = tax rate.

ols\_tax\_male <- lm(Mdata$hours ~ Mdata$taxrate)  
summary(ols\_tax\_male)

##   
## Call:  
## lm(formula = Mdata$hours ~ Mdata$taxrate)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -207.97 -56.45 -13.98 32.42 264.44   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1748.5 111.7 15.65 4.02e-11 \*\*\*  
## Mdata$taxrate -1122.4 369.2 -3.04 0.0078 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 117.7 on 16 degrees of freedom  
## Multiple R-squared: 0.3661, Adjusted R-squared: 0.3265   
## F-statistic: 9.241 on 1 and 16 DF, p-value: 0.007803

The value for beta hat 1 for women (53.46) estimates that for each additional hour worked, the tax rate increases by 53.46. The value for beta hat 1 for men (-1122.4) estimates that for each additional hour worked, the tax rate decreases by -1122.4. As seen above, there is a difference of 1175.86 between the beta hat 1 for women vs. men. This suggests that as men work more hours, their tax rate significantly decreases. On the other hand, the model suggests that as women work more hours, their tax rate increases.

5c. Fitted value and residual for men in U.S.

us\_men <- subset(Mdata, country == "United States")  
ols\_us\_men <- lm(us\_men$hours ~ us\_men$taxrate)

Fitted value for men in U.S.

ols\_us\_men$fitted.values

## 1   
## 1613

Residual for men in U.S.

summary(ols\_us\_men)

##   
## Call:  
## lm(formula = us\_men$hours ~ us\_men$taxrate)  
##   
## Residuals:  
## ALL 1 residuals are 0: no residual degrees of freedom!  
##   
## Coefficients: (1 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 1613 NA NA NA  
## us\_men$taxrate NA NA NA NA  
##   
## Residual standard error: NaN on 0 degrees of freedom

5d. Fitted value and residual for women in Italy

italy\_women <- subset(Wdata, country == "Italy")  
ols\_italy\_women <- lm(italy\_women$hours ~ italy\_women$taxrate)

Fitted value for women in Italy

ols\_italy\_women$fitted.values

## 1   
## 654.6

Residual for women in Italy

summary(ols\_us\_men)

##   
## Call:  
## lm(formula = us\_men$hours ~ us\_men$taxrate)  
##   
## Residuals:  
## ALL 1 residuals are 0: no residual degrees of freedom!  
##   
## Coefficients: (1 not defined because of singularities)  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 1613 NA NA NA  
## us\_men$taxrate NA NA NA NA  
##   
## Residual standard error: NaN on 0 degrees of freedom