Homework 4

Zachary Lazerick

7 October, 2021

##Problem 3 #Analysis of the Pharmacy Data

a) Read in the Data

```
Pharmacy = read.csv("Pharmacy.csv")
attach(Pharmacy)
```

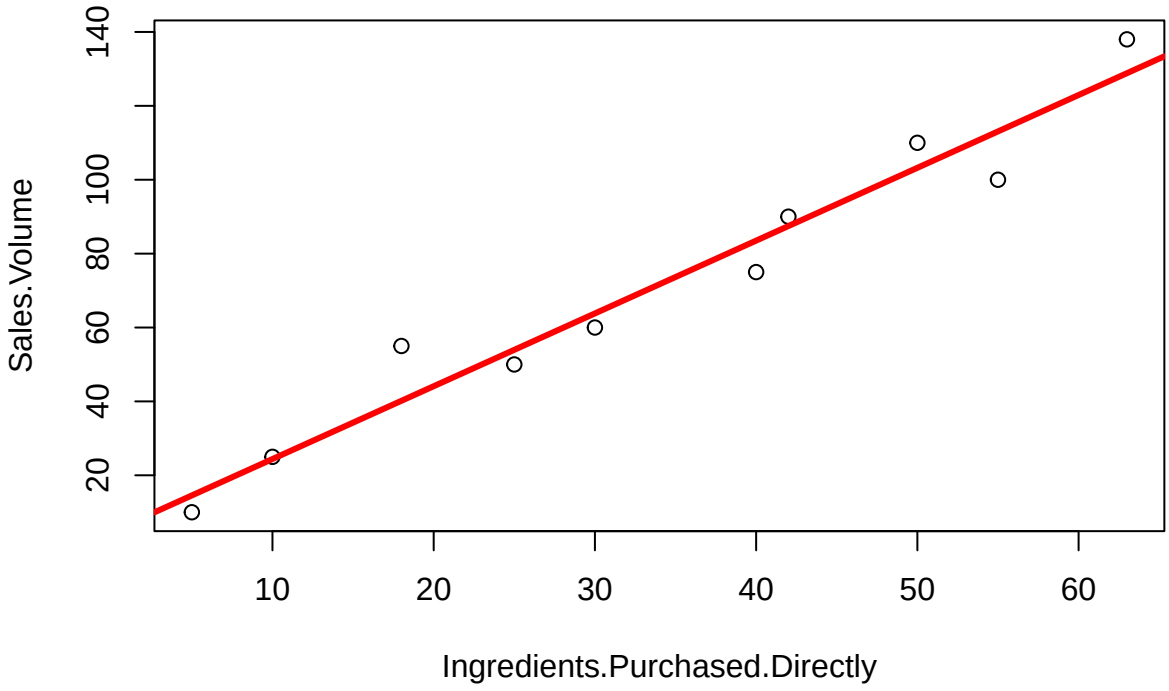
b) Fit a Linear Model with Predictor "Ingredients Purchase Directly" and Response "Sales Volume"

```
Pharmacy.fit <- lm(Sales.Volume~Ingredients.Purchased.Directly, data=Pharmacy)
summary(Pharmacy.fit)
```

```
##
## Call:
## lm(formula = Sales.Volume ~ Ingredients.Purchased.Directly, data = Pharmacy)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -13.074  -4.403  -1.607   5.719  14.834
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     4.6979     5.9520   0.789    0.453
## Ingredients.Purchased.Directly  1.9705     0.1545  12.750 1.35e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.022 on 8 degrees of freedom
## Multiple R-squared:  0.9531, Adjusted R-squared:  0.9472
## F-statistic: 162.6 on 1 and 8 DF,  p-value: 1.349e-06
```

c) Confirm Estimates for Beta_0 and Beta_1 from Problem 1

```
plot(Ingredients.Purchased.Directly, Sales.Volume)
abline(Pharmacy.fit, lwd = 3, col = "red")
```




```
coef(Pharmacy.fit)
```

```
##                   (Intercept) Ingredients.Purchased.Directly
##                      4.697852                       1.970478
```

These reported coefficients match the hand-calculated coefficients in Problem 1.

d) Use Output to Report the Estimate for V(e) = sigma^2

The output reports a Residual Std. Error of 9.022 on 8 Degrees of Freedom. Squaring this result gives the estimate for V(e), so (9.022)^2 = 81.39648.

e) How are the 8 Degrees of Freedom Obtained?

There are 8 Degrees of Freedom because the sample size, n = 10 and the formula for Degrees of Freedom for Simple Linear Regression is d.f. = n - 2, because we estimate Beta_0 and Beta_1. Substituting 10 for n yields 10 - 2 = 8 d.f.

##Problem 4 #Analysis of the Casino/Crime Data

a) Read in the Data

```
Casino = read.csv("Casino.csv")
attach(Casino)
```

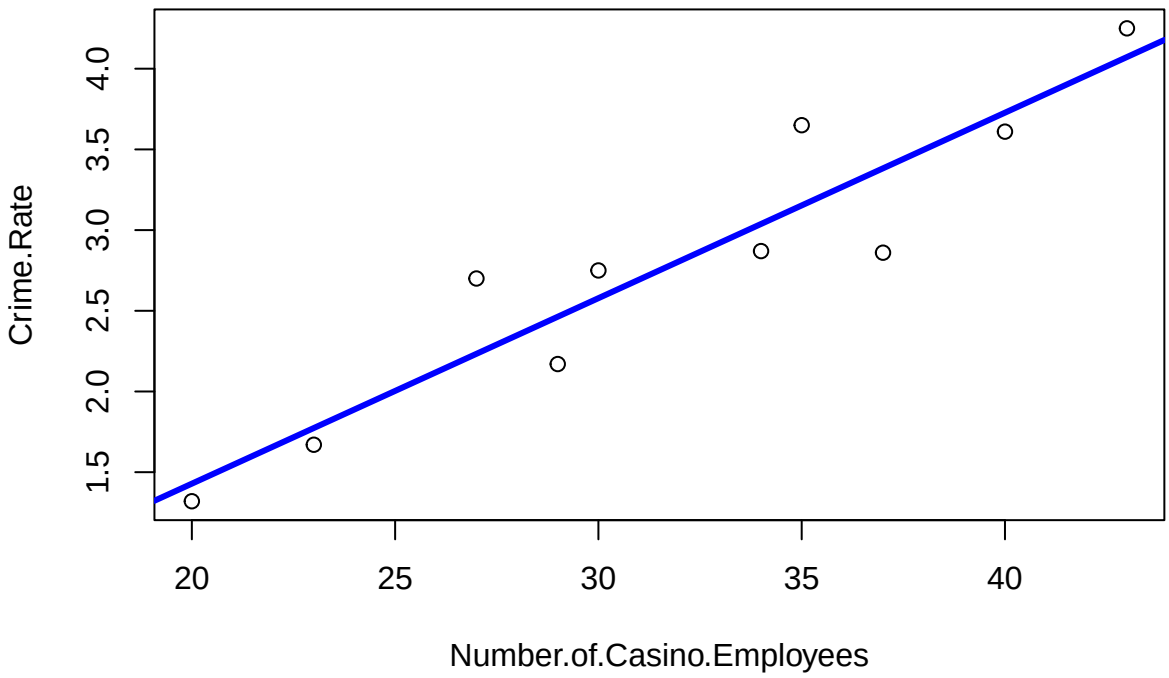
b) Fit a Linear Model with Predictor "Number of Casino Employees" and Response "Crime Rate"

```
Casino.fit <- lm(Crime.Rate~Number.of.Casino.Employees, data=Casino)
summary(Casino.fit)
```

```
##
## Call:
## lm(formula = Crime.Rate ~ Number.of.Casino.Employees, data = Casino)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5226 -0.1552 -0.1062  0.1763  0.4972
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                -0.86977    0.50903  -1.709    0.126
## Number.of.Casino.Employees  0.11493    0.01564   7.350 7.99e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3446 on 8 degrees of freedom
## Multiple R-squared:  0.871,  Adjusted R-squared:  0.8549
## F-statistic: 54.03 on 1 and 8 DF,  p-value: 7.992e-05
```

c) Plot the Data along the Regression Line

```
plot(Number.of.Casino.Employees, Crime.Rate)
abline(Casino.fit, lwd = 3, col="blue")
```



d) Use the Output to Report the Estimate for V(e) = sigma^2

The output reports a Residual Std. Error of 0.3446 on 8 Degrees of Freedom. Squaring this result gives the estimate for V(e) = (0.3446)^2 = 0.1187492.

e) Predict the Crime Rate when there are 25,000 Casino Employees

Given that there are 25,000 Casino Employees, this means that the 'Number.of.Casino.Employees' is 25. So, substituting 25 for X in our regression equation, we get Y = Beta_0 + Beta_1 * X = (-.86977) + (0.11493)(25) = 2.09896. Thus the expected Crime Rate when there are 25,000 Casino Employees is 2.098986 (per 1,000 population).