Homework 3

Jeffrey Chang

2/14/2020

Question 3.2:

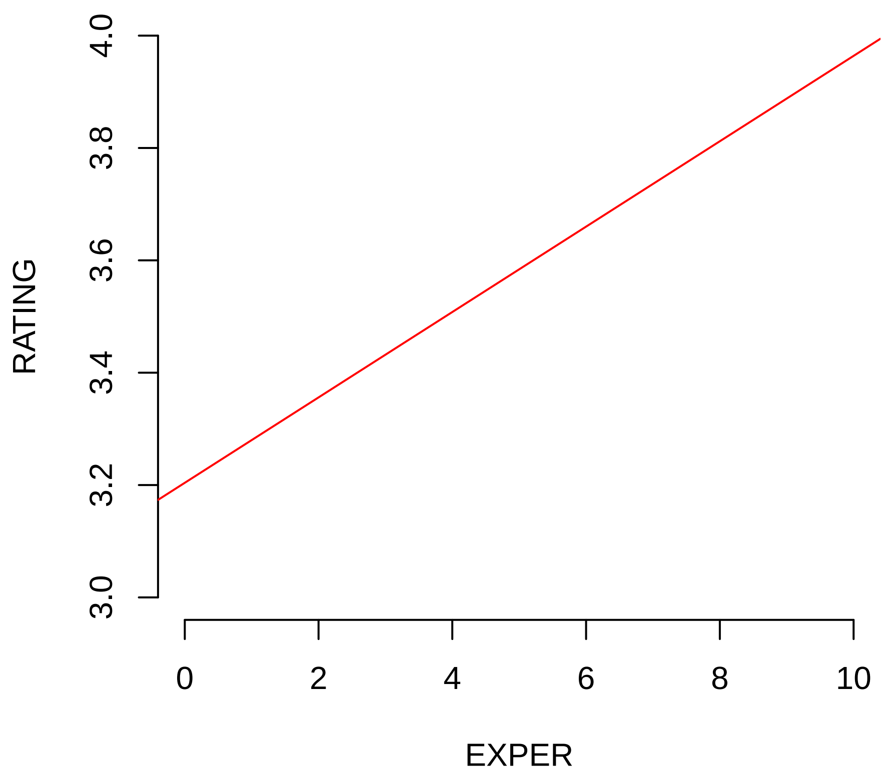
The general manager of an engineering firm wants to know whether a technical artist's experience influences the quality of his or her work. A random sample of 24 artists is selected and their years of work experience and quality rating (as assessed by their supervisors) recorded. Work experience (EXPER) is measured in years and quality rating (RATING) takes a value of 1 through 7, with 7 = excellent and 1 = poor: The simple regression model RATING = b1 + b2(EXPER) + e is proposed. The least squares estimates of the model, and the standard errors of the estimates, are

Rating = 3.204 + 0.076*(EXPER)

(SE) =(0.709) , (0.044)

(a)Sketch the estimated regression function. Interpret the coefficient of EXPER.

If the experience increased 1 year, the Rating of a technical artist would increase 0.076. And the sketch is shown below.



(b)Construct a 95% confidence interval for b2, the slope of the relationship between quality rating and experience. In what are you 95% confident?

```
round((qt(c(0.975), df=22)),3)   # 22 degrees of freedom
```

```
## [1] 2.074
```

To guarantee 95% confident, make t_c = t_(0.975, 22) = 2.074.

b2 ± t_c_se(b2) = 0.076 ± 2.074 * 0.044 => [-0.015, 0.167] = confident Interval

(c) Test the null hypothesis that b2 is zero against the alternative that it is not using a two-tail test and the a = 0.05 level of significance. What do you conclude?

Test the null hypothesis: H_0: b2 = 0 , H_1: b2 != 0,(!= means not equal to)

t = (b2-b_2)/se(b2) = (b2-0)/se(b2) = 0.076/0.044 = 1.727 When (N-2)=22

use two tail test, we have the result of 22 degrees of freedom is 2.074.

Because -2.074 < 1.727 < 2.074, so it is failed to reject the null hypothesis.

(d) Test the null hypothesis that b2 is zero against the one-tail alternative that it is positive at the a = 0.05 level of significance. What do you conclude?

```
# run one-tail, a=0.05, 22 degrees of freedom
round((qt(c(0.95), df=22)),3)
```

```
## [1] 1.717
```

Test the null hypothesis: H_0: b2 = 0 , H_1: b2 > 0,

t = (b2-b_2)/se(b2) = (b2-0)/se(b2) = 0.076/0.044 = 1.727 When (N-2)=22

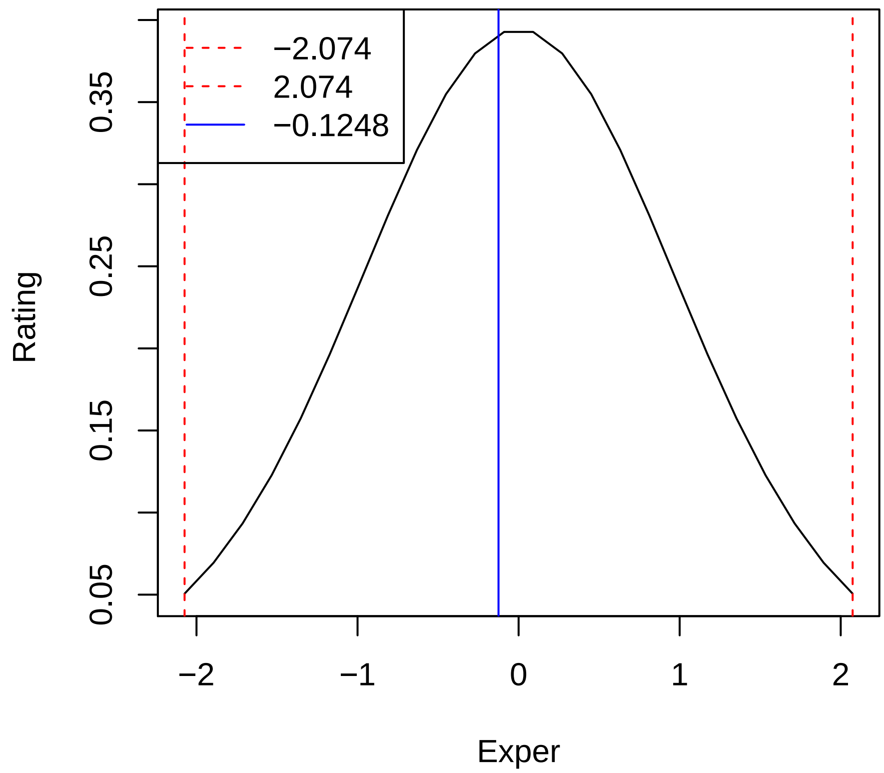
use one-tail test, we have the result of 22 degrees of freedom is 1.717. And, 1.717 < 1.727,

So H_0 was rejected and b2 is positive and have sufficient evidence to conclude that the EXPER would have a= positive influence to RATING.

(e) For the test in part (c), the p-value is 0.0982. If we choose the probability of a Type I error to be a = 0.05, do we reject the null hypothesis, or not, just based on an inspection of the p-value? Show, in a diagram, how this p-value is computed.

(c) is Two-tail test, so p-value is 0.0982 > a = 0.05 we do not reject H_0

```
tcr = 2.074
# t_(1-0.0982,22)= -.1248
t = -0.1248
pts = seq(-2.074,2.074,length=24)
plot(pts,dt(pts,df=22),col='black',type='l', ylab="Rating ", xlab="Exper")
abline(v=c(-tcr, tcr, t), col=c("red", "red", "blue"),lty=c(2,2,1))
legend("topleft", legend=c("-2.074 ", "2.074 ", "-0.1248"), col= c("red", "red", "blue"),lty=c(2,2,1))
```



Question 3.4:

Consider a simple regression in which the dependent variable MIM = mean income of males who are 18 years of age or older, in thousands of dollars. The explanatory variable PMHS = percent of males 18 or older who are high school graduates. The data consist of 51 observations on the 50 states plus the District of Columbia. Thus MIM andPMHS are "state averages" The estimated regression, along with standard errors and t-statistics, is

MIM = (a) + (0.180)*(PMHS)

(se) (2.174) (b)

(t) (1.257) (5.754)

(a) What is the estimated equation intercept? Show your calculation. Sketch the estimated regression function.

```
## [1] "The estimated equation intercept is b1 = t*se(b1) =2.7327"
```

(b) What is the standard error of the estimated slope? Show your calculation.

```
## [1] "the standard error of the estimated slope is se(b2) = b2/t =0.0313"
```

(c) What is the p-value for the two-tail test of the hypothesis that the equation intercept is zero? Draw a sketch to illustrate.

```
## [1] "the p-value for is 2*(1-P(t<1.257) =0.2148"
```

(d) State the economic interpretation of the estimated slope. Is the sign of the coefficient what you would expect from economic theory?

Estimated b2 = 0.18, means that high school graduates man whose age >= 18 has a 1% increase, the average income will increase $180.

(e) Construct a 99% confidence interval estimate of the slope of this relationship.

```
## [1] "t_c = t_(0.995, 49) = 2.68"
```

99% confidence interval b2 ± t_c * se(b2) = 0.18 ± 2.68 * 0.0313 => [0.096, 0.264]

(f) Test the hypothesis that the slope of the relationship is 0.2 against the alternative that it is not. State in words the meaning of the null hypothesis in the context of this problem.

t = (0.18-0.2)/0.0313=-0.639

5% significant level and 49 degrees of freedom => ± t_c = ± 2.01

Do not reject H0 because t = -0.634 is in interval [-2.01, 2.01].

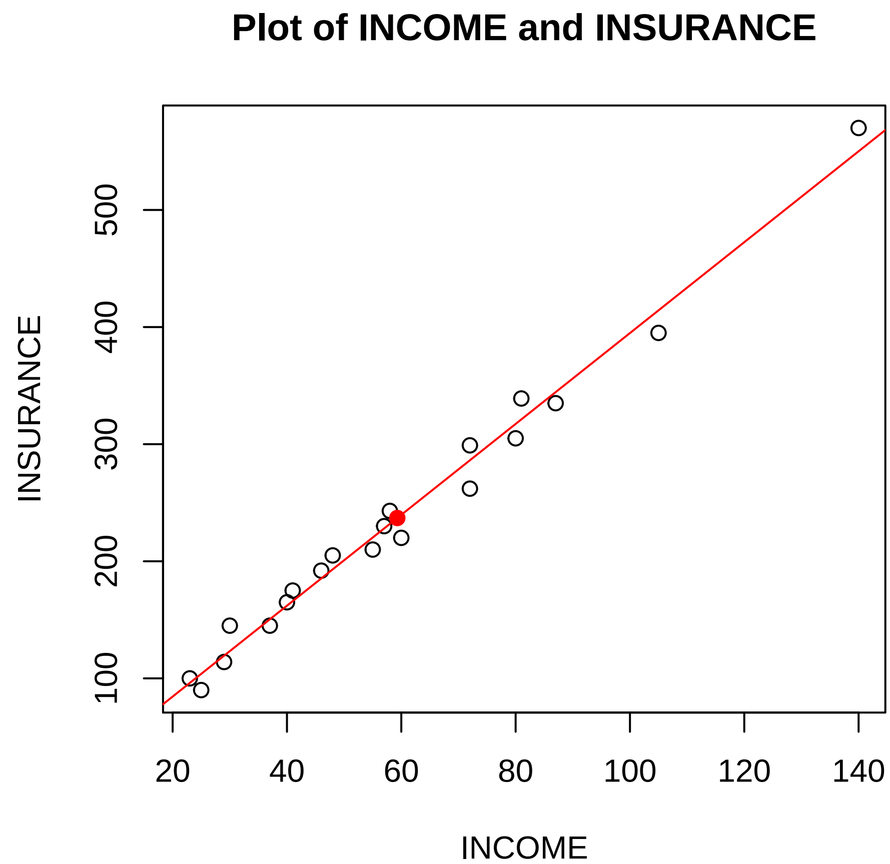
1% increase of man's age >= 18 will have increasing 0.2 * 1000 = $200 dollars.

Question 3.5:

A life insurance company wishes to examine the relationship between the amount of life insurance held by a family and family income. From a random sample of 20 households, the company collected the data in the file insur.dat. The data are in units of thousands of dollars.

(a) Estimate the linear regression with dependent variable INSURANCE and independent variable INCOME. Write down the fitted model and draw a sketch of the fitted function. Identify the estimated slope and intercept on the sketch. Locate the point of the means on the plot.

```
##
## Call:
## lm(formula = insur$insurance ~ insur$income, data = insur)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -24.228 -10.766   2.456  11.295  21.739
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    6.8550     7.3835   0.928    0.365
## insur$income   3.8802     0.1121  34.606   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.36 on 18 degrees of freedom
## Multiple R-squared:  0.9852, Adjusted R-squared:  0.9844
## F-statistic:  1198 on 1 and 18 DF,  p-value: < 2.2e-16
```




```
## [1] "INSURANCE = 6.855 + 3.8802 * INCOME"
```

(b) Discuss the relationship you estimated in (a). In particular.

As the plot show in (a), When INCOME increase will make INSURANCE increase. There are positive relationship

B1 = 6.855 means when INCOME is 0, INSURANCE will be 6.855*1000 = $6855.

(i) What is your estimate of the resulting change in the amount of life insurance when income increases by $1,000?

When INCOME increase 1000, an estimate of INSURANCE would increase 3.8802*1000 = $3880.2.

(ii) What is the standard error of the estimate in (i), and how do you use this standard error for interval estimation and hypothesis testing?

The standard error of b2 is 0.1121, test a hypothesis about b_2 the test statistic is

(b2-b_2)/se(b2), df = 20-2 =18.

An interval estimator of b_2 => [b2-tc* se(b2), b2+tc* se(b2)].

tc is the critical value for t with degrees of freedom = (N-2), a = significance level.

(c) One member of the management board claims that for every \$1,000 increase in income, the amount of life insurance held will go up by \$5,000. Choose an alternative hypothesis and explain your choice. Does your estimated relationship support this claim? Use a 5% significance level.

Set H_0 : b_2 = 5, H_1 : b_2 != 5 (!= means not equal to)

This is a two-tail test, so we can find the rejection area below:

```
## [1] "t_c = t_(0.975, 18) = 2.101"
```

```
## [1] "t_c = t_(0.025, 18) = -2.101"
```

now, caculate the t when b_2 = 5

```
## [1] "t = (b2-b_2)/se(b2) = (3.8802-5)/0.1121 = -9.989"
```

As t = -9.989 < -2.101, we reject the null hypothesis and conclude that the estimated relationship does not support the claim.

(d) Test the hypothesis that as income increases the amount of life insurance increases by the same amount. That is, test the hypothesis that the slope of the relationship is one.

the slope of the relationship is one = b_2 = 1

Set H_0 : b_2 = 1, H_1 : b_2 != 1 (!= means not equal to)

This is a two-tail test, so we can find the rejection area below:

[-2.101, 2.101]

now, caculate the t when b_2 = 1

```
## [1] "t = (b2-b_2)/se(b2) = (3.8802-1)/0.1121 = 25.693"
```

As t = 25.639 > 2.101, we reject the null hypothesis and conclude that the amount of life insurance does not increases by the same amount as income increases.

(e) Write a short report (200–250 words) summarizing your findings about the relationship between income and the amount of life insurance held.

Find the household characteristics that influence the amount of life insurance cover by different households. One important determinant of life insurance cover is INCOME. we set up the model to quantify its effect on insurance,

INSURANCEi = b_1+b_2*INCOMEi+ ei

1. INSURANCEi : Life insurance cover by the i-th household.
2. INCOMEi : It is household income.
3. b_1 and b_2: Describing the relationship and are unknown parameters.
4. ei : It is a random uncorrelated error. Assumed to have zero mean and constant variance VAR^2.

Using hypothesized relationship to estimate, Dataset have a random sample of 20 households and collecting observations on INSURANCE and INCOME and applying the least-squares estimation procedure. The estimated equation, with standard errors in parentheses, is

(INSURANCE) = 6.855 + 3.8802 * (INCOME)

(se) (7.3835) (0.1121)

Find out the feature below:

1. The point estimate for the response of life-insurance coverage. When income increase by $ 1000 the INSURANCE is $3880.2 (the slope*1000).
2. A 95% interval estimate for this quantity is [$ 3645, $ 4116]. This interval is relatively narrow.
3. The value of b1 = 6.8550, if a family has no income, they would purchase $6855 of insurance. However, the data lack of information in the region about INCOME = 0, this value is not reliable. So do not give this value a direct economic interpretation.
4. The estimated equation could be used to assess likely requests for life insurance and what changes may occur as a result of income changes.
5. the amount of life insurance does not increase by the same amount as income increases.

Question 3.7:

Consider the capital asset pricing model (CAPM) in Exercise 2.10. Use the data in capm4.dat to answer each of the following:

(a)Test at the 5% level of significance the hypothesis that each stock's "beta" value is 1 against the alternative that it is not equal to 1. What is the economic interpretation of a beta equal to 1?

This is a two-tail test, so we can find the rejection area: [t<-1.978, t>1.978]

```
## [1] "t_c = t_(0.975, 130) = 1.978"
```

```
## [1] "t_c = t_(0.025, 130) = -1.978"
```

The results for each company:

```r
#Disney
moddis <- lm( ydisrf ~ x, data = capm)
summary(moddis)
```

```
##
## Call:
## lm(formula = ydisrf ~ x, data = capm)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.184659 -0.029342 -0.007377  0.029038  0.277836
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.001149   0.005956  -0.193    0.847
## x            0.897838   0.123627   7.262 3.11e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06843 on 130 degrees of freedom
## Multiple R-squared:  0.2886, Adjusted R-squared:  0.2831
## F-statistic: 52.74 on 1 and 130 DF,  p-value: 3.11e-11
```

```
## [1] " Disney t = (b2-b_2)/se(b2) = (0.897838-1)/0.123627 = -0.826"
```

Since -1.978 < t < 1.978, fail to reject H_0.

```r
#GE
modge <- lm( ygerf ~ x, data = capm)
summary(modge)
```

```
##
## Call:
## lm(formula = ygerf ~ x, data = capm)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.161385 -0.037402 -0.003939  0.034422  0.182421
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.001167   0.004759  -0.245    0.807
## x            0.899260   0.098782   9.104 1.33e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.05468 on 130 degrees of freedom
## Multiple R-squared:  0.3893, Adjusted R-squared:  0.3846
## F-statistic: 82.87 on 1 and 130 DF,  p-value: 1.325e-15

## [1] " GE t = (b2-b_2)/se(b2) = (0.899260-1)/0.098782 = -1.02"
```

Since -1.978 < t < 1.978, fail to reject H_0.

```r
#GM
modgm <- lm( ygmrf ~ x, data = capm)
summary(modgm)
```

```
##
## Call:
## lm(formula = ygmrf ~ x, data = capm)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.40485 -0.05473 -0.00517  0.06035  0.29157
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.011550   0.009743  -1.185    0.238
## x            1.261411   0.202223   6.238 5.77e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1119 on 130 degrees of freedom
## Multiple R-squared:  0.2304, Adjusted R-squared:  0.2244
## F-statistic: 38.91 on 1 and 130 DF,  p-value: 5.773e-09

## [1] " GM t = (b2-b_2)/se(b2) = (1.261411-1)/0.202223 = 1.293"
```

Since -1.978 < t < 1.978, fail to reject H_0.

```r
#IBM
modibm <- lm( yibmrf ~ x, data = capm)
summary(modibm)
```

```
##
## Call:
## lm(formula = yibmrf ~ x, data = capm)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.266910 -0.039943 -0.002883  0.036759  0.265579
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.005851   0.006091   0.961    0.339
## x           1.188208   0.126433   9.398 2.52e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06998 on 130 degrees of freedom
## Multiple R-squared:  0.4045, Adjusted R-squared:    0.4
## F-statistic: 88.32 on 1 and 130 DF,  p-value: 2.518e-16

## [1] " IBM t = (b2-b_2)/se(b2) = (1.188208-1)/0.126433 = 1.489"
```

Since -1.978 < t < 1.978, fail to reject H_0.

```
#Microsoft
modmsft <- lm( ymsftrf ~ x, data = capm)
summary(modmsft)
```

```
##
## Call:
## lm(formula = ymsftrf ~ x, data = capm)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.27541 -0.05240 -0.00780  0.04239  0.35082
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.006098   0.007747   0.787    0.433
## x           1.318947   0.160790   8.203 1.98e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.089 on 130 degrees of freedom
## Multiple R-squared:  0.3411, Adjusted R-squared:  0.336
## F-statistic: 67.29 on 1 and 130 DF,  p-value: 1.979e-13

## [1] " Microsoft t = (b2-b_2)/se(b2) = (1.318947-1)/0.160790 = 1.984"
```

Since -1.978 < t < 1.978, reject H_0.

```
#Exxon-Mobil
modxom <- lm( yxomrf ~ x, data = capm)
summary(modxom)
```

```
##
## Call:
## lm(formula = yxomrf ~ x, data = capm)
##
```

```
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.117543 -0.029814 -0.001812  0.025259  0.213676
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.007880   0.004322   1.823   0.0706 .
## x           0.413969   0.089713   4.614 9.33e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.04966 on 130 degrees of freedom
## Multiple R-squared:  0.1407, Adjusted R-squared:  0.1341
## F-statistic: 21.29 on 1 and 130 DF,  p-value: 9.332e-06
```

```
## [1] "Exxon-Mobil t = (b2-b_2)/se(b2) = (0.413969-1)/0.089713 = -6.532"
```

Since -1.978 < t < 1.978, reject H_0.

For Disney, GE, GM and IBM fail to reject the null hypothesis, indicating that the sample data are consistent with the conjecture that the Disney, GE, GM, and IBM stocks have the same volatility as the market portfolio.

For Microsoft and Exxon-Mobil, reject the null hypothesis, these stocks do not have the same volatility as the market portfolio.

(b)Test at the 5% level of significance the null hypothesis that Mobil-Exxon's "beta" value is greater than or equal to 1 against the alternative that it is less than 1. What is the economic interpretation of a beta less than 1?

Set Mobil-Exxon H_0 : b_2 >= 1, H_1 : b_2 < 1

```
## [1] "t_c = t_(0.05, 130) = -1.657"
```

now, caculate the t when b_2 = 1

```
## [1] "Exxon-Mobil t = (b2-b_2)/se(b2) = (0.413969-1)/0.089713 = -6.532"
```

Since t = -6.532 < t_c = -1.657, we reject H_0 and conclude that Mobil-Exxon's beta is less than 1. A beta equal to 1 suggests a stock's variation is the same as the market variation. A beta less than 1, the stock is less volatile than the market; it is a defensive stock.

(c)Test at the 5% level of significance the null hypothesis that Microsoft's "beta" value is less than or equal to 1 against the alternative that it is greater than 1. What is the economic interpretation of a beta more than 1?

Set Microsoft H_0 : b_2 >= 1, H_1 : b_2 < 1

```
## [1] "t_c = t_(0.95, 130) = 1.657"
```

```
## [1] "Microsoft t = (b2-b_2)/se(b2) = (1.318947-1)/0.160790 = 1.984"
```

Since t = 1.984 > t_c = 1.657, we reject H_0 and conclude that Microsoft's beta is greater than 1. A beta equal to 1 suggests a stock's variation is the same as the market variation. A beta greater than 1 implies the stock is more volatile than the market; it is an aggressive stock.

(d)Construct a 95% interval estimate of Microsoft's "beta." Assume that you are a stockbroker. Explain this result to an investor who has come to you for advice.

A 95% interval estimator for Microsoft's beta is

```
## [1] "b_2+t_(0.975, 130)*se(b2) = 1.637"
```

```
## [1] "b_2+t_(0.025, 130)*se(b2) = 1.001"
```

the corresponding interval estimate is [1.001, 1.637].

When 95% confidence, Microsoft's beta falls in the interval 1.001 to 1.637. It is possible that Microsoft's beta falls outside this interval.According to result conclusion in both parts (a):b_2 not equal 1 and (c): b_2 >1. Show that Microsoft is an aggressive stock.

(e)Test (at a 5% significance level) the hypothesis that the intercept term in the CAPM model for each stock is zero, against the alternative that it is not. What do you conclude?

The results for each company:

Set H_0 = 0, H_1 != 0, (!= means not equal to)

```
## [1] " Disney t = a_dis/se(b1) = (-0.00115)/0.005956 = -0.193"
```

Since -1.978 < t < 1.978, fail to reject H_0.

```
## [1] " GE t = a_ge/se(b1) = (-0.001167)/0.004759 = -0.245"
```

Since -1.978 < t < 1.978, fail to reject H_0.

```
## [1] " GM t = a_gm/se(b1) = (-0.01155)/0.009743 = -1.185"
```

Since -1.978 < t < 1.978, fail to reject H_0.

```
## [1] " IBM t = a_ibm/se(b1) = (0.005851)/0.006091 = 0.961"
```

Since -1.978 < t < 1.978, fail to reject H_0.

```
## [1] " Microsoft t = a_msft/se(b1) = (0.006098)/0.007747 = 0.787"
```

Since -1.978 < t < 1.978, fail to reject H_0.

```
## [1] "Exxon-Mobil t = a_xom/se(b1) = (0.00788)/0.004322 = 1.823"
```

Since -1.978 < t < 1.978, fail to reject H_0.

The result show do not reject the null hypothesis for any of the stocks. This result indicates that the sample data is consistent with the guess from economic theory that the intercept term equals 0.

Question 3.9:

Reconsider the presidential voting data (fair4.dat) introduced in Exercise 2.14. Use the data from 1916 to 2008 for this exercise.

```r
fair4  <- read.dta("fair4.dta") # input Stata file
fair4_1<- fair4[-1:-9,] # from 1916 to 2008
```

(a)Using the regression model VOTE = b1 + b2*GROWTH + e, test (at a 5% significance level) the null hypothesis that economic growth has no effect on the percentage vote earned by the incumbent party. Select an alternative hypothesis and a rejection region. Explain your choice.

```r
# Run regression and check the summary
modfair4 <- lm(fair4_1$vote~ fair4_1$growth, data = fair4_1)
summary(modfair4)
```

```
##
## Call:
## lm(formula = fair4_1$vote ~ fair4_1$growth, data = fair4_1)
##
```

```
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.866 -3.334 -1.003  3.004 10.826
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      50.8484     1.0125  50.218  < 2e-16 ***
## fair4_1$growth    0.8859     0.1819   4.871  7.2e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.798 on 22 degrees of freedom
## Multiple R-squared:  0.5189, Adjusted R-squared:  0.497
## F-statistic: 23.73 on 1 and 22 DF,  p-value: 7.199e-05
```

Set H_0: b2 = 0, H_1: b2 >0. Assume that GROWTH does influence the VOTE in positive way.

The estimated regression model:

VOTE = 50.8484 + 0.8859 * GROWTH

(SE) b1:(1.0125), b2:(0.1819)

```
## [1] "The rejection region is t > t_(0.95, 22) = 1.717"
```

```
## [1] "t = b2-0/se(b2) = (0.8859-0)/0.1819 = 4.87"
```

Since t = 4.870 > 1.717, we reject the null hypothesis that b2 = 0 and accept the alternative that b2 > 0 . We conclude that economic GROWTH has a positive effect on the percentage VOTE earned by the incumbent party

(b)Using the regression model in part (a), construct a 95% interval estimate for b2, and interpret.

```
b2 <- 0.8859; seb2 <- 0.1819
a <- round((qt(c(0.975), df=22)),3); b <- round((qt(c(0.025), df=22)),3)
cat("A 95% interval estimate for b2 from the regression in part (a) is:[", round((b2+b*seb2),3), ",",ro
```

```
## A 95% interval estimate for b2 from the regression in part (a) is:[ 0.509 , 1.263 ]
```

When 95% confidence, the true value of b2 is between [0.509,1.263]. Since b2 = the change in percentage VOTE due to economic GROWTH, we expect that an increase in the growth rate of 1% will increase the percentage VOTE by an amount between [0.509,1.263].

(c)Using the regression model VOTE = b1 + b2*INFLATION + e, test the null hypothesis that inflation has no effect on the percentage vote earned by the incumbent party. Select an alternative hypothesis, a rejection region, and a significance level. Explain your choice.

Set the hypotheses H0: b2 = 0, H1: b2 <= 0 . The alternative b2 <=0 is chosen because we assume that INFLATION influence the VOTE in a negative way.

```
modfair41 <- lm(fair4_1$vote~ fair4_1$inflation, data = fair4_1)
summary(modfair41)
```

```
##
## Call:
## lm(formula = fair4_1$vote ~ fair4_1$inflation, data = fair4_1)
##
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max 
## -17.2887  -3.2734  -0.4371   5.2854  10.5206 
## 
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)    
## (Intercept)        53.4077     2.2500  23.737   <2e-16 ***
## fair4_1$inflation  -0.4443     0.5999  -0.741    0.467    
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 6.833 on 22 degrees of freedom
## Multiple R-squared:  0.02433,    Adjusted R-squared:  -0.02002 
## F-statistic: 0.5485 on 1 and 22 DF,  p-value: 0.4668
```

The estimated regression model:

VOTE = 53.4077 - 0.4443 * INFLATION

(SE) b1:(2.25) , b2:(0.5999)

```
## [1] "The rejection region is t > t_(0.05, 22) = -1.717"
```

```
## [1] "t = b2-0/se(b2) = (-0.4443-0)/0.5999 = -0.741"
```

Since -0.741 > -1.717 we do not reject the null hypothesis. There is not enough evidence to suggest inflation has a negative effect on the vote.

(d)Using the regression model in part (c), construct a 95% interval estimate for b2, and interpret.

```
b2 <- -0.4443; seb2 <- 0.5999
a <- round((qt(c(0.975), df=22)),3); b <- round((qt(c(0.025), df=22)),3)
cat("A 95% interval estimate for b2 from the regression in part (a) is:[", round((b2+b*seb2),3), ",",ro
```

```
## A 95% interval estimate for b2 from the regression in part (a) is:[ -1.688 , 0.8 ]
```

When 95% confidence, the true value of b2 is between [-1.688 , 0.8].It suggests that an increase in the INFLATION rate of 1% could increase or decrease or have no effect on the percentage VOTE earned by the incumbent party.

(e)Test the null hypothesis that if INFLATION = 0 the expected vote in favor of the incumbent party is 50%, or more. Select the appropriate alternative. Carry out the test at the 5% level of significance. Discuss your conclusion.

E(VOTE) = E(b1 - b2 * 0) = E(b1)

Set H_0: b1 >= 50 H_1: b1 < 50 The test statistic assume H0 is true at the point b1 = 50

```
## [1] "t = b1-50/se(b1) = (53.4077-50)/2.25 = 1.515"
```

t: 1.515 > -1.717, do not reject the null hypothesis. There is no evidence to suggest that the expected VOTE in favor of the incumbent party is less than 50% when there is no inflation.

(f)Construct a 95% interval estimate of the expected vote in favor of the incumbent party if INFLATION = 2%. Discuss the interpretation of this interval estimate.

```
## [1] "E(VOTE)_bar = b1+b2*2 = 53.4077 + -0.44431*2= 52.519"
```

The standard error of this estimate is the square root:

VAR(b1+b2 * 2)_bar = VAR(b1)_bar + VAR(b2)_bar * 2^2 + 2 * 2* COV(b1,b2)_bar

```
## [1] "5.0625+4*(0.3599)+4*(-1.0592) = 2.265"
```

So The 95% interval estimate is:

```
## [(b1+2*b2) + t_(0.025,22)*se(b1+2b2), (b1+2*b2) + t_(0.975,22)*se(b1+2b2)] =
## [ 49.398 , 55.64 ]
```

At 95% confidence that the expected VOTE in favor of the incumbent party when inflation is at 2% is between 49.40% and 55.64%.