Econometric Homework 3

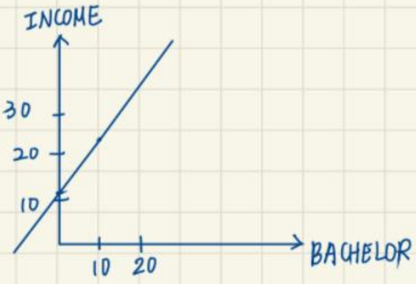
0613404 陸恭葦

3.7 $N=51$ $\widehat{INCOME} = (a) + 1.029\ BACHELOR = \hat{\beta}_0 + \hat{\beta}_1\ BACHELOR$, $\hat{\beta}_1 = 1.029$

se (2.672) (c)

t (4.31) (10.75)

(a.) $(a) = \hat{\beta}_0 = 4.31 \times 2.672 = 11.51632$



(b.) $\widehat{INCOME} = 11.51632 + 1.029\ BACHELOR$

The line is increasing.

The relationship is positive.

And increasing at a constant rate (1.029).

(c.) $(c) = \dfrac{1.029}{10.75} = 0.0957$

(d.) Let intercept parameter : $\hat{\beta}_0$, slope parameter : $\hat{\beta}_1$

$H_0: \hat{\beta}_0 = 10$

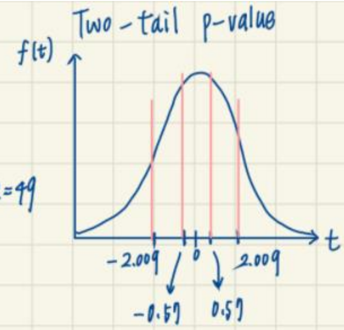
$H_0: \hat{\beta}_0 \neq 10$

$$t = \frac{\hat{\beta}_0 - 10}{se(\hat{\beta}_0)} = \frac{1.51632}{2.672} = 0.567 \sim t_{(N-2)},\quad N-2 = 49$$

(e.) p-value from (d.) is 0.572, $\alpha = 0.05$

$t(0.025, 49) = 2.009$

$t(-0.025, 49) = -2.009$

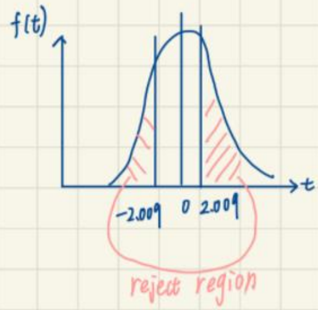


(f.) 99% confidence interval $\alpha = 0.01$, $df = 51-2 = 49$

$\hat{\beta}_1 \pm t_{(49, 0.005)} \times se(\hat{\beta}_1)$

$= 1.029 \pm 2.68 \times 0.0957$

$= 1.029 \pm 0.2565 = [0.7725, 1.2855]$



(g.) $H_0: \hat{\beta}_1 = 1$

$H_1: \hat{\beta}_1 \neq 1$

$$t = \frac{\hat{\beta}_1 - 1}{se(\hat{\beta}_1)} = \frac{0.029}{0.0957} = 0.303$$

$\alpha = 0.05$, two-tailed $t(0.025, 49) = 2.009$

$t(-0.025, 49) = -2.009$

$\therefore$ fail to reject the null hypothesis $H_0: \beta_2 = 1$, reject H_0

$\because -2.009 < 0.303 < 0.2009$

We are not able to conclude that the slople coefficient is one against the alternative.

18.(c) 99% CI $\Rightarrow \alpha = 0.01$, $N=20$

\$100,000 income $\rightarrow$ 100 (thousands of dollars)

$$\widehat{E(INSURANCE|INCOME=100)} = \beta_0 + \beta_1 100$$

The point estimate is

$$\widehat{E(INSURANCE|INCOME=100)} = b_0 + b_1 100 = 6.855 + 3.880(100)$$
$$= 6.855 + 388 = 394.855$$

$$\widehat{var}(b_0 + 100b_1) = \widehat{var}(b_0) + (100^2 \times \widehat{var}(b_1)) + (2 \times 100 \times \widehat{cov}(b_1, b_2))$$
$$= 7.383^2 + (100^2 \times 0.112^2) + (2 \times 100 \times -0.746)$$
$$= 54.5087 + 125.44 - 149.2 = 30.7487$$

$$se(b_0 + 100b_1) = \sqrt{\widehat{var}(b_0 + 100b_1)} = \sqrt{30.7487} = 5.5452$$

99% interval estimate:

$$(b_0 + 100b_1) \pm t_{(0.005,18)}\, se(b_0 + 100b_1)$$
$$= [394.855 \pm 2.8784 \times 5.5452]$$
$$= [378.8937 \quad , \quad 410.8163]$$

3.26

(a) Estimate the linear regression $\mathrm{WAGE} = \beta_1 + \beta_2 EXPER + e$

We can find that the p-value of intercept and exper are all significant. However, the R-squared of the estimator is small.

```
Call:
lm(formula = wage ~ exper, data = cps5_small)

Residuals:
    Min      1Q  Median      3Q     Max 
-22.113 -10.476  -3.911   5.970 196.198 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 21.62307    0.88789  24.353  < 2e-16 ***
exper        0.08629    0.03304   2.612  0.00912 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15.18 on 1198 degrees of freedom
Multiple R-squared:  0.005662,	Adjusted R-squared:  0.004832 
F-statistic: 6.822 on 1 and 1198 DF,  p-value: 0.009117
```

Table: Regression output showing the coefficients

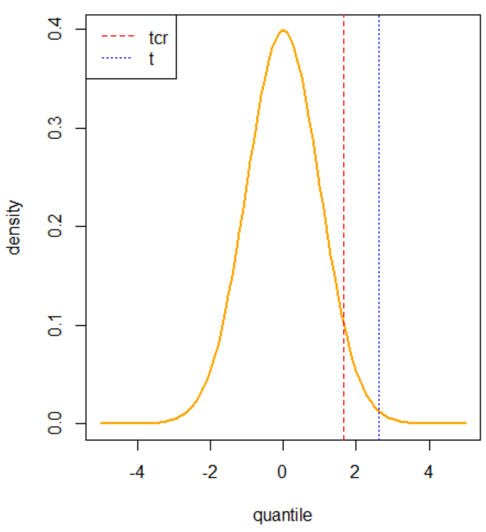
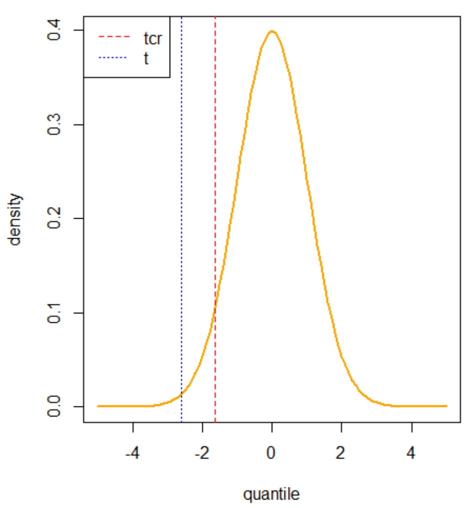
	Estimate	Std..Error	t.value	Pr...t..
(Intercept)	21.6230748	0.8878907	24.353307	0.0000000
exper	0.0862904	0.0330375	2.611895	0.0091168

Rcode

```
mod <- lm(wage~exper,data = cps5_small)
smod <- summary(mod)
smod
SStable <- data.frame(xtable(mod))
kable(table, caption="Regression output showing the coefficients")
```

(b)

I test right-tailed and left-tailed test with $\propto= 0.05$ and calculate t and tcr(t critical value),then conclude the result below:

	
The result if the right-tailed test	He result od the left-tailed test
$H_0: \beta_2 \geq 0$ $H_1: \beta_2 < 0$	$H_0: \beta_2 \leq 0$ $H_1: \beta_2 > 0$
tcr<t	t<tcr
Reject H_0	Reject H_0
We are able to conclude that each additional exper (potential experience) will increase the wage.	

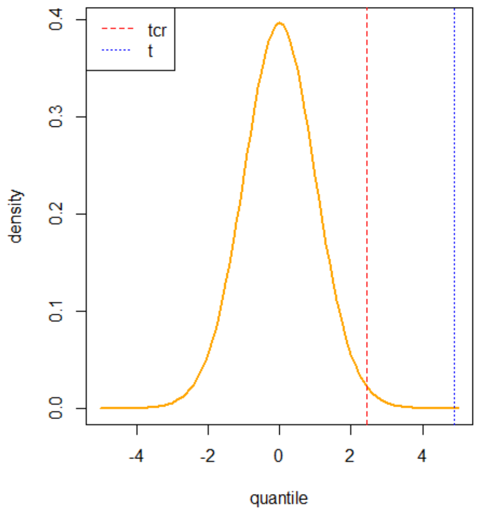
```
Rcode
b2 <- coef(mod)[[2]] # the coefficient on exper
seb <- sqrt(vcov(mod)[2,2]) #standard error of b2
df0 <- df.residual(mod) # degrees of freedom
c <- 0
alpha <- 0.05
# one tail test
t <- (b2-c)/seb
tcr <- qt(1-alpha, df0) # note: alpha is not divided by 2
```

```
t01 <- (b2-c)/seb
tcr01 <- qt(alpha, df0)
curve(dt(x, df0), from = -5, to = 5, col = "orange",
      xlab = "quantile", ylab = "density", lwd = 2)
abline(v=c(tcr,t), col=c("red", "blue"), lty=c(2, 3))
legend("topleft", legend=c("tcr", "t"), col=c("red", "blue"), lty=c(2, 3))
curve(dt(x, df0), from = -5, to = 5, col = "orange",
      xlab = "quantile", ylab = "density", lwd = 2)
abline(v=c(tcr01,-t01), col=c("red", "blue"), lty=c(2, 3))
legend("topleft", legend=c("tcr", "t"), col=c("red", "blue"), lty=c(2, 3))
```

(c) Estimate the linear regression $\mathrm{WAGE} = \beta_1 + \beta_2 EXPER + e \ and \ METRO = 1$

```
Call:
lm(formula = food_exp ~ income, data = food)

Residuals:
     Min       1Q   Median       3Q      Max 
-223.025  -50.816   -6.324   67.879  212.044 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   83.416     43.410   1.922   0.0622 .  
income        10.210      2.093   4.877 1.95e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 89.52 on 38 degrees of freedom
Multiple R-squared:  0.385,	Adjusted R-squared:  0.3688 
F-statistic: 23.79 on 1 and 38 DF,  p-value: 1.946e-05


 Table: Regression output showing the coefficients

|             | Estimate| Std..Error|  t.value|  Pr...t..|
|:------------|--------:|----------:|--------:|---------:|
|(Intercept)  | 83.41600|  43.410163| 1.921578| 0.0621824|
|income       | 10.20964|   2.093263| 4.877381| 0.0000195|
```

Then,I test right-tailed test with $\propto = 0.01$ and calculate t and tcr(t critical value),then conclude the result below:



The result if the right-tailed test
$H_0: \beta_2 \geq 0$ $H_1: \beta_2 < 0$
tcr<t
Reject H_0
We are able to conclude that each additional exper (potential experience) in metropolitan area will increase the wage.

The effect here is significant because we get the similar result with (b). We can conclude each additional exper (potential experience) in metropolitan area will increase the wage.

Rcode

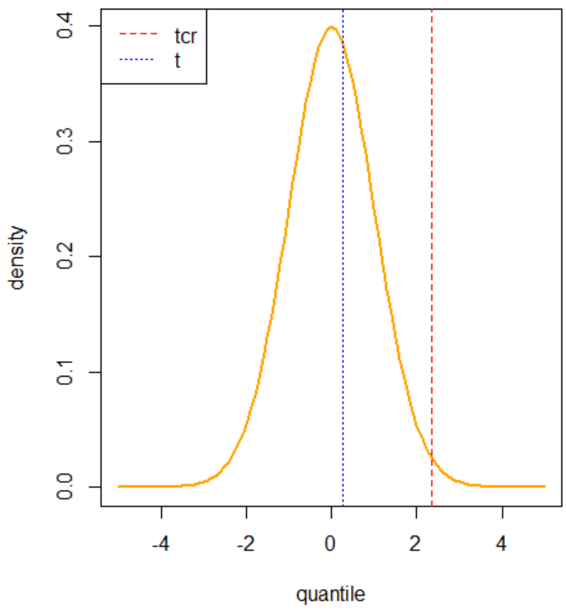
totoal <-which(cps5_small$metro==1)
cps2<-cps5_small[total,]
mod1 <- lm(wage~exper,data = cps2)
#mod1
smod1 <- summary(mod1)
#smod1
table <- data.frame(xtable(mod1))
kable(table, caption="Regression output showing the coefficients")
 	
b22 <- coef(mod1)[[2]] # the coefficient on exper
seb1 <- sqrt(vcov(mod1)[2,2]) #standard error of b2

```
df1 <- df.residual(mod1) # degrees of freedom
c <- 0
alpha <- 0.01
# one tail test
t1 <- (b22-c)/seb1
tcr1 <- qt(1-alpha, df1) # note: alpha is not divided by 2

curve(dt(x, df1), from = -5, to = 5, col = "orange",
        xlab = "quantile", ylab = "density", lwd = 2)
abline(v=c(tcr1,t1), col=c("red", "blue"), lty=c(2, 3))
legend("topleft", legend=c("tcr", "t"), col=c("red", "blue"), lty=c(2, 3))
```

(d) Estimate the linear regression $\mathrm{WAGE} = \beta_1 + \beta_2 EXPER + e \; and \; METRO = 0$

```
Call:
lm(formula = wage ~ exper, data = cps3)

Residuals:
    Min      1Q  Median      3Q     Max 
-15.954  -7.066  -2.673   5.013  52.535 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 19.41034    1.48782  13.046   <2e-16 ***
exper        0.01321    0.05239   0.252    0.801    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 10.39 on 212 degrees of freedom
Multiple R-squared:  0.0003,	Adjusted R-squared:  -0.004416 
F-statistic: 0.06363 on 1 and 212 DF,  p-value: 0.8011


Table: Regression output showing the coefficients

|            |   Estimate| Std..Error|    t.value| Pr...t..|
|:-----------|----------:|----------:|----------:|--------:|
|(Intercept) | 19.4103373|  1.4878173| 13.0461832| 0.000000|
|exper       |  0.0132143|  0.0523875|  0.2522426| 0.801098|
```

Then,I test right-tailed test with $\propto = 0.01$ and calculate t and tcr(t critical value),then conclude the result below:



The result if the right-tailed test

$H_0: \beta_2 \geq 0$

$H_1: \beta_2 < 0$

t<tcr

Don't reject H_0

We are not able to conclude that each additional exper (potential experience) in nonmetropolitan area will increase the wage.

From the result above, we can't safely say that experience has no effect on wages for individuals living in nonmetropolitan areas because we reject H_0. There is not sufficient evidence to make sure the relation between experience in nonmetropolitan and wage.

Rcode

```
total2 <-which(cps5_small$metro==0)
cps3<-cps5_small[total2,]
mod2 <- lm(wage~exper,data = cps3)
smod2 <- summary(mod2)
smod2
table <- data.frame(xtable(mod2))
kable(table, caption="Regression output showing the coefficients")
```

```
b23 <- coef(mod2)[[2]] # the coefficient on exper
seb2 <- sqrt(vcov(mod2)[2,2]) #standard error of b2
df2 <- df.residual(mod2) # degrees of freedom
c <- 0
alpha <- 0.01
# one tail test
t2 <- (b23-c)/seb2
tcr2 <- qt(1-alpha, df2) # note: alpha is not divided by 2

curve(dt(x, df2), from = -5, to = 5, col = "orange",
      xlab = "quantile", ylab = "density", lwd = 2)
abline(v=c(tcr2,t2), col=c("red", "blue"), lty=c(2, 3))
legend("topleft", legend=c("tcr", "t"), col=c("red", "blue"), lty=c(2, 3))
```