Solutions for Homework 6

Tong Zhou

Problems: 8.3, 8.4, 8.18, 8.19, 9.2, 9.20, 9.24, 9.25.

PROBLEM 8.3

Proof. The restricted regression can be written as

$$\boldsymbol{y} = (\boldsymbol{X}_2 - \boldsymbol{X}_1)\beta_2 + \boldsymbol{e}$$

Thus, we have estimators

$$\widehat{\beta}_2 = ((\boldsymbol{X}_2 - \boldsymbol{X}_1)'(\boldsymbol{X}_2 - \boldsymbol{X}_1))^{-1}(\boldsymbol{X}_2 - \boldsymbol{X}_1)'\boldsymbol{y}$$

and

$$\widehat{\beta}_1 = -\widehat{\beta}_2.$$

□

PROBLEM 8.4

Proof. (a) When $\boldsymbol{\beta} = 0$, the estimator of α is the average of Y, i.e.

$$\widehat{\alpha} = \bar{\boldsymbol{y}} = \frac{1}{n}\sum_{i=1}^{n} Y_i.$$

(b) By equation (8.25) in Hansen's text,

$$\widehat{\alpha}_{\mathsf{emd}} = \widehat{\alpha} - \widehat{\boldsymbol{V}}_{\alpha\beta}\widehat{\boldsymbol{V}}_{\beta}^{-1}\widehat{\beta}_{\mathsf{ols}},$$

where $\widehat{\boldsymbol{V}}_{\eta} = \begin{pmatrix} \widehat{V}_{\alpha} & \widehat{V}_{\alpha\beta} \\ \widehat{V}_{\beta\alpha} & \widehat{V}_{\beta} \end{pmatrix}$, $\boldsymbol{\eta} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$.

□

PROBLEM 8.18

Proof. We shall start with part (c), because part (a) and (b) are special cases.
(c) Mimicking the objective function in 8.19, when n_1 and n_2 are different,

$$\begin{aligned} J(\beta_1, \beta_2) &= \frac{1}{2}\begin{pmatrix} n_1\left(\widehat{\beta}_1 - \beta_1\right) \\ n_2\left(\widehat{\beta}_2 - \beta_2\right) \end{pmatrix}' \begin{pmatrix} \boldsymbol{V}_1^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{V}_2^{-1} \end{pmatrix} \begin{pmatrix} \widehat{\beta}_1 - \beta_1 \\ \widehat{\beta}_2 - \beta_2 \end{pmatrix} \\ &= \frac{1}{2}n_1\left(\widehat{\beta}_1 - \beta_1\right)'\boldsymbol{V}_1^{-1}\left(\widehat{\beta}_1 - \beta_1\right) + \frac{1}{2}n_2\left(\widehat{\beta}_2 - \beta_2\right)'\boldsymbol{V}_2^{-1}\left(\widehat{\beta}_2 - \beta_2\right) \end{aligned}$$

and the constraint is

$$\boldsymbol{R}'\beta = (\beta_1 - \beta_2) = \mathbf{0}$$

The Lagrangian function is

$$\mathscr{L}(\beta_1, \beta_2, \lambda) = J(\beta_1, \beta_2) + \lambda'(\beta_1 - \beta_2)$$

The FOC is:

$$\begin{cases} \frac{\partial \mathscr{L}}{\partial \beta_1} = -n_1 \boldsymbol{V}_1^{-1}\widehat{\beta}_1 + n_1 \boldsymbol{V}_1^{-1}\widetilde{\beta}_1 + \widetilde{\lambda} = \mathbf{0} \\ \frac{\partial \mathscr{L}}{\partial \beta_2} = -n_2 \boldsymbol{V}_2^{-1}\widehat{\beta}_2 + n_2 \boldsymbol{V}_2^{-1}\widetilde{\beta}_2 - \widetilde{\lambda} = \mathbf{0} \\ \frac{\partial \mathscr{L}}{\partial \lambda} = \widetilde{\beta}_1 - \widetilde{\beta}_2 = \mathbf{0} \end{cases}$$

From it, we have

$$\begin{aligned} \widetilde{\beta} &= \left(n_1 \boldsymbol{V}_1^{-1} + n_2 \boldsymbol{V}_2^{-1}\right)^{-1} \left(n_1 \boldsymbol{V}_1^{-1}\widehat{\beta}_1 + n_2 \boldsymbol{V}_2^{-1}\widehat{\beta}_2\right) \\ &= \left[\frac{n_1}{n_2}\boldsymbol{V}_1^{-1} + \boldsymbol{V}_2^{-1}\right]^{-1} \left[\frac{n_1}{n_2}\boldsymbol{V}_1^{-1}\widehat{\beta}_1 + \boldsymbol{V}_2^{-1}\widehat{\beta}_2\right] \end{aligned}$$

Therefore, the asymptotic behavior of $\widetilde{\beta}$ totally depends on the convergence rate $\frac{n_1}{n_2}$.

- **Case 1.** $n_1 = n_2 = n$

  This case reduces to part (a). It is straightforward that

  $$\widetilde{\beta} = \left(\boldsymbol{V}_1^{-1} + \boldsymbol{V}_2^{-1}\right)^{-1} \left(\boldsymbol{V}_1^{-1}\widehat{\beta}_1 + \boldsymbol{V}_2^{-1}\widehat{\beta}_2\right)$$

  and by

  $$\sqrt{n}\left(\widetilde{\beta} - \beta\right) = \left(\boldsymbol{V}_1^{-1} + \boldsymbol{V}_2^{-1}\right)^{-1} \boldsymbol{V}_1^{-1}\sqrt{n}\left(\widehat{\beta}_1 - \beta\right) + \left(\boldsymbol{V}_1^{-1} + \boldsymbol{V}_2^{-1}\right)^{-1} \boldsymbol{V}_2^{-1}\sqrt{n}\left(\widehat{\beta}_2 - \beta\right)$$

  Define $\Omega_1 = \left(\boldsymbol{V}_1^{-1} + \boldsymbol{V}_2^{-1}\right)^{-1}\boldsymbol{V}_1^{-1}$ and $\Omega_2 = \left(\boldsymbol{V}_1^{-1} + \boldsymbol{V}_2^{-1}\right)^{-1}\boldsymbol{V}_2^{-1}$

  Then we have

  $$\Omega_1\sqrt{n}\left(\widehat{\beta}_1 - \beta\right) \xrightarrow{\mathsf{d}} \mathcal{N}\left(\mathbf{0}, \left(\boldsymbol{V}_1^{-1} + \boldsymbol{V}_2^{-1}\right)^{-1} \boldsymbol{V}_1^{-1} \left(\boldsymbol{V}_1^{-1} + \boldsymbol{V}_2^{-1}\right)^{-1}\right)$$

  and

  $$\Omega_2\sqrt{n}\left(\widehat{\beta}_2 - \beta\right) \xrightarrow{\mathsf{d}} \mathcal{N}\left(\mathbf{0}, \left(\boldsymbol{V}_1^{-1} + \boldsymbol{V}_2^{-1}\right)^{-1} \boldsymbol{V}_2^{-1} \left(\boldsymbol{V}_1^{-1} + \boldsymbol{V}_2^{-1}\right)^{-1}\right)$$

  Since they are independent,

  $$\sqrt{n}\left(\widetilde{\beta} - \beta\right) \xrightarrow{\mathsf{d}} \mathcal{N}\left(\mathbf{0}, \left(\boldsymbol{V}_1^{-1} + \boldsymbol{V}_2^{-1}\right)^{-1}\right)$$

- **Case 2:** $\frac{n_1}{n_2} \to 0$.

  In this case, $\widetilde{\beta} \xrightarrow{\mathsf{P}} \widehat{\beta}_2$ and

  $$\sqrt{n_1 + n_2}\left(\widetilde{\beta} - \beta\right) \xrightarrow{\mathsf{P}} \mathcal{N}(\mathbf{0}, \boldsymbol{V}_2)$$

- **Case 3**: $\frac{n_1}{n_2} \to \infty$.
  Similarly,

$$\sqrt{n_1+n_2}\left(\widetilde{\beta}-\beta\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(\mathbf{0}, \boldsymbol{V}_1\right)$$

- **Case 4**: $\frac{n_1}{n_2} \to c > 0$.
  It is easy to show that

$$\sqrt{n_1+n_2}\left(\widetilde{\beta}-\beta\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(\mathbf{0}, (1+c)\left(c\boldsymbol{V}_1^{-1}+\boldsymbol{V}_2^{-1}\right)^{-1}\right)$$

  It is easily verified that when $c=1$, it reduces to **Case 1**.

□

PROBLEM 8.19

(a) (b) (c) See Table 1, Table 2 and Table 3
(d) The constraint is $\beta_2 \geqslant 0$ and $\beta_2+\beta_3 \geqslant 0$.
(e) Observe that $\widehat{\beta}_{2,\text{emd}} > 0$ and $\widehat{\beta}_{2,\text{emd}}+\widehat{\beta}_{3,\text{emd}} < 0$ in Table 3 , so the inequality constraint is equivalent to imposing a constraint $\beta_2+\beta_3=0$. See results in Table 4

TABLE 1
OLS Estimates and s.e.

	Estimates	se
education	0.089	0.003
experience	0.030	0.003
experience2	-0.037	0.006
married 1	0.181	0.025
married 2	-0.480	0.033
married 3	-0.040	0.056
widow	0.236	0.173
divorced	0.074	0.045
sep	0.016	0.053
intercept	1.192	0.046

TABLE 2
CLS Estimates and s.e.

	Estimates	se
education	0.089	0.003
experience	0.030	0.003
experience 2	-0.037	0.005
married 1	0.180	0.024
married 2	-0.479	0.572
married 3	-0.040	0.057
widow	0.180	0.024
divorced	0.055	0.038
sep	0.055	0.038
intercept	1.189	0.045

TABLE 3
EMD estimates and s.e.

	Estimates	se
education	0.089	0.003
experience	0.030	0.003
experience 2	-0.037	0.006
married 1	0.180	0.025
married 2	-0.480	0.033
married 3	-0.040	0.056
widow	0.180	0.025
divorced	0.050	0.038
sep	0.050	0.038
intercept	1.188	0.046

TABLE 4
EMD Estimates and s.e. Under Inequality Constraint

	Estimates	se
education	0.090	0.003
experience	0.024	0.002
experience 2	-0.024	0.002
married 1	0.190	0.025
married 2	-0.487	0.033
married 3	-0.035	0.056
widow	0.190	0.025
divorced	0.063	0.038
sep	0.063	0.038
intercept	1.216	0.044

PROBLEM 9.2

PROOF. We know $\widehat{\beta}_1 = (X_1'X_1)^{-1}X_1'y_1$ and $\widehat{\beta}_2 = (X_2'X_2)^{-1}X_2'y_2$. Under certain regularity conditions,

$$\sqrt{n}(\widehat{\beta}_1 - \beta_1) \xrightarrow{d} N(\mathbf{0}, V_{\beta_1}),$$
$$\sqrt{n}(\widehat{\beta}_2 - \beta_2) \xrightarrow{d} N(\mathbf{0}, V_{\beta_2}),$$

where $V_{\beta_1} = [\mathbb{E}(x_{1i}x_{1i}')]^{-1}\mathbb{E}(x_{1i}x_{1i}'e_{1i}^2)[\mathbb{E}(x_{1i}x_{1i}')]^{-1}$ and V_{β_2} is defined in the same way.

By the independence between the two data sets,

$$\sqrt{n}\left((\widehat{\beta}_1 - \beta_1) - (\widehat{\beta}_2 - \beta_2)\right) \xrightarrow{d} N(\mathbf{0}, V_{\beta_1} + V_{\beta_2})$$

(b)

We propose a Wald-type statistic:

$$\begin{aligned} \mathcal{W}_n &= n(\widehat{\beta}_1 - \widehat{\beta}_2)'\left(\widehat{V}_{\beta_1} + \widehat{V}_{\beta_2}\right)^{-1}(\widehat{\beta}_1 - \widehat{\beta}_2) \\ &= (\widehat{\beta}_1 - \widehat{\beta}_2)'\left(\widehat{V}_{\widehat{\beta}_1} + \widehat{V}_{\widehat{\beta}_2}\right)^{-1}(\widehat{\beta}_1 - \widehat{\beta}_2) \end{aligned}$$

where $\widehat{V}_{\beta_1} \xrightarrow{P} V_{\beta_1}$ and $\widehat{V}_{\beta_2} \xrightarrow{P} V_{\beta_2}$.

(c)

Under H_0, from (a) we can derive:

$$\sqrt{n}\left(\widehat{V}_{\beta_1} + \widehat{V}_{\beta_2}\right)^{-\frac{1}{2}}\left(\widehat{\beta}_1 - \widehat{\beta}_2\right) \xrightarrow{d} N(\mathbf{0}, I_k)$$

and thus,

$$\mathcal{W}_n \xrightarrow{d} \chi_k^2$$

□

PROBLEM 9.20

PROOF. It depends on how you perceive the sample size $n = 50$.

If you believe $n = 50$ is a small sample, then based on the given information, there is no hope of conducting a hypothesis testing procedure, even if the homoskedasticity assumption is made. However, under the normality condition, F test can be used to do the testing.

Recall that under the normality assumption, we have two independent χ^2-distributed statistics:

$$\frac{(n-k)s^2}{\sigma^2} \sim \chi_{n-k}^2$$

and

$$\left(\boldsymbol{R}'\widehat{\beta}-\boldsymbol{r}\right)'\left(\sigma^2\boldsymbol{R}'\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{R}\right)^{-1}\left(\boldsymbol{R}'\widehat{\beta}-\boldsymbol{r}\right)\sim\chi_q^2$$

Therefore, their ratio divided by respective degrees of freedom follows a F-distribution (Equation 9.12 of Hansen's book)

$$\frac{\left(\boldsymbol{R}'\widehat{\beta}-\boldsymbol{r}\right)'\left(\sigma^2\boldsymbol{R}'\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{R}\right)^{-1}\left(\boldsymbol{R}'\widehat{\beta}-\boldsymbol{r}\right)/q}{s^2/\sigma^2}=\frac{(\widetilde{\sigma}^2-\widehat{\sigma}^2)/q}{\widehat{\sigma}^2/(n-k)}=\frac{\left(R_L^2-R_S^2\right)/q}{\left(1-R_L^2\right)/(n-k)}\sim F_{(q,n-k)}$$

where R_L^2 and R_S^2 are R^2's of the long regression and the short regression respectively, and $n=50, k=8$ and $q=3$.

If you believe $n=50$ is a large sample, then only the homoskedasticity condition is sufficient for hypothesis testing. By **Theorem 9.6**, the above F-statistics

$$\frac{\left(\widetilde{\sigma}^2-\widehat{\sigma}^2\right)/q}{\widehat{\sigma}^2/(n-k)}\xrightarrow{\mathsf{d}}\frac{\chi_q^2}{q},$$

where the χ_q^2 distribution does not depend on the normality assumption.

Both tests give the same result that the null hypothesis $\boldsymbol{\beta}_2=\mathbf{0}$ cannot be rejected. □

PROBLEM 9.24

Proof. **(a)**

Since the two ratios do not rely on $\widehat{\alpha}$, setting $\alpha=0$ is irrelevant.

(b)

From the simulation (codes are provided below), the bias for $\widehat{\beta}$ and $\widehat{\theta}$ are

$$\mathsf{Bias}\left(\widehat{\beta}\right)=\mathbb{E}\left[\widehat{\beta}\right]-\beta=0.01$$
$$\mathsf{Bias}\left(\widehat{\theta}\right)=\mathbb{E}\left[\widehat{\theta}\right]-\theta=0.42.$$

So $\widehat{\theta}$ is biased. Though the continuous mapping theorem implies $\widehat{\theta}\xrightarrow{\mathsf{P}}\theta$, $n=50$ is a small sample size for a nonlinear transformation of β. Also, this upward bias can be justified by the Jensen's Inequality:

$$\mathbb{E}\left[\widehat{\theta}\right]=\mathbb{E}\left[e^{\widehat{\beta}}\right]>e^{\mathbb{E}\left[\widehat{\beta}\right]}=\theta$$

(c)

From the 1000 replications,

$$\mathbb{P}\left[t_\beta>1.645\right]=0.052$$
$$\mathbb{P}\left[t_\theta>1.645\right]=0.003$$

Since $t_\beta, t_\theta \xrightarrow{d} \mathcal{N}(0,1)$, both probabilities should be close to 0.05. Thanks to the small sample size, it is not surprising that t_θ performs quite worse.

Why sample size matters? We further run simulations by varying sample size n: 50, 100, 500, 1000, 3000, 5000, and investigate how the biases and probabilities evolve with those sample sizes.

From Fig. 1, both biases converge to 0, as sample sizes increase. In Fig. 2, the probability for t_β and t_θ hovers around 0.05, as the sample sizes become large.

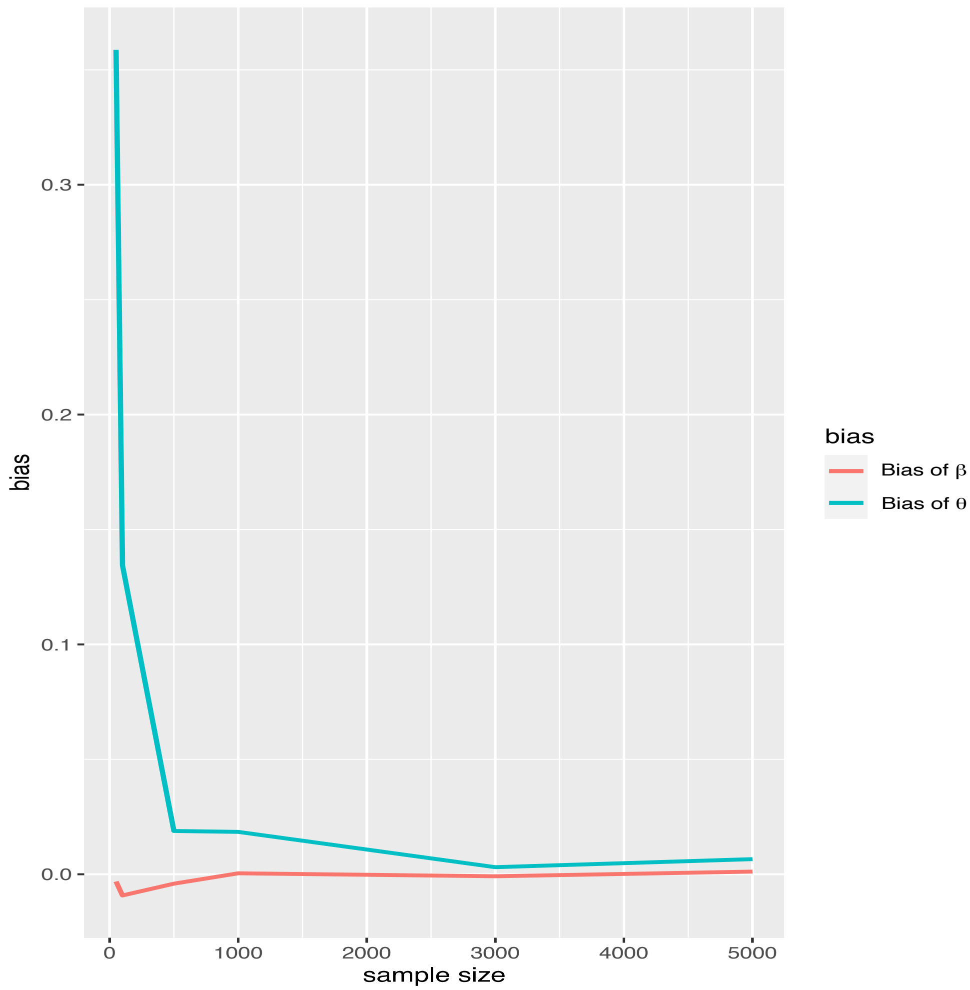

FIGURE 1. $Bias\left(\widehat{\beta}\right)$ and $Bias\left(\widehat{\theta}\right)$

□

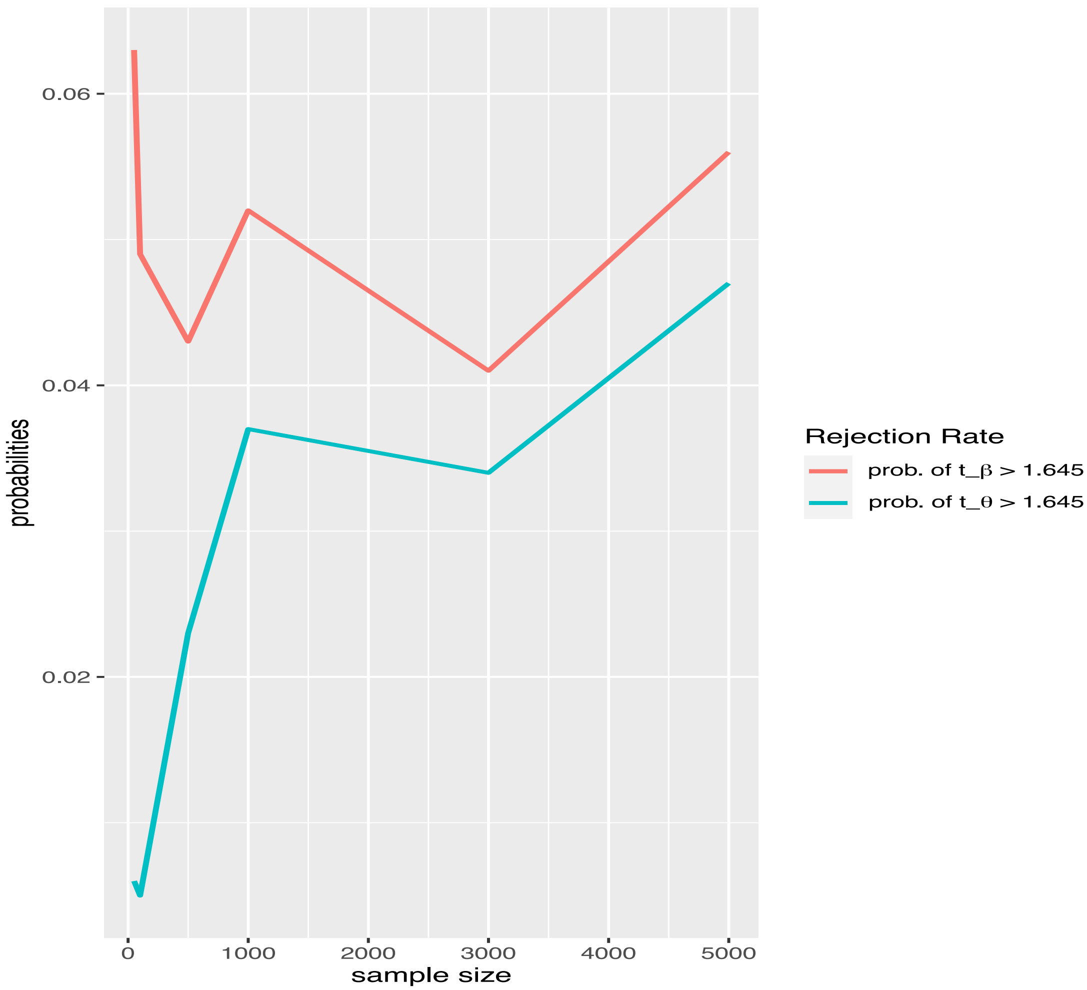

FIGURE 2. $\mathbb{P}\left[t_\beta > 1.645\right]$ *and* $\mathbb{P}\left[t_\theta > 1.645\right]$

PROBLEM 9.25

9.25.R

tongzhou

2021-03-17

```
rm(list = ls())

library(pacman)
p_load(tidyverse,gtsummary,jtools,car,haven)

df <- read_stata("https://www.ssc.wisc.edu/~bhansen/econometrics/Invest1993.dta")

#### (a), (b)

sub_1987 <- df %>% filter(year == 1987)

model <- sub_1987 %>%
  lm(inva ~ vala + cfa + debta, data = .)

summ(model, .robust = "HC3", digits=3, ci_level=0.95,confint=TRUE)
```

```
## MODEL INFO:
## Observations: 1028
## Dependent Variable: inva
## Type: OLS linear regression
##
## MODEL FIT:
## F(3,1024) = 12.057, p = 0.000
## R² = 0.034
## Adj. R² = 0.031
##
## Standard errors: OLS
## -----------------------------------------------------------
##                        Est.     2.5%   97.5%   t val.       p
## ------------------ ------- -------- ------- -------- -------
## (Intercept)          0.101    0.093   0.109   24.344   0.000
## vala                 0.003    0.002   0.004    4.774   0.000
## cfa                  0.004   -0.009   0.018    0.665   0.506
## debta                0.012    0.002   0.023    2.278   0.023
## -----------------------------------------------------------
```

```
#### (c)

coefs <- names(coef(model))

linearHypothesis(model,coefs[c(-1,-2)])
```

```
## Linear hypothesis test
```

```
##
## Hypothesis:
## cfa = 0
## debta = 0
##
## Model 1: restricted model
## Model 2: inva ~ vala + cfa + debta
##
##   Res.Df    RSS Df Sum of Sq      F  Pr(>F)
## 1   1026 10.072
## 2   1024 10.016  2  0.056019 2.8636 0.05752 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#### (d)

model2 <- sub_1987 %>% mutate( vala2 = vala^2,
                      cfa2 = cfa^2,
                      debta2 = debta^2,
                      QC = vala * cfa,
                      QD = vala * debta,
                      CD = cfa * debta) %>%
          lm(inva ~ vala + cfa + debta + vala2 + cfa2 + debta2 + QC + QD + CD,data=.)


summ(model2, .robust= "HC3",digits = 3)
```

```
## MODEL INFO:
## Observations: 1028
## Dependent Variable: inva
## Type: OLS linear regression
##
## MODEL FIT:
## F(9,1018) = 15.092, p = 0.000
## R² = 0.118
## Adj. R² = 0.110
##
## Standard errors: OLS
## ------------------------------------------------------
##                          Est.    S.E.   t val.       p
## ----------------- --------- ------- -------- -------
## (Intercept)            0.070   0.006   11.598   0.000
## vala                   0.012   0.001    8.625   0.000
## cfa                    0.032   0.013    2.493   0.013
## debta                  0.036   0.010    3.633   0.000
## vala2                 -0.000   0.000   -6.872   0.000
## cfa2                   0.005   0.005    1.070   0.285
## debta2                -0.004   0.004   -1.090   0.276
## QC                    -0.001   0.001   -2.333   0.020
## QD                    -0.001   0.002   -0.630   0.529
## CD                     0.001   0.008    0.121   0.904
## ------------------------------------------------------
```

```
coefs2 <- names(coef(model2))
linearHypothesis(model2,coefs2[c(-1,-2,-3,-4)])
```

```
## Linear hypothesis test
##
## Hypothesis:
## vala2 = 0
## cfa2 = 0
## debta2 = 0
## QC = 0
## QD = 0
## CD = 0
##
## Model 1: restricted model
## Model 2: inva ~ vala + cfa + debta + vala2 + cfa2 + debta2 + QC + QD +
##     CD
##
##   Res.Df    RSS Df Sum of Sq      F    Pr(>F)
## 1   1024 10.016
## 2   1018  9.149  6   0.86692 16.077 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

R CODE

8.19

```
data <- read.csv("cps09mar.csv" ,header = TRUE,sep=",")
log_inc <- log(data$earnings/(data$hours*data$week))
edu <- data$education
exp_1 <- data$age - data$education- 6
exp_2 <- exp_1^2/100
married_1 <- ifelse(data$marital ==1, 1, 0)
married_2 <- ifelse(data$marital ==2, 1, 0)
married_3 <- ifelse(data$marital ==3, 1, 0)
widow <- ifelse(data$marital ==4 , 1, 0)
div <- ifelse(data$marital == 5, 1, 0)
sep <- ifelse(data$marital ==6, 1, 0)
subsample <- data$race ==1 & data$female ==0 & data$hisp == 1

## (a)
library(stargazer)
library(matlib)

Y = cbind(edu,exp_1,exp_2,married_1,married_2,married_3,widow,div,sep)
X = subset(Y,subsample)
y = log_inc[subsample]
n = dim(X)[1]
ones = rep(1,n)
k = 10

X_new = cbind(X,ones)
beta_ols = inv(t(X_new)%*%X_new)%*%t(X_new)%*%y
e_ols = y - X_new%*%beta_ols
Omega = matrix(0,nrow=k,ncol=k)
for (i in 1:n){
  Omega = X_new[i,]%*%t(X_new[i,])*(e_ols[i]^2) + Omega
 }
v_ols = inv(t(X_new)%*%X_new)%*%Omega%*%inv(t(X_new)%*%X_new)*n/(n-k)
se_ols = sqrt(diag(v_ols))

ols <- cbind(beta_ols,se_ols)
colnames(ols) = c("Estimates","se")
rownames(ols) = c("education","experience","experience 2","married 1","married 2","
    married 3","widow","divorced","sep","intercept")
stargazer(ols)
## b

R = cbind(c(0,0,0,1,0,0,-1,0,0,0),c(0,0,0,0,0,0,0,1,-1,0))
beta_cls = beta_ols - inv(t(X_new)%*%X_new)%*%R%*%inv(t(R)%*%inv(t(X_new)%*%X_new)%*%R)%*
    %t(R)%*%beta_ols
e = y - X_new%*%beta_cls
e2 = e^2
s_cls2 = sum(e2)/(n-k-2)
X_inv = inv(t(X_new)%*%X_new)
```

```
R_inv = inv(t(R)%*%X_inv%*%R)
var_cls = (X_inv - X_inv%*%R%*%R_inv%*%t(R)%*%X_inv)*s_cls2
se_cls = sqrt(diag(var_cls))

cls <- cbind(beta_cls,se_cls)
colnames(cls) = c("Estimates","se")
rownames(cls) = c("education","experience","experience 2","married 1","married 2","
    married 3","widow","divorced","sep","intercept")
stargazer(cls)
## b

## c

beta_emd = beta_ols - v_ols%*%R%*%inv(t(R)%*%v_ols%*%R)%*%t(R)%*%beta_ols
v_emd = v_ols - v_ols%*%R%*%inv(t(R)%*%v_ols%*%R)%*%t(R)%*%v_ols
se_emd = sqrt(diag(v_emd))

emd <- cbind(beta_emd,se_emd)
colnames(emd) = c("Estimates","se")
rownames(emd) = c("education","experience","experience 2","married 1","married 2","
    married 3","widow","divorced","sep","intercept")
stargazer(emd)

## d

## e

R2 = cbind(R,c(0,1,1,0,0,0,0,0,0,0))

beta_inequal = beta_ols - v_ols%*%R2%*%inv(t(R2)%*%v_ols%*%R2)%*%t(R2)%*%beta_ols
v_inequal = v_ols - v_ols%*%R2%*%inv(t(R2)%*%v_ols%*%R2)%*%t(R2)%*%v_ols
se_inequal = sqrt(diag(v_inequal))

inequality <- cbind(beta_inequal,se_inequal)
colnames(inequality) = c("Estimates","se")
rownames(inequality) = c("education","experience","experience 2","married 1","married 2",
    "married 3","widow","divorced","sep","intercept")
stargazer(inequality)
```

9.24

```
rm(list = ls())
library(pacman)
p_load(tidyverse,ggplot2,data.table,sandwich)
set.seed(872)
n = 50
b_true = 1
theta_true = exp(b_true)
B = 1000

beta <- rep(NA,B)
theta <- rep(NA,B)
```

```
t_beta <- rep(NA,B)
t_theta <- rep(NA,B)
se_beta <- rep(NA,B)

for (i in 1:B){
  x = runif(n, 0, 1)
  e = rnorm(n)
  alpha = runif(1, -100,100)
  y <- alpha + b_true*x + e
  model <- lm(y ~ x)
  beta[i] <- coef(model)[2]
  se_beta[i] <- sqrt( vcovHC(model, type = "HC2")[2,2] )
  theta[i] = exp(beta[i])
  t_beta[i] = (beta[i] - b_true )/se_beta[i]
  se_theta = se_beta[i] * exp(beta[i])
  t_theta[i] = (theta[i] - theta_true )/se_theta
}

bias_beta <- mean(beta) - b_true
bias_theta <- mean(theta) - theta_true
sum(t_beta > 1.645)/B
sum(t_theta > 1.645)/B

#### asymptotic

beta = 1
theta = exp(beta)
B = 1000
beta_est = rep(NA,B)
theta_est = rep(NA,B)
t_beta = rep(NA,B)
t_theta = rep(NA,B)
sample_seq = c(50,100,500,1000,3000,5000)
Bias_beta = rep(NA,length(sample_seq))
Bias_theta = rep(NA,length(sample_seq))
Prob_beta = rep(NA,length(sample_seq))
Prob_theta = rep(NA,length(sample_seq))
for (i in 1:length(sample_seq)){
  n = sample_seq[i]
  for (j in 1:B){
    x = runif(n, min = 0, max = 1)
    e = rnorm(n,mean = 0, sd = 1)
    ones = rep(1,n)
    alpha = runif(1,-100,100)
    y = alpha*ones + beta*x + e
    fit <- lm(y~x)
    beta_est[j]= fit$coefficients[2]
    se_beta = sqrt(vcovHC(fit,type="HC1")[2,2])
    t_beta[j] = (beta_est[j] - beta)/se_beta
    theta_est[j] = exp(beta_est[j])
    se_theta = se_beta*exp(beta_est[j])
    t_theta[j] = (theta_est[j] - theta)/se_theta
```

```
  }
  Bias_beta[i] = mean(beta_est) - beta
  Bias_theta[i] = mean(theta_est) - theta
  Prob_beta[i] = sum(t_beta > 1.645)/B
  Prob_theta[i] = sum(t_theta>1.645)/B
}

df <- data.frame(sample_seq,Bias_theta,Bias_theta)

df %>% ggplot(aes(x = sample_seq)) +
             geom_line(aes(y = Bias_beta, color = "darkred"), size=1) +
             geom_line(aes(y = Bias_theta, color = "steelblue"),size=1) +
             scale_color_discrete( name = "bias" , labels = c(expression(paste("Bias of " , beta,)),expression(paste("Bias of ", theta,))))+
       xlab("sample size")

df2 <- data.frame(sample_seq,Prob_theta, Prob_beta)

df2 %>% ggplot(aes(x = sample_seq)) +
            geom_line( aes( y = Prob_beta, color = "darkred"), size = 1) +
            geom_line( aes( y = Prob_theta, color= "steelblue"), size = 1) +
            scale_color_discrete( name = "Rejection Rate", labels = c(expression(paste("prob. of ","t_",,beta, " > 1.645")),
                                                                      expression(paste("prob. of ","t_",,theta, " > 1.645")) ))+
            xlab("sample size")
```

9.25

```
rm(list = ls())

library(pacman)
p_load(tidyverse,gtsummary,jtools,car,haven)

df <- read_stata("https://www.ssc.wisc.edu/~bhansen/econometrics/Invest1993.dta")

#### (a), (b)

sub_1987 <- df %>% filter(year == 1987)

model <- sub_1987 %>%
  lm(inva ~ vala + cfa + debta, data = .)

summ(model, .robust = "HC3", digits=3, ci_level=0.95,confint=TRUE)

#### (c)

coefs <- names(coef(model))

linearHypothesis(model,coefs[c(-1,-2)])

```

```

#### (d)

model2 <- sub_1987 %>% mutate( vala2 = vala^2,
                    cfa2 = cfa^2,
                    debta2 = debta^2,
                    QC = vala * cfa,
                    QD = vala * debta,
                    CD = cfa * debta) %>%
         lm(inva ~ vala + cfa + debta + vala2 + cfa2 + debta2 + QC + QD + CD,data=.)

summ(model2, .robust= "HC3",digits = 3)

coefs2 <- names(coef(model2))
linearHypothesis(model2,coefs2[c(-1,-2,-3,-4)])
```