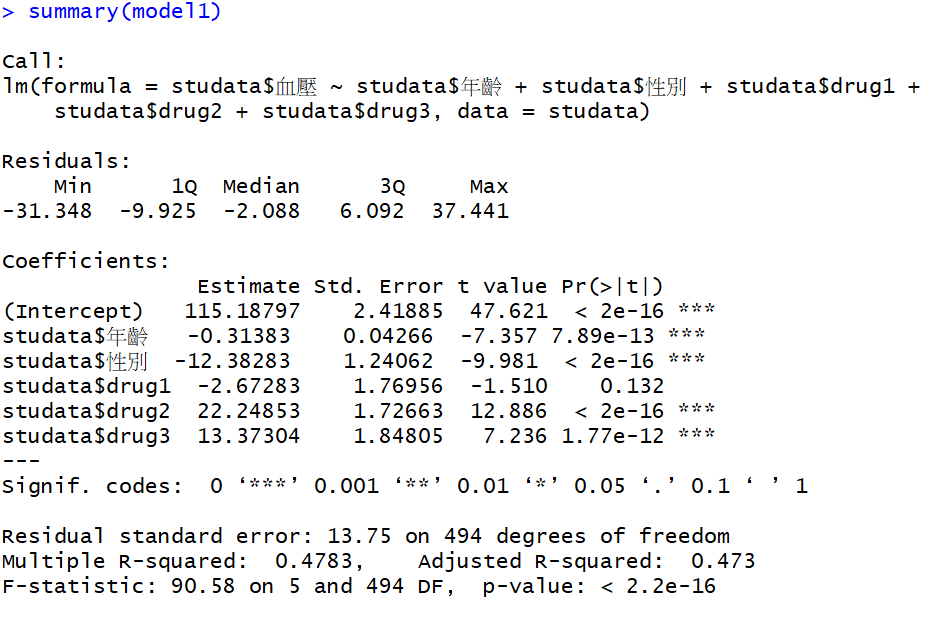
1.

生統二期中上機考

公衛二

梁嫚芳

B07801003



full model (model1)：Y = β0 + β1 X1 + β2 X2 + β3(1)X3(1) + β3(2)X3(2) + β3(3)X3(3) + ε, ε~𝑁(0, σ2)

Y：血壓

X1：年齡

X2：性別={1, 0}

|  |  |  |  |
| --- | --- | --- | --- |
| 劑量 | X3(1) | X3(2) | X3(3) |
| 1 | 1 | 0 | 0 |
| 2 | 0 | 1 | 0 |
| 3 | 0 | 0 | 1 |
| 4 | 0 | 0 | 0 |

β0: intercept in 血壓

β1: 年齡對血壓的斜率

β2: 性別對血壓的差異

β3(1) : 劑量1和劑量4對血壓的差異

β3(2) : 劑量2和劑量4對血壓的差異

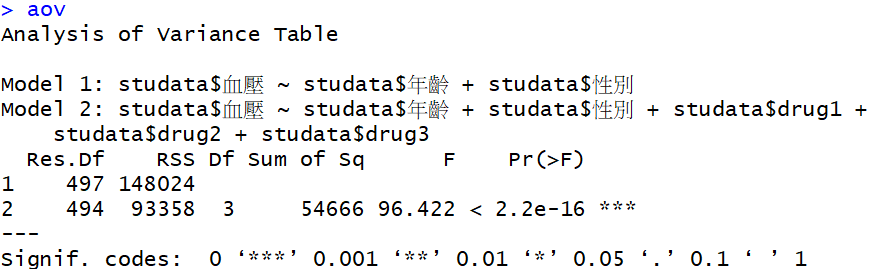
β3(3) : 劑量3和劑量4對血壓的差異

ε：模型的誤差

H0：β3(1)=β3(2)=β3(3)=0

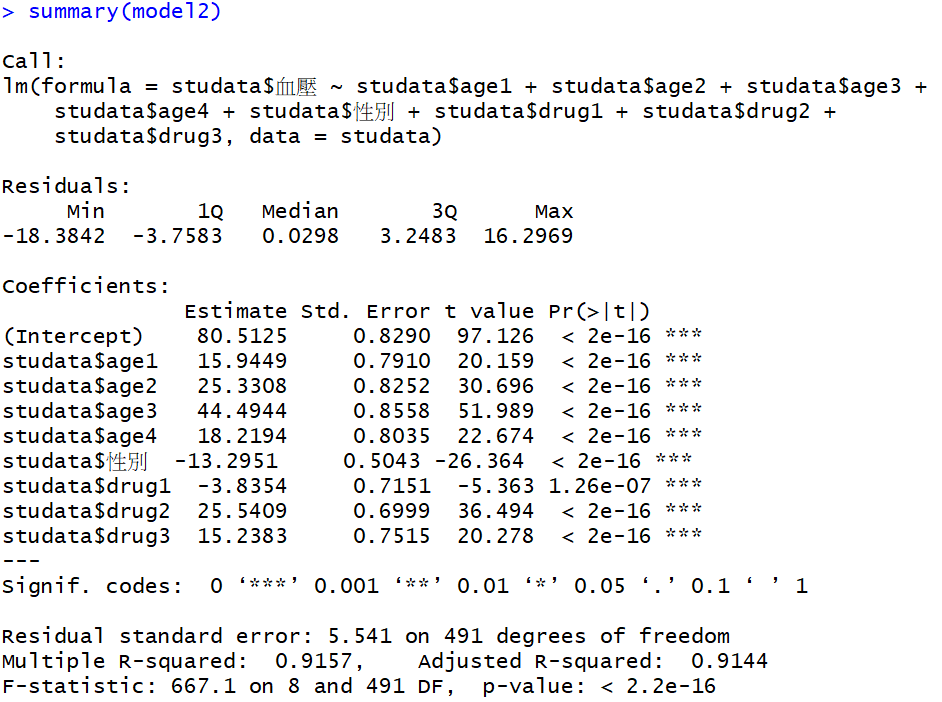
在α=5%下，進行Partial F test

Reduced model：Y=β0+β1X1+β2X2+ε , ε~𝑁(0, 𝜎2)



結論：P-value<2.2e-16 < α(0.05)，拒絕H0，則β3(1)、β3(2)、β3(3)不均為零，劑量對血壓造成統計顯著地影響。

2.



full model (model2)：Y = β0 + β1(1)X1(1) + β1(2)X1(2) + β1(3)X1(3) + β1(4)X1(4) + β2 X2 + β3(1)X3(1) + β3(2)X3(2) + β3(3)X3(3) + ε, ε~𝑁(0, σ2)

Y：血壓

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 年齡 | X1(1) | X1(2) | X1(3) | X1(4) |
| 20-29 | 1 | 0 | 0 | 0 |
| 30-39 | 0 | 1 | 0 | 0 |
| 40-49 | 0 | 0 | 1 | 0 |
| 50-59 | 0 | 0 | 0 | 1 |
| 60-69 | 0 | 0 | 0 | 0 |

X2：性別={1, 0}

|  |  |  |  |
| --- | --- | --- | --- |
| 劑量 | X3(1) | X3(2) | X3(3) |
| 1 | 1 | 0 | 0 |
| 2 | 0 | 1 | 0 |
| 3 | 0 | 0 | 1 |
| 4 | 0 | 0 | 0 |

β0: intercept in 血壓

β1(1) : 年齡20-29歲和年齡60-69歲對血壓的差異

β1(2) : 年齡30-39歲和年齡60-69歲對血壓的差異

β1(3) : 年齡40-49歲和年齡60-69歲對血壓的差異

β1(4) : 年齡50-59歲和年齡60-69歲對血壓的差異

β2: 性別對血壓的差異

β3(1) : 劑量1和劑量4對血壓的差異

β3(2) : 劑量2和劑量4對血壓的差異

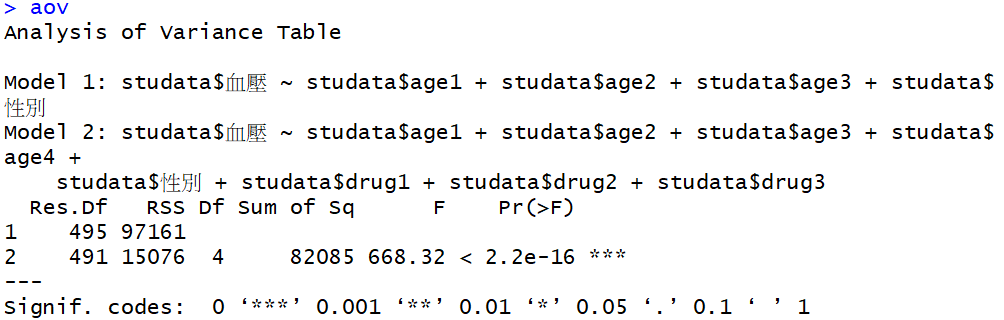
β3(3) : 劑量3和劑量4對血壓的差異

ε：模型的誤差

H0：β3(1)=β3(2)=β3(3)=0

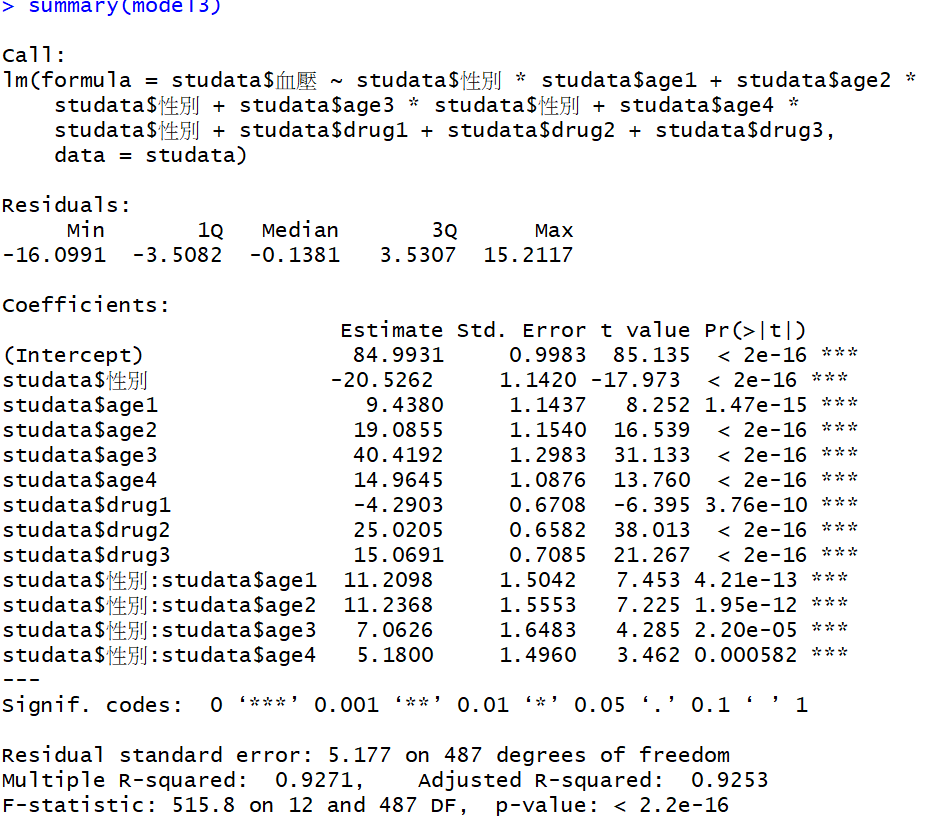
在α=5%下，進行Partial F test

Reduced model：Y=β0+ β1(1)X1(1) + β1(2)X1(2) + β1(3)X1(3) + β1(4)X1(4) +β2X2+ε , ε~𝑁(0, 𝜎2)



結論：P-value<2.2e-16 < α(0.05)，拒絕H0，則β3(1)、β3(2)、β3(3)不均為零，劑量對血壓造成統計顯著地影響。

3.



full model (model3)：Y = β0 + β1(1)X1(1) + β1(2)X1(2) + β1(3)X1(3) + β1(4)X1(4) + β2 X2 + β3(1)X3(1) + β3(2)X3(2) + β3(3)X3(3) + β4(1)(X1(1) X2) + β4(2) (X1(2) X2) + β4(3) (X1(3) X2) + β4(4) (X1(4) X2)+ ε, ε~𝑁(0, σ2)

Y：血壓

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 年齡 | X1(1) | X1(2) | X1(3) | X1(4) |
| 20-29 | 1 | 0 | 0 | 0 |
| 30-39 | 0 | 1 | 0 | 0 |
| 40-49 | 0 | 0 | 1 | 0 |
| 50-59 | 0 | 0 | 0 | 1 |
| 60-69 | 0 | 0 | 0 | 0 |

X2：性別={1, 0}

|  |  |  |  |
| --- | --- | --- | --- |
| 劑量 | X3(1) | X3(2) | X3(3) |
| 1 | 1 | 0 | 0 |
| 2 | 0 | 1 | 0 |
| 3 | 0 | 0 | 1 |
| 4 | 0 | 0 | 0 |

β0: intercept in 血壓

β1(1) : 年齡20-29歲和年齡60-69歲對血壓的差異

β1(2) : 年齡30-39歲和年齡60-69歲對血壓的差異

β1(3) : 年齡40-49歲和年齡60-69歲對血壓的差異

β1(4) : 年齡50-59歲和年齡60-69歲對血壓的差異

β2: 性別對血壓的差異

β3(1) : 劑量1和劑量4對血壓差異

β3(2) : 劑量2和劑量4對血壓差異

β3(3) : 劑量3和劑量4對血壓差異

β4(1) : 年齡20-29歲和性別交互作用後對血壓的變化

β4(2) : 年齡30-39歲和性別交互作用後對血壓的變化

β4(3) : 年齡40-49歲和性別交互作用後對血壓的變化

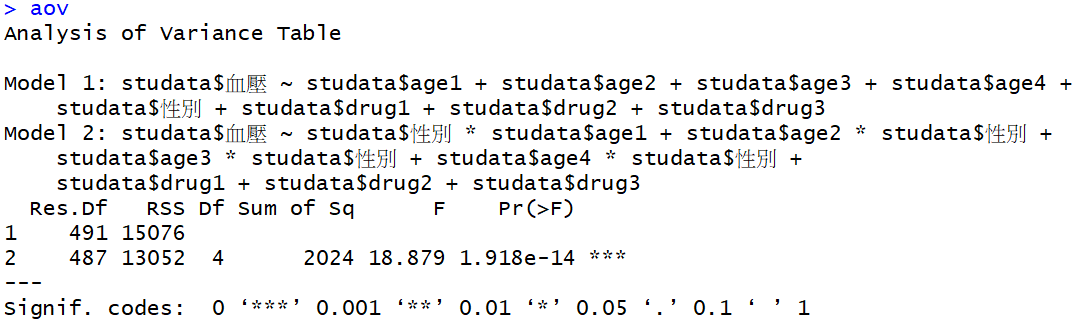
β4(4) : 年齡50-59歲和性別交互作用後對血壓的變化

ε：模型的誤差

H0：β4(1)=β4(2)=β4(3)=β4(4)=0

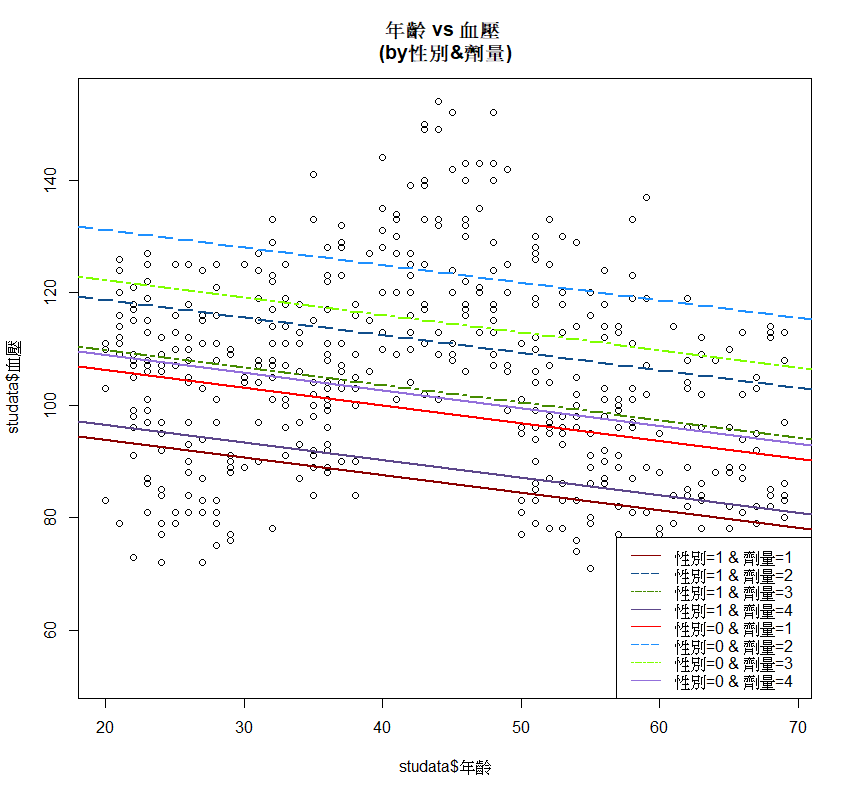
在α=5%下，進行Partial F test

Reduced model(model2)：Y=β0 + β1(1)X1(1) + β1(2)X1(2) + β1(3)X1(3) + β1(4)X1(4) + β2 X2 + β3(1)X3(1) + β3(2)X3(2) + β3(3)X3(3)+ε , ε~𝑁(0, 𝜎2)

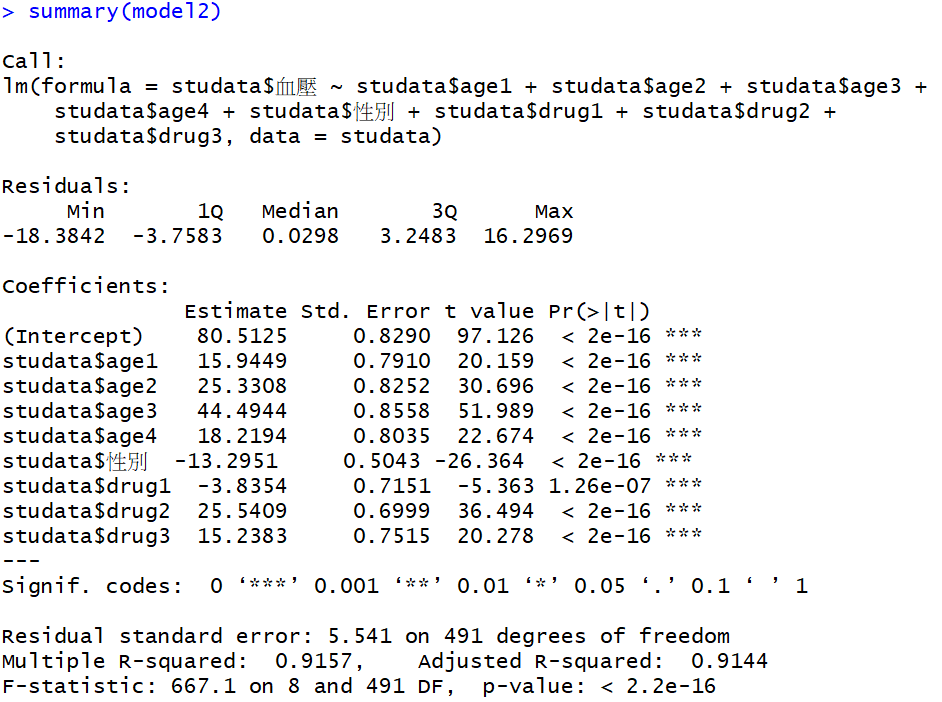


結論：P-value=1.918e-14 < α(0.05)，拒絕H0，則β4(1)、β4(2)、β4(3)、β4(4)不均為零，年齡和性別交互作用後對血壓有統計顯著地變化。

4.



5.

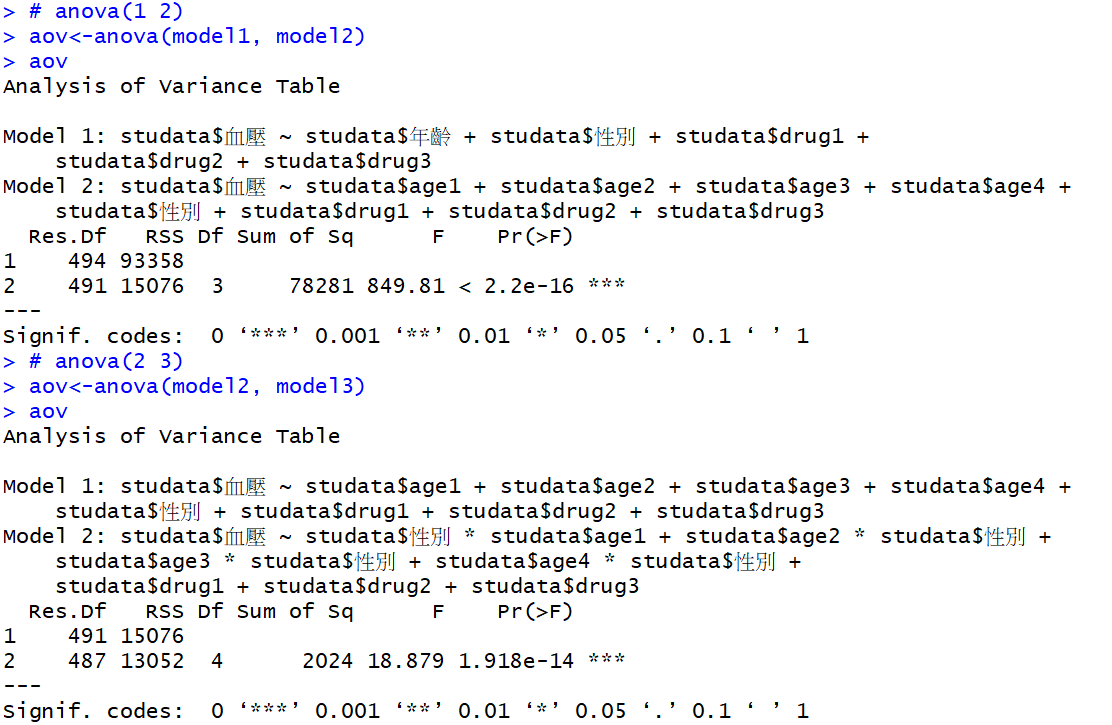


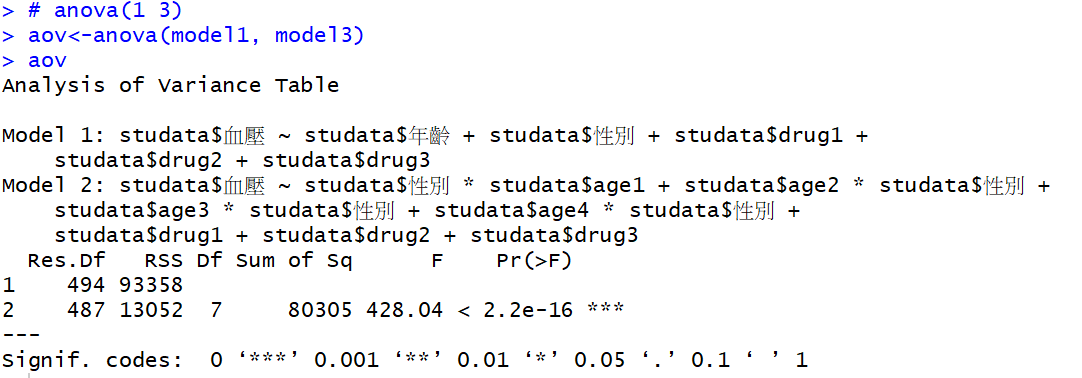
依照model2，β3(3)為劑量3(mediate)和劑量4(high)對血壓的差異

H0：β3(3)=0

p-value<2e-16，拒絕H0，β3(3)不為0，劑量3(mediate)和劑量4(high)對血壓有統計顯著的差異。

6.





model1：Y = β0 + β1 X1 + β2 X2 + β3(1)X3(1) + β3(2)X3(2) + β3(3)X3(3) + ε, ε~𝑁(0, σ2)

model2：Y = β0 + β1(1)X1(1) + β1(2)X1(2) + β1(3)X1(3) + β1(4)X1(4) + β2 X2 + β3(1)X3(1) + β3(2)X3(2) + β3(3)X3(3) + ε, ε~𝑁(0, σ2)

model3：Y = β0 + β1(1)X1(1) + β1(2)X1(2) + β1(3)X1(3) + β1(4)X1(4) + β2 X2 + β3(1)X3(1) + β3(2)X3(2) + β3(3)X3(3) + β4(1)(X1(1) X2) + β4(2) (X1(2) X2) + β4(3) (X1(3) X2) + β4(4) (X1(4) X2)+ ε, ε~𝑁(0, σ2)

分別將model1與model2、model2與model3、model1與model3在α=0.05下進行Partial F-test

model1(red) vs model2(full)

H0：β1(1)= β1(2)=β1(3)= β1(4)=0

p-value<2.2e-16 <α(0.05)，拒絕H0，β1(1)、β1(2)、β1(3)、β1(4)不均為零，由此推斷β1(1)、β1(2)、β1(3)、β1(4)對Y有一定影響力，model2較model1適合。

model2(red) vs model3(full)

H0：β4(1)= β4(2)=β4(3)= β4(4)=0

p-value=1.918e-14 <α(0.05)，拒絕H0，β4(1)、β4(2)、β4(3)、β4(4)不均為零，由此推斷β4(1)、β4(2)、β4(3)、β4(4)對Y有一定影響力，model3較model2適合。

model1(red) vs model3(full)

H0：β1(1)= β1(2)=β1(3)= β1(4)=β4(1)= β4(2)=β4(3)= β4(4)=0

p-value<2.2e-16 <α(0.05)，拒絕H0，β4(1)、β4(2)、β4(3)、β4(4)不均為零，由此推斷β4(1)、β4(2)、β4(3)、β4(4)對Y有一定影響力，model3較model1適合。

綜合以上三次檢定，model3最為適合，年齡(類別變項)和性別交互作用後對血壓有統計顯著地變化。

Code:

# data

studata<-read.csv(file.choose())

names(studata)

##1

# 建立model1

studata$drug1<-ifelse(studata$劑量==1,1,0)

studata$drug2<-ifelse(studata$劑量==2,1,0)

studata$drug3<-ifelse(studata$劑量==3,1,0)

model1<-lm(studata$血壓~ studata$年齡+studata$性別+studata$drug1+studata$drug2+studata$drug3, data=studata)

# 報表

summary(model1)

# Partial F test

## Reduced:

model\_red <- lm(studata$血壓~ studata$年齡+studata$性別,data=studata)

# anova

aov<-anova(model\_red, model1)

aov

##2

studata$age1<-ifelse(20<=studata$年齡 & studata$年齡<=29,1,0)

studata$age2<-ifelse(30<=studata$年齡 & studata$年齡<=39,1,0)

studata$age3<-ifelse(40<=studata$年齡 & studata$年齡<=49,1,0)

studata$age4<-ifelse(50<=studata$年齡 & studata$年齡<=59,1,0)

# 建立model2

model2<-lm(studata$血壓~studata$age1+studata$age2+studata$age3+studata$age4+studata$性別+studata$drug1+studata$drug2+studata$drug3, data=studata)

# 報表

summary(model2)

# Partial F test

## Reduced:

model\_red2 <- lm(studata$血壓~ studata$age1+studata$age2+studata$age3+studata$性別,data=studata)

# anova

aov<-anova(model\_red2, model2)

aov

##3

# 建立model3

model3<-lm(studata$血壓~ studata$性別\*studata$age1+studata$age2\*studata$性別+studata$age3\*studata$性別+studata$age4\*studata$性別+studata$drug1+studata$drug2+studata$drug3, data=studata)

# 報表

summary(model3)

# Partial F test

# anova

aov<-anova(model2, model3)

aov

##4

# Add mean response to scatter plot

plot(studata$年齡, studata$血壓, main="年齡 vs 血壓 \n(by性別&劑量)")

# left intercept

# right slope

# sex=(1)

# X3(1)

abline(model1$coefficients[1]+model1$coefficients[3]+model1$coefficients[4],model1$coefficients[2], col="red4", lty=1, lwd=2)

# X3(2)

abline(model1$coefficients[1]+model1$coefficients[3]+model1$coefficients[5],model1$coefficients[2], col="dodgerblue4", lty=5, lwd=2)

# X3(3)

abline(model1$coefficients[1]+model1$coefficients[3]+model1$coefficients[6],model1$coefficients[2], col="chartreuse4", lty=6, lwd=2)

# X3(4)

abline(model1$coefficients[1]+model1$coefficients[3],model1$coefficients[2], col="mediumpurple4", lty=7, lwd=2)

# sex=(0)

# X3(1)

abline(model1$coefficients[1]+model1$coefficients[4],model1$coefficients[2], col="red", lty=1, lwd=2)

# X3(2)

abline(model1$coefficients[1]+model1$coefficients[5],model1$coefficients[2], col="dodgerblue", lty=5, lwd=2)

# X3(3)

abline(model1$coefficients[1]+model1$coefficients[6],model1$coefficients[2], col="chartreuse", lty=6, lwd=2)

# X3(4)

abline(model1$coefficients[1],model1$coefficients[2], col="mediumpurple", lty=7, lwd=2)

# 加文字敘述於右下方

legend("bottomright", c("性別=1 & 劑量=1", "性別=1 & 劑量=2", "性別=1 & 劑量=3", "性別=1 & 劑量=4","性別=0 & 劑量=1", "性別=0 & 劑量=2", "性別=0 & 劑量=3", "性別=0 & 劑量=4"), col=c("red4", "dodgerblue4", "chartreuse4", "mediumpurple4","red", "dodgerblue", "chartreuse", "mediumpurple"), lty=c(1, 5, 6,7,1, 5, 6,7))

##5

summary(model2)

##6

# Partial F test

# anova(1 2)

aov<-anova(model1, model2)

aov

# anova(2 3)

aov<-anova(model2, model3)

aov

# anova(1 3)

aov<-anova(model1, model3)

aov