資料處理作業第十三週

統計二 108304010 趙啟宏

**EX\_A:**

library(arm)

library(car)

library(data.table)

library(foreign)

library(gmodels)

library(haven)

library(lm.beta)

library(huxtable)

##讀入資料##

mydata <- read\_sav('prestige.sav')

##調整受訪者教育年數##

mydata$edu <- recode(as.numeric(mydata$v13a),'1:2=0;3=6;4:5=9;6:9=12;10:11=14;12=15;13=14;14:16=16;17=18;18=22;else=NA')

##製作性別Dummy variable##

mydata$sex <- recode(as.numeric(mydata$v1),'1=1;2=0')

##製作交互作用項##

mydata$int <- mydata$edu\*mydata$sex

##製作線性模型##

outputa1 <- lm(mydata$v32b\_pre~mydata$edu+mydata$sex)

summary(lm.beta(outputa1))

outputa2 <- lm(mydata$v32b\_pre~mydata$edu+mydata$sex+mydata$int)

summary(lm.beta(outputa2))

##製作係數表格##

huxreg(outputa1,outputa2)

───────────────────────────────────────────────

(1) (2)

─────────────────────────────────

(Intercept) 39.899 \*\*\* 40.221 \*\*\*

(0.358) (0.432)

mydata$edu 1.062 \*\*\* 1.028 \*\*\*

(0.031) (0.040)

mydata$sex -1.311 \*\*\* -2.179 \*\*

(0.296) (0.715)

mydata$int 0.084

(0.063)

─────────────────────────────────

N 1951 1951

R2 0.376 0.377

logLik -6408.008 -6407.119

AIC 12824.016 12824.238

───────────────────────────────────────────────

##畫出座標圖##

a1\_m <- 39.899+mean(mydata$v26b\_pre,na.rm=T)-1.31

a1\_f <- 39.899+mean(mydata$v26b\_pre,na.rm=T)

ex\_a1 <- data.frame(初職聲望=round(c(a1\_m,a1\_f),3),性別=c('男','女'))

a2\_m <- 40.221+mean(mydata$v26b\_pre,na.rm=T)-2.179+0.084

a2\_f <- 40.221+mean(mydata$v26b\_pre,na.rm=T)

ex\_a2 <- data.frame(初職聲望=round(c(a2\_m,a2\_f),3),性別=c('男','女'))

ggplot() +

geom\_point(aes(x=ex\_a1$性別,y=ex\_a1$初職聲望,shape=ex\_a1$性別))+

geom\_line(aes(x=ex\_a1$性別,y=ex\_a1$初職聲望,group=1))+

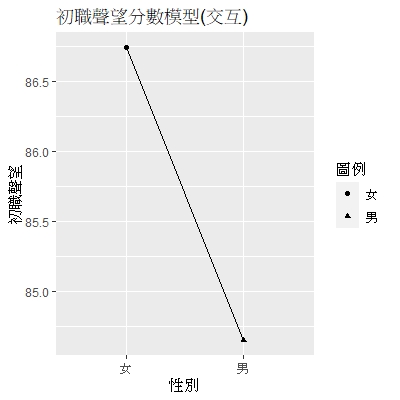
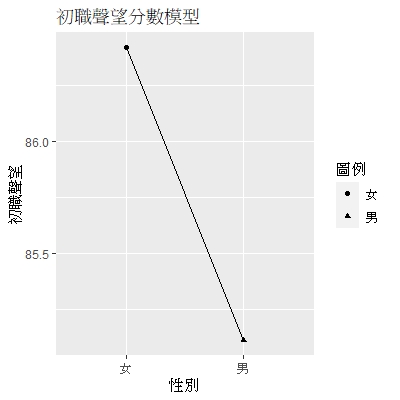
labs(title='初職聲望分數模型',x='性別',y='初職聲望',shape='圖例')

ggplot() +

geom\_point(aes(x=ex\_a2$性別,y=ex\_a2$初職聲望,shape=ex\_a2$性別))+

geom\_line(aes(x=ex\_a2$性別,y=ex\_a2$初職聲望,group=1))+

labs(title='初職聲望分數模型(交互)',x='性別',y='初職聲望',shape='圖例')



**結論:根據模型我們可以發現，在控制教育年數的情況下，男性的初職聲望比起女性低了大約1.31單位。而再加入交互作用項後，一樣在控制教育年數的情況下，男性的初職聲望比起女性低了大約2.179單位。**

**EX\_B(.1.):**

##EX\_B##

tscs971 <- read\_sav('tscs971\_l.sav')

tscs071 <- read\_sav('tscs071.sav')

tscs171 <- read\_sav('tscs171.sav')

##調整年紀&性別##

tscs171$age <- 106-tscs171$a2y+1

tscs971$sex <- factor(tscs971$v1,levels = c(1,2),labels = c('male','female'))

tscs071$sex <- factor(tscs071$a1,levels = c(1,2),labels = c('male','female'))

tscs171$sex <- factor(tscs171$a1,levels = c(1,2),labels = c('male','female'))

##調整教育年數##

tscs971$edu <- recode(as.numeric(tscs971$v13a),'1:2=0;3=6;4:5=9;6:9=12;10:11=14;12=15;13=14;14:16=16;17=18;18=22;else=NA')

tscs071$edu <- recode(as.numeric(tscs071$c1),'1:2=0;3=6;4:5=9;6:9=12;10=14;13=14;14:16=15;17:19=16;20=18;21=22;else=NA')

tscs171$edu <- recode(as.numeric(tscs171$b1),'1:2=0;3=6;4:5=9;6:9=12;10=14;13=14;14:16=15;17:19=16;20=18;21=22;else=NA')

##調整籍貫##

tscs971$home <- recode(tscs971$v5,'1=1;2=2;3=3;4=4;else=NA')

tscs071$home <- recode(tscs071$a5,'1=1;2=2;3=3;4=4;else=NA')

tscs171$home <- recode(tscs171$a6,'1=1;2=2;3=3;4:6=4;else=NA')

##調整主觀地位分數##

tscs971$score <- recode(tscs971$v86a,'97:98=NA')

tscs071$score <- recode(tscs071$f1,'95:98=NA')

tscs171$score <- recode(tscs171$e1,'97:98=NA')

##擷取成年人資料##

tscs971\_cut <- tscs971[which(tscs971$age>19 & tscs971$age<66 & tscs971$sex=='male'),]

tscs071\_cut <- tscs071[which(tscs071$age>19 & tscs071$age<66 & tscs071$sex=='male'),]

tscs171\_cut <- tscs171[which(tscs171$age>19 & tscs171$age<66 & tscs171$sex=='male'),]

##擷取地位分數+教育年數+籍貫+年度##

b1997 <- data.frame(score=as.numeric(tscs971\_cut$score),edu=as.numeric(tscs971\_cut$edu),home=as.numeric(tscs971\_cut$home),year=1997)

b2007 <- data.frame(score=as.numeric(tscs071\_cut$score),edu=as.numeric(tscs071\_cut$edu),home=as.numeric(tscs071\_cut$home),year=2007)

b2017 <- data.frame(score=as.numeric(tscs171\_cut$score),edu=as.numeric(tscs171\_cut$edu),home=as.numeric(tscs171\_cut$home),year=2017)

##合併資料##

ex\_b <- rbind(b1997,b2007,b2017)

##刪除含有遺漏值的樣本##

ex\_b <- ex\_b[complete.cases(ex\_b),]

##製作Dummy variables(1997,閩南人當對照組)##

ex\_b <- dummy\_columns(ex\_b,select\_columns =c('home','year'))

##建立回歸模型##

##地位分數vs教育年數##

outputb1 <- lm(ex\_b$score~ex\_b$edu)

summary(lm.beta(outputb1))

##地位分數vs教育年數+籍貫##

outputb2 <- lm(ex\_b$score~ex\_b$edu+ex\_b$home\_2+ex\_b$home\_3+ex\_b$home\_4)

summary(lm.beta(outputb2))

##地位分數vs教育年數+調查年度##

outputb3 <- lm(ex\_b$score~ex\_b$edu+ex\_b$year\_2007+ex\_b$year\_2017)

summary(lm.beta(outputb3))

##地位分數vs教育年數+籍貫(有互動項)##

ex\_b$int1 <- ex\_b$edu\*ex\_b$home\_2

ex\_b$int2 <- ex\_b$edu\*ex\_b$home\_3

ex\_b$int3 <- ex\_b$edu\*ex\_b$home\_4

outputb4 <- lm(ex\_b$score~ex\_b$edu+ex\_b$home\_2+ex\_b$home\_3+ex\_b$home\_4+ex\_b$int1+ex\_b$int2+ex\_b$int3)

summary(lm.beta(outputb4))

##地位分數vs教育年數+調查年度(有互動項)##

ex\_b$int4 <- ex\_b$edu\*ex\_b$year\_2007

ex\_b$int5 <- ex\_b$edu\*ex\_b$year\_2017

outputb5 <- lm(ex\_b$score~ex\_b$edu+ex\_b$year\_2007+ex\_b$year\_2017+ex\_b$int4+ex\_b$int5)

summary(lm.beta(outputb5))

##製作係數表格##

huxreg(outputb1,outputb2,outputb3,outputb4,outputb5)

##畫出座標圖##

##model2##

edu\_mean <- mean(ex\_b$edu)

M2a <- 4.31+0.08105\*edu\_mean

M2b <- 4.31+0.08105\*edu\_mean+0.13818

M2c <- 4.31+0.08105\*edu\_mean-0.14757

M2d <- 4.31+0.08105\*edu\_mean+0.32677

M2 <- data.frame(score=round(c(M2a,M2b,M2c,M2d),3),home=c('閩南','客家','原住民','大陸各省'))

ggplot() +

geom\_point(aes(x=M2$home,y=M2$score,shape=M2$home))+

geom\_line(aes(x=M2$home,y=M2$score,group=1))+

labs(title = '主觀社會地位分數vs籍貫',x='籍貫',y='地位分數',shape='圖例')

##model3##

M3a <- 4.53+0.05735\*edu\_mean

M3b <- 4.53+0.05735\*edu\_mean-0.60473

M3c <- 4.53+0.05735\*edu\_mean+0.85411

M3 <- data.frame(score=round(c(M3a,M3b,M3c),3),year=c('1997','2007','2017'))

ggplot() +

geom\_point(aes(x=M3$year,y=M3$score,shape=M3$year))+

geom\_line(aes(x=M3$year,y=M3$score,group=1))+

labs(title = '主觀社會地位分數vs調查年度',x='年度',y='地位分數',shape='圖例')

##model4##

M4a <- 4.33+0.07942\*edu\_mean

M4b <- 4.33-0.07942\*edu\_mean-0.27166

M4c <- 4.33+0.07942\*edu\_mean-0.9269

M4d <- 4.33+0.07942\*edu\_mean+2.10583

M4 <- data.frame(score=round(c(M4a,M4b,M4c,M4d),3),home=c('閩南','客家','原住民','大陸各省'))

ggplot() +

geom\_point(aes(x=M4$home,y=M4$score,shape=M4$home))+

geom\_point(aes(x=M2$home,y=M2$score,shape=M2$home))+

geom\_line(aes(x=M4$home,y=M4$score,group=1))+

geom\_line(aes(x=M2$home,y=M2$score,group=1))+

labs(title = '主觀社會地位分數vs籍貫(交互)',x='籍貫',y='地位分數',shape='圖例')

##model5##

M5a <- 3.24+0.17135\*edu\_mean

M5b <- 3.24+0.17135\*edu\_mean-0.43542

M5c <- 3.24+0.17135\*edu\_mean+5.48009

M5 <- data.frame(score=round(c(M5a,M5b,M5c),3),year=c('1997','2007','2017'))

ggplot() +

geom\_point(aes(x=M3$year,y=M3$score,shape=M3$year))+

geom\_line(aes(x=M3$year,y=M3$score,group=1))+

geom\_point(aes(x=M5$year,y=M5$score,shape=M5$year))+

geom\_line(aes(x=M5$year,y=M5$score,group=1))+

labs(title = '主觀社會地位分數vs調查年度(交互)',x='年度',y='地位分數',shape='圖例')

─────────────────────────────────────────────────────────────────────────────────────

(1) (2) (3) (4) (5)

───────────────────────────────────────────────────────────────────────

(Intercept) 4.323 \*\*\* 4.309 \*\*\* 4.526 \*\*\* 4.328 \*\*\* 3.244 \*\*\*

(0.128) (0.129) (0.125) (0.143) (0.176)

ex\_b$edu 0.082 \*\*\* 0.081 \*\*\* 0.057 \*\*\* 0.079 \*\*\* 0.171 \*\*\*

(0.010) (0.010) (0.010) (0.012) (0.015)

ex\_b$home\_2 0.138 -0.272

(0.112) (0.424)

ex\_b$home\_3 -0.148 -0.927

(0.141) (0.614)

ex\_b$home\_4 0.327 2.106 \*\*

(0.182) (0.718)

ex\_b$year\_2 -0.605 \*\*\* -0.435

007

(0.086) (0.285)

ex\_b$year\_2 0.854 \*\*\* 5.480 \*\*\*

017

(0.085) (0.301)

ex\_b$int1 0.033

(0.033)

ex\_b$int2 0.058

(0.045)

ex\_b$int3 -0.135 \*

(0.053)

ex\_b$int4 -0.022

(0.023)

ex\_b$int5 -0.365 \*\*\*

(0.023)

───────────────────────────────────────────────────────────────────────

N 2825 2825 2825 2825 2825

R2 0.022 0.024 0.100 0.028 0.183

logLik -5846.129 -5843.069 -5728.490 -5838.199 -5592.938

AIC 11698.257 11698.138 11466.980 11694.399 11199.876

─────────────────────────────────────────────────────────────────────────────────────

**結論:從模型一中我們可以得知，對於成年男生而言，在主觀社會地位分數為1~10分時，每增加一年教育年數，地位分數會增加約0.082分，若未受教育，則地位分數約為4.323分。**

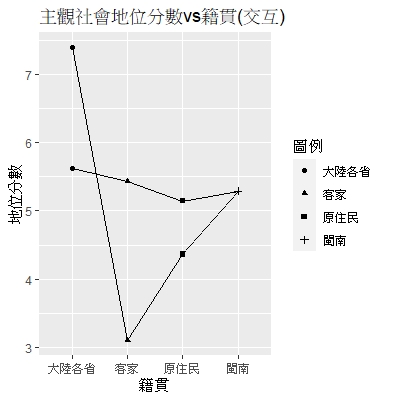
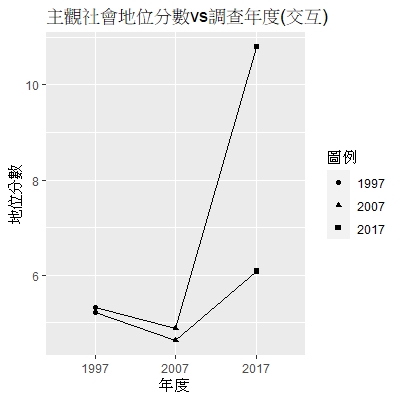
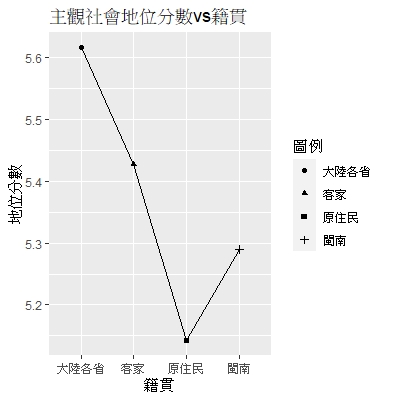
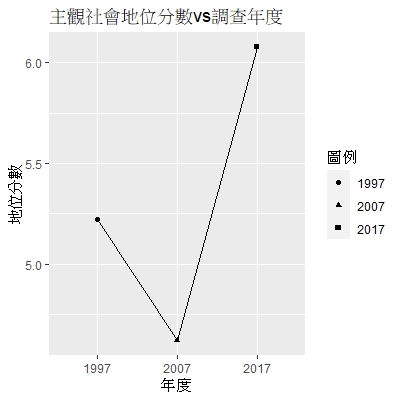
**從模型二中我們可以得知，對於成年男生而言，在主觀社會地位分數為1~10分且以台灣閩南人做為參考組時，在控制其他變數下，比較主觀社會地位分數。客家人比起閩南人高出約0.14分，原住民比起閩南人低了約0.147分，大陸外省人比起閩南人高出約0.326分。若控制籍貫，則每增加一年教育年數，地位分數會增加約0.08分。參考組的平均分數約為4.31分。**

**從模型三我們可以得知，對於成年男生而言，在主觀社會地位分數為1~10分且以1997年做為參考組時，在控制其他變數下，比較主觀社會地位分數。2007年比起1997年低了約0.605分，2017年比起1997年高出約0.854分。若控制調查年度，則每增加一年教育年數，地位分數會增加約0.057分。參考組的平均分數約為4.53分。**

**從模型四我們可以得知，對於成年男生而言，在主觀社會地位分數為1~10分且以台灣閩南人做為參考組時，在考慮交互作用以及控制其他變數下，比較主觀社會地位分數。客家人比起閩南人低了約0.272分，原住民比起閩南人低了約0.927分，大陸外省人比起閩南人高出約2.106分。若控制籍貫，則每增加一年教育年數，地位分數會增加約0.08分。參考組的平均分數約為4.33分。**

**從模型五可以得知，對於成年男生而言，在主觀社會地位分數為1~10分且以1997年做為參考組時，在考慮交互作用及控制其他變數下，比較主觀社會地位分數。2007年比起1997年低了約0.435分，2017年比起1997年高出約5.48分。若控制調查年度，則每增加一年教育年數，地位分數會增加約0.171分。參考組的平均分數約為3.24分。**

**EX\_B(.2.):**



**108304010@nccu.edu.tw**