1 H0:Iowa performs equally at home v.s. on road

p-value=0.0152, so reject it at the level of 0.10, means Iowa performs differently at home and on the road

**Codes**

scores = matrix(c(82,65,51,66,80,75,71,79,66,68,51,66,72,55,75,59,76,62,69,56,81,55,69,57,84,62,59,56,78,53,48,49),nrow=8,byrow=T)

d1= scores[,1]- scores[,3]

d2= scores[,2]- scores[,4]

d=cbind(d1,d2)

T.obs=sum((colMeans(d))^2)

M = 10000

T.perm = NULL

T.perm2 = NULL

for (i in 1:M) {

d.perm=d\*sample(c(-1,1),size=8,replace=T)

tstat=sum((colMeans(d.perm))^2)

T.perm = c(T.perm,tstat)}

p.val = mean(T.perm >= T.obs)

p.val

[1] 0.0152

2 a)

[,1] [,2]

[1,] 6.750 24.5

[2,] -2.125 4.0

b) [,1] [,2]

d1 6.802135 24.697865

d2 -2.243339 3.993339

**Codes**

d1=colMeans(d)

B = 10000

d.boot = matrix(0,nrow=B,ncol=2)

for (i in 1:B) {

bd= d[sample(1:8,replace=T),]

d.boot[i,] = colMeans(bd)}

alpha = 0.05

# Percentile-based CR

cbind(apply(d.boot,2,quantile,alpha/4),apply(d.boot,2,quantile,1-alpha/4))

# SE-based CR

boot.var = var(d.boot)

cbind(d1-qnorm(1-alpha/4)\*sqrt(diag(boot.var)),

d1+qnorm(1-alpha/4)\*sqrt(diag(boot.var)))

3 a)

#weight

H0: μ(smoker)= μ(non-smoker)

Response 1 :

Df Sum Sq Mean Sq F value Pr(>F)

smoke 1 20444 20444.3 46.603 1.038e-11 \*\*\*

Residuals 3152 1382739 438.7

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

#height

H0: μ(smoker)= μ(non-smoker)

Response 2 :

Df Sum Sq Mean Sq F value Pr(>F)

smoke 1 0.3 0.2552 0.0399 0.8417

Residuals 3152 20160.9 6.3962

#SBP

H0: μ(smoker)= μ(non-smoker)

Response 3 :

Df Sum Sq Mean Sq F value Pr(>F)

smoke 1 5 4.959 0.0217 0.8829

Residuals 3152 720600 228.617

b) It is not valid because the minimum of the p-value is random and does not have uniform distribution. Could use bonferroni method to adjust α

c)

H0: (μweight for smoker= μweight for non-smoker, μheight for smoker= μheight for non-smoker, μSBP for smoker= μSBP for non-smoker)

Use F-statistics:

Df Pillai approx F num Df den Df Pr(>F)

smoke 1 0.021753 23.349 3 3150 6.06e-15 \*\*\*

Residuals 3152

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

P-value is 6.06e-15, so reject the hypothesis at the level of 0.01

**Codes**

wcgs2=read.table("wcgs2.dat",header=T)

wcgs=wcgs2[,1:5]

wcgs=as.matrix(wcgs)

smoke=wcgs[,5]

Y=cbind(wcgs[,3], wcgs[,2], wcgs[,4])

fit=manova(Y~smoke)

summary.aov(fit)

summary(fit)