**UNI: Jh3561**

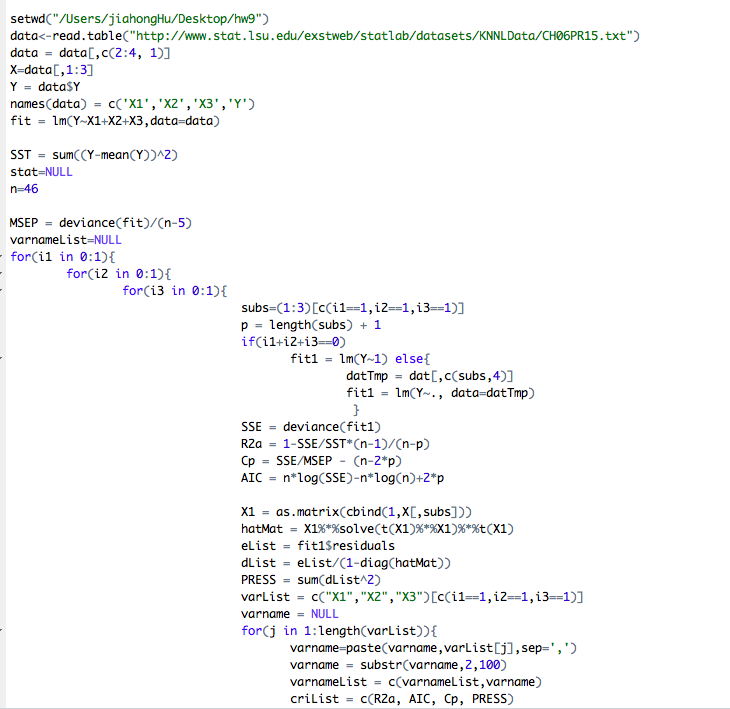
**Name: Jiahong Hu**

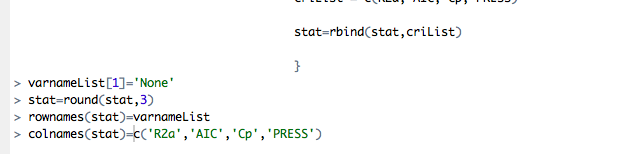
**HW9**

**Stat 4315**

9.9

Part a

****

****

The result is shown as follows:

R2a AIC Cp PRESS

None 0.000 262.915 85.010 13970.098

X3 0.402 240.214 33.406 8451.432

X2 0.349 244.131 40.110 9254.489

X2,X3 0.444 237.845 28.575 8115.912

X1 0.610 220.529 7.155 5569.562

X1,X3 0.661 215.061 1.788 4902.751

X1,X2 0.639 217.968 4.514 5235.192

X1,X2,X3 0.659 216.185 3.000 5057.886

All the four criteria support that we should use the model incorporating only X1 and X3.

Part b

Yes, the four criteria identify the same best subset in this case but this doesn’t always happen, since the criteria are different and evaluate model by different approach and measurements.

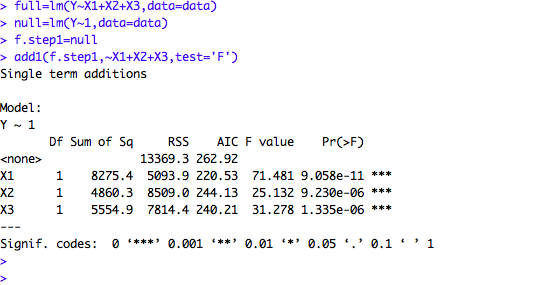
Part c

No. Forward stepwise regression does have advantages over the all-possible-regressions procedures when the number of subsets n is large enough. If 2n is much larger than n\*(n+1)/2, it will take much more time to build models by all-possible-regressions procedure. In this problem, the forward stepwise regression has no advantages over the all-possible-regressions procedure because the number of subset is small. In this problem, the number of subsets is 3, so the all-possible-regressions procedure needs to build 23=8 models, and the forward stepwise regression procedure needs 6 models. Hence, both models have the similar effect in determining the best model for predicting Y.

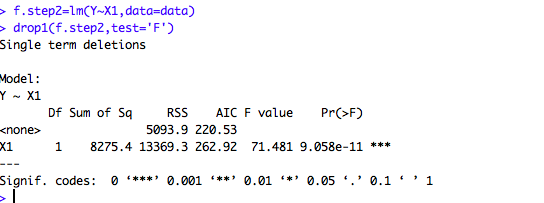
9.17

Part a

Step 1:

we can find that x1 has the largest F value, so x1 should be added in this model, namely, Y=X1

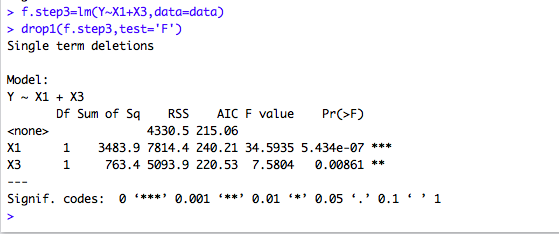
Step 2:



The result shows that x1 should not be dropped, and x3 should be added in this model, namely

Y=X1 +X3.

Step 3:



The result shows that x1 and x3 should not be eliminated from the model, and the F value of x2 is too small, so x2 should not be added in this model.

In conclusion, we can arrive at the conclusion that x1 and x3 are the best subset of variables that best for predicting Y (patient satisfaction) by forward stepwise regression procedure.

Y=X1+X3.

Part b

The level of significance in any individual test is the F limit of 3.0 for adding a

Variable here is approximately α =1-pF=0.0903.

Part c

Step1:

full=lm(y~x1+x2+x3,data=d1)

null=lm(y~1,data=d1)

f.step1=null

add1(f.step1,~x1+x2+x3,test='F')

Single term additions

Model:

y ~ 1

Df Sum of Sq RSS AIC F value Pr(>F)

<none> 13369.3 262.92

x1 1 8275.4 5093.9 220.53 71.481 9.058e-11 \*\*\*

x2 1 4860.3 8509.0 244.13 25.132 9.230e-06 \*\*\*

x3 1 5554.9 7814.4 240.21 31.278 1.335e-06 \*\*\*

---

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

X1 should be added in this model, namely, Y=X1

Step 2:

f.step2=lm(y~x1)

drop1(f.step2,test='F')

Single term deletions

Model:

y ~ x1

Df Sum of Sq RSS AIC F value Pr(>F)

<none> 5093.9 220.53

x1 1 8275.4 13369.3 262.92 71.481 9.058e-11 \*\*\*

---

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

Step 3:

f.step3=lm(y~x1+x3)

drop1(f.step3,test='F')

Single term deletions

Model:

y ~ x1 + x3

Df Sum of Sq RSS AIC F value Pr(>F)

<none> 4330.5 215.06

x1 1 3483.9 7814.4 240.21 34.5935 5.434e-07 \*\*\*

x3 1 763.4 5093.9 220.53 7.5804 0.00861 \*\*

---

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

The result is the same with part (a), and x1 and x3 is the best subset of variables that is selected as best by forward selection procedure. The steps in part (b) are similar with steps in part (a), the only difference is that we do not need to check whether we should delete a variable at each step in part c .

Part d.

**Step 1**

f.step1=full

drop1(f.step1,test='F')

Single term deletions

Model:

y ~ x1 + x2 + x3

Df Sum of Sq RSS AIC F value Pr(>F)

<none> 4248.8 216.19

x1 1 2857.55 7106.4 237.84 28.2471 3.81e-06 \*\*\*

x2 1 81.66 4330.5 215.06 0.8072 0.37407

x3 1 364.16 4613.0 217.97 3.5997 0.06468 .

---

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

pf(2.9,1,44)

[1] 0.9043653

We should delete x2 and the model is with the variables of x1 and x3.

**Step 2**

f.step2=lm(y~x1+x3)

drop1(f.step2,test='F')

Single term deletions

Model:

y ~ x1 + x3

Df Sum of Sq RSS AIC F value Pr(>F)

<none> 4330.5 215.06

x1 1 3483.9 7814.4 240.21 34.5935 5.434e-07 \*\*\*

x3 1 763.4 5093.9 220.53 7.5804 0.00861 \*\*

---

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

We get the conclusion that using the backward elimination procedure with an F limit of 2.9 to delete a variable, our best subset is x1 and x3.

Part e.

The results of three selection procedures are consistent. They are also consistent with those derived in Problem 9.9.