1、选择变量

null=glm(cbind(whole.satisfy<2,whole.satisfy>=2)~1,family=binomial,data=family.oamt)

full=glm(cbind(whole.satisfy<2,whole.satisfy>=2)~fam.num+fam.str+fam.city+income+money.charged+dinner.together+dinner.importance+

house.area+house.cost+house.type+wine+drink+fruit+dessert+other.food+midnight.snack,family=binomial,data=family.oamt)

step(full, scope=list(lower=null, upper=full), direction="backward")

采用backward选择模型，得到三个变量（酒水、共进晚餐次数、晚餐重要性）： whole.satisfy ~ dinner.together + dinner.importance + wine

summary(glm(formula = cbind(whole.satisfy < 2, whole.satisfy >= 2) ~

dinner.together + dinner.importance + wine, family = binomial,

data = family.oamt))

Call:

glm(formula = cbind(whole.satisfy < 2, whole.satisfy >= 2) ~

dinner.together + dinner.importance + wine, family = binomial,

data = family.oamt)

Deviance Residuals:

Min 1Q Median 3Q Max

-1.4534 -0.9283 -0.6017 1.1980 2.0290

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) 1.4528 0.9223 1.575 0.1152

dinner.together -0.2257 0.1251 -1.805 0.0711 .

dinner.importance -0.5983 0.3583 -1.670 0.0949 .

wine1 -0.9399 0.5707 -1.647 0.0996 .

---

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

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 122.21 on 95 degrees of freedom

Residual deviance: 113.93 on 92 degrees of freedom

AIC: 121.93

Number of Fisher Scoring iterations: 4

2、选择模型

Y<-wholesat

A<-wine

B<-dinner.importance

C<-dinner.together

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Residual/df | AIC | Null/df |
| A\*B\*C | 108.19/88 | 124.19 | 122.21/95 |
| A\*B+A\*C+B\*C | 113.41/89 | 127.41 |
| A\*B+A\*C | 113.79/90 | 125.79 |
| A\*B+B\*C | 113.50/90 | 125.5 |
| A\*C+B\*C | 113.54/90 | 125.54 |
| B\*C | 113.58/91 | 123.58 |
| A\*C | 113.91/91 | 123.91 |
| A\*B | 113.83/91 | 123.83 |
| A+B+C | 113.93/92 | 121.93 |

最终选择模型为： glm(Y~A+B+C,family=binomial)