

```

1  ###第8回判別分析###
2  install.packages("MASS",dependencies = T)
3  dat <- read.csv("data/cars93.csv" ,
4                header = T ,
5                row.names = 1 )
6  head(dat , 3)
7  library(psych)
8  describe(dat[,c(5,7,8,12,13,14,17,25)]) #基本統計量
9  table(dat$Origin) #目的変数をカウント
10 library(MASS)
11 da.out <- lda(Origin~Price+ #lda()関数:判別分析
12             MPG.city+EngineSize+
13             RPM+Fuel.tank.capacity,
14             data = dat ,
15             CV = F) #CV = T:交差妥当性
16 da.out
17 names(da.out)
18 round((da.out$scaling),3) #判別係数
19 round((apply(da.out$means%*%da.out$scaling, #定数項
20             2 , mean)),3)
21 round((da.out$mean),1) #群ごとの重心
22 predict.out <- predict(da.out) #予測値
23 tab <- table(dat$Origin ,
24             predict.out$class)
25 tab #予測精度の確認
26 round(((tab[1,1]+
27             tab[2,2])/sum(tab)*100),2) #的中率の計算
28 plot(da.out , dimen = 1) #分布確認,判別区分点:0より大か小か
29 ##Cross Validation(交差妥当性)##
30 cv.out <- lda(Origin~Price+ #lda()関数:判別分析
31             MPG.city+EngineSize+
32             RPM+Fuel.tank.capacity,
33             data = dat , CV = T) #CV = T:交差妥当性
34 cv.t <- table(dat$Origin ,
35             cv.out$class)
36 cv.t #予測精度の確認
37 round(((cv.t[1,1]+
38             cv.t[2,2])/sum(cv.t)*100),2) #的中率の計算
39
40 ##標準化データによる判別分析
41 z <- scale(dat[,c(5,7,8,12,13,14,17,25)]) #scale()関数:標準化
42 z.dat <- data.frame(dat , z ) #元dataとzをデータフレーム化
43 names(z.dat) #項目名確認:zは変数名.1に
44 z.out <- lda(Origin~Price.1+ #lda()関数:判別分析
45             MPG.city.1+
46             EngineSize.1+RPM.1+
47             Fuel.tank.capacity.1,
48             data = z.dat)
49 z.out
50 round((z.out$scaling),3)
51 round((apply(z.out$means%*%z.out$scaling, #定数項
52             2,mean)),3)
53 z.out$means
54 p.out <- predict(z.out) #予測値
55 t <- table(z.dat$Origin ,
56             p.out$class)
57 t #予測精度の確認
58

```