

جواب ۱

فراخوانی داده‌ها

```
> library(MASS)
> library(foreign)
> football<-read.spss("C:/Users/12345/Desktop/football.sav",to.data.frame=TRUE)
> attach(football)
```

پیدا کردن ماتریس E

```
> result<-manova(cbind(WDIM,CIRCUM,FBEYE,EYEHD,EARHD,JAW)~group , data=football)
> football.summry<-summary(result)
> football.summry$SS
$group
      WDIM      CIRCUM      FBEYE      EYEHD      EARHD      JAW
WDIM    2.1660 -6.651900 -1.700500 -12.179000 -5.909000 -2.641000
CIRCUM  -6.6519 20.428335  5.222325  37.402350 18.146850  8.110650
FBEYE   -1.7005  5.222325  1.335042   9.561583  4.639083  2.073417
EYEHD  -12.1790 37.402350  9.561583 68.480167 33.225167 14.849833
EARHD   -5.9090 18.146850  4.639083 33.225167 16.120167  7.204833
JAW     -2.6410  8.110650  2.073417 14.849833  7.204833  3.220167

$Residuals
      WDIM      CIRCUM      FBEYE      EYEHD      EARHD      JAW
WDIM    37.2740 49.69690 13.64550   6.129000 10.379000 19.781000
CIRCUM  49.6969 293.98215 91.77290  94.472650 44.420261 45.728794
FBEYE   13.6455  91.77290 47.98276  12.712917 13.561194 14.128694
EYEHD    6.1290  94.47265 12.71292 182.108833 56.859833  7.242167
EARHD   10.3790  44.42026 13.56119  56.859833 65.373722  2.190722
JAW     19.7810  45.72879 14.12869   7.242167  2.190722 32.838056

> E<-football.summry$SS$Residuals
```

تقسیم ماتریس E بر $k(n-1)=3*29=87$ جداسازی درایه‌های قطر اصلی ماتریس E/87

```
> (e<-E/87)
      WDIM      CIRCUM      FBEYE      EYEHD      EARHD      JAW
WDIM    0.42843678 0.5712287 0.1568448 0.07044828 0.11929885 0.22736782
CIRCUM  0.57122874 3.3791052 1.0548609 1.08589253 0.51057771 0.52561833
FBEYE   0.15684483 1.0548609 0.5515260 0.14612548 0.15587580 0.16239879
EYEHD   0.07044828 1.0858925 0.1461255 2.09320498 0.65356130 0.08324330
EARHD   0.11929885 0.5105777 0.1558758 0.65356130 0.75142209 0.02518072
JAW     0.22736782 0.5256183 0.1623988 0.08324330 0.02518072 0.37744891

> (err<-diag(e))
      WDIM      CIRCUM      FBEYE      EYEHD      EARHD      JAW
0.4284368 3.3791052 0.5515260 2.0932050 0.7514221 0.3774489
```

پیدا کردن بردار ویژه اول (ضرایب غیر استاندارد شده)

```
> model <- lda(group~., data = football)
> a1<-model$scaling[,1]
> a1
```

	WDIM	CIRCUM	FBEYE	EYEHD	EARHD	JAW
	0.948423100	-0.003639865	-0.006439599	-0.647483088	-0.504360916	-0.828535064

استاندارد سازی ضرایب

```
> a.star1<-t(sqrt(err)) * t(a1)
> a.star1
```

	WDIM	CIRCUM	FBEYE	EYEHD	EARHD	JAW
[1,]	0.6207911	-0.0066909	-0.0047823	-0.9367729	-0.4372033	-0.509026

میبینیم که استاندارد سازی کردن ضرایب باعث میشود که راحت تر تفسیر کنیم و متغیرهای موثر را پیدا و تفکیک کنیم.

متغیرهای مهم ما :

JAW , EARHD , EYEHD , WDIM

متغیرهای غیر مهم :

CIRCUM , FBEYE

جواب ۲.

اجرای تابع ممیزی گام به گام برای داده‌های فوتبال

```
> library(klar)
> greedy.wilks(group~., data = football, niveau = 0.1)
Formula containing included variables:
```

```
group ~ EYEHD + WDIM + JAW + EARHD
<environment: 0x000000018a2cd18>
```

Values calculated in each step of the selection procedure:

	vars	Wilks.lambda	F.statistics.overall	p.value.overall	F.statistics.diff	p.value.diff
1	EYEHD	0.4278892	58.16183	9.181503e-17	58.161826	9.181503e-17
2	WDIM	0.4003001	24.96348	2.604034e-16	2.963612	5.686808e-02
3	JAW	0.3382838	20.38104	6.677444e-18	7.791368	7.765645e-04
4	EARHD	0.3071512	16.89162	2.888334e-18	4.257081	1.729878e-02

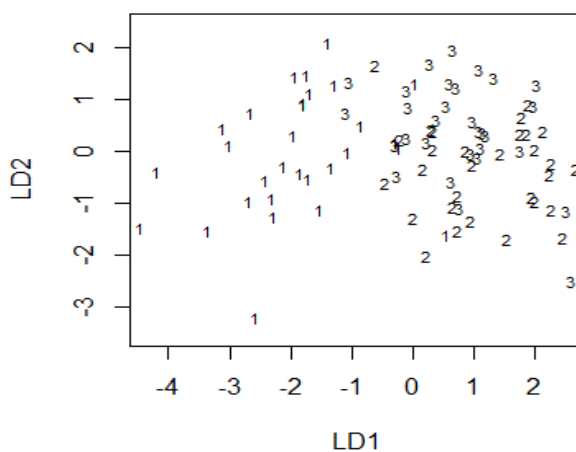
مشاهده میشود که تابع ممیزی گام به گام متغیرهای زیر را برای ما انتخاب میکند :

EYEHD , WDIM , JAW , EARHD

در ستون Wilks.lambda میبینیم با اضافه کردن هر متغیر به مدل مقدار Wilks.lambda کاهش میابد که یعنی مدل ما را رو به بهبودی رفته است .

رسم نمودار

```
> plot(model)
```



بنظر میرسد که LD1 بخوبی میتواند گروه شماره ۱ را تفکیک کند در واقع یعنی با چهار متغیر زیر به خوبی میتوان بازیکنان تیم فوتبال دبیرستان را از بقیه بازیکنان جدا کرد(تفکیک کرد).

EYEHD , WDIM , JAW , EARHD

جواب ۳.

فراخوانی داده‌ها

```
> Appel<-read.table("C:/Users/12345/Desktop/Apple.txt",header = T)
> head(Appel)
  group  y1    y2    y3    y4
1     1 1.11 2.569 3.58 0.760
2     1 1.19 2.928 3.75 0.821
3     1 1.09 2.865 3.93 0.928
4     1 1.25 3.844 3.94 1.009
5     1 1.11 3.027 3.60 0.766
6     1 1.08 2.336 3.51 0.726
```

اجرای تحلیل ممیزی

```
> #discriminant analysis
> library(MASS)
> modelAppel <- lda(group~., data = Appel)
> modelAppel
Call:
lda(group ~ ., data = Appel)

Prior probabilities of groups:
      1      2      3      4      5      6
0.1666667 0.1666667 0.1666667 0.1666667 0.1666667 0.1666667

Group means:
      y1      y2      y3      y4
1 1.13750 2.977125 3.73875 0.871125
2 1.15750 3.109125 4.51500 1.280500
3 1.10750 2.815250 4.45500 1.391375
4 1.09750 2.879750 3.90625 1.039000
5 1.08000 2.557250 4.31250 1.181000
6 1.03625 2.214625 3.59625 0.735000

Coefficients of linear discriminants:
      LD1      LD2      LD3      LD4
y1  3.0479952 -1.140083 -1.002448 23.419063
y2 -1.7025953 -1.215888  1.672714 -3.076804
y3  4.2332645  7.166403  3.045553 -2.011416
y4 -0.4785144 -11.520302 -5.506192  3.101660
```

```
Proportion of trace:
  LD1    LD2    LD3    LD4
0.6421 0.2707 0.0784 0.0089
```

در قسمت **Group means** میانگین هر متغیر را به تفکیک گروه نمایش داده است.

در قسمت **Coefficients of linear discriminants** میتوان بردارهای ویژه را مشاهده که **LD1** بردار ویژه اول و **LD2** بردار ویژه دوم ما است.

در قسمت **Proportion of trace** میزان تفکیک پذیری هر تابع ممیزی را نمایش میدهد که میبینم تابع ممیزی که به کمک ضرایب موجود در **LD1** تشکیل شده ۶۴٪ در تفکیک به ما کمک میکند و تابع ممیزی دوم ۲۷٪ در تفکیک به ما کمک میکند.

در تابع ممیزی اول متغیرهای **y1** , **y2** , **y3** سهم بیشتری در تفکیک پذیری دارند و متغیر **y4** زیاد متغیر مهمی نیست.

تابع ممیزی استاندارد شده

```
> result<-manova(cbind(y1,y2,y3,y4)~group , data=Appel)
> Appel.summry<-summary(result)
> Appel.summry$SS
$group
      y1      y2      y3      y4
y1 0.06407161 0.4623959 0.1599116 0.1139384
y2 0.46239591 3.3370472 1.1540599 0.8222773
y3 0.15991161 1.1540599 0.3991116 0.2843704
y4 0.11393836 0.8222773 0.2843704 0.2026163

$Residuals
      y1      y2      y3      y4
y1 0.3294763 1.771553 0.7264405 0.3116716
y2 1.7715530 13.005405 5.5649411 2.9250447
y3 0.7264405 5.564941 10.0056363 5.9783296
y4 0.3116716 2.925045 5.9783296 4.0129997

> E<-Appel.summry$SS$Residuals
>
> (e<-E/87)
      y1      y2      y3      y4
y1 0.003787084 0.02036268 0.008349891 0.003582433
y2 0.020362679 0.14948741 0.063964841 0.033621203
y3 0.008349891 0.06396484 0.115007314 0.068716433
y4 0.003582433 0.03362120 0.068716433 0.046126434
> (err<-diag(e))
      y1      y2      y3      y4
0.003787084 0.149487411 0.115007314 0.046126434
>
>
```

```
> modelAppel <- lda(group~., data = Appel)
> a1<-modelAppel$scaling[,1]
> a1
      y1      y2      y3      y4
3.0479952 -1.7025953 4.2332645 -0.4785144
>
> a.star1<-t(sqrt(err)) * t(a1)
> a.star1
      y1      y2      y3      y4
[1,] 0.1875715 -0.6582847 1.435615 -0.1027709
```

اکنون که ضرایب را استاندارد کردیم میبینیم که y_2 , y_3 نقش زیادی در تفکیک گروه‌ها دارند (بالاترین ضرایب استاندارد شده را دارند).

در اینجا y_4 و y_1 کمترین ضرایب را دارند پس میتوان از آنها چشم پوشی کرد .

تابع ممیزی گام به گام

```
> # step-wise discriminant analysis
> library(klaR)
> greedy.wilks(group~., data = Appel, niveau = 0.1)
Formula containing included variables:
```

```
group ~ y4 + y3 + y2
<environment: 0x00000002c1570d8>
```

Values calculated in each step of the selection procedure:

vars	Wilks.lambda	F.statistics.overall	p.value.overall	F.statistics.diff	p.value.diff
y4	0.4086057	12.157716	2.586901e-07	12.157716	2.586901e-07
y3	0.2655484	7.712631	1.428716e-08	4.417536	2.536579e-03
y2	0.1599170	6.964237	1.899753e-10	5.284314	7.761966e-04

در تابع ممیزی گام به گام ما به این نتیجه رسیدیم که متغیرهای y_2 , y_3 , y_4 مهم ترین متغیرهای ما در تفکیک گروه هستند.

رسم نمودار

```
> fitPredict=predict(modelAppel)
> LDscores=fitPredict$x
> LD1=LDscores[,1]
> plot(LD1,xlab="first linear discriminant",col=factor(group.A))
> legend("topleft", legend=levels(factor(group.A)),
+       text.col=seq_along(levels(factor(group.A))))
```

در نمودار زیر گروه‌های مختلف را بر اساس رنگ‌های مختلف مشاهده میکنید .

تفکیک پذیری که در مشاهده میکنید به کمک ضرایب **LD1** میباشد، که ضرایب استاندارد شده ما هستند.

مشاهده میشود که ما تفکیک پذیری به نسبت خوبی داریم .

