

# **Discriminant Analysis (DA)**

## **判别分析**

# 第五章 判别分析

## 1. 引言

## 2. 距离判别

(1) 马氏距离判别

(2) 两组距离判别函数

① 等协方差矩阵

a) 判别函数

b) 误判概率

c) 计算

② 协方差矩阵不相等

(3) 多组距离判别函数

## 3. Bayes判别

## 4. Fisher判别

# 1. 引言

- 基本概念：根据观测到的随机向量，判断该对象的所属类别(总体)。
- 判别函数： $Y=a_1x_1+a_2x_2+\dots+a_px_p$ ,

其中：

$Y$ 为判别分数(判别值)，

$x=(x_1\ x_2\dots x_p)'$ 为随机向量，

$a_1\ a_2\dots a_p$ 为判别方程系数。

- 三种判别方法

距离判别、 Bayes判别、 Fisher判别

## 2.距离判别

### (1) 马氏距离判别

计算马氏距离平方

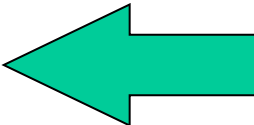
### (2) 两组距离判别函数

#### ① 等协方差矩阵

a) 判别函数

b) 误判概率

c) 计算



均数相等?  
协方差矩阵相等?

p142,例5.2.1

# 1<sup>st</sup>, p142, 例子5-2-3

\*DA, 《应用多元分析》 p142,例5\_2\_3\*.

New file.

Data list free/obs x1 to x6 class.

BEGIN DATA

1.00	3.60	4.30	82.30	70.00	90.00	18.52	1.00
------	------	------	-------	-------	-------	-------	------

^^^^^^^^^^^^^^

84.00	3.40	4.30	97.36	75.00	110.00	22.12	9.00
-------	------	------	-------	-------	--------	-------	------

End data.

Variable Lab Obs '运动员编号'/class '运动员级别' /x1 '30米跑' x2 '投掷小球'  
/x3 '挺举重量' x4 '抛实心球' x5 '前抛铅球' x6 '五级跳远'.

Value Lab class 1 '一级运动员' 2 '运动健将' 9 '未分组运动员'.

## Manova

x1 to x6 BY class(1 2)

/Print **CELLINFO**(all) **HOMOGENEITY** (all)

/Error **within+residual**

/Design .

# Univariate Homogeneity of Variance Tests

## 单变量

Variable .. x3

挺举重量

Cochrans C(26, 2) = .82986, P = .000 (approx.)

Bartlett-Box F(1, 7731) = 14.66531, P = .000

Variable .. x6

五级跳远

Cochrans C(26, 2) = .73884, P = .010 (approx.)

Bartlett-Box F(1, 7731) = 6.58583, P = .010

EFFECT .. class (Cont.)

Univariate F-tests with (1,51) D. F.

Variable	Hypoth. SS	Error SS	Hypoth. MS	Error MS	F	Sig. of F
x1	.08626	2.68427	.08626	.05263	1.63886	.206
x2	.50325	2.04782	.50325	.04015	12.53312	.001
x3	3588.88404	1695.56945	3588.88404	33.24646	107.94786	.000
x4	836.25067	4527.42857	836.25067	88.77311	9.42009	.003
x5	1801.36456	7024.10714	1801.36456	137.72759	13.07918	.001
x6	301.84348	89.10334	301.84348	1.74712	172.76588	.000

# 多变量

Multivariate test for Homogeneity of Dispersion matrices

Boxs M = 96.53204

F with (21, 9296) DF = 4.01074, P = .000 (Approx.)

Chi-Square with 21 DF = 84.44363, P = .000 (Approx.)

Multivariate Tests of Significance (S = 1, M = 2, N = 22)

Test Name	Value	Exact F Hypoth.	DF	Error DF	Sig. of F
Pillais	.84736	42.56122	6.00	46.00	.000
Hotellings	5.55146	42.56122	6.00	46.00	.000
Wilks	.15264	42.56122	6.00	46.00	.000
Rois	.84736				

判别函数：公式5.2.8

误判率计算：

训练样本计算误判率

新样本计算误判率

交叉验证 (cross-validation)

② 协方差矩阵不相等

判别方程 公式5.2.12

判别规则 公式5.2.13

**例5.2.2**



## 练习题

33. 《医学统计学（2ed）》 pp385,表18-1的距离判别和线性判别函数的两种判别结果？

## 32.大荔县19岁男青年有多少人符合仪仗兵 体格要求？

- 仪仗兵
- 总体均数向量
- (身高、体重)

- 175cm

- 66kg

- 协方差矩阵

- 2.9      1.8

- 1.8      10.0

大荔县19岁男青年

总体均数向量

(身高、体重)

168.34cm

57.20kg

协方差矩阵

27.1448      18.3603

18.3603      41.5441