5.5 根据经验，今天与昨天的湿度差x_1及今天的压温差(气压与温度之差)x_2是预报明天是否下雨的两个重要因素。今测得$x_1 = 0.6$，$x_2 = 3.0$，假定两组的协方差矩阵相等。

(1) 试给出判别规则，并预报明天是否会下雨及用回代法来估计误判概率；

$$\bar{\boldsymbol{x}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \boldsymbol{x}_{ij}$$

因为两组的协方差矩阵相等

$$\boldsymbol{S}_p = \frac{1}{n-k} \sum_{i=1}^{k} (n_i - 1) \boldsymbol{S}_i$$

$$n = n_1 + n_2 + \cdots + n_k$$

$$\boldsymbol{S}_i = \frac{1}{n_i - 1} \sum_{j=1}^{n_1} (\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)(\boldsymbol{x}_{ij} - \bar{\boldsymbol{x}}_i)'$$

因为

$$
\begin{aligned}
d^2(\boldsymbol{x}, \pi_i) &= (\boldsymbol{x} - \bar{\boldsymbol{x}}_i)' \boldsymbol{S}_p^{-1} (\boldsymbol{x} - \bar{\boldsymbol{x}}_i) \\
&= \boldsymbol{x}' \boldsymbol{S}_p^{-1} \boldsymbol{x} - 2\bar{\boldsymbol{x}}_i' \boldsymbol{S}_p^{-1} \boldsymbol{x} + \bar{\boldsymbol{x}}_i' \boldsymbol{S}_p^{-1} \bar{\boldsymbol{x}}_i \\
&= \boldsymbol{x}' \boldsymbol{S}_p^{-1} \boldsymbol{x} - 2(\hat{\boldsymbol{I}}_i' \boldsymbol{x} + \hat{c}_i)
\end{aligned}
$$

其中$\boldsymbol{I}_i = \boldsymbol{S}_p^{-1} \bar{\boldsymbol{x}}_i$，$\hat{c}_i = -\dfrac{1}{2} \bar{\boldsymbol{x}}_i' \boldsymbol{S}_p^{-1} \bar{\boldsymbol{x}}_i$

$$
\begin{cases}
\boldsymbol{x} \in \pi_1, & \boldsymbol{I}_1' \boldsymbol{x} + c_1 \geq \boldsymbol{I}_2' \boldsymbol{x} + c_2 \\
\boldsymbol{x} \in \pi_2, & \boldsymbol{I}_1' \boldsymbol{x} + c_1 < \boldsymbol{I}_2' \boldsymbol{x} + c_2
\end{cases}
$$

Linear Discriminant Function for type		
Variable	1	2
Constant	-0.06870	-1.02533
x1	0.03251	-0.04553
x2	0.05118	0.22976

由SAS输出的结果可知，

若$0.03251x_1 + 0.05118x_2 - 0.06870 \geq -0.04553 + 0.22976x_2 - 1.02533$，则判$x \in \pi_2$；

若$0.03251x_1 + 0.05118x_2 - 0.06870 < -0.04553 + 0.22976x_2 - 1.02533$，则判$x \in \pi_1$；

将$x_1 = 0.6, x_2 = 3.0$代入该判别规则，得$0.0104346 > -0.363368$，故判$x \in \pi_1$，即预报明天会下雨。

Error Count Estimates for type			
	1	2	Total
Rate	0.2000	0.1000	0.1500
Priors	0.5000	0.5000	

$\hat{P}(2|1) = 0.2, \hat{P}(1|2) = 0.1$。

(2) 假定两组的 $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2)'$ 均服从二元正态分布，且根据其他信息及经验给出先验概率 $p_1 = 0.3$，$p_2 = 0.7$，试预报明天是否会下雨；

Posterior Probability of Membership in type			
Obs	Classified into type	1	2
1	2	0.4062	0.5938

$P(\pi_1|\boldsymbol{x}) = 0.4062 < 0.5938 = P(\pi_2|\boldsymbol{x})$，故判 $\boldsymbol{x} \in \pi_2$，即预报明天不会下雨。

(3) 假如你现在考虑是否为明天安排一项活动，该活动不太适合在雨天进行，并在(2)中假定的基础上还认为 $c(2|1) = 3c(1|2)$，那么你今天是否应该为明天安排这项活动呢？

$$\hat{\boldsymbol{a}} = \boldsymbol{S}_p^{-1}(\bar{\boldsymbol{x}}_1 - \bar{\boldsymbol{x}}_2) = \begin{pmatrix} 0.078 \\ -0.179 \end{pmatrix}$$

$$\bar{\boldsymbol{x}} = \frac{1}{2}(\bar{\boldsymbol{x}}_1 + \bar{\boldsymbol{x}}_2) = \begin{pmatrix} 0.27 \\ 5.475 \end{pmatrix}$$

$$\hat{\boldsymbol{a}}'(\boldsymbol{x} - \bar{\boldsymbol{x}}) = 0.468 > -0.251 = \ln\left[\frac{c(1|2)p_2}{c(2|1)p_1}\right]$$

故判 $x \in \pi_1$，即预报明天会下雨，不应该安排这项活动。

SAS Code

```
1  data work.file_data;
2  input type x1 x2;
3  cards;
4  1 -1.9 3.2

5  1 -6.9 10.4

6  1 5.2 2.0

7  1 5.0 2.5
```

```sas
1 7.3 0.0
1 6.8 12.7
1 0.9 -15.4
1 -12.5 -2.5
1 1.5 1.3
1 3.8 6.8
2 0.2 6.2
2 -0.1 7.5
2 0.4 14.6
2 2.7 8.3
2 2.1 0.8
2 -4.6 4.3
2 -1.7 10.9
2 -2.6 13.1
2 2.6 12.8
2 -2.8 10.0
;
run;
proc print data=work.file_data;
run;
data work.data_test;
input x1 x2;
cards;
0.6 3.0
;
run;
proc discrim data=work.file_data testdata=work.data_test list
testlist;
class type;
var x1 x2;
run;
proc discrim data=work.file_data testdata=work.data_test list
testlist method=normal pool=yes;
priors '1'=0.3 '2'=0.7;
class type;
var x1 x2;
run;
```