一、 填空与选择填空（本题答案写在此试卷上， 30 分）

1、模式识别系统的基本构成单元包括： 模式采集 、 特征提取与选择 和 模式分类 。

2、统计模式识别中描述模式的方法一般使用 特真矢量 ；句法模式识别中模式描述方法一般有 串 、 树 、 网 。

3、聚类分析算法属于 （1） ；判别域代数界面方程法属于 （3） 。

（1）无监督分类 (2) 有监督分类 （3）统计模式识别方法（ 4）句法模式识别方法

4、若描述模式的特征量为 0-1 二值特征量，则一般采用 （4） 进行相似性度量。

（1）距离测度 （2）模糊测度 （3）相似测度 （4）匹配测度

5、 下列函数可以作为聚类分析中的准则函数的有 （1）（3）（4） 。

（1） $J = Tr[S_W^{-1} S_B]$ （2） $J = \left| S_W S_B^{-1} \right|$ (3) $J = \sum_{j=1}^{c} \sum_{i=1}^{n_j} \left\| \vec{x}_i^{(j)} - \vec{m}_j \right\|^2$

(4) $J = \sum_{j=1}^{c} (\vec{m}_j - \vec{m})'(\vec{m}_j - \vec{m})$

6、 Fisher 线性判别函数的求解过程是将 N 维特征矢量投影在 （2） 中进行 。

（1）二维空间 （2）一维空间 （3）N-1 维空间

7、下列判别域界面方程法中只适用于线性可分情况的算法有 （1） ；线性可分、不可分都适用的有 （3） 。

（1）感知器算法 （2）H-K 算法 （3）积累位势函数法

8 、下列四元组中满足文法定义的有 （1）（2）（4） 。

（1）({A, B}, {0, 1}, {A→01, A→0A1, A→1A0, B→BA, B→0}, A)

（2）({A}, {0, 1}, {A→0, A→0A}, A)

（3）({S}, {a, b}, {S→00S, S→11S, S→00, S→11}, S)

（4）({A}, {0, 1}, {A→01, A→0A1, A→1A0}, A)

9、影响层次聚类算法结果的主要因素有（ 计算模式距离的测度、（聚类准则、类间距离门限、预定的类别数目））。

10、欧式距离具有（ 1 、2 ）；马式距离具有（ 1、2、3、4 ）。

（1）平移不变性（ 2）旋转不变性（ 3）尺度缩放不变性（ 4）不受量纲影响的特性

11、线性判别函数的正负和数值大小的几何意义是（ 正（负）表示样本点位于判别界面法向量指向的正（负）半空间中；绝对值正比于样本点到判别界面的距离。 ）。

12、感知器算法 1 。

（1）只适用于线性可分的情况；（ 2）线性可分、不可分都适用。

13、积累势函数法较之于 H-K 算法的优点是（该方法可用于非线性可分情况 （也可用于线性可分情况） ）；

位势函数 $K(x,x_k)$ 与积累位势函数 $K(x)$ 的关系为（ $K(\bar{x})=\sum_{\bar{x}_k\in\tilde{X}}\alpha_k K(\bar{x},\bar{x}_k)$ ）。

14、在统计模式分类问题中，聂曼 - 皮尔逊判决准则主要用于（ 某一种判决错误较另一种判决错误更为重要 ）情况；最小最大判决准则主要用于（ 先验概率未知的 ）情况。

15、“特征个数越多越有利于分类”这种说法正确吗？（ 错误 ）。特征选择的主要目的是（ 从 n 个特征中选出最有利于分类的的 m 个特征（ m<n），以降低特征维数 ）。一般在（ 可分性判据对特征个数具有单调性 ）和（ $C_n^m>>n$ ）的条件下，可以使用分支定界法以减少计算量。

16、 散度 Jij 越大，说明 ω_i 类模式与 ω_j 类模式的分布（ 差别越大 ）；当 ω_i 类模式与 ω_j 类模式的分布相同时， Jij= （0）。

17、 已知有限状态自动机 Af=(Σ, Q, δ, q0, F)，Σ={0，1}；Q={q0，q1}；δ：δ(q0，0)= q1，δ(q0，1)= q1，δ(q1，0)=q0，δ(q1，1)=q0；q0=q0；F={q0}。现有输入字符串：(a) 00011101011，(b) 1100110011，(c) 101100111000，(d)0010011，试问，用 Af 对上述字符串进行分类的结果为（ ω 1:{a,d}; ω 2:{b,c} ）。

18、影响聚类算法结果的主要因素有（ __②③④ __）。

①已知类别的样本质量；②分类准则；③特征选取；④模式相似性测度。

19、模式识别中，马式距离较之于欧式距离的优点是（ ___③④ ____）。

①平移不变性；②旋转不变性；③尺度不变性；④考虑了模式的分布。

20、基于二次准则函数的 H-K 算法较之于感知器算法的优点是（ __①③ __）。

①可以判别问题是否线性可分；②其解完全适用于非线性可分的情况；

③其解的适应性更好；④计算量小。

21、影响基本 C 均值算法的主要因素有（ ____④①② ____）。

①样本输入顺序；②模式相似性测度；③聚类准则；④初始类心的选取。

22、位势函数法的积累势函数 K(x) 的作用相当于 Bayes 判决中的（ __②④ __）。

①先验概率；②后验概率；③类概率密度；④类概率密度与先验概率的乘积。

23、在统计模式分类问题中，当先验概率未知时，可以使用（ __②④ __）。

①最小损失准则；②最小最大损失准则；③最小误判概率准则；④ N-P 判决。

24、在（ ___①③ ___）情况下，用分支定界法做特征选择计算量相对较少。

① $C_n^d>>n$，（n 为原特征个数， d 为要选出的特征个数）；②样本较多；③选用的可分性判据 J 对特征数目单调不减；④选用的可分性判据 J 具有可加性。

25、 散度 J_D是根据（ ___③___ ）构造的可分性判据。

①先验概率；②后验概率；③类概率密度；④信息熵；⑤几何距离。

26、似然函数的概型已知且为单峰，则可用（ _①②③④⑤ _）估计该似然函数。

①矩估计；②最大似然估计；③ Bayes 估计；④ Bayes 学习；⑤ Parzen 窗法。

27、 Kn近邻元法较之 Parzen 窗法的优点是（ __②__）。

①所需样本数较少；②稳定性较好；③分辨率较高；④连续性较好。

28、从分类的角度讲，用 DKLT做特征提取主要利用了 DKLT的性质：（__①③ __）。

①变换产生的新分量正交或不相关； ②以部分新的分量表示原矢量均方误差最小； ③使变换后的矢量能量更趋集中；

29、一般，剪辑 k-NN 最近邻方法在（ __①__）的情况下效果较好。

①样本数较大；②样本数较小；③样本呈团状分布；④样本呈链状分布。

30、如果以特征向量的相关系数作为模式相似性测度， 则影响聚类算法结果的主要因素有 （__②③ __）。

①已知类别样本质量；②分类准则；③特征选取；④量纲。

二、(15 分) 简答及证明题

（1）影响聚类结果的主要因素有那些？

（2）证明马氏距离是平移不变的、非奇异线性变换不变的。

答：（ 1）分类准则，模式相似性测度，特征量的选择，量纲。

（2）证明：

$$d^2(\vec{x}_i,\vec{x}_j)=(\vec{x}_i-\vec{x}_j)'V^{-1}(\vec{x}_i-\vec{x}_j)$$ (2 分)

$$V=\frac{1}{m-1}\sum_{i=1}^{m}(\vec{x}_i-\bar{\vec{x}})(\vec{x}_i-\bar{\vec{x}})'$$ (2 分)

$$\bar{\vec{x}}=\frac{1}{m}\sum_{i=1}^{m}\vec{x}_i$$ (1 分)

三、 (8 分) 说明线性判别函数的正负和数值大小在分类中的意义并证明之。

答：（1）（4 分） $d(\vec{x})$ 的绝对值 $|d(\vec{x})|$ 正比于 $\vec{x}$ 到超平面 $d(\vec{x})=0$ 的距离 d_x 式(1-1) 的分子为判别函数绝对值， 上式表明， $d(\vec{x})$ 的值 $|d(\vec{x})|$ 正比于 $\vec{x}$ 到超平面 $d(\vec{x})=0$ 的距离 d_x，一个特征矢量代入判别函数后所得值的绝对值越大表明该特征点距判别界面越远。

（2）（4分） 判别函数值的正负表示出特征点位于哪个半空间中，或者换句话说，表示特征点位于界面的哪一侧。

四、(12 分，每问 4 分) 在目标识别中，假定有农田和装甲车两种类型，类型 ω_1和类型 ω_2分别代表农田和装甲车，它们的先验概率分别为 0.8 和 0.2，损失函数如表 1 所示。现在做了三次试验，获得三个样本的类概率密度如下：

$p(x/\omega_1)$：0.3， 0.1， 0.6

$p(x/\omega_2)$：0.7， 0.8， 0.3

（1） 试用贝叶斯最小误判概率准则判决三个样本各属于哪一个类型；

（2） 假定只考虑前两种判决，试用贝叶斯最小风险准则判决三个样本各属于哪一类；

（3） 把拒绝判决考虑在内，重新考核三次试验的结果。

表 1

判决 \ 损失 \ 类型	ω_1	ω_2
α_1	1	4
α_2	5	1
α_3	1	1

解：由题可知：$P(\omega_1)=0.7, P(\omega_2)=0.3$，$\dfrac{P(\omega_2)}{P(\omega_1)}=\dfrac{3}{7}$，$\dfrac{P(x_1|\omega_1)}{P(x_1|\omega_2)}=\dfrac{3}{7}$，$\dfrac{P(x_2|\omega_1)}{P(x_2|\omega_2)}=\dfrac{1}{8}$，$\dfrac{P(x_3|\omega_1)}{P(x_3|\omega_2)}=2$

（1）（4 分）根据贝叶斯最小误判概率准则知：

$\dfrac{P(x_1|\omega_1)}{P(x_1|\omega_2)}=\dfrac{P(\omega_2)}{P(\omega_1)}$，则可以任判；$\dfrac{P(x_2|\omega_1)}{P(x_2|\omega_2)}<\dfrac{P(\omega_2)}{P(\omega_1)}$，则判为 ω_2；$\dfrac{P(x_3|\omega_1)}{P(x_3|\omega_2)}>\dfrac{P(\omega_2)}{P(\omega_1)}$，则判为 ω_1；

（2）（4 分）由题可知：$\dfrac{P(\omega_2)(\lambda_{21}-\lambda_{22})}{P(\omega_1)(\lambda_{12}-\lambda_{11})}=\dfrac{0.3(5-1)}{0.7(4-1)}=\dfrac{4}{7}$

则 $\dfrac{P(x_1|\omega_1)}{P(x_1|\omega_2)}<\dfrac{4}{7}$，判为 ω_2；

$\dfrac{P(x_2|\omega_1)}{P(x_2|\omega_2)}<\dfrac{4}{7}$，判为 ω_2；

$\dfrac{P(x_3|\omega_1)}{P(x_3|\omega_2)}>\dfrac{4}{7}$，判为 ω_1；

（3）（4 分）对于两类问题，对于样本 x，假设 $P(x)$ 已知，有

$$R(\alpha_j|x)=\lambda(\alpha_j|\omega_1)P(\omega_1|x)+\lambda(\alpha_j|\omega_2)P(\omega_2|x)=$$
$$=\frac{\lambda(\alpha_j|\omega_1)P(x|\omega_1)P(\omega_1)+\lambda(\alpha_j|\omega_2)P(x|\omega_2)P(\omega_2)}{P(x)}$$

则对于第一个样本，

$R(\alpha_1|x)=\frac{5\times0.21}{P(x)},R(\alpha_2|x)=\frac{4\times0.21}{P(x)},R(\alpha_3|x)=\frac{2\times0.21}{P(x)}$，则拒判；

$R(\alpha_1|x)=\frac{1.03}{P(x)},R(\alpha_2|x)=\frac{0.59}{P(x)},R(\alpha_3|x)=\frac{0.24}{P(x)}$，则拒判；

$R(\alpha_1|x)=\frac{0.78}{P(x)},R(\alpha_2|x)=\frac{2.19}{P(x)},R(\alpha_3|x)=\frac{0.51}{P(x)}$，拒判。

五、 1. 监督学习与非监督学习的区别：

监督学习方法用来对数据实现分类，分类规则通过训练获得。该训练集由带分类号的数据集组成，因此监督学习方法的训练过程是离线的。

非监督学习方法不需要单独的离线训练过程，也没有带分类号（标号）的训练数据集，一般用来对数据集进行分析，如聚类，确定其分布的主分量等。

（实例：道路图）就道路图像的分割而言，监督学习方法则先在训练用图像中获取道路象素与非道路象素集，进行分类器设计，然后用所设计的分类器对道路图像进行分割。

使用非监督学习方法，则依据道路路面象素与非道路象素之间的聚类分析进行聚类运算，以实现道路图像的分割。

2. 线性分类器三种最优准则：

Fisher 准则：根据两类样本一般类内密集，类间分离的特点，寻找线性分类器最佳的法线向量方向，使两类样本在该方向上的投影满足类内尽可能密集，类间尽可能分开。

该种度量通过类内离散矩阵 Sw和类间离散矩阵 Sb 实现。

感知准则函数：准则函数以使错分类样本到分界面距离之和最小为原则。

其优点是通过错分类样本提供的信息对分类器函数进行修正，这种准则是人工神经元网络多层感知器的基础。

支持向量机：基本思想是在两类线性可分条件下，所设计的分类器界面使两类之间的间隔为最大，它的基本出发点是使期望泛化风险尽可能小。

一、 试说明 Mahalanobis 距离平方的定义，到某点的 Mahalanobis 距离平方为常数的轨迹的几何意义，它与欧氏距离的区别与联系。

答：Mahalanobis 距离的平方定义为：

$$r^2(x,u)=(x-u)^T\Sigma^{-1}(x-u)$$

其中 x，u 为两个数据，Σ^{-1}是一个正定对称矩阵（一般为协方差矩阵）。根据定义，距某一点的 Mahalanobis 距离相等点的轨迹是超椭球，如果是单位矩阵 Σ，则 Mahalanobis 距离就是通常的欧氏距离。

二、 试说明用监督学习与非监督学习两种方法对道路图像中道路区域的划分的基本做法，以说明这两种学习方法的定义与它们间的区别。

答：监督学习方法用来对数据实现分类，分类规则通过训练获得。该训练集由带分类号的数据集组成，因此监督学习方法的训练过程是离线的。

非监督学习方法不需要单独的离线训练过程，也没有带分类号（标号）的训练数据集，一般用来对数据集进行分析，如聚类，确定其分布的主分量等。就道路图像的分割而言，监督学习方法则先在训练用图像中获取道路象素与非道路象素集，进行分类器设计，然后用所设计的分类器对道路图像进行分割。使用非监督学习方法，则依据道路路面象素与非道路象素之间的聚类分析进行聚类运算，以实现道路图像的分割。

三、 试述动态聚类与分级聚类这两种方法的原理与不同。

答：动态聚类是指对当前聚类通过迭代运算改善聚类；分级聚类则是将样本个体，按相似度标准合并，随着相似度要求的降低实现合并。

四、 试说明以下问题求解是基于监督学习或是非监督学习：

1. 求数据集的主分量 2. 汉字识别 3. 自组织特征映射 4. CT 图像的分割

答： 1、求数据集的主分量是非监督学习方法；

2、汉字识别对待识别字符加上相应类别号——有监督学习方法；

3、自组织特征映射——将高维数组按保留近似度向低维映射——非监督学习；

4、CT图像分割——按数据自然分布聚类——非监督学习方法；

五、 试列举线性分类器中最著名的三种最佳准则以及它们各自的原理。

答：线性分类器三种最优准则：

Fisher 准则：根据两类样本一般类内密集，类间分离的特点，寻找线性分类器最佳的法线向量方向，使两类样本在该方向上的投影满足类内尽可能密集，类间尽可能分开。该种度量通过类内离散矩阵 Sw和类间离散矩阵 Sb 实现。

感知准则函数：准则函数以使错分类样本到分界面距离之和最小为原则。其优点是通过错分类样本提供的信息对分类器函数进行修正，这种准则是人工神经元网络多层感知器的基础。

支持向量机：基本思想是在两类线性可分条件下，所设计的分类器界面使两类之间的间隔为最大，它的基本出发点是使期望泛化风险尽可能小。

十、 对一副道路图像，希望把道路部分划分出来，可以采用以下两种方法：

1．在该图像中分别在道路部分与非道路部分画出一个窗口，把在这两个窗口中的象素数据作为训练集，用 Fisher 准则方法求得分类器参数，再用该分类器对整幅图进行分类。

2．将整幅图的每个象素的属性记录在一张数据表中，然后用某种方法将这些数据按它们的自然分

布状况划分成两类。 因此每个象素就分别得到相应的类别号， 从而实现了道路图像的分割。 试问以上两种方法哪一种是监督学习，哪个是非监督学习？

答： 第一种方法中标记了两类样本的标号，需要人手工干预训练过程，属于监督学习方法；

第二种方法只是依照数据的自然分布，把它们划分成两类，属于非监督学习方法。

十三、 试分析五种常用决策规则思想方法的异同。

答、五种常用决策是：

1. 基于最小错误率的贝叶斯决策， 利用概率论中的贝叶斯公式， 得出使得错误率最小的分类规则。
2. 基于最小风险的贝叶斯决策， 引入了损失函数， 得出使决策风险最小的分类。 当在 0－1 损失函数条件下，基于最小风险的贝叶斯决策变成基于最小错误率的贝叶斯决策。
3. 在限定一类错误率条件下使另一类错误率最小的两类别决策。
4. 最大最小决策：类先验概率未知，考察先验概率变化对错误率的影响，找出使最小贝叶斯奉献最大的先验概率，以这种最坏情况设计分类器。
5. 序贯分类方法，除了考虑分类造成的损失外，还考虑特征获取造成的代价，先用一部分特征分类，然后逐步加入性特征以减少分类损失，同时平衡总的损失，以求得最有效益。

十四、 假设在某个地区细胞识别中正常（ w_1）和异常（ w_2）两类先验概率分别为 $P(w_1)=0.9$ ， $P(w_2)=0.1$ ，现有一待识别的细胞，其观察值为 x，从类条件概率密度分布曲线上查得 $P(x/w_1)=0.2$，$P(x/w_2)=0.4$，并且已知 $\lambda_{11}=0$，$\lambda_{12}=6$，$\lambda_{21}=1$，$\lambda_{22}=0$

试对该细胞 x 用一下两种方法进行分类： 1. 基于最小错误率的贝叶斯决策； 2. 基于最小风险的贝叶斯决策；请分析两种结果的异同及原因。

解：利用贝叶斯公式，分别计算出 ω_1 及 ω_2 的后验概率。

$$P(\omega_1|\boldsymbol{x})=\frac{p(\boldsymbol{x}|\omega_1)P(\omega_1)}{\sum_{j=1}^{2}p(\boldsymbol{x}|\omega_j)P(\omega_j)}=\frac{0.2\times0.9}{0.2\times0.9+0.4\times0.1}=0.818$$

$$P(\omega_2|\boldsymbol{x})=1-p(\omega_1|\boldsymbol{x})=0.182$$

根据贝叶斯决策规则式(2-2)，有

$$P(\omega_1|\boldsymbol{x})=0.818>P(\omega_2|\boldsymbol{x})=0.182$$

所以合理的决策是把 $\boldsymbol{x}$ 归类于正常状态。

2.

解：已知条件为

$$P(\omega_1)=0.9, \qquad P(\omega_2)=0.1$$
$$p(\boldsymbol{x}|\omega_1)=0.2, \qquad p(\boldsymbol{x}|\omega_2)=0.4$$
$$\lambda_{11}=0, \qquad \lambda_{12}=6$$
$$\lambda_{21}=1, \qquad \lambda_{22}=0$$

根据 1 的计算结果可知后验概率为

$$P(\omega_1|\boldsymbol{x})=0.818, \qquad P(\omega_2|\boldsymbol{x})=0.182$$

再计算出条件风险

$$R(\alpha_1|\boldsymbol{x}) = \sum_{j=1}^{2}\lambda_{1j}P(\omega_j|\boldsymbol{x}) = \lambda_{12}P(\omega_2|\boldsymbol{x}) = 1.092$$
$$R(\alpha_2|\boldsymbol{x})=\lambda_{21}P(\omega_1|\boldsymbol{x})=0.818$$

由于 $R(\alpha_1|\boldsymbol{x})>R(\alpha_2|\boldsymbol{x})$

即决策为 ω_2 的条件风险小于决策为 ω_1 的条件风险，因此我们采取决策行动 α_2，即判断待识别的细胞 $\boldsymbol{x}$ 为 ω_2 类——异常细胞。

将 1与2 相对比，其分类结果正好相反，这是因为这里影响决策结果的因素又多了一个，即“损失”。而且两类错误决策所造成的损失相差很悬殊，因此“损失”就起了主导作用。

十五、 有线性判别函数，为什么还要引进非线性判别函数？分析由“ 线性判别函数 ”向“ 非线性判别函数 ”推广的思想和方法。

答：实际中有很多模式识别问题并不是线性可分的， 这时就需要采用非线性分类器， 比如当两类样本分不具有多峰性质并互相交错时， 简单的线性判别函数往往会带来较大的分类错误。 这时， 树分类器作为一种分段线性分类器，常常能有效地应用于这种情况。

十六、 1. 什么是特征选择？ 2. 什么是 Fisher 线性判别？

答： 1. 特征选择就是从一组特征中挑选出一些最有效的特征以达到降低特征空间维数的目的。

2. Fisher 线性判别：可以考虑把 d 维空间的样本投影到一条直线上，形成一维空间，即把维数压缩到一维， 这在数学上容易办到，然而， 即使样本在 d 维空间里形成若干紧凑的互相分得开的集群， 如果把它们投影到一条任意的直线上， 也可能使得几类样本混在一起而变得无法识别。 但是在一般情况下，总可以找到某个方向， 使得在这个方向的直线上， 样本的投影能分开得最好。 问题是如何根据实际情况找到这条最好的、最易于分类的投影线，这就是 Fisher 算法所要解决的基本问题。

十七、 写出两类和多类情况下最小风险贝叶斯决策判别函数和决策面方程。

两类别问题：判别函数

$$g_1(x) = \lambda_{11}p(\omega_1|x) + \lambda_{12}p(\omega_2|x)$$
$$g_2(x) = \lambda_{21}p(\omega_1|x) + \lambda_{22}p(\omega_2|x)$$

决策面方程：$g_1(x) = g_2(x)$

c 类别问题：判别函数

$$g_i(x) = \sum_{j=1}^{c} \lambda_{ji}p(\omega_j|x), \qquad i = 1, \cdots, c$$

决策面方程：$g_i(x) = g_j(x),\ i \neq j \quad i = 1, \cdots, c \quad j = 1, \cdots, c$

二十、 定性说明基于参数方法和非参数方法的概率密度估计有什么区别？

答： 基于参数方法：是由已知类别的样本集对总体分布的某些参数进行统计推断

非参数方法：已知样本所属类别，但未知总体概率密度函数形式

二十二、 简述支持向量机的基本思想。

答： SVM从线性可分情况下的最优分类面发展而来。

最优分类面就是要求分类线不但能将两类正确分开 (训练错误率为 0)，且使分类间隔最大。 SVM考虑寻找一个满足分类要求的超平面，并且使训练集中的点距离分类面尽可能的远，也就是寻找一个分类面使它两侧的空白区域 (margin) 最大。过两类样本中离分类面最近的点，且平行于最优分类面的超平面上 H_1，H_2的训练样本就叫支持向量。

3 对两类问题，若损失函数； $\lambda_{11} = \lambda_{22} = 0$， $\lambda_{12} \neq 0$， $\lambda_{21} \neq 0$，试求基于最小风险贝叶斯决策分界面处的两类错误率 $P(e)|_{x\in\omega_1} = 1 - P(\omega_1|x)$、 $P(e)|_{x\in\omega_2} = 1 - P(\omega_2|x)$ 与 λ_{12}、 λ_{21} 的关系

解：由于在基于最小风险贝叶斯决策分界面处有

$R(\alpha_1|x) = \lambda_{11}P(\omega_1|x) + \lambda_{12}P(\omega_2|x) = \lambda_{12}P(\omega_2|x)$ $R(\alpha_2|x) = \lambda_{21}P(\omega_1|x) + \lambda_{22}P(\omega_2|x) = \lambda_{21}P(\omega_1|x)$

$R(\alpha_1|x) = R(\alpha_2|x) \Rightarrow \lambda_{12}P(\omega_2|x) = \lambda_{21}P(\omega_1|x)$ 而在两类问题中， $P(e)|_{x\in\omega_1} = 1 - P(\omega_1|x) = P(\omega_2|x)$， $P(e)|_{x\in\omega_2} = 1 - P(\omega_2|x) = P(\omega_1|x)$ 故

$$\frac{P(e)|_{x\in\omega_1}}{P(e)|_{x\in\omega_2}} = \frac{\lambda_{21}}{\lambda_{12}}$$