06 年

一、答：

$$p(\omega_i|x)=\frac{p(x|\omega_i)P(\omega_i)}{\sum_{j=1}^{M}p(x|\omega_j)P(\omega_j)},\quad i=1,\cdots,M$$

$$R(\alpha_i|x)=\sum_{j=1}^{M}\lambda_{ij}p(\omega_j|x),\quad i=1,\cdots,M$$

如果

$$R(\alpha_i|x)=\min_{j=1,\cdots,M}R(\alpha_j|x)$$

则 x 属于ω_i

$$\lambda(\alpha_i,\omega_j)=\begin{cases}0, i=j,\\1, i\neq j,\end{cases}\quad i,j=1,2,\cdots,c \tag{2-18}$$

式中假定对于 c 类只有 c 个决策，即不考虑“拒绝”的情况。式(2-18)中 $\lambda(\alpha_i,\omega_j)$是对于正确决策(即 $i=j$)没有损失；而对于任何错误决策，其损失均为 1。这样定义的损失函数称为 0-1 损失函数。

根据式(2-15)，条件风险为

$$R(\alpha_i|\boldsymbol{x})=\sum_{j=1}^{c}\lambda(\alpha_i,\omega_j)P(\omega_j|\boldsymbol{x})=\sum_{\substack{j=1\\j\neq i}}^{c}P(\omega_j|\boldsymbol{x}) \tag{2-19}$$

$\sum_{\substack{j=1\\j\neq i}}^{c}P(\omega_j|\boldsymbol{x})$表示对 $\boldsymbol{x}$ 采取决策 ω_i 的条件错误概率。所以在 0-1 损失函数时，使

$$R(\alpha_k|\boldsymbol{x})=\min_{i=1,\cdots,c}R(\alpha_i|\boldsymbol{x})$$

的最小风险贝叶斯决策就等价于

$$\sum_{\substack{j=1\\j\neq k}}^{c}P(\omega_j|\boldsymbol{x})=\min_{i=1,\cdots,c}\sum_{\substack{j=1\\j\neq i}}^{c}P(\omega_j|\boldsymbol{x})$$

的最小错误率贝叶斯决策。

由此可见，最小错误率贝叶斯决策就是在 0-1 损失函数条件下的最小风险贝叶斯决策。

二、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。

三、答：

实际应用中，多元正态分布的更典型情况是，均值 $\boldsymbol{\mu}$ 和协方差矩阵 $\boldsymbol{\Sigma}$ 都未知。这样，参数向量 $\boldsymbol{\theta}$ 就是由这两个成分组成。我们首先考虑单变量的情况，其中参数向量 $\boldsymbol{\theta}$ 的组成成分是：$\theta_1=\mu,\theta_2=\sigma^2$。这样，对于单个训练样本的对数似然函数为：

$$\ln p(x_k|\boldsymbol{\theta})=-\frac{1}{2}\ln 2\pi\theta_2-\frac{1}{2\theta_2}(x_k-\theta_1)^2 \tag{12}$$

对上式关于变量 $\boldsymbol{\theta}$ 求导：

$$\nabla_{\boldsymbol{\theta}} l=\nabla_{\boldsymbol{\theta}}\ln p(x_k|\boldsymbol{\theta})=\begin{bmatrix}\dfrac{1}{\theta_2}(x_k-\theta_1)\\ -\dfrac{1}{2\theta_2}+\dfrac{(x_k-\theta_1)^2}{2\theta_2^2}\end{bmatrix} \tag{13}$$

运用式(7)，我们得到对于全体样本的对数似然函数的极值条件

$$\sum_{k=1}^{n}\frac{1}{\hat{\theta}_2}(x_i-\hat{\theta}_1)=0 \tag{14}$$

和

$$-\sum_{k=1}^{n}\frac{1}{\hat{\theta}_2}+\sum_{k=1}^{n}\frac{(x_k-\hat{\theta}_1)^2}{\hat{\theta}_2^2}=0 \tag{15}$$

其中的 $\hat{\theta}_1,\hat{\theta}_2$ 分别是对于 θ_1,θ_2 的最大似然估计。

把 $\hat{\theta}_1,\hat{\theta}_2$ 用 $\hat{\mu},\hat{\sigma}^2$ 代替，并进行简单的整理，我们得到下述的对于均值和方差的最大似然估计结果

$$\hat{\mu}=\frac{1}{n}\sum_{k=1}^{n}x_k \tag{16}$$

和

$$\hat{\sigma}^2=\frac{1}{n}\sum_{k=1}^{n}(x_k-\hat{\mu})^2 \tag{17}$$

四、答：

Fisher:

$$\mathrm{m}_1=\begin{bmatrix}\frac{2}{3}\\ \frac{2}{3}\end{bmatrix},\ \mathrm{m}_2=\begin{bmatrix}-\frac{1}{3}\\ -\frac{1}{3}\end{bmatrix},\ \Sigma_1=\begin{bmatrix}\frac{1}{3} & -\frac{1}{6}\\ -\frac{1}{6} & \frac{1}{3}\end{bmatrix},\ \ \Sigma_2=\begin{bmatrix}\frac{1}{3} & -\frac{1}{6}\\ -\frac{1}{6} & \frac{1}{3}\end{bmatrix}\ \mathrm{S_w}=\begin{bmatrix}\frac{1}{3} & -\frac{1}{6}\\ -\frac{1}{6} & \frac{1}{3}\end{bmatrix}$$

$$\mathrm{w}^*=\mathrm{S_w^{-1}}(\mathrm{m}_1-\mathrm{m}_2)=\begin{bmatrix}6\\6\end{bmatrix},\quad \mathrm{y}_0=\frac{\widetilde{\mathrm{m}_1}+\widetilde{\mathrm{m}_2}}{2}=\frac{\mathrm{w}^{*\mathrm{T}}\mathrm{m}_1+\mathrm{w}^{*\mathrm{T}}\mathrm{m}_1\mathrm{m}_2}{2}=2$$

所以判别函数为：$\mathrm{g(x)}=\begin{bmatrix}6\\6\end{bmatrix}^{\mathrm{T}}\mathrm{x}-2$

最小平方误差法：

$$Y=\begin{bmatrix}1 & 0\\ 0 & 1\\ 1 & 1\\ 0 & 0\\ 1 & 0\\ 0 & 1\end{bmatrix}\qquad Y^+=\left(Y^TY\right)^{-1}Y^T=\frac{1}{8}\begin{bmatrix}3 & -1 & 2 & 0 & 3-\\ -1 & 3 & 2 & 0- & 1\end{bmatrix}$$

令 $b=\begin{bmatrix}1 & 1 & 1 & 1 & 1 & 1\end{bmatrix}^T$，$a^*=Y^+b=\begin{bmatrix}\frac{3}{4}\\ \frac{3}{4}\end{bmatrix}$

五、答：

$$\mu=\begin{bmatrix}0\\0\end{bmatrix}$$

$$S_b=\frac{1}{2}\sum_{i=1}^{2}(\mu_i-\mu)(\mu_i-\mu)^T=\frac{1}{4}(\mu_1-\mu_2)(\mu_1-\mu_2)^T=\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}$$

$$S_w=\frac{1}{2}(S_1+S_2)=\frac{1}{2}\begin{bmatrix}2 & 0\\0 & 5\end{bmatrix},\qquad S_w^{-1}=\begin{bmatrix}5 & 0\\0 & 2\end{bmatrix}$$

$$S_w^{-1}S_b=\begin{bmatrix}5 & 5\\2 & 2\end{bmatrix}$$

因为$S_w^{-1}S_b$的秩为 1,所以$S_w^{-1}S_b$只有一个非零本征值,W 是 D×1 矩阵,即 W=w。为求$S_w^{-1}S_b$的本征值应解方程：

$$S_w^{-1}S_b w=\lambda_1 w$$

或

$$\frac{1}{4}S_w^{-1}(\mu_1-\mu_2)(\mu_1-\mu_2)^T w=\lambda_1 w$$

$\frac{1}{4}(\mu_1-\mu_2)^T w$是标量，所以

$$w=S_w^{-1}(\mu_1-\mu_2)=\begin{bmatrix}10\\4\end{bmatrix}$$

六、答：

$X=\emptyset, J(x_1)=0, J(x_2)=3, J(x_3)=3, J(x_4)=4$

所以$X=\{x_4\}, J(X+x_1)=4, J(X+x_2)=5.2785, J(X+x_3)=5.1569$

所以$X=\{x_2x_4\}, J(X+x_1)=5.2785, J(X+x_3)=6.2237$

所以$X=\{x_2x_3x_4\}$

七、答：

第一类：(-1,1), (-1,0), (-1,-1), (0,-1), (0,-2)

第二类：(0,1)

第三类：(1,1), (1,0), (2,0)