

一、单选题 (本题共 25 小题, 满分 50 分)

题号	1	2	3	4	5	6	7	8	9	10
答案	D	D	C	A	B	A	C	B	A	B
题号	11	12	13	14	15	16	17	18	19	20
答案	B	B	B	B	C	A	A	B	A	D
题号	21	22	23	24	25					
答案	B	A	D	A	B					

二、计算题 (本题共 5 小题, 满分 50 分)

1. 使用公式正确 (2 分) 计算正确 (2 分)

$$SSE = (-0.2)^2 + (0.4)^2 + (-0.8)^2 + (1.3)^2 + (-0.7)^2 = 3.02$$

2. (本小题 8 分) 已知逻辑回归模型得到一组逻辑回归结果, 要求:

(1) 假设阈值为 0.6, 写出预测结果。(2 分)

逻辑回归结果	逻辑回归预测结果	真实结果
0.40	0	1
0.65	1	0
0.20	0	0
0.80	1	1
0.70	1	1

(2) 计算出损失函数的值 (即真实值与预测值之间的损失值)。(6 分)

样本 1 为正样本, 损失为: $1 \cdot \log(0.4)$

样本 2 为负样本, 损失为: $(1-0) \cdot \log(1-0.65)$

样本 3 为负样本, 损失为: $(1-0) \cdot \log(1-0.2)$

样本 4 为正样本, 损失为: $1 \cdot \log(0.8)$

样本 5 为正样本, 损失为: $1 \cdot \log(0.7)$

(以上过程 4 分: 公式错误 1 处 1 分, 扣完 2 分为止, 计算错误 1 处 1 分, 扣完 2 分为止)

$$\text{总损失函数} = -\log(0.4) - \log(0.35) - \log(0.8) - \log(0.8) - \log(0.7)$$

$$\text{或总损失函数} = -(\log(0.4) + \log(0.35) + \log(0.8) + \log(0.8) + \log(0.7))$$

$$\log \text{ 以 } 10 \text{ 为底结果} = -1.2026 \text{ 或 } 1.2026$$

$$\log \text{ 以 } 2 \text{ 为底结果} = -3.9949 \text{ 或 } 3.9949$$

$$\log \text{ 以 } e \text{ 为底结果} = -2.7691 \text{ 或 } 2.7691$$

(2 分, 其中公式 1 分, 计算 1 分)

3. (本小题 10 分)

	真实好瓜 P	真实坏瓜 N
预测好瓜 P	TP = 14	FP = 2
预测坏瓜 N	FN = 1	TN = 3

(表格内每空 1 分, 共 4 分)

$$P = 14 / (14+2) = 7/8 = 0.875 \text{ (2 分) (公式和计算各 1 分)}$$

$$R = 14 / (14+1) = 14/15 = 0.933 \text{ (2 分) (公式和计算各 1 分)}$$

$$F1 = 2 \cdot P \cdot R / (P+R) = 28/31 = 0.903 \text{ (2 分) (公式和计算各 1 分)}$$

$$4. (1) \hat{\omega} = (X^T X)^{-1} X^T y \text{ (2 分)}$$

(2) def standRegres(xArr, yArr):

```

xMat = np.mat(xArr)          (1 分)
yMat = np.mat(yArr).T        (1 分)
xTx = xMat.T * xMat          (1 分)
if np.linalg.det(xTx) == 0.0: (1 分)
    print("矩阵为奇异矩阵, 不能求逆")
    return                    (1 分)
ws = xTx.I * (xMat.T*yMat)    (1 分)
return ws                     (1 分)

```

此外: 格式 3 分, 错一处扣一分, 扣完为止

5. 我们需要最大化 $P(X|C_i)P(C_i)$, $i=1, 2$. 每个类的先验概率 $P(C)$ 可以根据训练样本计算:

$P(C1)=9/14=0.643$

$P(C2)=5/14=0.357$

(2 分, 公式和计算各 1 分)

为计算 $P(X/C_i)$ $i=1, 2$, 我们计算下面的条件概率:

$P(\text{天气}=\text{"雨"} \mid C1)=3/9=0.333$

$P(\text{天气}=\text{"雨"} \mid C2)=2/5=0.400$

(2 分, 公式和计算各 1 分)

$P(\text{温度}=\text{"凉"} \mid C1)=3/9=0.333$

$P(\text{温度}=\text{"凉"} \mid C2)=1/5=0.200$

(2 分, 公式和计算各 1 分)

$P(\text{湿度}=\text{"高"} \mid C1)=3/9=0.333$

$P(\text{湿度}=\text{"高"} \mid C2)=4/5=0.800$

(2 分, 公式和计算各 1 分)

$P(\text{风力}=\text{"弱"} \mid C1)=6/9=0.667$

$P(\text{风力}=\text{"弱"} \mid C2)=2/5=0.400$

(2 分, 公式和计算各 1 分)

使用以上概率, 我们得到:

$P(X \mid C1)=0.333 \times 0.333 \times 0.333 \times 0.667=0.0247$

$P(X \mid C2)=0.400 \times 0.200 \times 0.800 \times 0.400=0.0256$

(2 分, 公式和计算各 1 分)

$P(X|C1)P(C1)=0.0247 \times 0.643=0.01588$

$P(X|C2)P(C2)=0.0256 \times 0.357=0.00914$

或者

$P(X|C1)P(C1)/P(X)=49/90=0.5444$

$P(X|C2)P(C2)/P(X)=196/625=0.3136$

(2 分, 公式和计算各 1 分)

因此, 对于样本 X, 朴素贝叶斯分类预测 C1 (2 分)

