一	二	三	四	五	六	七	八	总分

（请在框内进行答题，请标清题号）

一、1. $\begin{cases}(a-1)(b-1)=2\\ ab(m-1)=6\\ abm-1=11\\ B的自由度为1\end{cases} \Rightarrow \begin{cases}a=3\\ b=2\\ m=2\end{cases}$

故 $SS_A = 0.8333 \times (3-1) = 0.1666$.

2. 由上可知，A自由度为 $a-1=2$.

3. 水平数为 2.

4. $MSE = \frac{10}{6} = \frac{5}{3}$

5. $m=2$.

6. 由于 $\frac{SS_{AB}}{\sigma^2} \sim \chi^2((a-1)(b-1))$

故 $F_{AB} = \frac{SS_{AB}/(a-1)(b-1)}{SS_E/ab(m-1)} \sim F((a-1)(b-1), ab(m-1))$

故p值为 $P(F > F_{AB})$ 其中 $F \sim F((a-1)(b-1), ab(m-1))$

当p值小于0.05，我们拒绝原假设，认为交互效应显著，说明 F_{AB} 已经较大.

否则，我们接受原假设，认为交互效应不显著.

7. 证：$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{m}(y_{ijk} - \bar{y}_{...})^2$.

$= \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{m}((y_{ijk} - \bar{y}_{ij.}) + (\bar{y}_{i..} - \bar{y}_{...}) + (\bar{y}_{.j.} - \bar{y}_{...}) +$

$(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...}))^2$.

(请在框内进行答题，请标清题号)

欲证 $SS_T = SS_A + SS_B + SS_{AB} + SS_E$，只需证上述平方和展开式交互项为0。

令 $A = y_{ijk} - \bar{y}_{ij\cdot}$　$B = \bar{y}_{i\cdots} - \bar{y}_{\cdots}$　$C = \bar{y}_{j\cdot\cdot} - \bar{y}_{\cdots}$

$D = \bar{y}_{ij\cdot} - \bar{y}_{i\cdots} - \bar{y}_{j\cdot\cdot} + \bar{y}_{\cdots}$

"AB": $\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{m}(y_{ijk} - \bar{y}_{ij\cdot})(\bar{y}_{i\cdots} - \bar{y}_{\cdots})$

$= \sum_{i=1}^{a}(\bar{y}_{i\cdots} - \bar{y}_{\cdots})\sum_{j=1}^{b}\sum_{k=1}^{m}(y_{ijk} - \bar{y}_{ij\cdot})$

由于 $\sum_{k=1}^{m}(y_{ijk} - \bar{y}_{ij\cdot}) = 0$ 故该项为0。

~~"AC"~~: 同理: $\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{m}(y_{ijk} - \bar{y}_{ij\cdot})(\bar{y}_{j\cdot\cdot} - \bar{y}_{\cdots}) = 0$。

"AD": $\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{m}(y_{ijk} - \bar{y}_{ij\cdot})(\bar{y}_{ij\cdot} - \bar{y}_{i\cdots} - \bar{y}_{j\cdot\cdot} + \bar{y}_{\cdots})$

$= \sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{y}_{ij\cdot} - \bar{y}_{i\cdots} - \bar{y}_{j\cdot\cdot} + \bar{y}_{\cdots})\sum_{k=1}^{m}(y_{ijk} - \bar{y}_{ij\cdot})$

由于 $\sum_{k=1}^{m}(y_{ijk} - \bar{y}_{ij\cdot}) = 0$ 故该项为0。

"BC": $\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{m}(\bar{y}_{i\cdots} - \bar{y}_{\cdots})(\bar{y}_{j\cdot\cdot} - \bar{y}_{\cdots})$

$= m\sum_{i=1}^{a}(\bar{y}_{i\cdots} - \bar{y}_{\cdots})\sum_{j=1}^{b}(\bar{y}_{j\cdot\cdot} - \bar{y}_{\cdots})$

由于 $\sum_{i=1}^{a}(\bar{y}_{i\cdots} - \bar{y}_{\cdots}) = 0$ 故该项为0。

"BD": $\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{m}(\bar{y}_{i\cdots} - \bar{y}_{\cdots})(\bar{y}_{ij\cdot} - \bar{y}_{i\cdots} - \bar{y}_{j\cdot\cdot} + \bar{y}_{\cdots})$

$= m\sum_{i=1}^{a}(\bar{y}_{i\cdots} - \bar{y}_{\cdots})\sum_{j=1}^{b}(\bar{y}_{ij\cdot} - \bar{y}_{i\cdots} - \bar{y}_{j\cdot\cdot} + \bar{y}_{\cdots})$

(请在框内进行答题，请标清题号)

由于 $\sum_{j=1}^{q}(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot j\cdot}+\bar{y}_{\cdots})=0$. 故该项为0.

同理，$\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{m}(\bar{y}_{ij\cdot}-\bar{y}_{\cdots})(\bar{y}_{ij\cdot}-\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot j\cdot}+\bar{y}_{\cdots})=0$

故所有交错项为0. 故 $SS_T=SS_A+SS_B+SS_{AB}+SS_E$.

二. 1. $y=0.6954+1.6034x_1+\varepsilon,\ \varepsilon\sim N(0,\sigma^2)$

由于 t 检验的 p 值为 0.000. 故我们认为该模型显著（取 $\alpha=0.05$）

2. $x_1=0.5$ 时，$\hat{y}=0.6954+1.6034\times0.5=1.4971$.

在线性回归模型中，$\hat{y}_0\sim N(x_0'\beta,\sigma^2x_0'(X'X)^{-1}x_0)$

而 $y\sim N(x_0'\beta,\sigma^2)$

故 $y-\hat{y}_0\sim N(0,\sigma^2(1+x_0'(X'X)^{-1}x_0))$.

又：σ^2 未知，故用估计值 $\hat{\sigma}^2=\frac{SSE}{n-p-1}$ 代替.

故 $t=\frac{y-\hat{y}_0}{\sqrt{\hat{\sigma}^2(1+x_0'(X'X)^{-1}x_0)SSE/(n-p-1)}}\sim t(n-p-1)$

用 $t(n-p-1)$ 的分位数 $t_{1\pm\frac{\alpha}{2}}(n-p-1)$ 给出预测区间.

3. 这个结论普遍存在.

证：$R^2=\frac{SS_R}{SS_T}=\frac{y'(I-H_0)y}{y'(I-H_m)y}\quad\frac{y'(I-H_m)y}{y'(I-H)y}$

$=\cdots$

(请在框内进行答题，请标清题号)

4. 由3中的结论：放入模型的自变量越多则R^2越大，则若以R^2作为模型选择标准，当选模型不准确，该标准也会让我们选用全模型，从而造成过拟合，故不应用R^2作为标准。

改进方案：用 $\hat{\sigma} = \frac{SSE}{n-p-1}$ 作为标准。

~~当自变量较少，则 $\frac{1}{n-p-1}$ 可作为罚项，~~

~~当自变量过多，SSE要小很多~~

三、4 对 $X'X$ 特征分解：$X'X = V'\Lambda V$，其中V是对称正交阵，$\Lambda = diag(\lambda_1, \cdots \lambda_{p+1})$ $\lambda_1 \geq \cdots \geq \lambda_{p+1}$ 是 $X'X$ 的特征值。

当数据集中出现多重共线性，$|X'X| \approx 0$。

故 又 $X'X\alpha = \lambda\alpha$。故 $X'X$ 仍然有比较接近于0的特征值。

取 $k_j = \sqrt{\frac{\lambda_1}{\lambda_j}}$。当 λ_j 接近于0，条件数 k_j 会很大，当某个 k_j 大于某个特定的值，就认为该数据集存在多重共线性。

四、感知机损失函数为 $\sum_{x \in W} -y_i(wx_i + b)$

其中W是误分类点的集合，只要所有点均正确分类，损失函数即为0。

线性SVM的损失函数为合页损失函数

$\sum_{i=1}^{n} [y_i(x_i + b)]_+ + \alpha\xi_i$ 其中 $[w_+] = \begin{cases} w, & w > 0 \\ 0, & w \leq 0. \end{cases}$ $\xi_i \geq 0$.

(请在框内进行答题，请标清题号)

也就是说，在线性SVM中，不仅要完全分类正确，所有点离分类超平面的距离要不小于1(即最大程度上把正负实例点分开)，损失函数的值才为0，

五/1. 生成式模型就是在用给定的数据集在某种准则下逐步生成出来的模型；

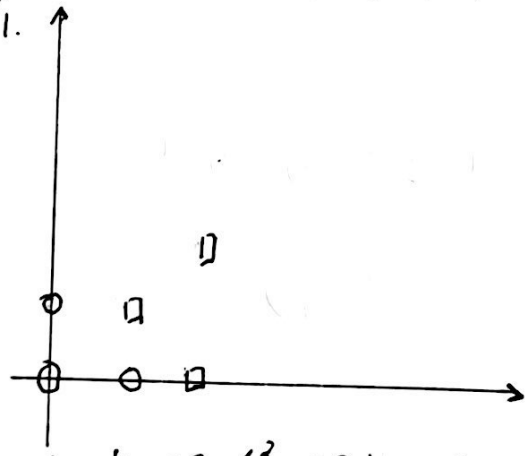
判别式模型就是在用给定的数据集学习出来，并用于判断未知点所属类别的模型

不同在于判别式模型无显式的学习过程，而生成式模型是通过多次的迭代逐渐"生成"出来的

二、判别模型：KNN，朴素贝叶斯，逻辑斯谛回归，

生成模型：支持向量机，决策树。

六/1.



显然，数据集线性可分，构造最优化问题

$$\min_{w} \frac{1}{2}\|w\|^2$$

$$s.t.\ y_i(wx_i+b)-1\geq 0,\quad i=1,\cdots,N$$

即：$\min_{w_1,w_2} w_1^2+w_2^2$

$$s.t.\ w_1+w_2+b-1\geq 0$$

(请在框内进行答题，请标清题号)

$2w_1+2w_2+b-1\geq 0$.

$2w_1+b-1\geq 0$

$-b-1\geq 0$.

$-w_1-b-1\geq 0$.

$-w_2-b-1\geq 0$.

约束条件化简后得：
$$\begin{cases} w_1+w_2-2\geq 0 \\ w_1+2w_2-2\geq 0 \\ 2w_1-w_2-2\geq 0 \\ w_2\geq 2 \\ w_1\geq 2 \\ 2w_1+w_2-2\geq 2 \end{cases}$$

解得：$w_1=w_2=2$.

2. (1,1)(2,0)(0,1)(1,0)都是支撑向量.

3. 构造对偶优化问题：

$$\min_{\alpha} \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j (x_i x_j) - \sum_{i=1}^{n}\alpha_i$$

$$s.t. \sum_{i=1}^{n}\alpha_i y_i = 0. \qquad i=1\cdots n.$$

$$\alpha_i \geq 0$$

即 $\min_{\alpha} \frac{1}{2}(4\alpha_1\alpha_2+2\alpha_1\alpha_3+4\alpha_2\alpha_3+2\alpha_1^2+8\alpha_2^2+4\alpha_3^2$

$+\alpha_5^2+\alpha_6^2+6\alpha_1\alpha_5+2\alpha_2\alpha_5+2\alpha_3\alpha_5+\alpha_1\alpha_6$

$+2\alpha_2\alpha_6) - \alpha_1-\alpha_2-\alpha_3-\alpha_4-\alpha_5-\alpha_6$

$s.t. \sum_{i=1}^{n}\alpha_1+\alpha_2+\alpha_3-\alpha_4-\alpha_5-\alpha_6=0$

$\alpha_i\geq 0.$

（请在框内进行答题，请标清题号）

令拉格朗日乘子为 l，则

$$L = \frac{1}{2}\alpha_1$$

$$\begin{cases} \frac{\partial L}{\partial \alpha_1} = 2\alpha_2 + \alpha_3 + 2\alpha_1 + 5\alpha_5 + \frac{1}{2}\alpha_6 + l - 1 = 0 \\ \frac{\partial L}{\partial \alpha_2} = 2\alpha_1 + 2\alpha_3 + 8\alpha_2 + \alpha_5 + \alpha_6 + l - 1 = 0 \\ \frac{\partial L}{\partial \alpha_3} = \alpha_1 + 2\alpha_2 + 4\alpha_3 + \alpha_5 + l - 1 = 0 \\ \frac{\partial L}{\partial \alpha_4} = -l - 1 = 0 \\ \frac{\partial L}{\partial \alpha_5} = \alpha_5 + 3\alpha_1 + \alpha_2 + -l - 1 = 0 \\ \frac{\partial L}{\partial \alpha_6} = \alpha_6 + \frac{1}{2}\alpha_1 + \alpha_2 - l - 1 = 0 \end{cases}$$

$$\begin{cases} \alpha_1 = -\frac{5}{19} \\ \alpha_2 = \frac{3}{19} \\ \alpha_3 = \frac{25}{76} \\ \alpha_5 = -3\alpha_1 - \alpha_2 = \frac{12}{19} \\ \alpha_6 = -\frac{1}{2}\alpha_1 - \alpha_2 = \frac{9}{38} \\ l = -1 \end{cases}$$

（请在框内进行答题，请标清题号）