# 期末考试回忆 (2024 秋)

## Problem 1

双因子方差分析，试求 SSA, SSB, SSAB 的期望.

**Solution:**
在双因子方差分析的背景下，我们有:

$$\bar{y}_{ij\cdot} := \frac{1}{n}\sum_{k=1}^{n} y_{ijk} = \mu_{ij} + \bar{\varepsilon}_{ij\cdot} \sim N\left(\mu_{ij}, \frac{\sigma^2}{n}\right)$$

$$\bar{y}_{i\cdot\cdot} := \frac{1}{b}\sum_{j=1}^{b} \bar{y}_{ij\cdot} = \mu_{i\cdot} + \bar{\varepsilon}_{i\cdot\cdot} \sim N\left(\mu_{i\cdot}, \frac{\sigma^2}{bn}\right)$$

$$\bar{y}_{\cdot j\cdot} := \frac{1}{a}\sum_{i=1}^{a} \bar{y}_{ij\cdot} = \mu_{\cdot j} + \bar{\varepsilon}_{\cdot j\cdot} \sim N\left(\mu_{\cdot j}, \frac{\sigma^2}{an}\right)$$

$$\bar{y}_{\cdots} := \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b} \bar{y}_{ij\cdot} = \mu_{\cdot\cdot} + \bar{\varepsilon}_{\cdots} \sim N\left(\mu_{\cdot\cdot}, \frac{\sigma^2}{abn}\right)$$

$$\mathrm{SSAB} := n\sum_{i,j}^{a,b} (\bar{y}_{ij\cdot} + \bar{y}_{\cdots} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot})^2$$

$$\mathrm{SSA} := bn\sum_{i=1}^{a} (\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots})^2$$

$$\mathrm{SSB} := an\sum_{j=1}^{b} (\bar{y}_{\cdot j\cdot} - \bar{y}_{\cdots})^2$$

定义 $\mathrm{E} := [\bar{\varepsilon}_{ij\cdot}] \in \mathbb{R}^{a\times b}$
用 $\mathrm{E}_{(i,:)}$ 代表 $\mathrm{E}$ 的第 $i$ 行，用 $\mathrm{E}_{(:,j)}$ 代表 $\mathrm{E}$ 的第 $j$ 列.
注意到:

$$\bar{\varepsilon}_{i\cdot\cdot} = \frac{1}{b}\sum_{j=1}^{b} \bar{\varepsilon}_{ij\cdot} = \mathrm{E}_{(i,:)} \cdot \frac{1}{b} 1_b$$

$$\bar{\varepsilon}_{\cdot j\cdot} = \frac{1}{a}\sum_{i=1}^{a} \bar{\varepsilon}_{ij\cdot} = \frac{1}{a} 1_a^{\mathrm{T}} \cdot \mathrm{E}_{(:,j)}$$

$$\bar{\varepsilon}_{\cdots} = \frac{1}{ab}\sum_{i=1}^{a}\sum_{j=1}^{b} \bar{\varepsilon}_{ij\cdot} = \frac{1}{a} 1_a^{\mathrm{T}} \cdot \mathrm{E} \cdot \frac{1}{b} 1_b$$

因此我们有:

$$\begin{bmatrix} \bar{\varepsilon}_{1\cdot\cdot} & \cdots & \bar{\varepsilon}_{1\cdot\cdot} \\ \vdots & & \vdots \\ \bar{\varepsilon}_{a\cdot\cdot} & \cdots & \bar{\varepsilon}_{a\cdot\cdot} \end{bmatrix} = \begin{bmatrix} \bar{\varepsilon}_{1\cdot\cdot} 1_b^{\mathrm{T}} \\ \vdots \\ \bar{\varepsilon}_{a\cdot\cdot} 1_b^{\mathrm{T}} \end{bmatrix} = \mathrm{E} \cdot \frac{1}{b} 1_b 1_b^{\mathrm{T}}$$

$$\begin{bmatrix} \bar{\varepsilon}_{\cdot 1\cdot} & \cdots & \bar{\varepsilon}_{\cdot b\cdot} \\ \vdots & & \vdots \\ \bar{\varepsilon}_{\cdot 1\cdot} & \cdots & \bar{\varepsilon}_{\cdot b\cdot} \end{bmatrix} = [\bar{\varepsilon}_{\cdot 1\cdot} 1_a \quad \cdots \quad \bar{\varepsilon}_{\cdot b\cdot} 1_a] = \frac{1}{a} 1_a 1_a^{\mathrm{T}} \cdot \mathrm{E}$$

$$\begin{bmatrix} \bar{\varepsilon}_{\cdots} & \cdots & \bar{\varepsilon}_{\cdots} \\ \vdots & & \vdots \\ \bar{\varepsilon}_{\cdots} & \cdots & \bar{\varepsilon}_{\cdots} \end{bmatrix} = 1_a \bar{\varepsilon}_{\cdots} 1_b^{\mathrm{T}} = \frac{1}{a} 1_a 1_a^{\mathrm{T}} \cdot \mathrm{E} \cdot \frac{1}{b} 1_b 1_b^{\mathrm{T}}$$

### (1) $\mathrm{E}[\mathrm{SSAB}]$

注意到:

$$
\begin{aligned}
&(I_a-\frac{1}{a}1_a1_a^{\mathrm{T}})\mathrm{E}(I_b-\frac{1}{b}1_b1_b^{\mathrm{T}}) \\
&=\mathrm{E}+\frac{1}{a}1_a1_a^{\mathrm{T}}\cdot\mathrm{E}\cdot\frac{1}{b}1_b1_b^{\mathrm{T}}-\mathrm{E}\cdot\frac{1}{b}1_b1_b^{\mathrm{T}}-\frac{1}{a}1_a1_a^{\mathrm{T}}\cdot\mathrm{E} \\
&=[\bar{\varepsilon}_{ij\cdot}]+[\bar{\varepsilon}_{\cdots}]-[\bar{\varepsilon}_{i\cdot\cdot}]-[\bar{\varepsilon}_{\cdot j\cdot}] \\
&=[\bar{\varepsilon}_{ij\cdot}+\bar{\varepsilon}_{\cdots}-\bar{\varepsilon}_{i\cdot\cdot}-\bar{\varepsilon}_{\cdot j\cdot}]
\end{aligned}
$$

记 $\begin{cases}P_a=I_a-\frac{1}{a}1_a1_a^{\mathrm{T}}\\ P_b=I_b-\frac{1}{b}1_b1_b^{\mathrm{T}}\end{cases}$

显然它们是投影算子 (即自伴且幂等)，记其谱分解为:

$$
\begin{cases}P_a=Q_a\Lambda_aQ_a^{\mathrm{T}}\\ P_b=Q_b\Lambda_bQ_b^{\mathrm{T}}\end{cases}\ \text{where}\ \begin{cases}\Lambda_a=\mathrm{diag}\{\underbrace{1,\ldots,1}_{a-1\text{ times}},0\}\\ \Lambda_b=\mathrm{diag}\{\underbrace{1,\ldots,1}_{b-1\text{ times}},0\}\end{cases}
$$

则我们有:

$$
\begin{aligned}
\sum_{i,j}^{a,b}(\bar{\varepsilon}_{ij\cdot}+\bar{\varepsilon}_{\cdots}-\bar{\varepsilon}_{i\cdot\cdot}-\bar{\varepsilon}_{\cdot j\cdot})^2&=\sum_{i,j}^{a,b}([P_a\mathrm{E}P_b]_{(i,j)})^2\quad(\text{note that }\bar{\varepsilon}_{ij\cdot}+\bar{\varepsilon}_{\cdots}-\bar{\varepsilon}_{i\cdot\cdot}-\bar{\varepsilon}_{\cdot j\cdot}=[P_a\mathrm{E}P_b]_{(i,j)}) \\
&=\|P_a\mathrm{E}P_b\|_{\mathrm{F}}^2\quad(\text{Frobenius norm}) \\
&=\mathrm{tr}\left((P_a\mathrm{E}P_b)^{\mathrm{T}}P_a\mathrm{E}P_b\right) \\
&=\mathrm{tr}\left(P_b^{\mathrm{T}}\mathrm{E}^{\mathrm{T}}P_a^{\mathrm{T}}P_a\mathrm{E}P_b\right) \\
&=\mathrm{tr}\left(\mathrm{E}^{\mathrm{T}}P_a^{\mathrm{T}}P_a\mathrm{E}P_bP_b^{\mathrm{T}}\right)\quad(\text{note that }\begin{cases}P_a^{\mathrm{T}}P_a=P_a^2=P_a\\ P_b^{\mathrm{T}}P_b=P_b^2=P_b\end{cases}) \\
&=\mathrm{tr}\left(\mathrm{E}^{\mathrm{T}}P_a\mathrm{E}P_b\right)\quad(\text{note that }\begin{cases}P_a=Q_a\Lambda_aQ_a^{\mathrm{T}}\\ P_b=Q_b\Lambda_bQ_b^{\mathrm{T}}\end{cases}) \\
&=\mathrm{tr}\left(\mathrm{E}^{\mathrm{T}}Q_a\Lambda_aQ_a^{\mathrm{T}}\mathrm{E}Q_b\Lambda_bQ_b^{\mathrm{T}}\right) \\
&=\mathrm{tr}\left(Q_b^{\mathrm{T}}\mathrm{E}^{\mathrm{T}}Q_a\Lambda_aQ_a^{\mathrm{T}}\mathrm{E}Q_b\Lambda_b\right)\quad(\text{denote }\widetilde{\mathrm{E}}:=Q_a^{\mathrm{T}}\mathrm{E}Q_b) \\
&=\mathrm{tr}\left(\widetilde{\mathrm{E}}^{\mathrm{T}}\Lambda_a\widetilde{\mathrm{E}}\Lambda_b\right) \\
&=\mathrm{tr}\left((\Lambda_a\widetilde{\mathrm{E}})^{\mathrm{T}}(\widetilde{\mathrm{E}}\Lambda_b)\right)\quad(\text{note that }\Lambda_a=\mathrm{diag}\{\underbrace{1,\ldots,1}_{a-1\text{ times}},0\}\text{ and }\Lambda_b=\mathrm{diag}\{\underbrace{1,\ldots,1}_{b-1\text{ times}},0\}) \\
&=\sum_{i=1}^{a-1}\sum_{j=1}^{b-1}\widetilde{\mathrm{E}}_{(i,j)}^2
\end{aligned}
$$

注意到 $\mathrm{E}:=[\bar{\varepsilon}_{ij\cdot}]\in\mathbb{R}^{a\times b}$ 的元素独立同分布:

$$
\{\bar{\varepsilon}_{ij\cdot}\}\overset{\text{iid}}{\sim}N\left(0,\frac{\sigma^2}{n}\right)
$$

我们可以将整个 $\mathrm{E}\in\mathbb{R}^{a\times b}$ 的分布表示为:

$$
\mathrm{vec}(\mathrm{E})\sim N\left(0_{ab},\frac{\sigma^2}{n}I_b\otimes I_a\right)
$$

其中 $\mathrm{vec}(\cdot)$ 是向量化操作符 (即将一个矩阵按列拉伸为向量)，而 $\otimes$ 代表 Kronecker 乘积.
于是我们有:

$$
\begin{aligned}
\mathrm{vec}(\widetilde{\mathrm{E}})&=\mathrm{vec}(Q_a^{\mathrm{T}}\mathrm{E}Q_b)\quad(\text{note that }\mathrm{vec}(AXB)=(B^{\mathrm{T}}\otimes A)\mathrm{vec}(X)) \\
&=(Q_b^{\mathrm{T}}\otimes Q_a^{\mathrm{T}})\mathrm{vec}(\mathrm{E}) \\
&\sim N\left((Q_b^{\mathrm{T}}\otimes Q_a^{\mathrm{T}})0_{ab},(Q_b^{\mathrm{T}}\otimes Q_a^{\mathrm{T}})\cdot\frac{\sigma^2}{n}I_b\otimes I_a\cdot(Q_b^{\mathrm{T}}\otimes Q_a^{\mathrm{T}})^{\mathrm{T}}\right) \\
&=N\left(0_{ab},(Q_b^{\mathrm{T}}\otimes Q_a^{\mathrm{T}})\cdot\frac{\sigma^2}{n}I_b\otimes I_a\cdot(Q_b\otimes Q_a)\right) \\
&=N\left(0_{ab},\frac{\sigma^2}{n}[(Q_b^{\mathrm{T}}\cdot I_b\cdot Q_b)\otimes(Q_a^{\mathrm{T}}\cdot I_a\cdot Q_a)]\right) \\
&=N\left(0_{ab},\frac{\sigma^2}{n}I_b\otimes I_a\right)
\end{aligned}
$$

因此 $\widetilde{\mathrm{E}}=Q_a^{\mathrm{T}}\mathrm{E}Q_b\in\mathbb{R}^{a\times b}$ 的元素独立同分布:

$$
\{\widetilde{\mathrm{E}}_{ij\cdot}\}\overset{\text{iid}}{\sim}N\left(0,\frac{\sigma^2}{n}\right)
$$

于是我们有:

$$\sum_{i=1}^{a-1}\sum_{j=1}^{b-1}\widetilde{\mathrm{E}}_{(i,j)}^2 \sim \frac{\sigma^2}{n}\chi^2_{(a-1)(b-1)}$$

因此我们有:

$$\begin{aligned}
&\mathrm{E}[\mathrm{SSAB}] \\
&= \mathrm{E}\left[n\sum_{i,j}^{a,b}(\bar{y}_{ij\cdot}+\bar{y}_{\cdots}-\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot j\cdot})^2\right] \\
&= n\cdot\mathrm{E}\left[\sum_{i,j}^{a,b}[(\mu_{ij}+\mu_{\cdot\cdot}-\mu_{i\cdot}-\mu_{\cdot j})+(\bar{\varepsilon}_{ij\cdot}+\bar{\varepsilon}_{\cdots}-\bar{\varepsilon}_{i\cdot\cdot}-\bar{\varepsilon}_{\cdot j\cdot})]^2\right] \\
&= n\left\{\sum_{i,j}^{a,b}(\mu_{ij}+\mu_{\cdot\cdot}-\mu_{i\cdot}-\mu_{\cdot j})^2+2\sum_{i,j}^{a,b}(\mu_{ij}+\mu_{\cdot\cdot}-\mu_{i\cdot}-\mu_{\cdot j})\mathrm{E}[\bar{\varepsilon}_{ij\cdot}+\bar{\varepsilon}_{\cdots}-\bar{\varepsilon}_{i\cdot\cdot}-\bar{\varepsilon}_{\cdot j\cdot}]+\mathrm{E}\left[\sum_{i,j}^{a,b}(\bar{\varepsilon}_{ij\cdot}+\bar{\varepsilon}_{\cdots}-\bar{\varepsilon}_{i\cdot\cdot}-\bar{\varepsilon}_{\cdot j\cdot})^2\right]\right\} \\
&= n\sum_{i,j}^{a,b}(\mu_{ij}+\mu_{\cdot\cdot}-\mu_{i\cdot}-\mu_{\cdot j})^2+0+n\cdot\frac{\sigma^2}{n}(a-1)(b-1) \\
&= n\sum_{i,j}^{a,b}(\mu_{ij}+\mu_{\cdot\cdot}-\mu_{i\cdot}-\mu_{\cdot j})^2+(a-1)(b-1)\sigma^2
\end{aligned}$$

## (2) $\mathrm{E}[\mathrm{SSA}]$ & $\mathrm{E}[\mathrm{SSB}]$

注意到:

$$\begin{aligned}
&(I_a-\frac{1}{a}1_a1_a^{\mathrm{T}})\cdot\mathrm{E}\cdot\frac{1}{b}1_b \\
&= \mathrm{E}\cdot\frac{1}{b}1_b-\frac{1}{a}1_a1_a^{\mathrm{T}}\cdot\mathrm{E}\cdot\frac{1}{b}1_b \\
&= \begin{bmatrix}\bar{\varepsilon}_{1\cdot\cdot}\\ \vdots \\ \bar{\varepsilon}_{a\cdot\cdot}\end{bmatrix}-\begin{bmatrix}\bar{\varepsilon}_{\cdots}\\ \vdots \\ \bar{\varepsilon}_{\cdots}\end{bmatrix} \\
&= \begin{bmatrix}\bar{\varepsilon}_{1\cdot\cdot}-\bar{\varepsilon}_{\cdots}\\ \vdots \\ \bar{\varepsilon}_{a\cdot\cdot}-\bar{\varepsilon}_{\cdots}\end{bmatrix}
\end{aligned}$$

记 $P_a=I_a-\frac{1}{a}1_a1_a^{\mathrm{T}}$
显然它是投影算子 (即自伴且幂等)，记其谱分解为:

$$P_a=Q_a\Lambda_aQ_a^{\mathrm{T}}\text{ where }\Lambda_a=\mathrm{diag}\{\underbrace{1,\dots,1}_{a-1\text{ times}},0\}$$

则我们有:

$$\begin{aligned}
\sum_{i=1}^{a}(\bar{\varepsilon}_{i\cdot\cdot}-\bar{\varepsilon}_{\cdots})^2 &= \sum_{i=1}^{a}\left(\left[P_a\mathrm{E}\cdot\frac{1}{b}1_b\right]_{(i)}\right)^2\quad(\text{note that }\bar{\varepsilon}_{i\cdot\cdot}-\bar{\varepsilon}_{\cdots}=\left[P_a\mathrm{E}\cdot\frac{1}{b}1_b\right]_{(i)}) \\
&= \left\|P_a\mathrm{E}\cdot\frac{1}{b}1_b\right\|_2^2 \\
&= \frac{1}{b}1_b^{\mathrm{T}}\mathrm{E}^{\mathrm{T}}P_a^{\mathrm{T}}P_a\mathrm{E}\cdot\frac{1}{b}1_b\quad(\text{note that }P_a^{\mathrm{T}}P_a=P_a^2=P_a) \\
&= \frac{1}{b^2}1_b^{\mathrm{T}}\mathrm{E}^{\mathrm{T}}P_a\mathrm{E}1_b\quad(\text{note that }P_a=Q_a\Lambda_aQ_a^{\mathrm{T}}) \\
&= \frac{1}{b^2}1_b^{\mathrm{T}}\mathrm{E}^{\mathrm{T}}Q_a\Lambda_aQ_a^{\mathrm{T}}\mathrm{E}1_b\quad(\text{denote }\tilde{e}:=Q_a^{\mathrm{T}}\mathrm{E}1_b\in\mathbb{R}^a) \\
&= \frac{1}{b^2}\tilde{e}^{\mathrm{T}}\Lambda_a\tilde{e}\quad(\text{note that }\Lambda_a=\mathrm{diag}\{\underbrace{1,\dots,1}_{a-1\text{ times}},0\}) \\
&= \frac{1}{b^2}\sum_{i=1}^{a}\tilde{e}_i^2
\end{aligned}$$

注意到 $\mathrm{E}:=[\bar{\varepsilon}_{ij\cdot}]\in\mathbb{R}^{a\times b}$ 的元素独立同分布:

$$\{\bar{\varepsilon}_{ij\cdot}\}\overset{\text{iid}}{\sim}N\left(0,\frac{\sigma^2}{n}\right)$$

我们可以将整个 $\mathrm{E} \in \mathbb{R}^{a\times b}$ 的分布表示为:

$$\mathrm{vec}(\mathrm{E}) \sim N\left(0_{ab}, \frac{\sigma^2}{n} I_b \otimes I_a\right)$$

其中 $\mathrm{vec}(\cdot)$ 是向量化操作符 (即将一个矩阵按列拉伸为向量)，而 $\otimes$ 代表 Kronecker 乘积.
于是我们有:

$$\begin{aligned}
\tilde{e} &= Q_a^{\mathrm{T}} \mathrm{E} 1_b \quad (\text{note that } \mathrm{vec}(AXB) = (B^{\mathrm{T}} \otimes A)\mathrm{vec}(X)) \\
&= (1_b^{\mathrm{T}} \otimes Q_a^{\mathrm{T}})\mathrm{vec}(\mathrm{E}) \\
&\sim N\left((1_b^{\mathrm{T}} \otimes Q_a^{\mathrm{T}})0_{ab}, (1_b^{\mathrm{T}} \otimes Q_a^{\mathrm{T}}) \cdot \frac{\sigma^2}{n} I_b \otimes I_a \cdot (1_b^{\mathrm{T}} \otimes Q_a^{\mathrm{T}})^{\mathrm{T}}\right) \\
&= N\left(0_{ab}, (1_b^{\mathrm{T}} \otimes Q_a^{\mathrm{T}}) \cdot \frac{\sigma^2}{n} I_b \otimes I_a \cdot (1_b \otimes Q_a)\right) \\
&= N\left(0_{ab}, \frac{\sigma^2}{n}[(1_b^{\mathrm{T}} \cdot I_b \cdot 1_b) \otimes (Q_a^{\mathrm{T}} \cdot I_a \cdot Q_a)]\right) \\
&= N\left(0_{ab}, \frac{\sigma^2 b}{n} I_a\right)
\end{aligned}$$

因此 $\tilde{e} = Q_a^{\mathrm{T}} \mathrm{E} 1_b \in \mathbb{R}^a$ 的元素独立同分布:

$$\{\tilde{e}_i\} \overset{\mathrm{iid}}{\sim} N\left(0, \frac{\sigma^2 b}{n}\right)$$

于是我们有:

$$\sum_{i=1}^a (\bar{\varepsilon}_{i\cdot\cdot} - \bar{\varepsilon}_{\cdot\cdot\cdot})^2 = \frac{1}{b^2}\sum_{i=1}^a \tilde{e}_i^2 \sim \frac{1}{b^2} \cdot \frac{\sigma^2 b}{n} \chi^2_{(a-1)} = \frac{\sigma^2}{bn}\chi^2_{(a-1)}$$

因此我们有:

$$\begin{aligned}
\mathrm{E}[\mathrm{SSA}] &= \mathrm{E}\left[bn\sum_{i=1}^a (\bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot\cdot\cdot})^2\right] \\
&= bn \cdot \mathrm{E}\left[\sum_{i=1}^a [(\mu_{i\cdot} - \mu_{\cdot\cdot}) + (\bar{\varepsilon}_{i\cdot\cdot} - \bar{\varepsilon}_{\cdot\cdot\cdot})]^2\right] \\
&= bn\left\{\sum_{i=1}^a (\mu_{i\cdot} - \mu_{\cdot\cdot})^2 + 2\sum_{i=1}^a (\mu_{i\cdot} - \mu_{\cdot\cdot})\mathrm{E}[\bar{\varepsilon}_{i\cdot\cdot} - \bar{\varepsilon}_{\cdot\cdot\cdot}] + \mathrm{E}\left[\sum_{i=1}^a (\bar{\varepsilon}_{i\cdot\cdot} - \bar{\varepsilon}_{\cdot\cdot\cdot})^2\right]\right\} \\
&= bn\sum_{i=1}^a (\mu_{i\cdot} - \mu_{\cdot\cdot})^2 + 0 + bn \cdot \frac{1}{bn}\sigma^2(a-1) \\
&= bn\sum_{i=1}^a (\mu_{i\cdot} - \mu_{\cdot\cdot})^2 + \sigma^2(a-1)
\end{aligned}$$

交换 $a, b$ 的记号，则我们有:

$$\mathrm{E}[\mathrm{SSB}] = an\sum_{j=1}^b (\mu_{\cdot j} - \mu_{\cdot\cdot})^2 + \sigma^2(b-1)$$

## Problem 2

animo acid level vs. 3 Species and 2 Genders (5 samples for each treatment)

- ① fill the ANOVA table with interaction and test whether there is Interaction.

| | SS | DF | MS |
|---|---|---|---|
| Regression | 200 | ? | ? |
| Species | 59 | ? | ? |
| Gender | 138 | ? | ? |
| Interaction | ? | ? | ? |
| Error | ? | ? | ? |
| Total | 240 | ? | ? |

- ② find the ANOVA table without interaction and test whether there is main effect for Species and Gender.

α=0.05

| Fα k1 k2 | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 12 | 24 | ∞ |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 161.4 | 199.5 | 215.7 | 224.6 | 230.2 | 234.0 | 238.9 | 243.9 | 249.0 | 254.3 |
| 2 | 18.51 | 19.00 | 19.16 | 19.25 | 19.30 | 19.33 | 19.37 | 19.41 | 19.45 | 19.50 |
| 3 | 10.13 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.84 | 8.74 | 8.64 | 8.53 |
| 4 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.04 | 5.91 | 5.77 | 5.63 |
| 5 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.82 | 4.68 | 4.53 | 4.36 |
| 6 | 5.99 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.15 | 4.00 | 3.84 | 3.67 |
| 7 | 5.59 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.73 | 3.57 | 3.41 | 3.23 |
| 8 | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.44 | 3.28 | 3.12 | 2.93 |
| 9 | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.23 | 3.07 | 2.90 | 2.71 |
| 10 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.07 | 2.91 | 2.74 | 2.54 |
| 11 | 4.84 | 3.98 | 3.59 | 3.36 | 3.20 | 3.09 | 2.95 | 2.79 | 2.61 | 2.40 |
| 12 | 4.75 | 3.88 | 3.49 | 3.26 | 3.11 | 3.00 | 2.85 | 2.69 | 2.50 | 2.30 |
| 13 | 4.67 | 3.80 | 3.41 | 3.18 | 3.02 | 2.92 | 2.77 | 2.60 | 2.42 | 2.21 |
| 14 | 4.60 | 3.74 | 3.34 | 3.11 | 2.96 | 2.85 | 2.70 | 2.53 | 2.35 | 2.13 |
| 15 | 4.54 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.64 | 2.48 | 2.29 | 2.07 |
| 16 | 4.49 | 3.63 | 3.24 | 3.01 | 2.85 | 2.74 | 2.59 | 2.42 | 2.24 | 2.01 |
| 17 | 4.45 | 3.59 | 3.20 | 2.96 | 2.81 | 2.70 | 2.55 | 2.38 | 2.19 | 1.96 |
| 18 | 4.41 | 3.55 | 3.16 | 2.93 | 2.77 | 2.66 | 2.51 | 2.34 | 2.15 | 1.92 |
| 19 | 4.38 | 3.52 | 3.13 | 2.90 | 2.74 | 2.63 | 2.48 | 2.31 | 2.11 | 1.88 |
| 20 | 4.35 | 3.49 | 3.10 | 2.87 | 2.71 | 2.60 | 2.45 | 2.28 | 2.08 | 1.84 |
| 21 | 4.32 | 3.47 | 3.07 | 2.84 | 2.68 | 2.57 | 2.42 | 2.25 | 2.05 | 1.81 |
| 22 | 4.30 | 3.44 | 3.05 | 2.82 | 2.66 | 2.55 | 2.40 | 2.23 | 2.03 | 1.78 |
| 23 | 4.28 | 3.42 | 3.03 | 2.80 | 2.64 | 2.53 | 2.38 | 2.20 | 2.00 | 1.76 |
| 24 | 4.26 | 3.40 | 3.01 | 2.78 | 2.62 | 2.51 | 2.36 | 2.18 | 1.98 | 1.73 |
| 25 | 4.24 | 3.38 | 2.99 | 2.76 | 2.60 | 2.49 | 2.34 | 2.16 | 1.96 | 1.71 |
| 26 | 4.22 | 3.37 | 2.98 | 2.74 | 2.59 | 2.47 | 2.32 | 2.15 | 1.95 | 1.69 |
| 27 | 4.21 | 3.35 | 2.96 | 2.73 | 2.57 | 2.46 | 2.30 | 2.13 | 1.93 | 1.67 |
| 28 | 4.20 | 3.34 | 2.95 | 2.71 | 2.56 | 2.44 | 2.29 | 2.12 | 1.91 | 1.65 |
| 29 | 4.18 | 3.33 | 2.93 | 2.70 | 2.54 | 2.43 | 2.28 | 2.10 | 1.90 | 1.64 |
| 30 | 4.17 | 3.32 | 2.92 | 2.69 | 2.53 | 2.42 | 2.27 | 2.09 | 1.89 | 1.62 |
| 40 | 4.08 | 3.23 | 2.84 | 2.61 | 2.45 | 2.34 | 2.18 | 2.00 | 1.79 | 1.51 |
| 60 | 4.00 | 3.15 | 2.76 | 2.52 | 2.37 | 2.25 | 2.10 | 1.92 | 1.70 | 1.39 |
| 120 | 3.92 | 3.07 | 2.68 | 2.45 | 2.29 | 2.17 | 2.02 | 1.83 | 1.61 | 1.25 |
| ∞ | 3.84 | 2.99 | 2.60 | 2.37 | 2.21 | 2.09 | 1.94 | 1.75 | 1.52 | 1.00 |

**Solution:**
根据题意我们有: $a=3, b=2, n=5$

## (1) 有交互效应的情形

在存在交互效应的假设下，双因子方差分析的 ANOVA TABLE 如下:

ANOVA TABLE (with interaction)

| Sum of Squares | | Degree of Freedom | Mean Squares |
|---|---|---|---|
| SST | $\sum_{i,j,k}^{a,b,n}(y_{ijk}-\bar{y}_{...})^2$ | $abn-1$ | $\text{MST}=\frac{\text{SST}}{abn-1}$ |
| SSE | $\sum_{i,j,k}^{a,b,n}(y_{ijk}-\bar{y}_{ij\cdot})^2$ | $ab(n-1)$ | $\text{MSE}=\frac{\text{SSE}}{ab(n-1)}$ |
| SAB | $n\sum_{i,j}^{a,b}\widehat{\text{ME}}_{(A_i,B_j)}^2$ | $ab-1$ | $\text{MAB}=\frac{\text{SAB}}{ab-1}$ |
| SSAB | $n\sum_{i,j}^{a,b}\widehat{\text{IA}}_{(A_i,B_j)}^2$ | $(a-1)(b-1)$ | $\text{MSAB}=\frac{\text{SSAB}}{(a-1)(b-1)}$ |
| SSA | $bn\sum_{i=1}^{a}\widehat{\text{ME}}_{A_i}^2$ | $a-1$ | $\text{MSA}=\frac{\text{SSA}}{a-1}$ |
| SSB | $an\sum_{j=1}^{b}\widehat{\text{ME}}_{B_j}^2$ | $b-1$ | $\text{MSB}=\frac{\text{SSB}}{b-1}$ |

$$\widehat{\text{ME}}_{A_i}=\bar{y}_{i\cdot\cdot}-\bar{y}_{...}$$

$$\widehat{\text{ME}}_{B_j}=\bar{y}_{\cdot j\cdot}-\bar{y}_{...}$$

$$\widehat{\text{ME}}_{(A_i,B_j)}=\bar{y}_{ij\cdot}-\bar{y}_{...}$$

$$\widehat{\text{IA}}_{(A_i,B_j)}=\widehat{\text{ME}}_{(A_i,B_j)}-\widehat{\text{ME}}_{A_i}-\widehat{\text{ME}}_{B_j}=\bar{y}_{ij\cdot}+\bar{y}_{...}-\bar{y}_{i\cdot\cdot}-\bar{y}_{\cdot j\cdot}$$

| | SS | DF | MS |
|---|---|---|---|
| Regression | $200$ | $ab-1=5$ | $40$ |
| Species | $59$ | $a-1=2$ | $29.5$ |
| Gender | $138$ | $b-1=1$ | $138$ |
| Interaction | $3$ | $(a-1)(b-1)=2$ | $1.5$ |
| Error | $40$ | $ab(n-1)=24$ | $1.67$ |
| Total | $240$ | $abn-1=29$ | $8.28$ |

零假设和备择假设分别为:

$$H_0:\widehat{\text{IA}}_{(A_i,B_j)}=0 \text{ for all } \begin{cases} i=1,\dots,a \\ j=1,\dots,b \end{cases}$$

$$\Updownarrow$$

$$H_1:\exists\, i,j \text{ such that } \widehat{\text{IA}}_{(A_i,B_j)}=0$$

SSAB 和 MSAB 的零假设分布为:

$$\text{SSAB}\overset{H_0}{\sim}\sigma^2\chi^2_{(a-1)(b-1)}$$

$$\text{MSAB}:=\frac{\text{SSAB}}{(a-1)(b-1)}\overset{H_0}{\sim}\frac{\sigma^2\chi^2_{(a-1)(b-1)}}{(a-1)(b-1)}$$

设第一类型错误概率界限为 $\alpha$
我们构造如下的 $F$ 检验统计量:

$$F:=\frac{\text{MSAB}}{\text{MSE}}\overset{H_0}{\sim}\frac{\sigma^2\chi^2_{(a-1)(b-1)}/(a-1)(b-1)}{\sigma^2\chi^2_{ab(n-1)}/ab(n-1)}=F_{(a-1)(b-1),ab(n-1)}$$

其中分子 MSAB 和分母 MSE 是相互独立的 (无论零假设 $H_0$ 是否成立)

查表得 $F_{2,24}(0.05)=3.40$
根据 $F=\frac{\text{MSAB}}{\text{MSE}}=\frac{1.5}{1.67}\approx 0.90<F_{2,24}(0.05)$ 不能拒绝零假设，可以认为交互效应较弱.

## (2) 无交互效应的情形

在不存在交互效应的假设下，双因子方差分析的 ANOVA TABLE 如下:

ANOVA TABLE (without interaction)

| Sum of Squares | | Degree of Freedom | Mean Squares |
|---|---|---|---|
| SST | $\sum_{i,j,k}^{a,b,n}(y_{ijk}-\bar{y}_{...})^2$ | $abn-1$ | $\text{MST}=\frac{\text{SST}}{abn-1}$ |
| $\text{SSE}_{\text{reduced}}$ | $\sum_{i,j,k}^{a,b,n}(y_{ijk}+\bar{y}_{...}-\bar{y}_{i..}-\bar{y}_{.j.})^2$ | $abn-a-b+1$ | $\text{MSE}_{\text{reduced}}=\frac{\text{SSE}_{\text{reduced}}}{abn-a-b+1}$ |
| SSA | $bn\sum_{i=1}^{a}\widehat{\text{ME}}_{A_i}^2$ | $a-1$ | $\text{MSA}=\frac{\text{SSA}}{a-1}$ |
| SSB | $an\sum_{j=1}^{b}\widehat{\text{ME}}_{B_j}^2$ | $b-1$ | $\text{MSB}=\frac{\text{SSB}}{b-1}$ |

$$\widehat{\text{ME}}_{A_i}=\bar{y}_{i..}-\bar{y}_{...}$$

$$\widehat{\text{ME}}_{B_j}=\bar{y}_{.j.}-\bar{y}_{...}$$

| | SS | DF | MS |
|---|---|---|---|
| Regression | 197 | $a+b-2=3$ | 65.67 |
| Species | 59 | $a-1=2$ | 29.5 |
| Gender | 138 | $b-1=1$ | 138 |
| Error | 43 | $abn-a-b+1=26$ | 1.65 |
| Total | 240 | $abn-1=29$ | 8.28 |

零假设和备择假设分别为:

$$H_0:\text{ME}_{A_1}=\cdots=\text{ME}_{A_a}=0$$
$$\Updownarrow$$
$$H_1:\exists\, i=1,\ldots,a \text{ such that } \text{ME}_{A_i}\neq 0$$

SSA 的零假设分布为:

$$\text{SSA}\overset{H_0}{\sim}\sigma^2\chi^2_{(a-1)}$$

设第一类型错误概率界限为 $\alpha$
我们构造如下的 $F$ 检验统计量:

$$F:=\frac{\text{MSA}}{\text{MSE}_{\text{reduced}}}\overset{H_0}{\sim}\frac{\sigma^2\chi^2_{(a-1)}/(a-1)}{\sigma^2\chi^2_{abn-a-b+1}/(abn-a-b+1)}=F_{a-1,abn-a-b+1}$$

其中分子 MSA 和分母 $\text{MSE}_{\text{reduced}}$ 是相互独立的 (无论零假设 $H_0$ 是否成立)

查表得 $F_{2,26}(0.05)=3.37$
根据 $F=\frac{\text{MSA}}{\text{MSE}_{\text{reduced}}}=\frac{29.5}{1.65}\approx 17.88>F_{2,26}(0.05)$ 可以拒绝零假设，即 "物种" 因子存在主效应.

---

零假设和备择假设分别为:

$$H_0:\text{ME}_{B_1}=\cdots=\text{ME}_{B_b}=0$$
$$\Updownarrow$$
$$H_1:\exists\, j=1,\ldots,b \text{ such that } \text{ME}_{B_j}\neq 0$$

SSB 的零假设分布为:

$$\text{SSB}\overset{H_0}{\sim}\sigma^2\chi^2_{(b-1)}$$

设第一类型错误概率界限为 $\alpha$
我们构造如下的 $F$ 检验统计量:

$$F:=\frac{\text{MSB}}{\text{MSE}_{\text{reduced}}}\overset{H_0}{\sim}\frac{\sigma^2\chi^2_{(b-1)}/(b-1)}{\sigma^2\chi^2_{abn-a-b+1}/(abn-a-b+1)}=F_{b-1,abn-a-b+1}$$

其中分子 MSB 和分母 $\text{MSE}_{\text{reduced}}$ 是相互独立的 (无论零假设 $H_0$ 是否成立)

查表得 $F_{1,26}(0.05)=4.22$
根据 $F=\frac{\text{MSB}}{\text{MSE}_{\text{reduced}}}=\frac{138}{1.65}\approx 83.64>F_{1,26}(0.05)$ 可以拒绝零假设，即 "性别" 因子存在主效应.

## Problem 3

For the one-way ANOVA model:

$$y_{ij} = \mu_i + \varepsilon_{ij} \quad \begin{cases} i = 1, \dots, a \\ j = 1, \dots, n_i \end{cases} \text{where } \{\varepsilon_{ij}\} \overset{\text{i.i.d}}{\sim} N(0, \sigma^2)$$

For given $c_1, \dots, c_a$, propose a test statistic for:

$$H_0 : \frac{\mu_1}{c_1} = \cdots = \frac{\mu_a}{c_a}$$

then find its null distribution and rejection region.

**Solution:**
考虑单因子方差分析的多元线性回归形式:

$$y = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(a)} \end{bmatrix} = \begin{bmatrix} 1_{n_1} & & & \\ & 1_{n_2} & & \\ & & \ddots & \\ & & & 1_{n_a} \end{bmatrix}_{n \times a} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_a \end{bmatrix} + \begin{bmatrix} \varepsilon^{(1)} \\ \varepsilon^{(2)} \\ \vdots \\ \varepsilon^{(a)} \end{bmatrix} = X\mu + \varepsilon$$

其中:

$$y^{(i)} = \begin{bmatrix} y_{i,1} \\ y_{i,2} \\ \vdots \\ y_{i,n_i} \end{bmatrix} \quad \varepsilon^{(i)} = \begin{bmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \\ \vdots \\ \varepsilon_{i,n_i} \end{bmatrix} \quad (i = 1, \dots, a)$$

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

$$\{\varepsilon_{i,j}\} \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2)$$

设计矩阵 $X = 1_{n_1} \oplus \cdots \oplus 1_{n_a} \in \mathbb{R}^{n \times a}$
我们定义投影矩阵为:

$$\begin{aligned} H &:= X(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}} \\ &= \begin{bmatrix} 1_{n_1} & & & \\ & 1_{n_2} & & \\ & & \ddots & \\ & & & 1_{n_a} \end{bmatrix} \begin{bmatrix} n_1 & & & \\ & n_2 & & \\ & & \ddots & \\ & & & n_a \end{bmatrix}^{-1} \begin{bmatrix} 1_{n_1} & & & \\ & 1_{n_2} & & \\ & & \ddots & \\ & & & 1_{n_a} \end{bmatrix}^{\mathrm{T}} \\ &= \begin{bmatrix} \frac{1}{n_1} 1_{n_1} 1_{n_1}^{\mathrm{T}} & & & \\ & \frac{1}{n_2} 1_{n_2} 1_{n_2}^{\mathrm{T}} & & \\ & & \ddots & \\ & & & \frac{1}{n_a} 1_{n_a} 1_{n_a}^{\mathrm{T}} \end{bmatrix} \end{aligned}$$

于是我们有:

$$
\begin{aligned}
\text{SSE}_{\text{full}} &= \|y-\hat{y}\|^2 \\
&= \|y-Hy\|^2 \\
&= \left\| \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(a)} \end{bmatrix} - \begin{bmatrix} 1_{n_1} & & & \\ & 1_{n_2} & & \\ & & \ddots & \\ & & & 1_{n_a} \end{bmatrix}_{n\times a} \begin{bmatrix} \bar{y}_{1\cdot} \\ \bar{y}_{2\cdot} \\ \vdots \\ \bar{y}_{a\cdot} \end{bmatrix} \right\|^2 \\
&= \left\| \begin{bmatrix} y^{(1)} - \bar{y}_{1\cdot}1_{n_1} \\ y^{(2)} - \bar{y}_{2\cdot}1_{n_2} \\ \vdots \\ y^{(a)} - \bar{y}_{a\cdot}1_{n_a} \end{bmatrix} \right\|^2 \\
&= \sum_{i=1}^{a} \|y^{(i)} - \bar{y}_{i\cdot}1_{n_i}\|^2 \\
&= \sum_{i=1}^{a}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{i\cdot})^2 \sim \chi^2_{(n-a)} \quad (\text{where } \bar{y}_{i\cdot} = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij})
\end{aligned}
$$

注意到零假设 $H_0: \frac{\mu_1}{c_1} = \cdots = \frac{\mu_a}{c_a}$ 可以等价表示为:

$$
H_0 : C\mu = \begin{bmatrix} \frac{1}{c_1} & -\frac{1}{c_2} & & & \\ \frac{1}{c_1} & & -\frac{1}{c_3} & & \\ \vdots & & & \ddots & \\ \frac{1}{c_1} & & & & -\frac{1}{c_a} \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_a \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0_{a-1}
$$

根据线性约束检验的结论可知:

$$
\text{ESS} = (C\hat{\mu}_{\text{full}})^{\mathrm{T}}[C(X^{\mathrm{T}}X)^{-1}C^{\mathrm{T}}]^{-1}(C\hat{\mu}_{\text{full}}) \sim \sigma^2\chi^2_{(a-1)}
$$

$$
\text{where } \hat{\mu}_{\text{full}} := \begin{bmatrix} \bar{y}_{1\cdot} \\ \bar{y}_{2\cdot} \\ \cdots \\ \bar{y}_{a\cdot} \end{bmatrix} \text{ and } C = \begin{bmatrix} \frac{1}{c_1} & -\frac{1}{c_2} & & & \\ \frac{1}{c_1} & & -\frac{1}{c_3} & & \\ \vdots & & & \ddots & \\ \frac{1}{c_1} & & & & -\frac{1}{c_a} \end{bmatrix} \in \mathbb{R}^{(a-1)\times a}
$$

$$
\text{ESS} \perp \text{SSE}_{\text{full}}
$$

于是检验统计量为:

$$
F := \frac{\text{EES}/(a-1)}{\text{SSE}_{\text{full}}/(n-a)} \overset{H_0}{\sim} \frac{\sigma^2\chi^2_{(a-1)}/(a-1)}{\sigma^2\chi^2_{(n-a)}/(n-a)} = F_{a-1,n-a}
$$

当 $F > F_{a-1,n-a}(\alpha)$ 时，我们拒绝零假设 $H_0: C\mu = 0$ (即 $H_0: \frac{\mu_1}{c_1} = \cdots = \frac{\mu_a}{c_a}$)

# Problem 4

Simultaneous confidence interval for linear regression $y = X\beta + \varepsilon$ where $\varepsilon \sim N(0_n, \sigma^2 I_n)$:
Find $M_\alpha > 0$ such that:

$$
\mathrm{P}\left\{ \max_{c\in\mathbb{R}^{p+1}} \frac{|c^{\mathrm{T}}(\hat{\beta}-\beta)|}{s\sqrt{c^{\mathrm{T}}(X^{\mathrm{T}}X)^{-1}c}} \geq M_\alpha \right\} = \alpha \quad (\text{where } \alpha \in (0,1))
$$

Hint: use Cauchy-Schwarz Inequality.

**Solution:**
根据 Cauchy-Schwarz 不等式我们有:

$$
\begin{aligned}
|c^{\mathrm{T}}(\hat{\beta}-\beta)| &= |c^{\mathrm{T}}(X^{\mathrm{T}}X)^{-\frac{1}{2}}(X^{\mathrm{T}}X)^{\frac{1}{2}}(\hat{\beta}-\beta)| \\
&\leq \|(X^{\mathrm{T}}X)^{-\frac{1}{2}}c\|\|(X^{\mathrm{T}}X)^{\frac{1}{2}}(\hat{\beta}-\beta)\| \\
&= \sqrt{c^{\mathrm{T}}(X^{\mathrm{T}}X)^{-1}c} \cdot \|(X^{\mathrm{T}}X)^{\frac{1}{2}}(\hat{\beta}-\beta)\|
\end{aligned}
$$

当且仅当 $(X^{\mathrm{T}}X)^{-\frac{1}{2}}c$ 与 $(X^{\mathrm{T}}X)^{\frac{1}{2}}(\hat{\beta}-\beta)$ 线性相关时取等.
因此我们有:

$$\max_{c\in\mathbb{R}^{p+1}}\frac{|c^{\mathrm{T}}(\hat\beta-\beta)|}{s\sqrt{c^{\mathrm{T}}(X^{\mathrm{T}}X)^{-1}c}}=\frac{\|(X^{\mathrm{T}}X)^{\frac{1}{2}}(\hat\beta-\beta)\|}{s}$$

根据多元线性回归的结论我们有:

$$\hat\beta=(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y\sim N(\beta,\sigma^2(X^{\mathrm{T}}X)^{-1})$$
$$s^2\sim\sigma^2\frac{\chi^2_{(n-p-1)}}{n-p-1}$$
$$\hat\beta\perp s^2$$

于是我们有:

$$(X^{\mathrm{T}}X)^{\frac{1}{2}}(\hat\beta-\beta)\sim N(0_{p+1},\sigma^2I_{p+1})$$
$$\frac{\|(X^{\mathrm{T}}X)^{\frac{1}{2}}(\hat\beta-\beta)\|^2}{s^2}\sim\frac{\sigma^2\chi^2_{(p+1)}}{\sigma^2\chi^2_{(n-p-1)}/(n-p-1)}=(p+1)\frac{\chi^2_{(p+1)}/(p+1)}{\chi^2_{(n-p-1)}/(n-p-1)}=(p+1)F_{p+1,n-p-1}$$

因此我们有:

$$\begin{aligned}\alpha&=\mathrm{P}\left\{\max_{c\in\mathbb{R}^{p+1}}\frac{|c^{\mathrm{T}}(\hat\beta-\beta)|}{s\sqrt{c^{\mathrm{T}}(X^{\mathrm{T}}X)^{-1}c}}\geq M_\alpha\right\}\\&=\mathrm{P}\left\{\frac{\|(X^{\mathrm{T}}X)^{\frac{1}{2}}(\hat\beta-\beta)\|}{s}\geq M_\alpha\right\}\\&=\mathrm{P}\left\{\frac{\|(X^{\mathrm{T}}X)^{\frac{1}{2}}(\hat\beta-\beta)\|^2}{s^2}\geq M_\alpha^2\right\}\\&=\mathrm{P}\left\{(p+1)F_{p+1,n-p-1}\geq M_\alpha^2\right\}\\&=\mathrm{P}\left\{F_{p+1,n-p-1}\geq\frac{1}{p+1}M_\alpha^2\right\}\end{aligned}$$

这表明:

$$\frac{1}{p+1}M_\alpha^2=F_{p+1,n-p-1}(\alpha)$$
$$\Updownarrow$$
$$M_\alpha=\sqrt{(p+1)F_{p+1,n-p-1}(\alpha)}$$

# Problem 5

Suppose that $Y_i$ is generated from PDF $f_{\lambda_i}(y)=e^{\lambda_iy-r(\lambda_i)}f_0(y)$ where $\lambda_i\in\Gamma$

$$\mathrm{E}[Y_i]=g(x_i^{\mathrm{T}}\beta)$$

- ① find $\mathrm{E}[Y_i]$ (Hint: use MGF)
- ② find link function $g(\cdot)$ such that $\lambda_i=x_i^{\mathrm{T}}\beta$
- ③ find the log-likelihood function (utilize ①②)
- ④ find the MLE

**Solution:**
容易验证 $r(\lambda_i)=\log\left(\int e^{\lambda_iy}f_0(y)\mathrm{d}y\right)$
这可以保证 PDF $f_{\lambda_i}(y)$ 在 $(-\infty,\infty)$ 上的积分为 $1$:

$$\begin{aligned}\int f_{\lambda_i}(y)\mathrm{d}y&=\int e^{\lambda_iy-r(\lambda_i)}f_0(y)\mathrm{d}y\\&=e^{-r(\lambda_i)}\int e^{\lambda_iy}f_0(y)\mathrm{d}y\\&=\exp\left\{-\log\left(\int e^{\lambda_iy}f_0(y)\mathrm{d}y\right)\right\}\int e^{\lambda_iy}f_0(y)\mathrm{d}y\\&=1\end{aligned}$$

## (1) Find $\mathrm{E}[Y_i]$

$Y_i$ 的矩母函数为:

$$
\begin{aligned}
M_{Y_i}(t) &= \mathrm{E}[e^{tY_i}] \\
&= \int e^{ty} f_{\lambda_i}(y)\mathrm{d}y \\
&= \int e^{ty} e^{\lambda_i y - r(\lambda_i)} f_0(y)\mathrm{d}y \\
&= e^{-r(\lambda_i)} \int e^{(\lambda_i + t)y} f_0(y)\mathrm{d}y \\
&= e^{-r(\lambda_i)} e^{r(\lambda_i + t)}
\end{aligned}
$$

于是我们有:

$$
\begin{aligned}
\mathrm{E}[Y_i] &= \frac{\mathrm{d}}{\mathrm{d}t} M_{Y_i}(t)\Big|_{t=0} \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\left\{e^{-r(\lambda_i)} e^{r(\lambda_i+t)}\right\}\Big|_{t=0} \\
&= \left\{e^{-r(\lambda_i)} e^{r(\lambda_i+t)} r'(\lambda_i + t)\right\}\Big|_{t=0} \\
&= r'(\lambda_i)
\end{aligned}
$$

## (2) Find $g(\cdot)$ such that $\lambda_i = x_i^{\mathrm{T}}\beta$

根据 $(1)$ 的结论可知 $\mathrm{E}[Y_i] = r'(\lambda_i) = g(x_i^{\mathrm{T}}\beta)$
欲令 $\lambda_i = x_i^{\mathrm{T}}\beta$，我们需要对于任意 $\lambda_i \in \Gamma$ 都有 $g(\lambda_i) = r'(\lambda_i)$ 成立.
因此 $g(\cdot) = r'(\cdot)$

## (3) Find the log-likelihood function

$Y_i$ 的似然函数为:

$$
L(\beta) := \prod_{i=1}^{n} f_{\lambda_i}(y_i) = \prod_{i=1}^{n} e^{\lambda_i y_i - r(\lambda_i)} f_0(y_i)
$$

于是 $Y_i$ 的对数似然函数为:

$$
\begin{aligned}
l(\beta) &:= \log\left(L(\beta)\right) \\
&= \sum_{i=1}^{n}[\lambda_i y_i - r(\lambda_i) + \log\left(f_0(y_i)\right)] \quad (\text{substitute } \lambda_i = x_i^{\mathrm{T}}\beta) \\
&= \sum_{i=1}^{n}[x_i^{\mathrm{T}}\beta y_i - r(x_i^{\mathrm{T}}\beta) + \log\left(f_0(y_i)\right)]
\end{aligned}
$$

丢弃与 $\beta$ 的无关项，记 $l(\beta) := \sum_{i=1}^{n}[x_i^{\mathrm{T}}\beta y_i - r(x_i^{\mathrm{T}}\beta)]$

## (4) Find the MLE

$\hat{\beta}_{\mathrm{MLE}}$ 是驻点方程 $\nabla_\beta l(\beta) = \sum_{i=1}^{n}[y_i - r'(x_i^{\mathrm{T}}\beta)]x_i = \sum_{i=1}^{n}(y_i - \mathrm{E}[Y_i])x_i = 0$ 的解.
我们可以使用梯度法进行数值逼近.

**The End**