1、对X进行标准化，对y做中心化.

$$x_{ij}^{**} = \frac{x_{ij}^{*}}{\sqrt{L_{jj}}} = \frac{x_{ij} - \bar{x}_j}{\sqrt{L_{jj}}}, \quad \bar{x}_j = \sum_{i=1}^{n} x_{ij} / n$$

$$y_i^{*} = y_i - \bar{y} \qquad \bar{y} = n^{-1} \sum^{n} y_i$$

$$L_{jj} = \sum_{i}^{n} (x_{ij} - \bar{x}_j)^2$$

所以 $X_S = X_C L$，$L = \mathrm{diag}\left\{ \frac{1}{\sqrt{L_{11}}}, \cdots, \frac{1}{\sqrt{L_{pp}}} \right\}$

对X标准化，对y中心化后的回归系数：

$$\hat{\beta}_{s,slope} = (X_s' X_s)^{-1} X_s' y^{*}$$

将 $X_S = X_C L$ 代入.

$$\begin{aligned} \text{所以 } \hat{\beta}_{s,slope} &= \left((X_C L)' (X_C L) \right)^{-1} (X_C L)^T y^{*} \\ &= (L^T X_C^T X_C L)^{-1} L^T X_C^T y^{*} \\ &= L^{-1} (X_C' X_C)^{-1} X_C' y^{*} \end{aligned}$$

又 $\hat{\beta}_{c,slope} = (X_C^T X_C)^{-1} X_C^T y^{*}$

故 $\hat{\beta}_{s,slope} = L^{-1} \hat{\beta}_{c,slope}$

(2) $\hat{\beta}_{s,j} = \sqrt{L_{jj}}\, \hat{\beta}_{cj}$，根据上式求分量的期望和方差：

$$E[\hat{\beta}_{sj}] = \sqrt{L_{jj}}\, E[\hat{\beta}_{cj}] = \sqrt{L_{jj}}\, \beta_j$$

$$Var[\hat{\beta}_{sj}] = L_{jj}\, Var[\hat{\beta}_{cj}] = L_{jj}\, \sigma^2 (X^T X)^{-1}_{jj}$$

表示 $\sigma^2 (X^T X)^{-1}$ 的j行j列元素

2、用回归解释 one-way ANOVA：

$y_{ij}=\mu_i+\epsilon_{ij}$，$i=1,\dots,a$；$j=1,\dots,m$ （均值模型），$\epsilon_{ij}\overset{iid}{\sim}N(0,\sigma^2)$

(1) 设计矩阵 X：

	样本均值
Population 1： $y_{11}, y_{12}, \dots, y_{1m}=\vec{y}_1^{\,T}$	$\bar{y}_1$
Population 2： $y_{21}, y_{22}, \dots, y_{2m}=\vec{y}_2^{\,T}$	$\bar{y}_2$
⋯	⋯
Population a： $y_{a1}, y_{a2}, \dots, y_{am}=\vec{y}_a^{\,T}$	$\bar{y}_a$

这里将列向量 $\vec{y}_i$ 堆叠起来，拉成 $n=ma$ 维的列向量即可构造出 design matrix X.

$$Y=\begin{pmatrix}\vec{y}_1\\ \vec{y}_2\\ \vdots\\ \vec{y}_a\end{pmatrix}\quad \text{设计矩阵 } X=\begin{pmatrix}\vec{1}_m & 0 & 0 & \cdots & 0\\ 0 & \vec{1}_m & 0 & \cdots & 0\\ 0 & 0 & \vec{1}_m & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & \vec{1}_m\end{pmatrix}$$

，X是 $n\times a$ 维，列满秩矩阵.

(2) 从数据层面写出回归模型：$Y=X\beta+\epsilon$

由(1)的设计矩阵可得：

$$\underbrace{\begin{pmatrix}\vec{y}_1\\ \vec{y}_2\\ \vdots\\ \vec{y}_a\end{pmatrix}_{n\times 1}}_{\text{【响应变量】}}=\underbrace{\begin{pmatrix}\vec{1}_m & 0 & 0 & \cdots & 0\\ 0 & \vec{1}_m & 0 & \cdots & 0\\ 0 & 0 & \vec{1}_m & \cdots & 0\\ \cdots & \cdots & \cdots & \cdots & \cdots\\ 0 & 0 & 0 & \cdots & \vec{1}_m\end{pmatrix}_{n\times a}}_{\text{【设计矩阵】}}\underbrace{\begin{pmatrix}\mu_1\\ \mu_2\\ \mu_3\\ \cdots\\ \mu_a\end{pmatrix}_{a\times 1}}_{\text{【回归参数】}}+\underbrace{\vec{\epsilon}}_{\text{【误差】}}$$

（这里用 $\vec{\mu}$ 代替 $\vec{\beta}$ 符号表示，以符合 one-way ANOVA 均值模型的 Notation 的选择）

$\vec{y}_i^{\,T}=(y_{i1}, y_{i2}, \cdots, y_{im})$

$\vec{\epsilon}=(\epsilon_{11},\epsilon_{12},\epsilon_{13},\dots,\epsilon_{ij},\dots\ \epsilon_{am})^T$，$\epsilon_{ij}\overset{iid}{\sim}N(0,\sigma^2)$

(3) 从回归OLSE的角度估计回归参数 $\hat{\vec{\mu}}$

$$\frac{d\sum_{i=1}^{a}\sum_{j=1}^{m}(y_{ij}-\mu_i)^2}{d\mu_i}=0$$

，所以 $\hat{\vec{\mu}}$ 的每一个分量 $\hat{\mu}_i=\sum_{j=1}^{m}\frac{y_{ij}}{m}=\bar{y}_{i\cdot}$

与 one-way ANOVA 中应用MLE的参数估计相同

(4) 因为已将 one-way anova 写成 $Y=X\beta+\epsilon$ 了，所以从回归F检验的角度可得：

$F_0=\dfrac{SSR/p}{SSE/(n-p-1)}$，这里 $p=a-1$，~~$F_0\sim F$~~ 在H_0成立的条件下 F_0服从自由度为$(a-1, n-a)$

$$F_0=\frac{SSR/p}{SSE/(n-p-1)}=\frac{\sum^a\sum^m(\hat{y_{ij}}-\bar{y})^2/a}{\sum^a\sum^m(y_{ij}-\hat{y_{ij}})^2/(n-a-1)}$$ （证明二者等价）

$$\frac{\sum^a\sum^m(\hat{y_{ij}}-\bar{y})^2/a}{\sum^a\sum^m(y_{ij}-\bar{y_{ij}})^2/(n-a-1)}=\frac{\sum^a\sum^m(\hat{y_{ij}}-\bar{y})^2}{\sum^a\sum^m(y_{ij}-\bar{y_{ij}})^2}\cdot\frac{n-a-1}{a}$$

$$=\frac{\sum^a\sum^m(\hat{y_{ij}}-\bar{y})^2}{\sum^a\sum^m(y_{ij}-\bar{y_{ij}})^2}\cdot\frac{n-a}{a-1}\cdot\frac{a-1}{n-a}\cdot\frac{n-a-1}{a}.$$

one-way ANOVA中，检验统计量为：$\frac{\sum^a\sum^m(\hat{y_{ij}}-\bar{y})^2/a-1}{\sum^a\sum^m(y_{ij}-\bar{y_{ij}})^2/n-a}=F$

所以 $F_0=\frac{(a-1)(n-a-1)}{(n-a)a}F$，这里F为one-way ANOVA.

所以二者等价。