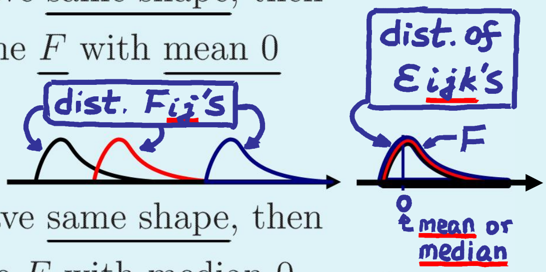
**Definition 4** (Alternative expressions of the model ($\nabla$) in LNp.38)

cf. LNp.2 & LN, CH11, p.2

- expression 1: Let $\epsilon_{ijk} = Y_{ijk} - \mu_{ij}$. Then, for $i = 1, \ldots, I$, $j = 1, \ldots, J$, $k = 1, \ldots, K_{ij}$, 規律 (parameters) $Y_{ijk} = \mu_{ij} + \epsilon_{ijk}$, 隨機 (random variables)
  where $\epsilon_{ijk}$'s are independent errors and $\epsilon_{ij1}, \ldots \epsilon_{ijK_{ij}} \sim$ i.i.d. $F_{ij}(x + \mu_{ij})$
  - If $\mu_{ij}$'s are the means of $F_{ij}$'s, then
    * $E(\epsilon_{ijk}) = 0$ for any $i, j, k$, and
    * if we further assume that all $F_{ij}$'s have same shape, then
      $\epsilon_{ijk}$'s $\sim$ i.i.d. from some $F$ with mean 0
  - If $\mu_{ij}$'s are the medians of $F_{ij}$'s, then
    * all $\epsilon_{ijk}$'s have median 0, and
    * if we further assume that all $F_{ij}$'s have same shape, then
      $\epsilon_{ijk}$'s $\sim$ i.i.d. from some $F$ with median 0

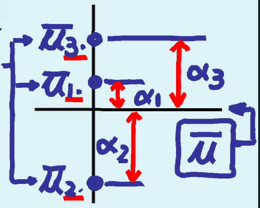
joint dist. of $\epsilon_{ijk}$'s is like the 1-sample data




cf. expression 2 in LNp.3

- expression 2
  - Define $\bar{\mu} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\mu_{ij}$ : grand average of all $\mu_{ij}$'s
  - For $i = 1, \ldots, I$, define (for A)

$\left(\frac{1}{I}\sum_{i=1}^{I}\bar{\mu}_{i\cdot} = \bar{\mu}\right)$ $\quad \bar{\mu}_{i\cdot} = \frac{1}{J}\sum_{j=1}^{J}\mu_{ij}$ and $\alpha_i = \bar{\mu}_{i\cdot} - \bar{\mu}$ $(\Leftrightarrow \bar{\mu}_{i\cdot} = \bar{\mu} + \alpha_i)$.



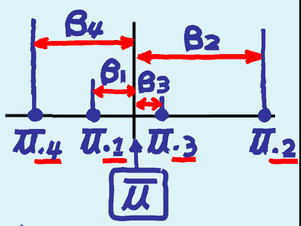

  * $\alpha_i$ : the **main effect** of the $i$th level of factor $A$



  * $\sum_{i=1}^{I}\alpha_i = \left(\sum_{i=1}^{I}\bar{\mu}_{i\cdot}\right) - I\bar{\mu} = 0$ → $\alpha_i$'s: $I$ parameters, but, dim($\alpha_i$'s) = $I-1$
- For $j = 1, \ldots, J$, define (for B)

$\left(\frac{1}{J}\sum_{j=1}^{J}\bar{\mu}_{\cdot j} = \bar{\mu}\right)$ $\quad \bar{\mu}_{\cdot j} = \frac{1}{I}\sum_{i=1}^{I}\mu_{ij}$ and $\beta_j = \bar{\mu}_{\cdot j} - \bar{\mu}$ $(\Leftrightarrow \bar{\mu}_{\cdot j} = \bar{\mu} + \beta_j)$.

  * $\beta_j$: the **main effect** the $j$th level of factor $B$
  * $\sum_{j=1}^{J}\beta_j = \left(\sum_{j=1}^{J}\bar{\mu}_{\cdot j}\right) - J\bar{\mu} = 0$ → $\beta_j$'s: $J$ parameters, but, dim($\beta_j$'s) = $J-1$




# of factors
0 $\quad \bar{\mu}$
1 $\quad \alpha_i$'s $(\bar{\mu}_{i\cdot} - \bar{\mu})$ $\quad \beta_j$'s $(\bar{\mu}_{\cdot j} - \bar{\mu})$
2 $\quad \delta_{ij}$'s $(\mu_{ij} - \alpha_i - \beta_j - \bar{\mu})$

- For $i = 1, \ldots, I$, $j = 1, \ldots, J$, define

$$\begin{aligned}\delta_{ij} &= (\mu_{ij} - \bar{\mu}) - (\alpha_i + \beta_j) \\ &= (\mu_{ij} - \bar{\mu}) - [(\bar{\mu}_{i\cdot} - \bar{\mu}) + (\bar{\mu}_{\cdot j} - \bar{\mu})] = \mu_{ij} - \bar{\mu}_{i\cdot} - \bar{\mu}_{\cdot j} + \bar{\mu}\end{aligned}$$

$(\Leftrightarrow \mu_{ij} = \bar{\mu} + \alpha_i + \beta_j + \delta_{ij})$

$\bar{\mu}_{\cdot j}$'s, $\beta_j$'s: parameters

parameters → * $\delta_{ij}$: the **interaction** (effect) between the $i$th level of factor $A$ and the $j$th level of factor $B$

$\Sigma_i \Sigma_j \delta_{ij} = 0$

  * For $j = 1, \ldots, J$, $\sum_{i=1}^{I}\delta_{ij} = (I\bar{\mu}_{\cdot j} - I\bar{\mu}) - I\beta_j = 0$, and
    for $i = 1, \ldots, I$, $\sum_{j=1}^{J}\delta_{ij} = (J\bar{\mu}_{i\cdot} - J\bar{\mu}) - J\alpha_i = 0$

→ $\delta_{ij}$'s: $I \times J$ parameters but, dim($\delta_{ij}$'s) $= IJ - I - J + 1 = (I-1)(J-1)$

- Then, we have the model: 規律

1-way layout $\xleftrightarrow{\text{cf.}}$ $Y_{ijk} = \mu_{ij} + \epsilon_{ijk} = \bar{\mu} + \overbrace{\alpha_i}^{\text{I}} + \overbrace{\beta_j}^{\text{II}} + \overbrace{\delta_{ij}}^{\text{III}} + \underbrace{\epsilon_{ijk}}_{\text{隨機}}$,

where $\epsilon_{ijk}$'s are distributed as above. In this model, all of them are parameters

* there are $I$ different $\alpha_i$'s, but they are of dimension $I-1$,
* there are $J$ different $\beta_j$'s, but they are of dimension $J-1$,
* there are $IJ$ different $\delta_{ij}$'s, but they are of dimension

(×)

dim($\mu_{ij}$'s) $IJ - I - J + 1 = IJ - (I-1) - (J-1) - 1 = (I-1)(J-1)$.

dim($\alpha_i$'s) dim($\beta_j$'s) dim($\bar{\mu}$)

table in LNp.39 cf. — connections between the parameters

$\bar{\mu}_{i_1\cdot} - \bar{\mu}_{i_2\cdot} = \alpha_{i_1} - \alpha_{i_2}$

| $A$ \ $B$ | 1 | 2 | $\cdots$ | $J$ | |
|---|---|---|---|---|---|
| 1 | $\mu_{11} = \bar{\mu} + \boxed{\alpha_1} + \boxed{\beta_1} + \boxed{\delta_{11}}$ | $\mu_{12} = \bar{\mu} + \boxed{\alpha_1} + \beta_2 + \boxed{\delta_{12}}$ | | $\mu_{1J} = \bar{\mu} + \boxed{\alpha_1} + \beta_J + \boxed{\delta_{1J}}$ | $\bar{\mu}_{1\cdot} = \bar{\mu} + \alpha_1$ |
| $\vdots$ | $\vdots$ | $\vdots$ | | $\vdots$ | $\vdots$ |
| $I$ | $\mu_{I1} = \bar{\mu} + \alpha_I + \boxed{\beta_1} + \boxed{\delta_{I1}}$ | $\mu_{I2} = \bar{\mu} + \alpha_I + \beta_2 + \delta_{I2}$ | $\cdots$ | $\mu_{IJ} = \bar{\mu} + \alpha_I + \beta_J + \delta_{IJ}$ | $\bar{\mu}_{I\cdot} = \bar{\mu} + \alpha_I$ |
| | $\bar{\mu}_{\cdot 1} = \bar{\mu} + \beta_1$ | $\bar{\mu}_{\cdot 2} = \bar{\mu} + \beta_2$ | $\cdots$ | $\bar{\mu}_{\cdot J} = \bar{\mu} + \beta_J$ | |

average

average

(+) = 0

(+) = 0

(+) = 0

(+) = 0

$\bar{\mu}_{\cdot j_1} - \bar{\mu}_{\cdot j_2} = \beta_{j_1} - \beta_{j_2}$

**Example 6** (Interpretation of the main effects of $A$ and $B$ and their interactions)

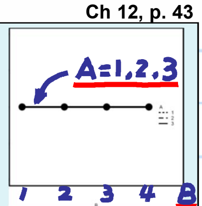
Consider a 3-level factor $A$ ($I=3$) and a 4-level factor $B$ ($J=4$).

- There are $3 \times 4 = 12$ level combinations $\Rightarrow$ 12 means ($\mu_{ij}$'s)
- Consider the following models (1)-(5) and

12 samples



compare the *interaction plots* of their $\mu_{ij}$'s:




1-sample data cf. (1). $Y_{ijk} = \bar{\mu} + \epsilon_{ijk}$ ($\Leftrightarrow$ all $\alpha_i$'s $=0$, all $\beta_j$'s $=0$, all $\delta_{ij}$'s $=0$)

– $\dim(E(Y_{ijk})\text{'s}) = 1$

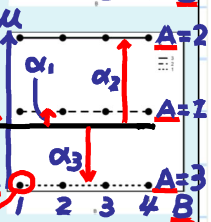
model for 1-way layout

cf. (2). $Y_{ijk} = \bar{\mu} + \alpha_i + \epsilon_{ijk}$ ($\Leftrightarrow$ all $\beta_j$'s $=0$, all $\delta_{ij}$'s $=0$)

– $\dim(E(Y_{ijk})\text{'s}) = 1 + (3-1) = 3$

cf. (3). $Y_{ijk} = \bar{\mu} + \beta_j + \epsilon_{ijk}$ ($\Leftrightarrow$ all $\alpha_i$'s $=0$, all $\delta_{ij}$'s $=0$)

– $\dim(E(Y_{ijk})\text{'s}) = 1 + (4-1) = 4$




$\bar{\mu}_{2\cdot}$ $\bar{\mu}$ $\bar{\mu}_{1\cdot}$ $\bar{\mu}_{3\cdot}$ $\mu_{31}$

(4). $Y_{ijk} = \bar{\mu} + \alpha_i + \beta_j + \epsilon_{ijk}$ ($\Leftrightarrow$ all $\delta_{ij}$'s $=0$)

– $\dim(E(Y_{ijk})\text{'s}) = 1 + (3-1) + (4-1) = 6$

– This model is called a *main-effect(-only)* model (MEM), a simple **additive model**

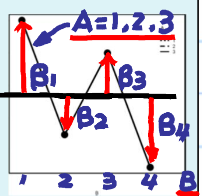
– Under a MEM, the lines in the interaction plot of $\mu_{ij}$'s are parallel.

– Under a MEM,

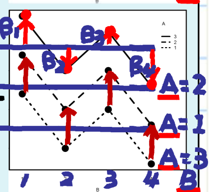
  * $\alpha_{i_1} - \alpha_{i_2} = \bar{\mu}_{i_1\cdot} - \bar{\mu}_{i_2\cdot} = \mu_{i_1 j} - \mu_{i_2 j}$ for any $j$
  * $\beta_{j_1} - \beta_{j_2} = \bar{\mu}_{\cdot j_1} - \bar{\mu}_{\cdot j_2} = \mu_{i j_1} - \mu_{i j_2}$ for any $i$

if we fix A at level i and change the level of B

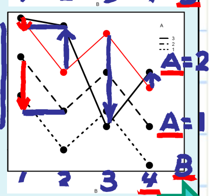
$\bar{\mu}_{\cdot 1}$ $\bar{\mu}_{\cdot 3}$ $\bar{\mu}$ $\bar{\mu}_{\cdot 2}$ $\bar{\mu}_{\cdot 4}$




$\alpha_i$'s $\bar{\mu}_{2\cdot}$ $\bar{\mu}_{1\cdot}$ $\bar{\mu}_{3\cdot}$



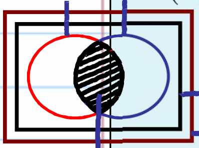

add 3 (J−1) parameters




interaction

**Q**: How to modify the MEM to allow for nonparallel lines, an indication that the effects of $A$ depends on $B$?

(2) (3)



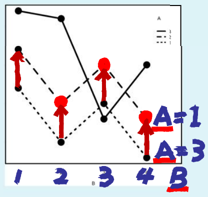
(4)

(5)

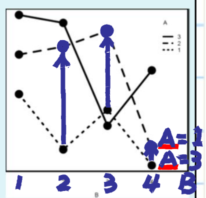
(1)

(5). $Y_{ijk} = \overline{\mu} + \alpha_i + \beta_j + \delta_{ij} + \epsilon_{ijk}$

– $\dim(E(Y_{ijk})\text{'s}) = 1 + (3-1) + (4-1) + (3-1)\times(4-1) = 12$









- **Note.** Model (1) $\subset \left\{ \begin{array}{l} \text{Model (2)} \\ \text{Model (3)} \end{array} \right\} \subset$ Model (4) $\subset$ Model (5)

(2)∩(3)

add $2\times 3 = 6$ $((I-1)(J-1))$ parameters

**Definition 5** (Interaction plot based on data)

- Notice that $\mu_{ij}$'s are parameters, which are unknown in practice.
- Because $\overline{Y}_{ij\cdot} = \frac{1}{K_{ij}} \sum_{k=1}^{K_{ij}} Y_{ijk} \xrightarrow{e} \mu_{ij}$, we can replace $\mu_{ij}$ by $\overline{Y}_{ij\cdot}$.
- If the lines in the interaction plot of $\overline{Y}_{ij\cdot}$'s are nearly parallel, it is an indication of no interaction between $A$ and $B$.

**Definition 6** (no effects) This is $H_0$ of ANOVA if treat $IJ$ samples as 1-way layout

Define When all the 3 conditions hold $\Rightarrow$ all $\mu_{ij}$'s $= \overline{\mu}$ (model (1))

- factor $A$ has no (main) effect on $Y$'s

  $\Leftrightarrow \quad \overline{\mu}_{1\cdot} = \cdots = \overline{\mu}_{I\cdot} = \overline{\mu} \quad \Leftrightarrow \quad \alpha_1 = \cdots = \alpha_I = 0$ $(\alpha_i = \overline{\mu}_{i\cdot} - \overline{\mu})$
- factor $B$ has no (main) effect on $Y$'s

  $\Leftrightarrow \quad \overline{\mu}_{\cdot 1} = \cdots = \overline{\mu}_{\cdot J} = \overline{\mu} \quad \Leftrightarrow \quad \beta_1 = \cdots = \beta_J = 0$ $(\beta_j = \overline{\mu}_{\cdot j} - \overline{\mu})$
- factors $A$ and $B$ have no interaction on $Y$'s $\quad \Leftrightarrow \quad$ all $\delta_{ij}$'s $= 0$



**Definition 7** (balanced data)

In the two-way layout, a set of data is called balanced if all the $IJ$ samples have equal sample size $K$, i.e.,

$$K_{11} = \cdots = K_{1J} = \cdots = K_{I1} = \cdots = K_{IJ} = K.$$

bring in "orthogonality" between the SS's for $\alpha_i$'s, $\beta_j$'s, $\delta_{ij}$'s.

- **Normal theory for the Two-Way Layout**

Consider the model ($\nabla$) in LNp.38.

- Assume that
  - $F_{ij}$'s are normal distributions
  - $F_{ij}$'s have same variance $\sigma^2$
- Thus, the statistical model is: for $1 \leqslant i \leqslant I$, $1 \leqslant j \leqslant J$, $1 \leqslant k \leqslant K_{ij}$,

  $(i,j)$th sample $\rightarrow$ $Y_{ijk} \sim$ i.i.d. $N(\mu_{ij}, \sigma^2)$, allow different means for different samples $\quad (\square)$

  and all $Y_{ijk}$'s are independent.

  (# of parameters $= IJ + 1$, including all $\mu_{ij}$'s and $\sigma^2$)

can write down joint pdf & likelihood

- Alternative expressions of this model: for $1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J, 1 \leqslant k \leqslant K_{ij}$,
  - expression 1: $Y_{ijk} \sim N(\overline{\mu}+\alpha_i+\beta_j+\delta_{ij}, \sigma^2)$, and all $Y_{ijk}$'s are independent,
  - expression 2: $Y_{ijk} = \underbrace{\overline{\mu} + \alpha_i + \beta_j + \delta_{ij}}_{規律} + \underbrace{\epsilon_{ijk}}_{隨機}$, and $\epsilon_{ijk}$'s $\sim$ i.i.d. $N(0, \sigma^2)$,

  where

  * $\sum_{i=1}^{I} \alpha_i = 0$ and $\sum_{j=1}^{J} \beta_j = 0$,

bring in "orthogonality"

* $\sum_{j=1}^{J}\delta_{ij}=0$, for $i=1,\ldots,I$, and $\sum_{i=1}^{I}\delta_{ij}=0$, for $j=1,\ldots,J$.

( in these two expressions, # of parameters $=2+I+J+IJ$, including $\overline{\mu}$, all $\alpha_i$'s, all $\beta_j$'s, all $\delta_{ij}$'s, and $\sigma^2$, but they are of dimension $IJ+1$ )

◉ (**Balanced condition**) For simplicity, assume that any sample sizes $K_{ij}=K$. Then, the number of all observations is $N\equiv\sum_{i=1}^{I}\sum_{j=1}^{J}K_{ij}=IJK$.

**Theorem 13** (log-likelihood of the model (□)) normal pdf : $(2\pi\sigma^2)^{-1/2}\exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right]$

Under the model (□), the log-likelihood is proportional to (exercise)

$$l(\mu_{11},\ldots,\mu_{IJ},\sigma^2)\ \propto\ -\frac{IJK}{2}\log(\sigma^2)-\frac{1}{2\sigma^2}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk}-\mu_{ij})^2$$

$$=-\frac{IJK}{2}\log(\sigma^2)-\frac{1}{2\sigma^2}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk}^2-2\,\mu_{ij}\,Y_{ijk}+\mu_{ij}^2)$$

$$=\underbrace{-\frac{1}{2\sigma^2}}_{\dim=1}\overbrace{\left(\sum_i\sum_j\sum_k Y_{ijk}^2\right)}^{\text{data}}+\underbrace{\sum_i\sum_j\left[\frac{\mu_{ij}}{\sigma^2}\right]}_{\text{parameters};\ \dim=I*J}\overbrace{\left(\sum_k Y_{ijk}\right)}^{\text{data}}-\underbrace{K\sum_i\sum_j\frac{\mu_{ij}^2}{2\sigma^2}-\frac{IJK}{2}\log(\sigma^2)}_{\text{parameters}}$$

$\in(IJ+1)$-parameter exponential family

**Note 8** (Some notes about the log-likelihood of model (□) in LNp.45)

- The log-likelihood can be written as a function of $\overline{\mu}$, $\alpha_i$'s, $\beta_j$'s, $\delta_{ij}$'s and $\sigma^2$ by substituting $\overline{\mu}+\alpha_i+\beta_j+\delta_{ij}$ for $\mu_{ij}$, but note that the $\alpha_i$'s, $\beta_j$'s, $\delta_{ij}$'s must meet some linear constraints (given in LNp.41).



- From the log-likelihood, we have
  - for any $1\leqslant i\leqslant I, 1\leqslant j\leqslant J$, $\dfrac{\partial l}{\partial\mu_{ij}}=\dfrac{1}{\sigma^2}\left[\left(\sum_{k=1}^{K}Y_{ijk}\right)-K\times\mu_{ij}\right]=0 \rightarrow \hat{\mu}_{ij,MLE}$
  - $\dfrac{\partial l}{\partial\sigma^2}=-\dfrac{\overbrace{IJK}^{N}}{2\sigma^2}+\dfrac{1}{2\sigma^4}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(Y_{ijk}-\mu_{ij})^2=0\rightarrow\hat{\sigma}^2_{MLE}$

12/18

**Theorem 14** (UMVUE and MLE of the parameters in the model (□) in LNp.45)

- Let $R_{ij}=\sum_k Y_{ijk}$ for $1\leq i\leq I, 1\leq j\leq J$, and $R_\sigma=\sum_i\sum_j\sum_k Y_{ijk}^2$. Then,

$$(R_{11},\ldots,R_{IJ},R_\sigma)$$

is sufficient and complete (**Hint**. $(IJ+1)$-parameter exponential family) LNp.46

◉ $\overline{Y}_{ij\cdot}=R_{ij}/K$ is the UMVUE (by Lehmann-Scheffe Thm) and MLE of $\mu_{ij}$
  - Note that $\hat{\mu}_{ij}\equiv\overline{Y}_{ij\cdot}$ is the *sample mean* of the $(i,j)$th sample $Y_{ij1},\ldots,Y_{ijK}$

◉ By the *invariance property* of MLE, we have

check LN,CH10,p.10 $\hat{\theta}_{MLE}\xrightarrow{e}\theta\rightarrow g(\theta)$ $\Rightarrow$ MLE of $g(\theta)$ is $g(\hat{\theta}_{MLE})$

check LNp.40

⊖ The MLE of $\overline{\mu}$ is

$$\hat{\overline{\mu}}=\frac{1}{IJ}\left(\sum_i\sum_j\hat{\mu}_{ij}\right)=\frac{1}{IJK}\left(\sum_i\sum_j\sum_k Y_{ijk}\right)\equiv\overline{Y}_{\cdots}$$ check tables in LNp.42 & 39

⊖ The MLE of $\overline{\mu}_{i\cdot}$ is $\hat{\overline{\mu}}_{i\cdot}=\frac{1}{J}\left(\sum_j\hat{\mu}_{ij}\right)=\frac{1}{JK}\left(\sum_j\sum_k Y_{ijk}\right)\equiv\overline{Y}_{i\cdot\cdot}$, $i=1,\ldots,I$

The MLE of $\alpha_i$ is $\hat{\alpha}_i=\hat{\overline{\mu}}_{i\cdot}-\hat{\overline{\mu}}=\overline{Y}_{i\cdot\cdot}-\overline{Y}_{\cdots}$, $i=1,\ldots,I$ check table in LNp.42