ST-302
Design, Planning and Analysis of Experiments

Dr. (Mrs.) Kirtee Kiran Kamalja
Department of Statistics, School of Mathematical Sciences,
Kavayitri Bahinbai Chaudhari North Maharashtra University,
Jalgaon

Topic 4

Two-way Classification Model with interaction

About the experiments for two-way classification model

Objective: To study the effect of two factors on the response variable and test whether the two factors have joint effect.

Factor	No. of Levels	Levels
A	p	$A_1, A_2, \ldots, A_p$
B	q	$B_1, B_2, \ldots, B_q$

Experiments: Corresponding to all possible combinations of levels of factor A and B are to be performed. That is the combination of $(A_i, B_j), i = 1,2, \ldots, p, j = 1,2, \ldots, q$ each is allocated randomly to experimental units and pq experiments are performed in random order.

Data…

Factor	B_1	B_2	…	B_q	Sums	Averages
A_1	y_{11}	y_{12}	…	y_{1q}	$y_{1.}$	$\bar{y}_{1.}$
A_2	y_{21}	y_{22}	…	y_{2q}	$y_{2.}$	$\bar{y}_{2.}$
:	:	:	…	:	…	…
A_p	y_{p1}	y_{p2}	…	y_{pq}	$y_{p.}$	$\bar{y}_{p.}$
Sums	$y_{.1}$	$y_{.2}$	…	$y_{.q}$	$y_{..}$	
Averages	$\bar{y}_{.1}$	$\bar{y}_{.2}$	…	$\bar{y}_{.q}$		$\bar{y}_{..}$

Two-way classification model

Model:

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ij}, \quad i = 1,2, \ldots, p;$$
$$j = 1,2, \ldots, q$$

y_{ij}:

μ:

α_i:

β_j:

γ_{ij}:

ε_{ij}:

Assumptions: $\varepsilon_{ij} \sim IID Normal(0, \sigma^2)$

Assumptions

- Mean error is zero i.e $E(\varepsilon_{ij}) = 0$
- Variance of errors is constant i.e $var(\varepsilon_{ij}) = \sigma^2 \; for \; all \; i, j$
- Covariances between errors is zero $var(\underline{\varepsilon}) = \sigma^2 I_n$
- $\underline{\varepsilon} \sim N_n(\underline{0}, \sigma^2 I_n)$

Implications

- $E(y_{ij}) = \mu + \alpha_i + \beta_j + \gamma_{ij}$
- $var(y_{ij}) = \sigma^2$
- y_{ij} are independently distributed but not identical
- $y_{ij} \sim N(\mu + \alpha_i + \beta_j + \gamma_{ij}, \sigma^2)$

Model details

- Number of observations=$n = pq$
- Number of parameters=$1 + p + q + pq$

$$\mu,$$

$$\alpha_1, \alpha_2, \dots, \alpha_p,$$

$$\beta_1, \beta_2, \dots, \beta_q,$$

$$\gamma_{11}, \gamma_{12}, \cdots, \gamma_{1q}$$

$$\gamma_{21}, \gamma_{22}, \cdots, \gamma_{2q}$$

$$\vdots$$

$$\gamma_{p1}, \gamma_{p2}, \cdots, \gamma_{pq}$$

Model speciality

- Here note that number of observations are less than number of parameters

$$n = pq < 1 + p + q + pq$$

- The assumption $n > p$ in GLM is not valid.
- Hence only estimation of parameters can be done.
- Degrees of freedom for error sum of squares would be zero.
- For this model error variance cannot be estimated.
- None of the hypothesis can be tested for this model.

Derivation of normal equations

- $\hat{y}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_{ij}$ is called fitted values (by model)
- Define errors as:

  Residuals/error, $e_{ij} = y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}$

- Obtain $\hat{\mu}, \hat{\alpha}_i$, $\hat{\beta}_j, \hat{\gamma}_{ij}$ such that the errors/error sum of squares is minimum.
- It will lead to $\hat{\mu}, \hat{\alpha}_i$, $\hat{\beta}_j$,$\hat{\gamma}_{ij}$ such that y_{ij} (observed value) and $\hat{y}_{ij}$ (fitted value) are close to each other in least square sense.

Derivation of normal equations…

Minimize function ϕ with respect to $\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_j, \hat{\gamma}_{ij}$

$$\phi = \sum_{i=1}^{p} \sum_{j=1}^{q} \left(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}\right)^2$$

$$\frac{d\phi}{d\hat{\mu}} = -2 \sum_{i=1}^{p} \sum_{j=1}^{q} \left(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}\right) \qquad (1)$$

$$\frac{d\phi}{d\hat{\alpha}_i} = -2 \sum_{j=1}^{q} \left(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}\right), \qquad i = 1{:}\,p \qquad (A)$$

$$\frac{d\phi}{d\hat{\beta}_j} = -2 \sum_{i=1}^{p} \left(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}\right), \qquad j = 1{:}\,q \qquad (B)$$

$$\frac{d\phi}{d\hat{\gamma}_{ij}} = -2\left(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}\right), \; i = 1{:}\,p, j = 1{:}\,q \qquad (C)$$

Derivation of normal equations…

Minimize function ϕ with respect to $\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_j, \hat{\gamma}_{ij}$

$$\phi = \sum_{i=1}^{p}\sum_{j=1}^{q}\left(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}\right)^2$$

$$\frac{d\phi}{d\hat{\mu}} = 0$$

$$\Rightarrow \quad -2\sum_{i=1}^{p}\sum_{j=1}^{q}\left(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}\right) = 0$$

$$\Rightarrow \quad y_{..} = pq\hat{\mu} + q\sum_{i=1}^{p}\hat{\alpha}_i + p\sum_{j=1}^{q}\hat{\beta}_j + \sum_{i=1}^{p}\sum_{j=1}^{q}\hat{\gamma}_{ij} \qquad (1)$$

Derivation of normal equations…

Minimize function ϕ with respect to $\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_j, \hat{\gamma}_{ij}$

$$\phi = \sum_{i=1}^{p} \sum_{j=1}^{q} \left(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}\right)^2$$

$$\frac{d\phi}{d\hat{\alpha}_i} = 0$$

$$\Rightarrow \qquad -2 \sum_{j=1}^{q} \left(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}\right), \qquad i = 1:p$$

$$\Rightarrow \qquad y_{i.} = q\hat{\mu} + q\hat{\alpha}_i + \sum_{j=1}^{q} \hat{\beta}_j + \sum_{j=1}^{q} \hat{\gamma}_{ij}, \qquad i = 1:p \qquad \text{(A)}$$

Derivation of normal equations…

Minimize function ϕ with respect to $\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_j, \hat{\gamma}_{ij}$

$$\phi = \sum_{i=1}^{p} \sum_{j=1}^{q} \left(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}\right)^2$$

$$\frac{d\phi}{d\hat{\beta}_j} = 0$$

$$\Rightarrow \qquad -2\sum_{i=1}^{p} \left(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}\right), \qquad j = 1:q$$

$$\Rightarrow \qquad y_{.j} = p\hat{\mu} + \sum_{i=1}^{p} \hat{\alpha}_i + p\hat{\beta}_j + \sum_{i=1}^{p} \hat{\gamma}_{ij}, \quad j = 1:q \quad \text{(B)}$$

Derivation of normal equations…

Minimize function ϕ with respect to $\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_j, \hat{\gamma}_{ij}$

$$\phi = \sum_{i=1}^{p} \sum_{j=1}^{q} \left(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}\right)^2$$

$$\frac{d\phi}{d\hat{\gamma}_{ij}} = 0$$

$$\Rightarrow \qquad -2\left(y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_{ij}\right), \qquad i = 1{:}\,p, j = 1{:}\,q$$

$$\Rightarrow \qquad y_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_{ij}, \qquad i = 1{:}\,p, j = 1{:}\,q \quad \text{(C)}$$

Derivation of normal equations…

$$y_{..} = pq\hat{\mu} + q\sum_{i=1}^{p}\hat{\alpha}_i + p\sum_{j=1}^{q}\hat{\beta}_j + \sum_{i=1}^{p}\sum_{j=1}^{q}\hat{\gamma}_{ij} \qquad (1)$$

$$y_{i.} = q\hat{\mu} + q\hat{\alpha}_i + \sum_{j=1}^{q}\hat{\beta}_j + \sum_{j=1}^{q}\hat{\gamma}_{ij}, \qquad i = 1{:}\,p \qquad (A)$$

$$y_{.j} = p\hat{\mu} + \sum_{i=1}^{p}\hat{\alpha}_i + p\hat{\beta}_j + \sum_{1=1}^{p}\hat{\gamma}_{ij}, \qquad j = 1{:}\,q \qquad (B)$$

$$y_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_{ij}, \qquad i = 1{:}\,p, j = 1{:}\,q \quad (C)$$

- Observe that (1), (A), (B) and (C) are $1 + p + q + pq$ equations in $1 + p + q + pq$ variables
- These equations are called normal equations.
- But all are not linearly independent.

Solution of normal equations

- **Observe the following inter-relations among the normal equations.**
  - $\sum_{i=1}^{p}(A) = (1)$ i.e. addition of p equations in (A) gives equation (1)
  - $\sum_{j=1}^{q}(B) = (1)$ i.e. addition of q equations in (B) gives equation (1)
  - $\sum_{i=1}^{p}\sum_{j=1}^{q}(C) = (1)$ i.e. addition of pq equations in (C) gives equation (1)
  - $\sum_{j=1}^{q}(C) = (A)$ for $i = 1{:}p$ i.e. for fixed i if we add q equations in (C) we get set of equations in (A)
  - $\sum_{i=1}^{p}(C) = (B)$ for $j = 1{:}q$ i.e. for fixed j if we add q equations in (C) we get set of equations in (B)

Solution of normal equations

- (1) is linearly dependent as $\sum_{i=1}^{p}(A) = (1)$
- (A) are linearly dependent as $\sum_{j=1}^{q}(C) = (A)$
- (B) are linearly dependent as $\sum_{i=1}^{p}(C) = (B)$
- Only equations in (C) are linearly independent.
- Thus only pq out of $1 + p + q + pq$ equations are linearly independent.
- Thus rank of estimation space i.e. number of linearly independent normal equations for this model is pq.

Solution of normal equations

- To solve the normal equations we need additional $1 + p + q$ equations which are linearly independent with (1), (A), (B) and (C).
- Let these two equations be:
  - $\sum_{i=1}^{p} \hat{\gamma}_{ij} = 0, \quad j = 1{:}\, q$ (A1)
  - $\sum_{j=1}^{q} \hat{\gamma}_{ij} = 0, \quad i = 1{:}\, p$ (A2)
  - $\sum_{i=1}^{p} \hat{\alpha}_{i} = 0$
  - $\sum_{j=1}^{q} \hat{\beta}_{j} = 0$
- These seem to be $p + q + 2$ equations.

Solution of normal equations

- Observe that

❖ $\sum_{i=1}^{p} \hat{\gamma}_{ij} = 0, \quad j = 1{:}q \Rightarrow \sum_{i=1}^{p} \sum_{j=1}^{q} \hat{\gamma}_{ij} = 0$

❖ $\sum_{j=1}^{q} \hat{\gamma}_{ij} = 0, \quad i = 1{:}p \Rightarrow \sum_{i=1}^{p} \sum_{j=1}^{q} \hat{\gamma}_{ij} = 0$

❖ Hence (A1) and (A2) in fact $p + q - 1$ linearly independent equations and not $p + q$.

❖ Thus (A1) and (A2) together with $\sum_{i=1}^{p} \hat{\alpha}_i = 0$ and $\sum_{j=1}^{q} \hat{\beta}_j = 0$ are $p + q + 1$ linearly independent equations

Derivation of normal equations…

$$y_{..} = pq\hat{\mu} + q\sum_{i=1}^{p}\hat{\alpha}_i + p\sum_{j=1}^{q}\hat{\beta}_j + \sum_{i=1}^{p}\sum_{j=1}^{q}\hat{\gamma}_{ij} \qquad (1)$$

$$\Rightarrow \qquad \hat{\mu} = \bar{y}_{..}$$

$$y_{i.} = q\hat{\mu} + q\hat{\alpha}_i + \sum_{j=1}^{q}\hat{\beta}_j + \sum_{j=1}^{q}\hat{\gamma}_{ij}, \qquad i = 1:p \quad (A)$$

$$\Rightarrow \qquad \hat{\alpha}_i = \bar{y}_{i.} - \hat{\mu} = \bar{y}_{i.} - \bar{y}_{..}, \qquad i = 1,2,\dots,p$$

$$y_{.j} = p\hat{\mu} + \sum_{i=1}^{p}\hat{\alpha}_i + p\hat{\beta}_j + \sum_{1=1}^{p}\hat{\gamma}_{ij}, \qquad j = 1:q \quad (B)$$

$$\Rightarrow \qquad \hat{\beta}_j = \bar{y}_{.j} - \hat{\mu} = \bar{y}_{.j} - \bar{y}_{..}, \qquad j = 1,2,\dots,q$$

Derivation of normal equations…

$$y_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_{ij}, \qquad i = 1{:}\,p, j = 1{:}\,q \qquad \text{(C)}$$

$$\Rightarrow \quad \hat{\gamma}_{ij} = y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j$$

$$= y_{ij} - \bar{y}_{..} - (\bar{y}_{i.} - \bar{y}_{..}) - (\bar{y}_{.j} - \bar{y}_{..})$$

$$= y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$$

Solution of normal equations…

$$\hat{\mu} = \bar{y}_{..}$$

$$\hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..}, \qquad i = 1,2, \dots , p$$

$$\hat{\beta}_j = \bar{y}_{.j} - \bar{y}_{..}, \qquad j = 1,2, \dots , q$$

$$\hat{\gamma}_{ij} = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}, \qquad i = 1{:}\,p, j = 1{:}\,q$$

Note:

- These are known as **solutions of normal equations** and **not estimates of the respective parameters**.
- Since the model in NFRM, individual parameters are not estimable.
- Only few linear parametric functions are estimable

Rank of Estimation Space

- Estimation space: It is collection of all *lpfs* which are estimable.

  $\rho(estimation\ space)$

  =number of linearly independent normal equations

  = Number of *linearly independent estimable lpfs*

  $=pq$

- It means that there would be only pq *epfs* which would be *l.i.*

Rank of error space

- Error space: This is the space which is orthogonal to estimation space and contain all unbiased estimators of zero (representing errors)

  $\rho(Error\ space)$

  $= n - \rho(estimation\ space)$

  $= pq - pq$

  $=0$

- Here $n = pq$ which are number of observations.

Normal equations

$$y_{..} = pq\hat{\mu} + q\sum_{i=1}^{p}\hat{\alpha}_i + p\sum_{j=1}^{q}\hat{\beta}_j + \sum_{i=1}^{p}\sum_{j=1}^{q}\hat{\gamma}_{ij} \qquad (1)$$

$$y_{i.} = q\hat{\mu} + q\hat{\alpha}_i + \sum_{j=1}^{q}\hat{\beta}_j + \sum_{j=1}^{q}\hat{\gamma}_{ij}, \qquad i = 1:p \qquad (A)$$

$$y_{.j} = p\hat{\mu} + \sum_{i=1}^{p}\hat{\alpha}_i + p\hat{\beta}_j + \sum_{i=1}^{p}\hat{\gamma}_{ij}, \qquad j = 1:q \qquad (B)$$

$$y_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_{ij}, \qquad i = 1:p, j = 1:q \qquad (C)$$

Expectations of normal equations:

Equations (1), (A), (B) and (C) can also be written as follows.

$$E(y_{..}) = pq\mu + q\sum_{i=1}^{p}\alpha_i + p\sum_{j=1}^{q}\beta_j + \sum_{i=1}^{p}\sum_{j=1}^{q}\gamma_{ij}$$

$$\mathrm{E}(y_{i.}) = q\mu + q\alpha_i + \sum_{j=1}^{q}\beta_j + \sum_{j=1}^{q}\gamma_{ij}, \qquad i = 1{:}\,p$$

$$\mathrm{E}(y_{.j}) = p\mu + \sum_{i=1}^{p}\alpha_i + p\beta_j + \sum_{i=1}^{p}\gamma_{ij}, \qquad j = 1{:}\,q$$

$$E(y_{ij}) = \mu + \alpha_i + \beta_j + \gamma_{ij}, \qquad i = 1{:}\,p, j = 1{:}\,q$$

- **Note:** $E(y_{ij}) = \mu + \alpha_i + \beta_j + \gamma_{ij}$ is simply the model

  $\Rightarrow \mu + \alpha_i + \beta_j + \gamma_{ij}$ are all estimable functions

Estimability condition of linear parametric functions

Equations (1), (A), (B) and (C) can also be written as follows.

$$E(y_{..}) = pq\mu + q\sum_{i=1}^{p}\alpha_i + p\sum_{j=1}^{q}\beta_j + \sum_{i=1}^{p}\sum_{j=1}^{q}\gamma_{ij}$$

$$\mathrm{E}(y_{i.}) = q\mu + q\alpha_i + \sum_{j=1}^{q}\beta_j + \sum_{j=1}^{q}\gamma_{ij}, \qquad i = 1:p$$

$$\mathrm{E}(y_{.j}) = p\mu + \sum_{i=1}^{p}\alpha_i + p\beta_j + \sum_{i=1}^{p}\gamma_{ij}, \qquad j = 1:q$$

$$E(\bar{y}_{..}) = \mu + \alpha_i + \beta_j + \gamma_{ij}, \qquad i = 1:p, j = 1:q$$

Estimability of parametric functions in terms of all parameters

Using first normal equation,

$$E(y_{..}) = pq\mu + q\sum_{i=1}^{p}\alpha_i + p\sum_{j=1}^{q}\beta_j + \sum_{i=1}^{p}\sum_{j=1}^{q}\gamma_{ij}$$

$$E(\bar{y}_{..}) = \mu + \frac{1}{p}\sum_{i=1}^{p}\alpha_i + \frac{1}{q}\sum_{j=1}^{q}\beta_j + \frac{1}{pq}\sum_{i=1}^{p}\sum_{j=1}^{q}\gamma_{ij}$$

$$E(\bar{y}_{..}) = \mu + \bar{\alpha} + \bar{\beta} + \bar{\gamma}_{..}$$

$\Rightarrow \ \mu + \frac{1}{p}\sum_{i=1}^{p}\alpha_i + \frac{1}{q}\sum_{j=1}^{q}\beta_j + \frac{1}{pq}\sum_{i=1}^{p}\sum_{j=1}^{q}\gamma_{ij}$ is an epf

$\Rightarrow \ \mu + \bar{\alpha} + \bar{\beta} + \bar{\gamma}_{..}$ is an estimable parametric function

Estimability of linear parametric functions in $\alpha' s$

$$\mathrm{E}(y_{i.}) = q\mu + q\alpha_i + \sum_{j=1}^{q} \beta_j + \sum_{j=1}^{q} \gamma_{ij}, \qquad i = 1{:}\,p$$

$$\mathrm{E}(\bar{y}_{i.}) = \mu + \alpha_i + \bar{\beta} + \bar{\gamma}_{i.}, \qquad i = 1,2,\ldots,p$$

Consider the pair of equations from (A) for $i \neq u$ as follows.

$$\mathrm{E}(\bar{y}_{i.}) = \mu + \alpha_i + \bar{\beta} + \bar{\gamma}_{i.}$$

$$\mathrm{E}(\bar{y}_{u.}) = \mu + \alpha_u + \bar{\beta} + \bar{\gamma}_{u.}$$

Subtracting these we get:

$$\mathrm{E}(\bar{y}_{i.} - \bar{y}_{u.}) = (\alpha_i + \bar{\gamma}_{i.}) - (\alpha_u + \bar{\gamma}_{u.}) \qquad i \neq u$$

$$= \alpha_i^* - \alpha_u^* \qquad \text{where } \alpha_i^* = (\alpha_i + \bar{\gamma}_{i.})$$

- Thus $\alpha_i^* - \alpha_u^*$ is estimable for all $i \neq u$.

Estimability of linear parametric function in $\alpha's$…

- All elementary contrasts in $\alpha^{*'}s$ are estimable.
- All contrasts in$\alpha^{*'}s$ are estimable
- Thus, $\sum_{i=1}^{p} c_i\alpha_i^*$ is estimable if $\sum_{i=1}^{p} c_i = 0$
- $\alpha_1^* - 2\alpha_2^* + \alpha_3^*$ and $\alpha_1^* - 2\alpha_2^* - \alpha_3^* + 2\alpha_4^*$ are estimable.
- While $\alpha_1^* + \alpha_2^*$, $\alpha_1^* - 2\alpha_3^*$ are not estimable.

Estimability of linear parametric functions in $\beta' s$

(B)⇒ $E(y_{.j}) = p\mu + \sum_{i=1}^{p} \alpha_i + p\beta_j + \sum_{i=1}^{p} \gamma_{ij}$

$$E(\bar{y}_{.j}) = \mu + \bar{\alpha} + \beta_j + \bar{\gamma}_{.j}, \qquad j = 1,2,\ldots,q$$

Consider the pair of equations from (B) for $j \neq v$ as follows.

$$E(\bar{y}_{.j}) = \mu + \bar{\alpha} + \beta_j + \bar{\gamma}_{.j}$$

$$E(\bar{y}_{.v}) = \mu + \bar{\alpha} + \beta_v + \bar{\gamma}_{.v}$$

Subtracting these we get:

$$E(\bar{y}_{.j} - \bar{y}_{.v}) = (\beta_j + \bar{\gamma}_{.j}) - (\beta_v + \bar{\gamma}_{.v}), \text{ where } \beta_j^* = \beta_j + \bar{\gamma}_{.j}$$

Thus $\beta_j^* - \beta_v^*$, is estimable for all $j \neq v$.

Estimability of linear parametric function in$\beta's$…

- All elementary contrasts in $\beta^{*'}s$ are estimable.
- All contrasts in $\beta^{*'}s$ are estimable
- Thus,$\sum_{j=1}^{q} d_j\beta_j^*$is estimable if $\sum_{j=1}^{q} d_j = 0$
- $\beta_1^* - 2\beta_2^* + \beta_3^*$ and $\beta_1^* - 2\beta_2^* - \beta_3^* + 2\beta_4^*$are estimable.
- While $\beta_1^* + \beta_2^*$, $\beta_1^* - 2\beta_3^*$ are *not* estimable.

Estimability of parametric function in interaction terms

Performing $(1) - (A) - (B) + (C)$ as follows.

$$E(\bar{y}_{..}) = \mu + \bar{\alpha} + \bar{\beta} + \bar{\gamma}_{..}$$

$$- \quad \mathrm{E}(\bar{y}_{i.}) = \mu + \alpha_i + \bar{\beta} + \bar{\gamma}_{i.}$$

$$- \quad \mathrm{E}\left(\bar{y}_{.j}\right) = \mu + \bar{\alpha} + \beta_j + \bar{\gamma}_{.j}$$

$$+ \quad \mathrm{E}\left(y_{ij}\right) = \mu + \alpha_i + \beta_j + \gamma_{ij}$$

$$E\left(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}\right) = \gamma_{ij} - \bar{\gamma}_{i.} - \bar{\gamma}_{.j} + \bar{\gamma}_{..}$$

Estimability of parametric function in interaction terms…

- Thus $\gamma_{ij}^{*} = \gamma_{ij} - \bar{\gamma}_{i.} - \bar{\gamma}_{.j} + \bar{\gamma}_{..}$ are all estimable for all $i = 1\!:\!p, j = 1\!:\!q$
- Further all pq γ_{ij}^{*} are not linearly independent as the following relations exists between them.
  - ➢ $\sum_{j=1}^{q} \gamma_{ij}^{*} = 0$, for $i = 1\!:\!p$
  - ➢ $\sum_{i=1}^{p} \gamma_{ij}^{*} = 0$, for $j = 1\!:\!q$
  - ➢ $\sum_{i=1}^{p} \sum_{j=1}^{q} \gamma_{ij}^{*} = 0$
- Thus only $(p-1)(q-1)$ of the γ_{ij}^{*} are linearly independent.

Summary of estimability conditions of l. parametric functions

1. $\mu + \bar{\alpha} + \bar{\beta} + \bar{\gamma}_{..}$ is estimable (A *lpf* which involve all parameters)

2. All contrasts in $\alpha^{*'}s$ are estimable.

3. All contrasts in $\beta^{*'}s$ are estimable.

4. $\gamma^{*}_{ij} = \gamma_{ij} - \bar{\gamma}_{i.} - \bar{\gamma}_{.j} + \bar{\gamma}_{..}$ are all estimable for all $i = 1{:}\,p, j = 1{:}\,q$

One set of linearly independent e.p.f.

- $\mu + \bar{\alpha} + \bar{\beta} + \bar{\gamma}_{..}$ (1)

$$\left.\begin{array}{l} \alpha_1^* - \alpha_2^* \\ \alpha_1^* - \alpha_3^* \\ \vdots \\ \alpha_1^* - \alpha_p^* \end{array}\right\} (p-1)$$

$$\left.\begin{array}{l} \beta_1^* - \beta_2^* \\ \beta_1^* - \beta_3^* \\ \vdots \\ \beta_1^* - \beta_q^* \end{array}\right\} (q-1)$$

$\gamma_{ij}^*, i = 1{:}\, p, j = 1{:}\, q$

$(p-1)(q-1)$

linearly independent e.p.fs. pq

BLUEs and Variance(BLUE) of *epf*

Result: LHS of normal equations are BLUE of expected value of their RHS

- In GLM $\underline{Y} = X\underline{\beta} + \underline{\varepsilon}$ normal equations are

$$X'\underline{Y} = X'X\underline{\hat{\beta}}$$

- $X'\underline{Y}$ =LHS of normal equations and
- $X'X\underline{\hat{\beta}}$ =RHS of normal equations and
- E(RHS of normal equations)=E($X'\underline{Y}$) $= X'X\underline{\beta}$
- Thus $X'\underline{Y} = X'X\underline{\hat{\beta}}$ is BLUE of its expected value i.e. $X'X\underline{\beta}$

BLUEs and Variance(BLUE) of *epf*

- BLUE of $\mu + \bar{\alpha} + \bar{\beta} + \bar{\gamma}_{..}$ which is estimable.

$$\text{Hence } \widehat{\mu + \bar{\alpha} + \bar{\beta} + \bar{\gamma}_{..}} = \bar{y}_{..} \quad \text{as } E(\bar{y}_{..}) = \mu + \bar{\alpha} + \bar{\beta} + \bar{\gamma}_{..}$$

- Variance

Variance(BLUE)

$$= \mathrm{var}\left(\widehat{\mu + \bar{\alpha} + \bar{\beta} + \bar{\gamma}_{..}}\right)$$

$$= \mathrm{var}(\bar{y}_{..})$$

$$= \frac{\sigma^2}{pq}$$

BLUEs and Variance(BLUE) of *epf*

- $\beta_j^* - \beta_v^*$, is estimable for all $j \neq v$.
- Further $\mathrm{E}\left(\bar{y}_{.j} - \bar{y}_{.v}\right) = \beta_j^* - \beta_v^*, \qquad j \neq v$
- Hence for $j \neq v$

BLUE of $\beta_j^* - \beta_v^*$

$= \widehat{\beta_j^* - \beta_v^*}$

$= \bar{y}_{.j} - \bar{y}_{.v}$

BLUEs and Variance(BLUE) of *epf*

- Variance(BLUE)

$$= \text{var}(\widehat{\beta_j^* - \beta_v^*})$$

$$= \text{var}\left(\bar{y}_{.j} - \bar{y}_{.v}\right)$$

$$= \text{var}\left(\bar{y}_{.j}\right) + \text{var}(\bar{y}_{.v}) - 2cov\left(\bar{y}_{.j}, \bar{y}_{.v}\right)$$

$$= \frac{\sigma^2}{p} + \frac{\sigma^2}{p} - 2 \times 0$$

$$= \frac{2\sigma^2}{p}$$

BLUEs and Variance(BLUE) of *epf*

- In general $\sum_{j=1}^{q} d_j\beta_j^{\,*}$ is estimable if $\sum_{j=1}^{q} d_j = 0$
- BLUE of $\sum_{j=1}^{q} d_j\beta_j^{\,*}$

$$= \widehat{\sum_{j=1}^{q} d_j}\beta_j^*$$

$$= \sum_{j=1}^{q} d_j\,\widehat{\beta_j}^{\,*}$$

$$= \sum_{j=1}^{q} d_j\left(\bar{y}_{.j} - \bar{y}_{..}\right)$$

$$= \sum_{j=1}^{q} d_j\,\bar{y}_{.j}$$

BLUEs and Variance(BLUE) of *epf*

- $\alpha_i^* - \alpha_u^*$, is estimable for all $i \neq u$.
- Further $\mathrm{E}(\bar{y}_{i.} - \bar{y}_{u.}) = \alpha_i^* - \alpha_u^*$, $\quad i \neq u$
- Hence for $i \neq u$

BLUE of $\alpha_i^* - \alpha_u^*$

$= \widehat{\alpha_i^* - \alpha_u^*}$

$= \bar{y}_{i.} - \bar{y}_{u.}$

BLUEs and Variance(BLUE) of *epf*

- Variance(BLUE)

$$= \text{var}(\widehat{\alpha_i^* - \alpha_u^*})$$

$$= \text{var}(\bar{y}_{i.} - \bar{y}_{u.})$$

$$= \text{var}(\bar{y}_{i.}) + \text{var}(\bar{y}_{u.}) - 2cov(\bar{y}_{i.}, \bar{y}_{u.})$$

$$= \frac{\sigma^2}{q} + \frac{\sigma^2}{q} - 2 \times 0$$

$$= \frac{2\sigma^2}{q}$$

BLUEs and Variance(BLUE) of *epf*

- In general $\sum_{i=1}^{p} c_i \alpha_i^*$ is estimable if $\sum_{i=1}^{p} c_i = 0$
- BLUE of $\sum_{i=1}^{p} c_i \alpha_i^*$

$$= \widehat{\sum_{i=1}^{p} c_i \alpha_i^*}$$

$$= \sum_{i=1}^{p} c_i \widehat{\alpha_i}^*$$

$$= \sum_{i=1}^{p} c_i (\bar{y}_{i.} - \bar{y}_{..})$$

$$= \sum_{i=1}^{p} c_i \bar{y}_{i.}$$

BLUEs and Variance(BLUE) of *epf*

- Variance(BLUE of $\sum_{i=1}^{p} c_i \alpha_i^*$)

$$= \text{var}\left(\widehat{\sum_{i=1}^{p} c_i \alpha_i^*}\right)$$

$$= \text{var}\left(\sum_{i=1}^{p} c_i \bar{y}_{i.}\right)$$

$$= \sum_{i=1}^{p} c_i^2 \,\text{var}(\bar{y}_{i.})$$

$$= \sum_{i=1}^{p} c_i^2 \frac{\sigma^2}{q}$$

$$= \frac{\sigma^2}{q} \sum_{i=1}^{p} c_i^2$$

BLUE of parametric function in interaction terms

- $\gamma^*_{ij} = \gamma_{ij} - \bar{\gamma}_{i.} - \bar{\gamma}_{.j} + \bar{\gamma}_{..}$ are all estimable
- BLUE of γ^*_{ij}

$$= \hat{\gamma}^*_{ij}$$

$$= \widehat{\gamma_{ij} - \bar{\gamma}_{i.} - \bar{\gamma}_{.j} + \bar{\gamma}_{..}}$$

$$= y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..} \quad \text{as } E\left(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}\right) = \gamma^*_{ij}$$

- $\hat{\gamma}^*_{ij} = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$

Variance of BLUE of parametric function in interaction terms

$$var(\hat{\gamma}^{*}_{ij})$$

$$= var(\gamma_{ij} - \widehat{\bar{\gamma}_{i.} - \bar{\gamma}}_{.j} + \bar{\gamma}_{..})$$

$$= var\left(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}\right)$$

$$= var\left(y_{ij}\right) + var(\bar{y}_{i.}) + var\left(\bar{y}_{.j}\right) + var(\bar{y}_{..})$$

$$-2cov\left(y_{ij}, \bar{y}_{i.}\right) - 2cov\left(y_{ij}, \bar{y}_{.j}\right) + 2cov\left(y_{ij}, \bar{y}_{..}\right)$$

$$+ 2cov\left(\bar{y}_{i.}, \bar{y}_{.j}\right) - 2cov(\bar{y}_{i.}, \bar{y}_{..}) \; - 2cov\left(\bar{y}_{.j}, \bar{y}_{..}\right)$$

$$= \sigma^2 + \frac{\sigma^2}{q} + \frac{\sigma^2}{p} + \frac{\sigma^2}{pq} - 2\,\frac{\sigma^2}{q} - 2\,\frac{\sigma^2}{p} + 2\,\frac{\sigma^2}{pq}$$

$$+2\,\frac{\sigma^2}{qp} - 2\,\frac{q\sigma^2}{q.pq} - 2\,\frac{p\sigma^2}{ppq}$$

Variance of BLUE of parametric function in interaction terms

$$var(\hat{\gamma}^*_{ij})$$

$$= \sigma^2 - \frac{\sigma^2}{q} - \frac{\sigma^2}{p} + \frac{\sigma^2}{pq}$$

$$= \sigma^2 \left(1 - \frac{1}{q} - \frac{1}{p} + \frac{1}{pq}\right)$$

$$= \frac{\sigma^2}{pq}(pq - p - q + 1)$$

$$= \frac{\sigma^2}{pq}(p-1)(q-1)$$

Summary of BLUEs and Variance(BLUE) of *epf*

Estimable parametric functions	BLUE	Variance(BLUE)
$\mu + \bar{\alpha} + \bar{\beta} + \bar{\gamma}_{..}$	$\bar{y}_{..}$	$\dfrac{\sigma^2}{pq}$
$\alpha_i^* - \alpha_u^*,\ i \neq u$	$\bar{y}_{i.} - \bar{y}_{u.}$	$\dfrac{2\sigma^2}{q}$
$\beta_j^* - \beta_v^*\ ,\ \ j \neq v$	$\bar{y}_{.j} - \bar{y}_{.\mathrm{v}}$	$\dfrac{2\sigma^2}{p}$
$\gamma_{ij}^* = \gamma_{ij} - \bar{\gamma}_{i.} - \bar{\gamma}_{.j} + \bar{\gamma}_{..}$	$y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$	$\dfrac{\sigma^2}{pq}(p-1)(q-1)$

Model value and error

- **Model value of y_{ij}**

$$\hat{y}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \hat{\gamma}_{ij}$$

$$= \bar{y}_{..} + (\bar{y}_{i.} - \bar{y}_{..}) + \left(\bar{y}_{.j} - \bar{y}_{..}\right) + (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})$$

$$= y_{ij}$$

Fitted and observed value are equal for this model.

- **Error sum of squares**

$$\text{SSE} = \sum_{i=1}^{p} \sum_{j=1}^{q} \left(y_{ij} - \hat{y}_{ij}\right)^2 = 0$$

Testing of hypothesis

- The following hypothesis are not testable.

$$H_0 : \gamma_{ij} = 0 \text{ for all } i, j$$

$$H_{01} : \alpha_1 = \alpha_2 = \cdots = \alpha_p$$

$$H_{02} : \beta_1 = \beta_2 = \cdots = \beta_q$$

$$H_{03} : \alpha_i = \alpha_u \ , i \neq u$$

$$H_{04} : \beta_j = \beta_v, \ j \neq v$$

Testing of hypothesis

- The following hypothesis are testable.

$$H_0: \gamma_{ij}^* = 0 \text{ for all } i = 1:p, j = 1:q$$

$$H_{01}: \alpha_1^* = \alpha_2^* = \cdots = \alpha_p^*$$

$$H_{02}: \beta_1^* = \beta_2^* = \cdots = \beta_q^*$$

$$H_{03}: \alpha_i^* = \alpha_u^* \text{ , } i \neq u$$

$$H_{04}: \beta_j^* = \beta_v^*, \; j \neq v$$

Hypothesis testing

- SSE for this model is zero.
- Error variance cannot be estimated.
- Hence for this model even though many of the hypothesis are testable, none of them can not be tested.

Remedy on the problem

- **Solution 1:** As $n <$ number of parameters, try to increase n, that is, replicate the experiments at least twice.
- That is consider the model with $r\ (r > 1)$ observations per cell with interaction terms.
- **Solution 2:** If the above is not possible, then reduce number of parameters, that is, drop the interaction terms from the model.

Sample questions

- BLUE of elementary contrast in treatment effects for two-way classification model with interaction effects.
- What would be the rank of estimation space for two-way model with interaction and with p=3 and q=4?
- For two-way classification model with interaction, what is the variance of BLUE of contrast in block effects?
- State the BLUE and its variance for contrasts in column effects in two-way classification model with interactions.

Sample questions…

- For two-way classification model with interactions, write the linear model and answer the following.
  - Obtain normal equations for the model.
  - What is rank of error space and estimation space?
  - Specify one complete set of linearly independent estimable parametric functions and their BLUEs and variances of BLUEs.
  - What is fitted value of y_{ij}?

What we have studied

- Two-way classification model with interactions
- Normal equations and their solutions
- Estimability conditions for linear parametric functions
- Estimable parametric functions, their BLUEs and variances of BLUEs.
- Rank of error space and estimation space
- A set of linearly independent estimable parametric functions
- Fitted value of y_{ij}