Analysis of Variance and Design of Experiments

General Linear Hypothesis and Analysis of Variance

:::

Lecture 8

Analysis of Variance in One Way Fixed Effect Model

Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology Kanpur

Slides can be downloaded from http://home.iitk.ac.in/~shalab/sp

One-way classification with fixed effect linear models of full rank:

Let y_{ij}(j = 1,2,…, n_i) be a random sample from the i^{th} normal population with mean β_i **and variance** $\sigma^2, i = 1, 2, \ldots, p$**, i.e.,**

$$Y_{ij} \sim N(\beta_i, \sigma^2), \; j = 1, 2, \ldots, n_i; \; i = 1, 2, \ldots, p.$$

The random samples from different populations are assumed to be independent of each other.

These observations follow the set up of linear model

$$Y = X\beta + \varepsilon$$

One-way classification with fixed effect linear models of full rank:

These observations follow the set up of linear model

$$Y = X\beta + \varepsilon$$

where

$$Y = (Y_{11}, Y_{12}, \ldots, Y_{1n_1}, Y_{21}, \ldots, Y_{2n_2}, \ldots, Y_{p1}, Y_{p2}, \ldots, Y_{pn_p})'$$

$$y = (y_{11}, y_{12}, \ldots, y_{1n_1}, y_{21}, \ldots, y_{2n_2}, \ldots, y_{p1}, y_{p2}, \ldots, y_{pn_p})'$$

$$\beta = (\beta_1, \beta_2, \ldots, \beta_p)'$$

$$\varepsilon = (\varepsilon_{11}, \varepsilon_{12}, \ldots, \varepsilon_{1n_1}, \varepsilon_{21}, \ldots, \varepsilon_{2n_2}, \ldots, \varepsilon_{p1}, \varepsilon_{p2}, \ldots, \varepsilon_{pn_p})'$$

One-way classification with fixed effect linear models of full rank:

where

$$X = \begin{pmatrix} \left.\begin{matrix} 1\ 0 \dots 0 \\ \vdots\ \ddots\ \vdots \\ 1\ 0\ \ 0 \end{matrix}\right\} n_1 \text{ values} \\ \left.\begin{matrix} 0\ 1 \dots 0 \\ \vdots\ \ddots\ \vdots \\ 0\ 1 \dots 0 \end{matrix}\right\} n_2 \text{ values} \\ \vdots\ \ \vdots\ \ \vdots \\ \left.\begin{matrix} 0\ 0 \dots 1 \\ \vdots\ \ddots\ \vdots \\ 0\ 0 \dots 1 \end{matrix}\right\} n_p \text{ values} \end{pmatrix}$$

$$x_{ij} = \begin{cases} 1 & \textbf{if } \beta_i \textbf{ occurs in the } j^{th} \textbf{ observation } x_j \\ & \textbf{or if effect } \beta_i \textbf{ is present in } x_j \\ 0 & \textbf{if effect } \beta_i \textbf{ is absent in } x_j \end{cases}$$

$$n = \sum_{i=1}^{p} n_i.$$

Obviously, $rank(X) = p,\ E(Y) = X\beta$

and $Cov(Y) = \sigma^2 I.$

This completes the representation of a fixed effect linear model of full rank.

One-way classification with fixed effect linear models of full rank:

The null hypothesis of interest is $H_0 : \beta_1 = \beta_2 = ... = \beta_p = \beta$ **(say)**

and H_1 **: At least one** $\beta_i \neq \beta_j (i \neq j)$

where β **and** σ^2 **are unknown.**

One-way classification with fixed effect linear models of full rank:

The whole parametric space Ω is a (*p* + 1) dimensional space

$$\Omega = \left\{ (\beta, \sigma^2) : -\infty < \beta_i < \infty, \sigma^2 > 0, i = 1, 2, ..., p \right\}$$

Note that there are (*p* + 1) parameters are $\beta_1, \beta_2, ..., \beta_p$ and σ^2.

Under *H*$_0$, Ω reduces to two dimensional space ω as

$$\omega = \left\{ (\beta, \sigma^2); -\infty < \beta < \infty, \ \sigma^2 > 0 \right\}.$$

One-way classification with fixed effect linear models of full rank:

The likelihood function under Ω is

$$L(y\,|\,\beta,\sigma^2)=\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\beta_i)^2\right]$$

$$L=\ln\ L(y\,|\,\beta,\sigma^2)=-\frac{n}{2}\ \ln\ (2\pi\sigma^2)-\frac{1}{2\sigma^2}\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\beta_i)^2.$$

One-way classification with fixed effect linear models of full rank:

The normal equations and the estimators are obtained as follows:

$$\frac{\partial L}{\partial \beta_i} = 0 \qquad \Rightarrow \hat{\beta}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} y_{ij} = \overline{y}_{io}$$

$$\frac{\partial L}{\partial \sigma^2} = 0 \qquad \Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij} - \overline{y}_{io})^2.$$

The dot sign (o) in $\overline{y}_{io}$ indicates that the average has been taken over the second subscript *j*.

One-way classification with fixed effect linear models of full rank:

The Hessian matrix of second-order partial derivation of $\ln L$ with respect to β_i and σ^2 is negative definite at $\beta = \bar{y}_{io}$ and $\sigma^2 = \hat{\sigma}^2$ which ensures that the likelihood function is maximized at these values.

One-way classification with fixed effect linear models of full rank:

Thus the maximum value of $L(y|\beta,\sigma^2)$ **over** Ω **is**

$$\underset{\Omega}{Max}\ L(y|\beta,\sigma^2) = \left(\frac{1}{2\pi\hat{\sigma}^2}\right)^{\frac{n}{2}} \exp\left[-\frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\hat{\beta}_i)^2\right]$$

$$= \left[\frac{n}{2\pi\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\overline{y}_{io})^2}\right]^{n/2} \exp\left(-\frac{n}{2}\right).$$

One-way classification with fixed effect linear models of full rank:

The likelihood function under ω is

$$L(y|\beta,\sigma^2)=\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}}\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\beta)^2\right]$$

$$\ln L(y|\beta,\sigma^2)=-\frac{n}{2}\ln(2\pi\sigma^2)-\frac{1}{2\sigma^2}\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\beta)^2$$

One-way classification with fixed effect linear models of full rank:

The normal equations and the estimators are obtained as follows:

$$\frac{\partial \ln L(y|\beta, \sigma^2)}{\partial \beta} = 0 \qquad \Rightarrow \hat{\beta} = \frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{n_i} y_{ij} = \bar{y}_{oo}$$

$$\frac{\partial \ln L(y|\beta, \sigma^2)}{\partial \sigma^2} = 0 \qquad \Rightarrow \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{p}\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{oo})^2.$$

One-way classification with fixed effect linear models of full rank:

The maximum value of the likelihood function over ω under H_0 is

$$\underset{\omega}{Max}\ L(y|\beta,\sigma^2) = \left(\frac{1}{2\pi\hat{\sigma}^2}\right)^{\frac{n}{2}} \exp\left[-\frac{1}{2\hat{\sigma}^2}\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\hat{\beta})^2\right]$$

$$= \left[\frac{n}{2\pi\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\overline{y}_{oo})^2}\right]^{n/2} \exp\left(-\frac{n}{2}\right).$$

One-way classification with fixed effect linear models of full rank:

The likelihood ratio test statistic is $\lambda = \dfrac{\underset{\omega}{Max}\, L(y \mid \beta, \sigma^2)}{\underset{\Omega}{Max}\, L(y \mid \beta, \sigma^2)}$

$$= \left[\frac{\sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{io})^2}{\sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{oo})^2} \right]^{n/2}$$

We have that

$$\sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{oo})^2 = \sum_{i=1}^{p} \sum_{j=1}^{n_i} \left[(y_{ij} - \bar{y}_{io}) + (\bar{y}_{io} - \bar{y}_{oo}) \right]^2$$

$$= \sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{io})^2 + \sum_{i=1}^{p} n_i (\bar{y}_{io} - \bar{y}_{oo})^2$$

One-way classification with fixed effect linear models of full rank:

Thus

$$\lambda = \left[\frac{\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_i)^2 + \sum_{I=1}^{p} n_i(\bar{y}_{io}-\bar{y}_{oo})^2}{\sum_{i=1}^{p}\sum_{j=1}^{n_i}(y_{ij}-\bar{y}_{io})^2} \right]^{-\frac{n}{2}}$$

$$= \left[1 + \frac{q_1}{q_2} \right]^{-\frac{n}{2}}$$

where

$$q_1 = \sum_{i=1}^{p} n_i(\bar{y}_{io}-\bar{y}_{oo})^2,$$ **and** $$q_2 = \sum_{i=1}^{p}\sum_{j=1}^{n_i} (y_{ij}-\bar{y}_{io})^2$$

One-way classification with fixed effect linear models of full rank:

Note that if the least-squares principal is used, then

q_1 : sum of squares due to deviations from H_0 or the between population sum of squares,

q_2 : sum of squares due to error or the within-population sum of squares,

$q_1 + q_2$: sum of squares due to H_0 or the total sum of squares.

One-way classification with fixed effect linear models of full rank:

Using theorems 6 and 7, let

$$Q_1 = \sum_{i=1}^{p} n_i (\bar{Y}_{io} - \bar{Y}_{oo})^2, \qquad Q_2 = \sum_{i=1}^{p} S_i^2$$

where

$$S_i^2 = \sum_{i=1}^{n_i} (Y_{ij} - \bar{Y}_{io})^2, \quad \bar{Y}_{oo} = \frac{1}{n} \sum_{i=1}^{p} \sum_{j=1}^{n_i} Y_{ij}, \quad \bar{Y}_{io} = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij},$$

then under H_0

$$\frac{Q_1}{\sigma^2} \sim \chi^2(p-1)$$

$$\frac{Q_2}{\sigma^2} \sim \chi^2(n-p)$$

and $\frac{Q_1}{\sigma^2}$ and $\frac{Q_2}{\sigma^2}$ are independently distributed.

One-way classification with fixed effect linear models of full rank:

Thus under H_0

$$\frac{\left(\dfrac{\frac{Q_1}{\sigma^2}}{p-1}\right)}{\left(\dfrac{\frac{Q_2}{\sigma^2}}{n-p}\right)} \sim F(p-1, n-p).$$

The likelihood ratio test reject H_0 whenever

$$\frac{q_1}{q_2} > C$$

where the constant $C = F_{1-\alpha}(p-1, n-p).$

If $F > F_{1-\alpha}(p-1, n-p)$, **then** $H_0 : \beta_1 = \beta_2 = ... = \beta_p$ **is rejected.**

One-way classification with fixed effect linear models of full rank:

The analysis of variance table for the one-way classification in fixed effect model is

Source of variation	Degrees of freedom	Sum of squares	Mean squares	*F* - value
Between populations	$p-1$	q_1	$\frac{q_1}{p-1}$	$\left(\frac{n-p}{p-1}\right)\cdot\frac{q_1}{q_2}$
Within populations	$n-p$	q_2	$\frac{q_2}{n-p}$	
Total	$n-1$	q_1+q_2		

One-way classification with fixed effect linear models of full rank:

Note that

$$E\left[\frac{Q_2}{n-p}\right]=\sigma^2$$

$$E\left[\frac{Q_1}{p-1}\right]=\sigma^2+\frac{\sum_{i=1}^{p}(\beta_i-\bar{\beta})^2}{p-1};$$

$$\bar{\beta}=\frac{1}{p}\sum_{i=1}^{p}\beta_i$$