Analysis of Variance and Design of Experiments-I

MODULE – III

LECTURE - 17

EXPERIMENTAL DESIGN MODELS

Dr. Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology Kanpur
Tukey’s test for nonadditivity

Consider the set up of two-way classification with one observation per cell and interaction as

\[ y_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ij}, \quad i = 1, 2, \ldots, I, \quad j = 1, 2, \ldots, J \]

with

\[ \sum_{j=1}^{J} \alpha_i = 0, \quad \sum_{i=1}^{I} \beta_j = 0. \]

The distribution of degrees of freedom in this case is as follows:

<table>
<thead>
<tr>
<th>Source</th>
<th>Degrees of freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>( I - 1 )</td>
</tr>
<tr>
<td>B</td>
<td>( J - 1 )</td>
</tr>
<tr>
<td>AB(interaction)</td>
<td>( (I - 1)(J - 1) )</td>
</tr>
<tr>
<td>Error</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>( (IJ - 1) )</td>
</tr>
</tbody>
</table>

There is no degree of freedom for error. The problem is that the two factor interaction effect and random error component are subsumed together and cannot be separated out. There is no estimate for \( \sigma^2 \).
If no interaction exists, then $H_0: \gamma_{ij} = 0$ for all $i, j$ is accepted and the additive model $y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$ is well enough to test the hypothesis $H_0: \alpha_i = 0$ and $H_0: \beta_j = 0$ with error having $(I-1)(J-1)$ degrees of freedom.

If interaction exists, then $H_0: \gamma_{ij} = 0$ is rejected. In such a case, if we assume that the structure of interaction effect is such that it is proportional to the product of individual effects, i.e.,

$$
\gamma_{ij} = \lambda \alpha_i \beta_j
$$

then a test for testing $H_0: \lambda = 0$ can be constructed. Such a test will serve as a test for nonadditivity. It will help in knowing the effect of presence of interact effect and whether the interaction enters into the model additively. Such a test is given by Tukey’s test for nonadditivity which requires one degree of freedom leaving $(I-1)(J-1)-1$ degrees of freedom for error.

Let us assume that departure from additivity can be specified by introducing a product term and writing the model as

$$
E(y_{ij}) = \mu + \alpha_i + \beta_j + \lambda \alpha_i \beta_j; \quad i = 1, 2, \ldots, I, \quad j = 1, 2, \ldots, J \quad \text{with} \quad \sum_{i=1}^{I} \alpha_i = 0, \quad \sum_{j=1}^{J} \beta_j = 0.
$$

When $\lambda \neq 0$, the model becomes nonlinear model and the least squares theory for linear models is not applicable.
Note that using $\sum_{i=1}^{I} \alpha_i = 0$, $\sum_{j=1}^{J} \beta_j = 0$, we have

\[
\bar{y}_{oo} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} y_{ij} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} \left[ \mu + \alpha_i + \beta_j + \lambda \alpha_i \beta_j + \epsilon_{ij} \right]
\]

\[
= \mu + \frac{1}{I} \sum_{i=1}^{I} \alpha_i + \frac{1}{J} \sum_{j=1}^{J} \beta_j + \frac{\lambda}{IJ} \left( \sum_{i=1}^{I} \alpha_i \right) \left( \sum_{j=1}^{J} \beta_j \right) + \bar{\epsilon}_{oo}
\]

\[
= \mu + \bar{\epsilon}_{oo}
\]

$E(\bar{y}_{oo}) = \mu$

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

Next

\[
\bar{y}_{io} = \frac{1}{J} \sum_{j=1}^{J} y_{ij} = \frac{1}{J} \sum_{j=1}^{J} \left[ \mu + \alpha_i + \beta_j + \lambda \alpha_i \beta_j + \epsilon_{ij} \right]
\]

\[
= \mu + \alpha_i + \frac{1}{J} \sum_{j=1}^{J} \beta_j + \lambda \alpha_i + \frac{1}{J} \sum_{j=1}^{J} \beta_j + \bar{\epsilon}_{io}
\]

\[
= \mu + \alpha_i + \bar{\epsilon}_{io}
\]

$E(\bar{y}_{io}) = \mu + \alpha_i$

$\Rightarrow \hat{\alpha}_i = \bar{y}_{io} - \hat{\mu} = \bar{y}_{io} - \bar{y}_{oo}$.
Similarly,
\[ \bar{y}_{oj} = \mu + \beta_j \]
\[ \Rightarrow \hat{\beta}_j = \bar{y}_{oj} - \hat{\mu} = \bar{y}_{oj} - \bar{y}_{oo}. \]

Thus \( \hat{\mu}, \hat{\alpha}_i \) and \( \hat{\beta}_j \) remain the unbiased estimators of \( \mu, \alpha_i \) and \( \beta_j \), respectively irrespective of whether \( \lambda = 0 \) or not.

Also
\[ E\left[y_{ij} - \bar{y}_{io} - \bar{y}_{aj} + \bar{y}_{oo}\right] = \lambda \alpha_i \beta_j \]

or
\[ E\left[(y_{ij} - \bar{y}_{oo}) - (\bar{y}_{io} - \bar{y}_{oo}) - (\bar{y}_{oj} - \bar{y}_{oo})\right] = \lambda \alpha_i \beta_j. \]

Consider the estimation of \( \mu, \alpha_i, \beta_j, \) and \( \lambda \) based on the minimization of
\[ S = \sum_i \sum_j (y_{ij} - \mu - \alpha_i - \beta_j - \lambda \alpha_i \beta_j)^2 \]
\[ = \sum_i \sum_j S_{ij}^2. \]
The normal equations are solved as

\[
\frac{\partial S}{\partial \mu} = 0 \Rightarrow \sum_{i=1}^{I} \sum_{j=1}^{J} S_{ij} = 0
\]

\[
\Rightarrow \hat{\mu} = \bar{y}_{oo}
\]

\[
\frac{\partial S}{\partial \alpha_i} = 0 \Rightarrow \sum_{j=1}^{J} (1 + \lambda \beta_j) S_{ij} = 0
\]

\[
\frac{\partial S}{\partial \beta_j} = 0 \Rightarrow \sum_{i=1}^{I} (1 + \lambda \alpha_i) S_{ij} = 0
\]

\[
\frac{\partial S}{\partial \lambda} = 0 \Rightarrow \sum_{i=1}^{I} \sum_{j=1}^{J} \alpha_i \beta_j S_{ij} = 0
\]

or \[
\sum_{i=1}^{I} \sum_{j=1}^{J} \alpha_i \beta_j (y_{ij} - \mu - \alpha_i - \beta_j - \lambda \alpha_i \beta_j) = 0
\]

\[
\lambda = \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} \alpha_i \beta_j y_{ij}}{\left( \sum_{i=1}^{I} \alpha_i^2 \right) \left( \sum_{j=1}^{J} \beta_j^2 \right)} = \tilde{\lambda} \text{ (say)}
\]

which can be estimated provided \( \alpha_i \) and \( \beta_j \) are assumed to be known.
Since $\alpha_i$ and $\beta_j$ can be estimated by
\[ \hat{\alpha}_i = \bar{y}_{io} - \bar{y}_{oo} \]
and
\[ \hat{\beta}_j = \bar{y}_{oj} - \bar{y}_{oo} \]
irrespective of whether $\lambda \neq 0$ or $\lambda = 0$ so we can substitute them in place of $\alpha_i$ and $\beta_j$ in $\hat{\lambda}$ which gives

\[ \hat{\lambda} = \frac{\sum_{i=1}^{I} \sum_{j=1}^{J} \hat{\alpha}_i \hat{\beta}_j y_{ij}}{\left( \sum_{i=1}^{I} \hat{\alpha}_i^2 \right) \left( \sum_{j=1}^{J} \hat{\beta}_j^2 \right)} \]

\[ = \frac{(IJ)\sum_{i=1}^{I} \sum_{j=1}^{J} \hat{\alpha}_i \hat{\beta}_j y_{ij}}{\left( I \sum_{i=1}^{I} \hat{\alpha}_i^2 \right) \left( J \sum_{j=1}^{J} \hat{\beta}_j^2 \right)} \]

\[ = \frac{IJ \sum_{i=1}^{I} \sum_{j=1}^{J} (\bar{y}_{io} - \bar{y}_{oo})(\bar{y}_{oj} - \bar{y}_{oo}) y_{ij}}{S_A S_B} \]

where $S_A = J \sum_{i=1}^{I} \hat{\alpha}_i^2 = J \sum_{i=1}^{I} (\bar{y}_{io} - \bar{y}_{oo})^2$

$S_B = I \sum_{j=1}^{J} \hat{\beta}_j^2 = I \sum_{j=1}^{J} (\bar{y}_{oj} - \bar{y}_{oo})^2$. 

Assuming $\alpha_i$ and $\beta_j$ to be known, we find

$$Var(\hat{\lambda}) = \left(\frac{1}{\sum_{i=1}^{I} \alpha_i^2 \sum_{j=1}^{J} \beta_j^2}\right)^2 \left[\sum_{i=1}^{I} \sum_{j=1}^{J} \alpha_i^2 \beta_j^2 Var(y_{ij}) + 0\right]$$

$$= \frac{\sigma^2 \left(\sum_{i=1}^{I} \alpha_i^2\right) \left(\sum_{j=1}^{J} \beta_j^2\right)}{\left(\sum_{i=1}^{I} \alpha_i^2\right)^2 \left(\sum_{j=1}^{J} \beta_j^2\right)^2}$$

$$= \frac{\sigma^2}{\left(\sum_{i=1}^{I} \alpha_i^2\right) \left(\sum_{j=1}^{J} \beta_j^2\right)}$$

using $Var(y_{ij}) = \sigma^2$, $Cov(y_{ij}, y_{jk}) = 0$ for all $i \neq k$. 

Since \( \alpha_i \) and \( \beta_j \) can be estimated by \( \hat{\alpha}_i \) and \( \hat{\beta}_j \) then substitute them back in the expression of \( \text{Var}(\hat{\lambda}) \) and treating it as \( \text{Var}(\hat{\lambda}) \) gives

\[
\text{Var}(\hat{\lambda}) = \frac{\sigma^2}{\left( \sum_{i=1}^{I} \hat{\alpha}_i \right) \left( \sum_{j=1}^{J} \hat{\beta}_j \right)}
\]

\[
= \frac{IJ \sigma^2}{S_A S_B}
\]

for given \( \hat{\alpha}_i \) and \( \hat{\beta}_j \).