Dr. Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology Kanpur
Now we find the expectations of SSTr and SSE.

\[
E(\text{SSTr}) = E \left[ \sum_{i=1}^{p} n_i (\bar{y}_{io} - \bar{y}_{oo})^2 \right] \\
= E \left[ \sum_{i=1}^{p} n_i \left\{ (\mu + \alpha_i + \bar{e}_{io}) - (\mu + \bar{e}_{oo}) \right\}^2 \right]
\]

where

\[
\bar{e}_{io} = \frac{1}{n_i} \sum_{j=1}^{n_i} e_{ij},
\]

\[
\bar{e}_{oo} = \frac{1}{n} \sum_{i=1}^{p} \sum_{j=1}^{n_i} e_{ij}
\]

and

\[
\sum_{i=1}^{p} \frac{n_i \alpha_i}{n} = 0.
\]

\[
E(\text{SSTr}) = E \left[ \sum_{i=1}^{p} n_i \left\{ \alpha_i + (\bar{e}_{io} - \bar{e}_{oo}) \right\}^2 \right] \\
= \sum_{i=1}^{p} n_i E(\alpha_i^2) + \sum_{i=1}^{p} n_i E(\bar{e}_{io} - \bar{e}_{oo})^2 + 0.
\]
Since

\[ E(\bar{\varepsilon}_{io}^2) = Var(\bar{\varepsilon}_{io}) = Var\left(\frac{1}{n} \sum_{j=1}^{n_i} \varepsilon_{ij}\right) = \frac{1}{n_i} n_i \sigma^2 = \frac{\sigma^2}{n_i} \]

\[ E(\bar{\varepsilon}_{oo}^2) = Var(\bar{\varepsilon}_{oo}) = Var\left(\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n_i} \varepsilon_{ij}\right) = \frac{1}{n^2} n \sigma^2 = \frac{\sigma^2}{n} \]

\[ E(\bar{\varepsilon}_{io} \bar{\varepsilon}_{oo}) = Cov(\bar{\varepsilon}_{io}, \bar{\varepsilon}_{oo}) \]

\[ = \frac{1}{n_i n} Cov\left(\sum_{j=1}^{n_i} \varepsilon_{ij}, \sum_{i=1}^{p} \sum_{j=1}^{n_i} \varepsilon_{ij}\right) \]

\[ = \frac{n_i \sigma^2}{n_i n} = \frac{\sigma^2}{n}. \]

\[ E(SSTr) = \sum_{i=1}^{p} n_i \alpha_i^2 + \sigma^2 \sum_{i=1}^{p} n_i \left(\frac{1}{n_i} - \frac{1}{n}\right) \]

\[ = \sum_{i=1}^{p} n_i \alpha_i^2 + (p-1) \sigma^2 \]

or

\[ E\left(\frac{SSTr}{p-1}\right) = \sigma^2 + \frac{\sum_{i=1}^{p} n_i \alpha_i^2}{p-1} \]

or

\[ E(MSTr) = \sigma^2 + \frac{\sum_{i=1}^{p} n_i \alpha_i^2}{p-1}. \]
Next

\[ E(SSE) = E \left[ \sum_{i=1}^{p} \sum_{j=1}^{n_i} (y_{ij} - \overline{y}_{io})^2 \right] \]

\[ = E \left[ \sum_{i=1}^{p} \sum_{j=1}^{n_i} \left( (\mu + \alpha_i + \epsilon_{ij}) - (\mu + \alpha_i + \overline{\epsilon}_{io}) \right)^2 \right] \]

\[ = E \left[ \sum_{i=1}^{p} \sum_{j=1}^{n_i} (\epsilon_{ij} - \overline{\epsilon}_{io})^2 \right] \]

\[ = \sum_{i=1}^{p} \sum_{j=1}^{n_i} E(\epsilon_{ij}^2 + \overline{\epsilon}_{io}^2 - 2\epsilon_{ij}\overline{\epsilon}_{io}) \]

\[ = \sum_{i=1}^{p} \sum_{j=1}^{n_i} \left( \sigma^2 + \frac{\sigma^2}{n_i} - \frac{2\sigma^2}{n_i} \right) \]

\[ = \sigma^2 \sum_{i=1}^{p} \sum_{j=1}^{n_i} \left( \frac{n_i - 1}{n_i} \right) \]

\[ = \sigma^2 \sum_{i=1}^{p} n_i(n_i - 1) \]

\[ = \sigma^2 \sum_{i=1}^{p} (n_i - 1) \]

\[ = (n - p)\sigma^2 \]

or \[ E\left(\frac{SSE}{n - p}\right) = \sigma^2 \]

or \[ E(MSE) = \sigma^2 \]

Thus \(MSE\) is an unbiased estimator of \(\sigma^2\).
Suppose the response of an outcome is affected by the two factors – $A$ and $B$. For example, suppose $I$ varieties of mangoes are grown on $I$ different plots of same size in each of the $J$ different locations. All plots are given same treatment like equal amount of water, equal amount of fertilizer etc. So there are two factors in the experiment which affect the yield of mangoes:

- Location ($A$)
- Variety of mangoes ($B$).

Such an experiment is called two-factor experiment. The different locations correspond to the different levels of $A$ and the different varieties correspond to the different levels of factor $B$. The observations are collected on the basis of per plot.

The combined effect of the two factors ($A$ and $B$ in our case) is called the interaction effect (of $A$ and $B$).

Mathematically, let $a$ and $b$ be the levels of factors $A$ and $B$ respectively then a function $f(a,b)$ is called a function of no interaction if and only if there exists functions $g(a)$ and $h(b)$ such that $f(a,b) = g(a) + h(b)$. Otherwise the factors are said to interact.

For a function $f(a,b)$ of no interaction,

\[
\begin{align*}
    f(a_1,b) &= g(a_1) + h(b) \\
    f(a_2,b) &= g(a_2) + h(b) \\
    \Rightarrow f(a_1,b) - f(a_2,b) &= g(a_1) - g(a_2)
\end{align*}
\]

and so it is independent of $b$. Such no interaction functions are called additive functions.
Now there are two options:

- Only one observation per plot is collected.
- More than one observations per plot are collected.

If there is only one observation per plot then there cannot be any interaction effect among the observations and we assume it to be zero.

If there are more than one observations per plot then interaction effect among the observations can be considered.

We consider here two cases:

1. One observation per plot in which the interaction effect is zero.
2. More than one observations per plot in which the interaction effect is present.
Two-way classification without interaction

Let $y_{ij}$ be the response of observation from the $i^{th}$ level of first factor, say $A$ and the $j^{th}$ level of second factor, say $B$. So assume $y_{ij}$ are independently distributed as $N(\mu, \sigma^2)$ $i = 1, 2, \ldots, I, j = 1, 2, \ldots, J$.

This can be represented in the form of a linear model as

$$E(Y_{ij}) = \mu_{ij}$$

$$= \mu_{oo} + (\mu_{io} - \mu_{oo}) + (\mu_{oj} - \mu_{oo}) + (\mu_{ij} - \mu_{io} - \mu_{oj} + \mu_{oo})$$

$$= \mu + \alpha_i + \beta_j + \gamma_{ij}$$

where

$$\mu = \mu_{oo}$$

$$\alpha_i = \mu_{io} - \mu_{oo}$$

$$\beta_j = \mu_{oj} - \mu_{oo}$$

$$\gamma_{ij} = \mu_{ij} - \mu_{io} - \mu_{oj} + \mu_{oo}$$

with

$$\sum_{i=1}^{I} \alpha_i = \sum_{i=1}^{I} (\mu_{io} - \mu_{oo}) = 0$$

$$\sum_{j=1}^{J} \beta_j = \sum_{j=1}^{J} (\mu_{oj} - \mu_{oo}) = 0.$$
Here

$\alpha_i : \text{effect of } i^{th} \text{ level of factor } A$

or excess of mean of $i^{th}$ level of $A$ over the general mean.

$\beta_j : \text{effect of } j^{th} \text{ level of } B$

or excess of mean of $j^{th}$ level of $B$ over the general mean.

$\gamma_{ij} : \text{Interaction effect of } i^{th} \text{ level of } A \text{ and } j^{th} \text{ level of } B.$

Here we assume $\gamma_{ij} = 0$ as we have only one observation per plot.

We also assume that the model $E(Y_{ij}) = \mu_{ij}$ is a full rank model so that $\mu_{ij}$ and all linear parametric functions of $\mu_{ij}$ are estimable.

The total number of observations are $I \times J$ which can be arranged in a two way classified $I \times J$ table where the rows corresponds to the different levels of $A$ and the column corresponds to the different levels of $B.$
The observations on $Y$ design matrix $X$ in this case are

<table>
<thead>
<tr>
<th>$Y$</th>
<th>$\mu$</th>
<th>$\alpha_1$</th>
<th>$\alpha_2$</th>
<th>$\cdots$</th>
<th>$\alpha_i$</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\cdots$</th>
<th>$\beta_j$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y_{11}$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
</tr>
<tr>
<td>$y_{12}$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>$\cdots$</td>
<td>0</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\ddots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\ddots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$y_{1,j}$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>1</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\ddots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\ddots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$y_{i1}$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$\cdots$</td>
<td>0</td>
</tr>
<tr>
<td>$y_{i2}$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>$\cdots$</td>
<td>0</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\ddots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\ddots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$y_{i,j}$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>$\cdots$</td>
<td>1</td>
</tr>
</tbody>
</table>

If the design matrix is not of full rank, then the model can be reparameterized. In such a case, we can start the analysis by assuming that the model $E(Y_{ij}) = \mu + \alpha_i + \beta_j$ is obtained after reparameterization.

There are two null hypothesis of interest:

$$H_{0\alpha} : \alpha_1 = \alpha_2 = \cdots = \alpha_i = 0$$
$$H_{0\beta} : \beta_1 = \beta_2 = \cdots = \beta_j = 0$$
Now we derive the least squares estimators (or equivalently the maximum likelihood estimator) of $\mu, \alpha_i$ and $\beta_j, \ i = 1, 2, \ldots, I, \ j = 1, 2, \ldots, J$ by minimizing the error sum of squares

$$E = \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \mu - \alpha_i - \beta_j)^2.$$ 

The normal equations are obtained as

$$\frac{\partial E}{\partial \mu} = 0 \Rightarrow -2 \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \mu - \alpha_i - \beta_j) = 0$$

$$\frac{\partial E}{\partial \alpha_i} = 0 \Rightarrow -2 \sum_{j=1}^{J} (y_{ij} - \mu - \alpha_i - \beta_j) = 0, \ i = 1, 2, \ldots, I$$

$$\frac{\partial E}{\partial \beta_j} = 0 \Rightarrow -2 \sum_{i=1}^{I} (y_{ij} - \mu - \alpha_i - \beta_j) = 0, \ j = 1, 2, \ldots, J.$$ 

Solving the normal equations and using $\sum_{i=1}^{I} \alpha_i = 0$ and $\sum_{j=1}^{J} \beta_j = 0$, the least squares estimator are obtained as

$$\hat{\mu} = \frac{1}{IJ} \sum_{i=1}^{I} \sum_{j=1}^{J} y_{ij} = \frac{G}{IJ} = \bar{y}_{oo}$$

$$\hat{\alpha}_i = \frac{1}{J} \sum_{j=1}^{J} y_{ij} - \bar{y}_{oo} = \frac{T_i}{J} - \bar{y}_{oo} = \bar{y}_{io} - \bar{y}_{oo}, \ i = 1, 2, \ldots, I$$

$$\hat{\beta}_j = \frac{1}{I} \sum_{i=1}^{I} y_{ij} - \bar{y}_{oo} = \frac{B_j}{I} - \bar{y}_{oo} = \bar{y}_{oj} - \bar{y}_{oo}, \ j = 1, 2, \ldots, J$$

where

$T_i$: treatment totals due to $i^{th}$ $\alpha$ effect, i.e., sum of all the observations receiving the $i^{th}$ treatment.

$B_j$: block totals due to $j^{th}$ $\beta$ effect, i.e., sum of all the observations in the $j^{th}$ block.
Thus the error sum of squares is

\[
SSE = \text{Min}_{\mu, \beta_j} \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \mu - \beta_j)^2
\]

\[
= \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \hat{\mu}_i - \hat{\alpha}_i - \hat{\beta}_j)^2
\]

\[
= \sum_{i=1}^{I} \sum_{j=1}^{J} \left[ (y_{ij} - \bar{y}_{oo}) - (\bar{y}_{io} - \bar{y}_{oo}) - (\bar{y}_{oj} - \bar{y}_{oo}) \right]^2
\]

\[
= \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \bar{y}_{io} - \bar{y}_{oj} + \bar{y}_{oo})^2
\]

\[
= \sum_{i=1}^{I} \sum_{j=1}^{J} (y_{ij} - \bar{y}_{oo})^2 - J \sum_{i=1}^{I} \sum_{j=1}^{J} (\bar{y}_{io} - \bar{y}_{oo})^2 - I \sum_{i=1}^{I} \sum_{j=1}^{J} (\bar{y}_{oj} - \bar{y}_{oo})^2
\]

which carries

\[
IJ - (I - 1) - (J - 1) - 1 = (I - 1)(J - 1)
\]

degrees of freedom.