In field experiments, the plot size for experimentation is selected for achieving a prescribed degree of precision for measurement of the character of primary interest, usually the yield of crop. The character under study is measured on the whole of the plot. The number of plants per plot is generally large. As such it is not possible to measure the characters like plant height, leaf area, rate of growth, period of flowering, extent of damage by pests and diseases etc. for all the plants in each plot due to paucity of resources (time and expenses). Hence, in order to save expenses and time, sampling methodology is adopted to select sampling units from each plot and to collect data from the selected units that can provide necessary information of the respective plots. For clarity consider the following examples.

Example 1.1: An experiment with 30 indigenous and exotic collections / cultivars of garden pea was conducted at Division of Vegetable crops, IARI, New Delhi during 1998. There were approximately 300 plants per plot. The characters observed were the day of appearance of first flower, first flowering node number, length of internodes, the day of 50% flowering, the day of first green pod harvest, number of primary branches, the length of pod, the number of seeds per pod, the number of pods per plant, the pod yield per plant, seed yield per plant, plant height, etc. It was not possible to take observations on 300 plants for each character. For convenience, the measurements were made only from 10 selected plants out of 300 plants in each of the plots for characters mentioned above.

Example 1.1.2: An experiment was conducted with 24 different combinations of fertilizers and their methods of application on the Chickpea at the Division of Agronomy, IARI, New Delhi. There are approximately 200 plants in each plot. The characters observed were the number of pods per plant, number of primary branches per plant, number of secondary branches per plant, number of grains per plant, grain weight per plant, etc. It was not feasible to take observations on 200 plants for each character. For convenience, the observations were collected only from 5 of the 200 plants in a plot for the characters mentioned above.

One may also count tiller numbers only 1 m² of the 15 m² plot, for leaf area, measure from only 20 of the approximately 2000 leaves in the plot etc.

The plants that are selected for collecting the measurements of characteristic under study constitute the sample, and the process of selection of some plants from a plot is termed as plot sampling technique.

For plot sampling, number of plants in each plot is a population. The character under study is estimated from some selected plants. The selected plants are known as a sample. An appropriate sample is one that provides an estimate, or a sample value, that is as close as possible to the value that would have been obtained had all plant in the plot been measured –
the plot value. The difference between the sample value and the plot value constitutes the sampling error. Thus a good sampling technique is one that gives small sampling error. In this lecture, we shall discuss the basic features of sampling technique as applied to field trials i.e. Plot Sampling. To develop a plot sampling technique for measurement of a character in a given trial, we must specify the sampling unit, the sample size, and the sampling design.

The sampling unit is the unit on which actual measurement is made. Some commonly used sampling units in a replicated field trial are a leaf, a plant, a group of plants, a unit area etc. The important features of an appropriate sampling unit are:

- **Ease of Identification**
- **Ease of Measurement**
- **High Precision**
- **Low Cost**

The number of sampling units taken from the population is known as sample size. In a replicated field trial where each plot is a population, sample size could be the number of plants per plot used for measuring plant height, the number of leaves per plot used for measuring leaf area, or the number of hills per plot used for counting tillers, etc. The required sample size for a particular experiment is governed by:

(i) The size of the variability among sampling units within the same plot (sampling variance).
(ii) The degree of precision desired for the character of interest.

In practice, the size of the sampling variance for most plant characters is generally not known. However, its value can be estimated or assumed to be known. The desired level of precision can, however, be prescribed by the researcher based on experimental objective and previous experience, in terms of the margin of error, either of the plot mean or of the treatment mean.

The sample size for a simple random sampling design that can satisfy a prescribed margin of error of the plot mean is computed as:

\[ n = \frac{(Z_{\alpha/2})^2 v_s}{d^2 \bar{X}^2} \]

where \( n \) is the required sample size, \( Z_{\alpha/2} \) is the value of the standardized normal variate corresponding to the level of significance \( \alpha \), \( v_s \) is the sampling variance, \( \bar{X} \) is the mean value, and \( d \) is the margin of error expressed as a fraction of the plot mean.

The information of primary interest to the researcher is usually the treatment means (the average over all plots receiving the same treatment) or actually the difference of means, rather than the plot mean (the value from a single plot). Thus, the desired degree of precision is usually specified in terms of the margin of error of the treatment mean rather than that of the plot mean. In such a case, sample size is computed as:
\[ n = \frac{(Z_{\alpha/2})^2 \left( v_s \right)}{r \left( D^2 \right) \left( X^2 \right) - (Z_{\alpha/2})^2 \left( v_p \right)} \]

where \( n \) is the required sample size, \( r \) is the number of replications, \( Z_{\alpha/2} \) and \( v_s \) are as defined earlier, \( v_p \) is the variance between plots of the same treatment \( (i.e. \) experimental error), and \( D \) is the prescribed margin of error expressed as a fraction of the treatment mean. In this case, additional information on the size of the experimental error \( (v_p) \) is needed to compute sample size.

A sampling design specifies the manner in which the \( n \) sampling units are to be selected from the whole plot. The most commonly used sampling designs in replicated field trials are:

1. Simple random sampling design.
2. Multistage random sampling design.
3. Stratified random sampling design.
4. Stratified multistage random sampling design.
5. Sub-sampling with an auxiliary variable.

It is well known that precision of a sample estimate generally increases with the size of the sampling unit, sample size, and the complexity of the sampling design used. However, an increase in either the size or the number of sampling units almost always increases the cost. Therefore, a proper balance has to be maintained between the size of the sampling unit, sample size and a sampling design on one hand and the cost at the other. This task requires the information on the variability of the character of interest so that the precision that will result from the various types of sampling techniques can be estimated. There are three sources of the data from which the required information can be obtained: (i) from previous trials, (ii) additional data from ongoing experiments or (iii) specifically planned sampling studies.

**Analysis of Data from Previous Trials**

The various steps involved in the analysis of sampled data is described here considering a block design setting. Consider an experimental situation in which \( v \) treatments are to be compared via \( N = vr \) experimental units (plots) arranged in \( r \) blocks each of size \( v \) such that each treatment occurs exactly once in each block \( i.e. \) the experiment is conducted using a randomized complete block (RCB) design. Let \( n \) plants be selected from each plot and observations are made from \( n \) selected plants. The response variable can be represented by a linear, additive, fixed effect model as

\[ Y_{ijt} = \mu + \tau_i + \beta_j + e_{ij} + \eta_{ijt} \quad (1.1) \]

where \( Y_{ijt} \) is the observation pertaining to the \( t^{th} \) sampling unit for the \( i^{th} \) treatment in the \( j^{th} \) block \( (i = 1, 2, ..., v; j = 1, 2, ..., r; t = 1, 2, ..., n) \), \( \mu \) is the general mean effect; \( \tau_i \) is \( i^{th} \) treatment effect, \( \beta_j \) is the effect \( j^{th} \) block effect, \( e_{ij} \) is the plot error distributed as \( N(0, \sigma_e^2) \), \( \eta_{ijt} \) is the sampling error distributed as \( N(0, \sigma_s^2) \). The analysis of variance (ANOVA) for the above is given as
Sampling in Field Experiments

### ANOVA for a RCB design
(Based on individual observations)

<table>
<thead>
<tr>
<th>Sources of variation</th>
<th>Degrees of freedom (DF)</th>
<th>Sum of squares (SS)</th>
<th>Mean Square (MS)</th>
<th>Expected Mean Square (MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocks</td>
<td>$r-1$</td>
<td>SSB</td>
<td></td>
<td>$\sigma_s^2 + n\sigma_e^2 + \frac{rn}{v-1} \sum_{i=1}^{v} \tau_i^2$</td>
</tr>
<tr>
<td>Treatments</td>
<td>$v-1$</td>
<td>SST</td>
<td></td>
<td>$\sigma_s^2 + n\sigma_e^2$</td>
</tr>
<tr>
<td>Treatments x blocks</td>
<td>$(v-1)(r-1)$</td>
<td>SSBT</td>
<td>$MSBT$</td>
<td>$\sigma_s^2 + n\sigma_e^2$</td>
</tr>
<tr>
<td>(experimental error)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sampling Error</td>
<td>$rv(n-1)$</td>
<td>SSSE</td>
<td>$MSSE$</td>
<td>$\sigma_s^2$</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>$bvn-1$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The sum of squares due to different components of ANOVA can be obtained as follows:

Form a $r \times v$ two-way table between blocks and treatments, each cell figure being the total over all samples from a plot.

<table>
<thead>
<tr>
<th>Blocks</th>
<th>Treatments</th>
<th>Block Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$T_{11}$</td>
<td>$B_{11}$</td>
</tr>
<tr>
<td>:</td>
<td>$T_{1i}$</td>
<td>$B_{i1}$</td>
</tr>
<tr>
<td>:</td>
<td>$T_{i1}$</td>
<td>$B_{12}$</td>
</tr>
<tr>
<td>:</td>
<td>$T_{ij}$</td>
<td>$B_{ij}$</td>
</tr>
<tr>
<td>:</td>
<td>$T_{ir}$</td>
<td>$B_{ir}$</td>
</tr>
<tr>
<td>$r$</td>
<td>$T_{1r}$</td>
<td>$B_{r}$</td>
</tr>
<tr>
<td>Treatment Totals</td>
<td>$T_1$, $T_2$, $\ldots$, $T_i$, $\ldots$, $T_v$, Grand Total</td>
<td></td>
</tr>
</tbody>
</table>

The sum of squares (S.S) due to different components of ANOVA can be obtained as follows:

- Grand Total (G.T.) = $\sum_{i=1}^{v} \sum_{j=1}^{r} \sum_{t=1}^{n} y_{ijt}$
- Correction factor (C.F.) = $\frac{(G.T.)^{2}}{rvn}$
- Total S.S. of the table (TSS) = $\sum_{i=1}^{v} \sum_{j=1}^{r} (\sum_{t=1}^{n} y_{ijt})^2 / n - C.F$
- $T_i = i^{th}$ Treatment total = $\sum_{j=1}^{r} \sum_{t=1}^{n} y_{ijt}$
Sampling In Field Experiments

\[ B_i = j^{th} \text{ Block total} = \sum_{i=1}^{v} \sum_{t=1}^{n} y_{ijt} \]

Treatment S.S (SST) = \[ \sum_{i=1}^{v} T_i^2 \] / \[ nr - C.F. \]

Block S.S (SSB) = \[ \sum_{j=1}^{r} B_j^2 \] / \[ nv - C.F. \]

Block x Treatment S.S (SSBT) = TSS – SST – SSB

Total S.S. of the entire data = \[ \sum_{i=1}^{v} \sum_{j=1}^{r} \sum_{t=1}^{n} y_{ijt}^2 - C.F. \]

Sum of squares due to the sampling error (SSSE) = Total S.S of the entire data – SSB – SST – SSB

Using the expressions of expected mean squares in the above ANOVA table, it is clear that the null hypothesis regarding the equality of treatment effects is tested against the experimental error. From the ANOVA, it is also clear that the sampling error is estimated as \( \hat{\sigma}_s^2 = s^2 \).

The experimental error (variance between plots of the same treatment) is estimated as \( \hat{\sigma}_e^2 = \frac{s_1^2 - s_2^2}{n} \). When \( \hat{\sigma}_e^2 \) is negative, it is taken as zero.

The variance of the \( i^{th} \) treatment mean (\( \bar{Y}_{i..} \)) based on \( r \)-replications and \( n \)-samples per plot

\[ = \frac{\sigma_s^2 + n\sigma_e^2}{rn} \]

The estimated variance of (\( \bar{Y}_{i..} \)) = \( \frac{(\hat{\sigma}_s^2 + n\hat{\sigma}_e^2)}{rn} \)

Taking the number of sampling units in a plot to be large (infinite), the estimated variance of a treatment mean when there is complete recording (i.e. the entire plot is harvested) = \( \frac{\hat{\sigma}_e^2}{r} \)

The efficiency of sampling as compared to complete enumeration

\[ \frac{\hat{\sigma}_e^2 / r}{(\hat{\sigma}_s^2 + n\hat{\sigma}_e^2) / rn} \]

The standard error of a treatment mean (\( \bar{Y}_{i..} \)) with \( n \) samples per plot and \( r \) replications is
The coefficient of variation is

\[
p = \left( \frac{\hat{\sigma}_s^2}{r n} + \frac{\hat{\sigma}_e^2}{r} \right)^{1/2} \left( \frac{100}{\bar{Y}_{i..}} \right)
\]

Thus, \( n \) is given as

\[
n = \frac{\hat{\sigma}_s^2}{r} \left\{ \frac{1}{p^2 \left( \frac{\bar{Y}_{i..}}{100} \right)^2 - \frac{\hat{\sigma}_e^2}{r}} \right\}
\]

Generally, the margin of error \( \bar{d} \) or \( \bar{D} \) is \( Z_{\alpha/2} \) times the value of coefficient of variation of \( \bar{Y}_{i..} \) based on the concept of 100(1-\( \alpha \)%), confidence intervals. Therefore,

\[
n = \frac{\hat{\sigma}_s^2}{r} \frac{Z_{\alpha/2}^2}{D^2 \left( \frac{\bar{Y}_{i..}}{100} \right)^2 - Z_{\alpha/2}^2 \frac{\hat{\sigma}_e^2}{r}}
\]

This formula can easily be simplified to

\[
n = \frac{(Z_{\alpha/2}^2 \hat{\sigma}_s^2)}{r (D^2 \left( \frac{\bar{Y}_{i..}}{100} \right)^2) - (Z_{\alpha/2}^2 \hat{\sigma}_e^2)}
\]

For any given \( r \) and \( p(D) \), there will be \( t \) values for \( n \) corresponding to the \( t \) treatment means. The maximum \( n \) will ensure the estimation of any treatment mean with a standard error not exceeding \( p \) percent or margin of error not exceeding \( D \).

**Example 1.1:** [Nigam and Gupta (1979)]. To study the effect of the differences in the number of plants per hill on the growth of Maize crop, a RCB design was laid out at the Agricultural College Farm, Poona. The treatments tried were A – one plant per hill, B – two plant per hill, C – three plant per hill, D – four plants per hill. The table below gives the data on the length (in inches) of 5 cobs randomly selected from each plot. The table below gives the data on the length (in inches) of 5 cobs randomly selected from each plot.
Sampling In Field Experiments

Data on cob lengths (in inches)

<table>
<thead>
<tr>
<th>Blocks</th>
<th>Cob Number</th>
<th>Treatments</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>9.30</td>
<td>9.00</td>
<td>8.60</td>
<td>6.40</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8.80</td>
<td>9.00</td>
<td>7.00</td>
<td>7.20</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>9.00</td>
<td>10.50</td>
<td>8.40</td>
<td>6.80</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>8.80</td>
<td>8.90</td>
<td>9.10</td>
<td>7.70</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>8.60</td>
<td>9.20</td>
<td>8.20</td>
<td>6.00</td>
</tr>
<tr>
<td>II</td>
<td>1</td>
<td>10.20</td>
<td>9.70</td>
<td>9.00</td>
<td>6.40</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>9.20</td>
<td>10.00</td>
<td>8.00</td>
<td>7.40</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>9.40</td>
<td>9.20</td>
<td>8.10</td>
<td>6.80</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>9.60</td>
<td>10.50</td>
<td>8.20</td>
<td>6.80</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>9.80</td>
<td>10.30</td>
<td>7.00</td>
<td>6.60</td>
</tr>
<tr>
<td>III</td>
<td>1</td>
<td>9.90</td>
<td>8.40</td>
<td>7.50</td>
<td>6.30</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>10.40</td>
<td>9.40</td>
<td>7.50</td>
<td>6.70</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>11.00</td>
<td>8.20</td>
<td>8.50</td>
<td>6.00</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>10.80</td>
<td>9.10</td>
<td>8.00</td>
<td>7.00</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>10.00</td>
<td>9.80</td>
<td>8.60</td>
<td>7.30</td>
</tr>
<tr>
<td>IV</td>
<td>1</td>
<td>10.60</td>
<td>8.80</td>
<td>7.00</td>
<td>8.40</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>9.20</td>
<td>9.30</td>
<td>7.30</td>
<td>7.80</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>9.90</td>
<td>9.90</td>
<td>7.60</td>
<td>8.00</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>10.40</td>
<td>9.00</td>
<td>6.70</td>
<td>8.40</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>9.90</td>
<td>8.00</td>
<td>6.50</td>
<td>7.50</td>
</tr>
<tr>
<td>V</td>
<td>1</td>
<td>10.40</td>
<td>11.00</td>
<td>9.90</td>
<td>7.70</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>9.00</td>
<td>10.40</td>
<td>9.00</td>
<td>7.00</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>9.70</td>
<td>9.00</td>
<td>8.90</td>
<td>7.00</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>9.30</td>
<td>10.20</td>
<td>8.90</td>
<td>6.70</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>9.60</td>
<td>9.60</td>
<td>9.40</td>
<td>7.20</td>
</tr>
</tbody>
</table>

(a) Analyze the data and find the standard error of treatment means.
(b) Find out the relative efficiency of sampling.
(c) Obtain minimum number of sampling units per plot necessary to estimate the treatment means with 4 and 5 per cent standard error or approximately 8% and 10% of margin of error when the number of replications are 5 and 6.

Calculations

Step 1: Obtain various sum of squares using the formulae given above and the procedure given by Nigam and Gupta (1979) and Gomez and Gomez (1984) and the Analysis of Variance table can be obtained. The ANOVA Table is given as

ANOVA

323
Sampling in Field Experiments

(Based on individual observations)

<table>
<thead>
<tr>
<th>Sources of variation</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Prob&gt;F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blocks</td>
<td>4</td>
<td>4.913</td>
<td>1.228</td>
<td>0.59</td>
<td>0.6747</td>
</tr>
<tr>
<td>Treatments</td>
<td>3</td>
<td>112.091</td>
<td>36.364</td>
<td>18.03</td>
<td>0.0001</td>
</tr>
<tr>
<td>Blocks × Treatments (Experimental error)</td>
<td>12</td>
<td>24.8734</td>
<td>2.0728</td>
<td>6.66</td>
<td>0.0001</td>
</tr>
<tr>
<td>Sampling error</td>
<td>80</td>
<td>24.912</td>
<td>0.311</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>99</td>
<td>166.790</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Step 2: Obtain the Standard Error of the difference between two treatment means as

\[ S.E_d = \sqrt{\frac{2s_1^2}{rn}} = \sqrt{\frac{2 \times 2.07}{5 \times 5}} = 0.4069 \text{ inches.} \]

Step 3: Obtain the estimate of sampling error as \( \hat{s}_e^2 = s_2^2 = 0.311 \)

The estimate of the experimental error is

\[ \hat{s}_e^2 = \frac{s_1^2 - s_2^2}{n} = \frac{2.07 - 0.31}{5} = 0.3520 \]

The estimated variance of

\[ \bar{Y}_{i.} = \frac{\hat{s}_s^2}{rn} + \frac{\hat{s}_e^2}{r} = \frac{2.070}{25} = 0.0828 \]

Estimated variance in case of complete recording \( \frac{\hat{s}_e^2}{r} = \frac{0.352}{5} = 0.0704 \).

Step 3: Efficiency of sampling as compared to complete recording

\[ \frac{\hat{s}_e^2 / r}{(\hat{s}_s^2 + n\hat{s}_e^2) / rn} = \frac{0.0704}{(2.0728) / 25} = 0.85 \]

Coefficient of variance for different treatments is given by

Treatment 1 = 2.965; Treatment 2 = 3.043; Treatment 3 = 3.545; Treatment 4 = 4.062.

Step 4: Estimation of sampling units per plot using the formula

\[ n = \frac{(Z_{\alpha/2}^2)\hat{s}_s^2}{r(D^2)(\bar{Y}_{i.}^2) - (Z_{\alpha/2}^2)\hat{s}_e^2} \]

From the formulae it is clear that the optimum sample size for a given character under study is just the sample size for the treatment with smallest mean. Therefore, one may do the computations only for the treatment with smallest mean instead of computing the sample size for each of the treatments. The treatment means in this case are


The number of sampling units required to measure the treatment means with 4 and 5 per cent coefficient of variation or 8 and 10 % margin of error when the number of replication are 5
and 6 for the treatment with smallest mean for replication 5 is 5 (4% standard error or 8% margin of error and 2 (5% standard error or 10% margin of error).

Now if the margin of error as obtained is satisfactory, then the present sampling procedure can be followed. However, if the margin of error is too high, then the researcher could take either or both of the following approaches
- Increases the number of replications
- Use a different sampling procedure with a change in sampling design type of sampling unit or sample size.

Sometimes, the data on more than one character is observed from the same sampling unit. In such situations, one may obtain the sample size for each of the characters and take the largest sample size among the optimum sample sizes of each of the characters.

**Note:** The usual procedure adopted for the analysis of plot sampled data from field experiments in the National Agricultural Research System (NARS) is based on the plot means i.e. the simple averages of the sampled observations of each plot. It is feared that this may cause loss in information. Hence, the above procedure is generally recommended. As above, this method enables one to get the estimate of sampling error along with experimental error. This estimate of sampling error can be used for determining the optimum sample size. However, one can easily see that in case of RCB design there is no effect on the significance of equality of treatment and block effects, if the data is analyzed using plot means or individual observations. For clarity, the ANOVA based on plot means is presented as

\[
\begin{array}{lcccccc}
\text{Sources of variation} & \text{DF} & \text{SS} & \text{MS} & F & \text{Prob>F} \\
\text{Blocks} & 4 & 0.983 & 0.2456 & 0.59 & 0.6747 \\
\text{Treatments} & 3 & 22.410 & 7.4728 & 18.03 & 0.0001 \\
\text{Error} & 12 & 4.975 & 0.4146 & & \\
\text{Total} & 19 & 28.375 & & & \\
\end{array}
\]

Now one can easily see that the probability level of significance of equality of treatment and block effects is same as with individual observations. The only advantage of performing analysis of variance based on individual observations is that one gets the estimate of the sampling error that can further be used for determining the optimum sample size. One can observe that the estimate of sampling error i.e. mean square due to sampling error is same as the pooled variance of the plot sampled observations. For illustration purpose, the variances of the plot sampled observations in the case of example 1.1 are given below:

\[
\begin{array}{lcccc}
\text{Variance of the plot sampled observations for Example 1.1} \\
\hline
\text{Blocks} & \text{Treatments} & \text{A} & \text{B} & \text{C} & \text{D} \\
\hline
\end{array}
\]

325
It can easily be seen that the pooled variance of these plot variances is 0.311, which is same as mean square of the sampling error. So the procedure based on the individual observation does not add much to information available from the analysis of variance based on the plot means in case of a RCB design. From the above Table, however, it is clear that the plot variances are different and hence, may violate the assumption of constant variance of observations. Therefore, one can test the homogeneity of sample variances based on the residuals. If the plot variances based on residuals are found to be different, then one can apply the appropriate transformations and analyze the data after transformation. For more details on this one may refer to Kumar (2002).

**Additional data from on going experiments**

To evaluate the efficiency of the various types of sampling units, additional data may be collected from on-going experiments, say, if previously, we have taken individual plants as sampling units say a group of plants may be taken up. It may require, more of researcher’s time and effort, there is generally sufficient flexibility for planning the collection to suit available resources for plots in an on-going experiment, not all plots need to be included in the additional data collection scheme. Or, if the resources are limited, the type of sampling unit could be limited to only a few. For example, let there are ‘p’ number of plots per treatment were selected, and ‘s’ denotes the large sampling units per plot and k is the number of small sampling units within a large sampling unit. Then for each treatment, one can perform the analysis of variance as given below:

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Degrees of freedom (D.F.)</th>
<th>Sum of Squares (S.S.)</th>
<th>Mean Square (M.S.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Plots</td>
<td>p-1</td>
<td>SS1</td>
<td>MS1</td>
</tr>
<tr>
<td>Between large sampling units within plots</td>
<td>p(s-1)</td>
<td>SS2</td>
<td>MS2</td>
</tr>
<tr>
<td>Between small sampling units within large sampling unit</td>
<td>ps(k-1)</td>
<td>SS3</td>
<td>MS3</td>
</tr>
</tbody>
</table>

Then, the estimates of sampling variances corresponding to small and large sampling units are as under

\[ s_1^2 = \frac{(N-1)MS_2 + N(k-1)MS_3}{Nk-1} ; \quad s_2^2 = MS_2 \]
where $s_1^2$ and $s_2^2$ are the estimates of the sampling variances based on small and large sampling units, $N$ is the total number of the large sampling units in the plot.

After this the relative efficiency of the two alternate sampling units can be obtained both with and without cost consideration and accordingly the decision may be taken on the type of sampling unit to be used.

Some experiments may also be planned specifically for sampling studies, related to the development of a sampling technique. Aside from providing information to evaluate the different types of sampling units, different sample sizes and different sampling designs, such experiments may also be used to identify some of the important factors that should be considered in developing a sampling technique, determine whether factors such as varieties and fertilizer rates influence the efficiencies of the different sampling techniques.

References