An introduction to data management and analysis for participatory varietal selection trials

DS VIRK AND JR WITCOMBE

Centre for Arid Zone Studies, University of Wales, Bangor, Gwynedd, LL57 2UW, UK.

Abstract

Data from on-farm varietal trials are usually not considered as amenable to statistical analysis as treatment of the data from research station experiments. This is because variability is believed to be higher in farmers’ fields; and the data from simple experiments are considered to be inappropriate for statistical analyses. However, in this paper, we show, with examples, how data from simple participatory varietal selection (PVS) experiments using Mother-Baby designs can be usefully analysed using existing statistical methods. The data collected on quantitative traits and farmers’ perceptions demand the application of parametric and non-parametric methods.

Introduction

We explored means of analysing and interpreting data from simple on-farm experiments, and have found many existing statistical methods that are directly relevant to participatory varietal trials. This paper aims at demonstrating the application of various statistical methods for analysing quantitative and qualitative data from PVS trials.

Creating spreadsheets

PVS data are often poorly documented. A researcher should be able to prepare abstracts of information in many ways and subject data to statistical analyses without much extra effort. To do so, data from Mother and Baby trials should be entered in spreadsheets so that they can be analysed without re-entry of data. This requires data to be organised in columns where each column is a single descriptor, or trait, so that a two-way ANOVA is possible in statistical programmes such as Minitab (Table 1).

Design of PVS trials and type of data

A simplified Mother-Baby trial design (Snapp, 1999; J de Meyer and M Bänziger [pers. comm.]) that does not require complex layout plans is used for PVS trials (Annexe 1).

Table 1. Extract from a spreadsheet for a rice mother trial with 3 farmers in 2 villages of 2 clusters in GVT east in 2001. Data for varieties 2 to 5 not shown for farmers 2 to 4. The data require no reformating for Minitab.

<table>
<thead>
<tr>
<th>Farmer</th>
<th>Village</th>
<th>Cluster</th>
<th>Var.</th>
<th>Area (m²)</th>
<th>Plant ht (cm)</th>
<th>Grain yd (kg/plot)</th>
<th>Straw yd (kg/plot)</th>
<th>Grain yd (kg/ha)</th>
<th>Straw yd (kg/ha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>95</td>
<td>8.7</td>
<td>8</td>
<td>1740</td>
<td>1600</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>50</td>
<td>82</td>
<td>7.3</td>
<td>7</td>
<td>1460</td>
<td>1400</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>50</td>
<td>87</td>
<td>3.7</td>
<td>4</td>
<td>740</td>
<td>800</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>50</td>
<td>82</td>
<td>3</td>
<td>3.1</td>
<td>600</td>
<td>620</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>50</td>
<td>87</td>
<td>7.4</td>
<td>7</td>
<td>1480</td>
<td>1400</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>50</td>
<td>70</td>
<td>3.4</td>
<td>3.5</td>
<td>680</td>
<td>700</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>50</td>
<td>79</td>
<td>11.3</td>
<td>11</td>
<td>2260</td>
<td>2200</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>50</td>
<td>76</td>
<td>5.2</td>
<td>5</td>
<td>1040</td>
<td>1000</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>54</td>
<td>100</td>
<td>7.3</td>
<td>8.1</td>
<td>1352</td>
<td>1500</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2</td>
<td>6</td>
<td>54</td>
<td>97</td>
<td>6</td>
<td>6.7</td>
<td>1111</td>
<td>1241</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>48</td>
<td>94</td>
<td>7.2</td>
<td>7.5</td>
<td>1500</td>
<td>1563</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>48</td>
<td>94</td>
<td>6.3</td>
<td>6.5</td>
<td>1313</td>
<td>1354</td>
</tr>
</tbody>
</table>
allow direct comparison among all varieties. Farmers (sites) serve as replications. Mother trials are kept to a minimum, as they are research intensive. The following types of data are usually collected in these trials:

- Quantitative data on traits such as grain yield, straw yield, days to flowering.
- Ranking of all the varieties for traits such as cooking quality, and market price.

2. For the Baby trials (single-replicate, single-variety) each farmer provides a replication and there are more baby trials than Mother trials. Household level questionnaires (HLQs) are collected from each farmer to record their perceptions about a new variety as ‘more’, ‘same’ or ‘less’ in comparison to the local check for many traits including yield and quality (Appendix 1). Sometimes quantitative data for yield are obtained, although considerable practical difficulties have to be overcome.

1. Analysis of data from Mother trials

Analysis of quantitative data.
The analysis of a two-way data table, with varieties on one side and farmers (replications) on the other, is straightforward and can be computed as a randomised complete block design using standard statistical packages.

We analysed the data of single-replicate all-entry Mother trials of six varieties of upland rice conducted by 12 farmers in Jharkhand, in the rainy season of 2001. There was significant variation among varieties and farmers (Table 2). These significances are conservative, as the error term to test the main effects is the varieties x farmers’ interaction that includes a large field-to-field variation.

Analysis of matrix ranks

Two-way ANOVA. The six varieties in the nine-replicate mother trial were ranked for grain yield by consensus in focus group discussions. The data were analysed by a two-way ANOVA having nine replications and six varieties. The ANOVA revealed significant differences among varieties and farmers for ranks at P<0.001.

The mean ranks for varieties of interest were: Ashoka 200F = 4.3; Ashoka 228 = 4.7; Kalinga III = 3.0; and BG 102 = 2.4. The lsd (5%) = 0.71 was computed. The mean scores of Ashoka 200F and Ashoka 228 do not differ but both have significantly higher score than the check varieties Kalinga III and BG 102.

Friedman’s method for randomised blocks. Alternatively, Friedman’s non-parametric method can be used. The matrix scores of the six varieties within each block (farmer) are assigned ranks (1=lowest, 6=highest) for making the paired comparisons. Friedman’s method yields a c² value for each comparison. The results are similar to that of the ANOVA, since c² for 1 d.f. for Ashoka 200F vs. Ashoka 228 = 1.78 (non-significant); and for Ashoka 200F vs. Kalinga III = 9.00 (P<0.01); Ashoka 200F vs. BG 102 = 9.00 (P<0.01); Ashoka 228 vs. Kalinga III = 5.44 (P<0.05); and Ashoka 228 vs. BG102 = 9.00 (P<0.01)

2. Analysis of data from Baby trials

Analysis of quantitative data

Grain yield data when collected can be analysed as follows:

Analysis of paired-plots by randomised block design or paired t-test: In Baby trials, a pair is composed of new and local cultivars sharing the common field. These data can by analysed as a two-way ANOVA. This design is a special case of the randomised block design where there are only two treatments and the n farmers represent the blocks.

An alternative analysis of paired plots is provided by the paired t-test (Snedecor and Cochran, 1973). Which of these two analyses should be preferred varies with the interest of the researcher. For example, when it is desired to assess if the varieties were tested over variable fields, then the ANOVA technique will provide a separate sum of squares for testing variation among farmers’ fields that is not possible with a paired t-test.

We analysed grain yield t ha⁻¹ of the upland rice variety

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sites (s, Farmers)</td>
<td>(s-1) =11</td>
<td>2973</td>
<td>270.27</td>
<td>35.95</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Varieties (a)</td>
<td>(a-1) =5</td>
<td>1046</td>
<td>209.20</td>
<td>27.83</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sites x Varieties (error)</td>
<td>(s-1) (a-1) =55</td>
<td>413</td>
<td>7.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>(as-1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Standard error of difference of variety mean yield = 0.04 t ha⁻¹; lsd = 0.07 t ha⁻¹. Ashoka 200F tested alongside
the local check variety, BG102, in Jharkhand, India, in the rainy season of 2001 in 22 Baby trials.

Both analyses similarly concluded that there were significant differences between varieties for grain yield (Table 3). Since there is 1 degree of freedom for varieties the computed $F = t^2$. The advantage of the ANOVA technique is clear as it discerns significant differences among the farmers’ fields over which the varieties were tested. Note that the $F$-test assumes the farmers x varieties interaction to be non-significant and it cannot be tested in the present case. The paired $t$-test assumes that the pair of observations is not independent but related; if one member of the pair is high, so is the other - i.e. they are positively correlated. This can be verified by performing a correlation or regression analysis.

Regression analysis: Regression of grain yield (t ha$^{-1}$) of Ashoka 200F on to the grain yield of the local variety was significant ($a = 0.123 \pm 0.106; b = 1.44 \pm 0.08$) with a very high $R^2 = 0.94$. The regression coefficient of the new variety is more than one indicating its increased response to improving environmental conditions. However, the yield of the new variety in the poor environment is still higher than the local check. The new variety is thus adaptable to the sample of target environments and can be a potential replacement for the local variety.

Non-parametric methods: We also investigated the use of some alternative non-parametric methods that make no assumptions for specific distribution for the population.

Friedman’s method for randomised blocks: This is an alternative to the two-way ANOVA. In this method the varieties ($a = 2$) are ranked within each block or farmer; $b = 22$ in our case. If there are ties, they are treated in the conventional manner (by average ranks). The ranks are summed for each of the variety and the statistic $S$ or $c^2$ is computed (Sokal and Rohlff, 1995). $c^2$ is distributed as $c^2(a-1)$, where a is the number of varieties. Our computed value of $c^2 = 21.98$ is significant at $P<0.001$. We conclude, as for farmers are computed and ranked without regard to the signs so that the smallest absolute difference is ranked 1 and the largest is ranked 22. Tied ranks are computed as averages as usual. The original signs of differences are then assigned to the corresponding ranks. The sum of the positive and the negative ranks are computed, and whichever is smaller in absolute value is labelled as $T_s$. This is compared with the critical value of $T$ in statistical tables for the corresponding sample size.

In our case the value of $T_s = 0$. Since this is less than the table value at the two-tailed 1% level (see Table A9 in Snedecor and Cochran, 1991) the null hypothesis is rejected. Thus Ashoka 200F significantly yields more than the local variety.

Chi-squared test: Alternatively, a chi-squared test may be applied for the perception of 40 farmers of which 31 recorded ‘more’, nine the ‘same’, and 0 ‘less’ for the grain yield of Ashoka 200F over the local check. A $c^2(2 \text{ d.f.}) = 38.15$ ($P<0.01$) was computed on a 1:1:1 hypothesis. It shows that significantly more farmers perceive Ashoka 200F to be higher yielding than the check variety.

Analysis of perception (qualitative) data

Z-test for comparing per cent perceptions: Standard errors can be computed for per cent perceptions and compared for testing the significance of differences. In HLQs returned by 40 farmers comparing the Ashoka 200F variety of rice to the local check in 2000 in Jharkhand, 78% perceived Ashoka 200F to have ‘more’, 23% to have ‘same’, and none ‘less’ yield than the local variety. In this case $N = 40$; variance of proportion ($p$) is $pq/n$ where $q = (1-p)$; SE for ‘more’ $= \sqrt{pq/n} = \sqrt{(0.78 \times 0.23)/40} = \pm 0.07$. We may compare ‘more’ with ‘same’ by computing, $Z = (p1-p2)/ \sqrt{pq/n1 + p2q2/n2}$. The computed $Z = 0.78-0.23/ \sqrt{0.78 \times 0.23}/40} = 5.81^{*\ast}$. The null hypothesis is rejected and we can conclude that a significantly larger number of farmers perceived Ashoka 200F to be higher yielding than the check variety.

Table 3. ANOVA and paired $t$-test for grain yield (t ha$^{-1}$) for 22 baby trials of A200F vs. local in Jharkhand, India, rainy

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>ANOVA</th>
<th>Paired t-test</th>
<th>Source</th>
<th>df</th>
<th>ANOVA</th>
<th>Paired t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Farmers (F)</td>
<td>21</td>
<td>2728</td>
<td>18.4</td>
<td>&lt;0.001</td>
<td>Mean</td>
<td>1.43</td>
<td>0.91</td>
</tr>
<tr>
<td>Varieties</td>
<td>1</td>
<td>3028</td>
<td>20.4</td>
<td>&lt;0.001</td>
<td>Mean</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>F x V</td>
<td>21</td>
<td>148</td>
<td>--</td>
<td>--</td>
<td>Mean</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

the ANOVA and the paired $t$-test, that the grain yield of Ashoka 200F is significantly higher than the local variety.

Wilcoxon’s signed-ranks test: This method is more widely used than the Friedman’s test for paired-plots. The differences between the grain yields of the varieties across farmers are computed and ranked without regard to the signs so that the smallest absolute difference is ranked 1 and the largest is ranked 22. Tied ranks are computed as averages as usual. The original signs of differences are then assigned to the corresponding ranks. The sum of the positive and the negative ranks are computed, and whichever is smaller in absolute value is labelled as $T_s$. This is compared with the critical value of $T$ in statistical tables for the corresponding sample size.

In our case the value of $T_s = 0$. Since this is less than the table value at the two-tailed 1% level (see Table A9 in Snedecor and Cochran, 1991) the null hypothesis is rejected. Thus Ashoka 200F significantly yields more than the local variety.
Discussion

We have shown how simple experiments (Mother-Baby design) by farmers provide data that can be statistically analysed by existing, robust parametric and non-parametric statistical techniques.

We recommend that researchers:

- Conduct some Mother trials and a large number of Baby trials.

- Analyse the data from Mother trials to compare all the new varieties among themselves.

- Analyse the matrix rank data from Mother trials using parametric or non-parametric methods.

- Analyse Baby trials as paired-plots, i.e. use a paired t-test or, preferably, a randomised block design analysis.

- Do a regression analysis of paired-variety data to test the response of the new variety in relation to the check variety.

- Alternatively, non-parametric methods such as Friedman’s test, Wilcoxon’s signed-rank test can be used to compare paired varieties.

- Analyse farmers’ perceptions from HLQs by non-parametric methods.

Implications for drought-prone station trials

While performing a combined analysis of multilocalational research station trials, only balanced sets are usually considered. The trials from drought-prone locations where only a few entries may survive in the trial are rejected as ‘failed trials’. However, the few entries that survive, in fact, provide the most valuable information for marginal areas. This demands modification of multilocalational analysis. If data from partly successful trials are analysed by the paired-plot techniques, described in this paper, there seems little need to reject a drought prone-site from the overall analysis.

Acknowledgement

This publication is an output from projects R7281 and R7542 funded by the UK Department for International Development (DFID) Plant Sciences Research Programme for the benefit of developing countries. The views expressed are not necessarily those of DFID.

References

