

QUANTITATIVE ANALYSIS APPROACHES TO QUALITATIVE DATA: WHY, WHEN AND HOW - *Savitri Abeyasekera*

*Statistical Services Centre, University of Reading, P.O. Box 240,
Harry Pitt Building, Whiteknights Rd., Reading RG6 6FN, UK.
Phone 0118 931 8459, e-mail s.abeyasekera@rdg.ac.uk*

1. Introduction

In many research studies involving the use of participatory tools, much of the information gathered is of a qualitative nature. Some of this will contribute to addressing specific research questions, while other parts provide a general understanding of peoples' livelihoods and constraints. The aim of this paper is to focus on the former. The paper concentrates on some quantitative analysis approaches that can be applied to qualitative data. A major objective is to demonstrate how qualitative information gathered during PRA work can be analysed to provide conclusions that are applicable to a wider target population. Appropriate sampling is of course essential for this purpose, and it will be assumed in what follows that any sampling issues have been satisfactorily addressed to allow generalization of results from the data analysis to be meaningful.

Most emphasis in this paper will be given to the analysis of data that can be put in the form of ranks, but some analysis approaches suitable for other types of qualitative data will also be considered. The general questions of *why* and *when* are discussed first, but the main focus will be on issues relating to the *how* component of data analysis. It is not the intention to present implementation details of any statistical analysis procedures, nor to discuss how output resulting from the application of statistical software could be interpreted. The aim is to highlight a few types of research questions that can be answered on the basis of qualitative information, to discuss the types of data format that will lend themselves readily to appropriate data analysis procedures and to emphasise how the data analysis can be benefited by recognizing the data structure and paying attention to relevant sources of variation.

2. Why use quantitative approaches?

Quantitative methods of data analysis can be of great value to the researcher who is attempting to draw meaningful results from a large body of qualitative data. The main beneficial aspect is that it provides the means to separate out the large number of confounding factors that often obscure the main qualitative findings. Take for example, a study whose main objective is to look at the role of non-wood tree products in livelihood strategies of smallholders. Participatory discussions with a number of focus groups could give rise to a wealth of qualitative information. But the complex nature of inter-relationships between factors such as the marketability of the products, distance from the road, access to markets, percent of income derived from sales, level of women participation, etc., requires some degree of quantification of the data and a subsequent analysis by quantitative methods. Once such quantifiable components of the data are separated, attention can be focused on characteristics that are of a more individualistic qualitative nature.

Quantitative analytical approaches also allow the reporting of summary results in numerical terms to be given with a specified degree of confidence. So for example, a statement such as "45% of households use an unprotected water source for drinking" may be enhanced by providing 95% confidence limits for the true proportion using unprotected water as ranging from 42% to 48%. Here it is possible to say with more than 95% confidence that about half the households had no access to a protected water supply, since the confidence interval lies entirely below 50%.

Likewise, other statements which imply that some characteristic differed across two or more groups, e.g. that "infant mortality differed significantly between households with and without access to a community based health care clinic", can be accompanied by a statement giving the chance (probability) of error (say $p=0.002$) in this statement, i.e. the chance that the conclusion is incorrect. Thus the use of quantitative procedures in analysing qualitative information can also lend greater credibility to the research findings by providing the means to quantify the degree of confidence in the research results.

3. When are quantitative analysis approaches useful?

Quantitative analysis approaches are meaningful only when there is a need for data summary across many repetitions of a participatory process, e.g. focus group discussions leading to seasonal calendars, venn diagrams, etc. Data summarisation in turn implies that some common features do emerge across such repetitions. Thus the value of a quantitative analysis arises when it is possible to identify features that occur frequently across the many participatory discussions aimed at studying a particular research theme. If there are common strands that can be extracted and subsequently coded into a few major categories, then it becomes easier to study the more interesting qualitative aspects that remain.

For example, suppose it is of interest to learn about peoples' perceptions of what poverty means for them. It is likely that the narratives that result from discussions across several communities will show some frequently occurring answers like experiencing periods of food shortage, being unable to provide children with a reasonable level of education, not owning a radio, etc. Such information can be extracted from the narratives and coded. Quantitative approaches provide the opportunity to study this coded information first and then to turn to the remaining qualitative components in the data. These can then be discussed more easily, unhindered by the quantitative components.

Quantitative analysis approaches are particularly helpful when the qualitative information has been collected in some structured way, even if the actual information has been elicited through participatory discussions and approaches. An illustration is provided by a daily activity diary study (Abeyasekera and Lawson-McDowall, 2000) conducted as part of the activities of a Farming Systems Integrated Pest Management Project in Malawi. This study was aimed at determining how household members spend their time throughout the year. The information was collected in exercise books in text format by a literate member of a household cluster and was subsequently coded by two research assistants by reading through many of the diaries and identifying the range of different activities involved. Codes 1, 2, 3, 4, ... were then allocated to each activity. In this study, the information was collected in quite a structured way since the authors of the diaries were asked to record daily activities of every household member by dividing the day into four quarters, i.e. morning, mid morning, afternoon and late evening, and recording the information separately within each quarter.

4. Data Structure

The data structure plays a key role in conducting the correct analysis with qualitative data through quantitative methods. The process is greatly facilitated by some attention to the data structure at the time of data collection. This does not imply any major change to the numerous excellent methodologies that researchers undertake when gathering qualitative information. Data structure refers to the way in which the data can be visualized and categorized in different ways, largely as a result of the method of data collection. For example, many research studies involve a wealth ranking exercise, and data may be collected from each wealth group. Your data is then structured by wealth categories. During data collection, additional structure may arise, for example, by villagers' level of access to natural resources.

You may also find that your data are structured in many other ways, for instance, (a) by community level variables such as by the presence/absence of a community school or a health care clinic; (b) by focus group variables such as the degree of women participation in the discussions, the degree of agreement with respect to specific issues, etc; and (c) by household level variables such as their primary source of income, gender of household head, etc.

Thinking about the data structure forces the researcher to focus on what constitutes replicates for data summarization, and it helps to identify the numerous factors that may have a bearing on those components of the qualitative information that cannot be coded. Often the replicates may be several focus group discussions. If the data are to be later summarised over all groups, then some effort is needed to ensure that the information is collected in the same systematic way each time. For example, a member of the research team may record the information that emerges from any participatory discussions in a semi-structured way. This systematization helps in regarding the sample, consisting of many focus groups, as a valid sample for later statistical analysis.

Considering the data structure also helps the researcher to recognize the different hierarchical levels at which the data resides, e.g. whether at community level, focus group level or household level. The data hierarchy plays a key role in data analysis as well as in computerizing the information collected. If spreadsheets are to be used, then the data at each level of the hierarchy has to be organized in a separate sheet in a rectangular array. However, it must be noted that hierarchical data structures are more appropriately computerized using a suitable database rather than as a series of spreadsheets.

An example of a simple data structure is provided in Table 1. Here there is structure between women since they come from five villages, they fall into one of four wealth groups, their household size is known, and they have been identified according to whether or not they earn wages by some means. The data should also be recognized as being hierarchical since the information resides, both at the “between women” level (e.g. whether wage earner) and at the “within woman” level, e.g. preference for the four oils. The data structure comes into play when quantitative analysis approaches are used to analyse the data. An example using ranks was chosen here since this paper focuses on ranks and scores to illustrate some quantitative data analysis procedures. One or other is often used in participatory work to address similar objectives.

Table 1. An example data set showing ranked preference to four types of oil by a number of women. (The full data set extends over 5 villages, with 6, 8, 5, 11 and 14 women interviewed per village)

<i>Village</i>	<i>Wealth group</i>	<i>Household size</i>	<i>Wage earner</i>	<i>Covo</i>	<i>Superstar</i>	<i>Market</i>	<i>Moringa</i>
1	2	6	Yes	3	1	4	2
1	1	3	No	2	4	3	1
1	1	7	Yes	4	3	2	1
1	1	3	Yes	2	4	3	1
1	3	4	No	4	2	1	3
.
.

5. Objectives that may be addressed through ranking/scoring methods

The “*how*” component of any analysis approach must be driven by the need to fulfill the research objectives, so we consider here some examples that may form sub-component objectives of a wider set of objectives. The following list gives a number of objectives that may be addressed by eliciting information in the form of ranks or scores.

- (a) Identifying the most important constraints faced by parents in putting their children through school;
- (b) Identifying womens’ most preferred choice of oil for cooking;
- (c) Assessing the key reasons for depletion of a fisheries resource;
- (d) Identifying elements of government policies that cause most problems for small traders;
- (e) Assessing participants’ perceptions of the value of health information sources.

Part of a typical data set, corresponding to example (b), was shown in Table 1. The primary aim was to compare womens’ preference for four different oil types. Such data may arise if the oils are presented to a number of women in each of several villages for use in cooking, and they are asked, several weeks later, to rank the items in order of preference. Here, the results merely represent an ordering of the items and no numerical interpretation can be associated with the digits 1, 2, 3, 4 representing the ranks.

An alternative to ranking is to conduct a scoring exercise. To determine the most preferred choice from a given set of items, respondents may be asked to allocate a number of counters (e.g. pebbles, seeds), say out of a maximum of five counters per item, to indicate their views on the importance of that item. The number allocated then provides a score, on a 0-5 scale with 0 being regarded as being “worst” or “of no importance”. An example data set is shown in Table 2. The “scores” in this table represent farmers’ own perception of the severity of the pest.

Table 2. An example data set showing scores given by farmers to the severity of pest attack on beans, large scores indicating greater severity.

Farmer	Ootheca	Pod borers	Bean stem maggot	Aphids
1	4	2	1	2
2	5	4	1	3
3	4	1	2	1
4	4	5	1	4
5	1	2	1	1
6	1	4	1	2
7	5	1	1	5
8	2	5	5	3
Mean	3.3	3.0	1.6	2.6

Addressing objectives of the type presented in this section often involve either a ranking or a scoring exercise. The discussion that follows is aimed at researchers who may want to understand and appreciate the advantages and limitations of these two forms of elicitation when addressing such objectives.

6. Ranks or scores?

Does it matter whether the information is extracted from ranks or scores? Generally, ranks are better for elicitation as it is always easier to judge whether one item is better or worse, more or less important, than another item. However, the ease with which the information can be collected must be balanced against the fact that the information cannot be directly analysed through quantitative means.

The main difficulty in using ranks is that they give no idea of “distance”. Say for example that two respondents each give a higher rank for item E than item B. However, the first respondent might have thought E was only slightly better than B while the second thought E was a lot better than B. This information is not elicited with ranks. Thus it is not possible to attribute a “distance” measure to differences between numerical values given to the ranks in Table 1.

Scores on the other hand have a numerical meaning. Usually “best” or “good” in some respect is associated with larger scores, whereas in ranking exercises, “best” is always associated with a rank of 1. In studies concerning livelihood constraints or problem identification, high scores or a rank of 1 are associated with the “most severe” constraint or problem.

A second point is that scores can have an absolute meaning while ranks are always relative to the other items under consideration. So a rank of 1 is not necessarily a favoured item, it is just “better than the rest”. For example, four items A, B, C, D, ranked as 2, 3, 1, 4 by a respondent could receive scores 2, 1, 3, 0 by the same respondent on a 0-10 scale where 10 represents “best”. The lack of a standard scale for ranks makes the task of combining ranks over several respondents difficult unless effort is made to ask supplementary questions to elicit respondents’ absolute views on the “best” and “worst” ranked items. Such additional information may give a basis on which to convert ranks to a meaningful set of scores, so that the resulting set of scores can be analysed. Sometimes both the ranks and the (approximate) scores can be analysed. If the results are similar, then this indicates that the ranks may be usefully processed on their own.

Ranks thus represent an ordering of a list of items according to their importance for the particular issue under consideration. In interpreting ranks, it is therefore necessary to keep in mind that the digits 1, 2, 3, ... etc., allocated to represent ranks, have little numerical significance. Ties should normally be allowed, i.e. permitting two or more items to occupy "equal" positions in the ordered list. This is because it is usually unrealistic to oblige the respondent to make a forced choice between two items if she/he has no real preference for one over the other. Each item involved in a tie can be given the average value of ranks that would have been allocated to these items had they not been tied.

Thus for example, suppose six items A, B, C, D, E and F are to be ranked. Suppose item B is said to be the best; item C the worst; item F second poorest but not as bad as C; and the remaining items are about the same. Then the ranks for the six items A, B... F should be 3, 1, 6, 3, 3, 5. This set of ranks is obtained by using 3 as the average of ranks 2, 3 and 4, i.e. the ranks that items A, D and E would have got if the respondent had perceived some difference in these items. One reason for using the full range from 1 to 6 rather than ranking the items as 2, 1, 4, 2, 2, 3 is that misleading results will otherwise be obtained in any further data summaries which combine information across respondents. Fielding *et al* (1998) give a fuller discussion on the use of ties

The above discussion assumes that the ranking or scoring exercise is done on the basis of one identified criterion. More typically, a number of criteria are first identified to form the basis on which to compare and evaluate a set of items. For example, respondents may identify yield, seed size, cooking time, disease resistance and marketability as criteria for evaluating a number of pigeonpea varieties. Once suitable criteria have been chosen, items for evaluation are scored with respect to each criterion in turn. This is what is essentially referred to as matrix scoring (Pretty *et al*, 1995). It is common to use scores from 1 to 5 although a wider range can be useful because it will give better discrimination between the items.

During scoring exercises, the greatest challenge is for the researcher facilitating the process. He/she needs to ensure that respondents clearly understand what exactly the scores are meant to represent. For example in using scores in a variety evaluation trial, the facilitator must ensure respondents are clear between views such as

"I will give this variety a high score because I have seen it is high yielding under fertiliser application although I could never afford the seed or the required amount of fertiliser"; and

"I will give this variety a high score because it is performing very well compared to the local variety I used to grow".

Assuming that scores have been elicited unambiguously by a good facilitator, their major advantage over ranks is that they are numerically meaningful. Differences between scores given to different items show the strength of the preference for one item over the other. Scores provide a ranking of the items but also something extra - a usable distance measure between preferences for different items. The availability of a meaningful measurement scale also means that the resulting scores usually correspond easily to the objectives of the study, and can usefully be summarised across respondents.

Various forms of scoring methods exist (Maxwell and Bart, 1995). Fully open scoring, where each item to be scored can be given any value within a particular range (say 1-5) is the most flexible since it leads to observations that are "independent" of each other - a requirement for most simple statistical analysis procedures. However, restricted forms of scoring, where a fixed number of "units" (e.g. bean seeds, pebbles) are distributed between a set of items, has to be used with some care, recognising that it is only a little improvement over ranks.

Say, for example, that a respondent is asked to allocate 10 seeds among four bean varieties, giving more seeds to the variety she/he most prefers. If the respondent allocates 5 seeds to variety C and 3 seeds to variety A because C and A are the varieties she/he most prefers, then the respondent is left with just 2 seeds to allocate to the remaining varieties B and D. This is a problem if she/he has only a marginal preference for A over B and D. Such a forced choice suffers the same difficulties faced by ranks in not being able to give a meaningful interpretation to the "distance" between the scores. A larger number of seeds, say 100, partially gets over this problem because now the respondent has more flexibility in expressing his/her strength of preference for one variety over another. One hundred

seeds may however still be too few if the number of cells into which the seeds must be distributed is large (e.g. in matrix scoring exercises).

7. Some approaches to the analysis of ranked data

We illustrate analysis approaches to ranked data through some situation specific examples. There is no attempt to give a full coverage of analysis methods since the most appropriate form of analysis will depend on the objectives of the study.

Whatever the objective, it is usually advisable to begin by thinking carefully about the data structure and then to produce a few simple graphs or summary statistics so that the essential features of the data are clear. Often this form of summary may be all that is needed if extending the results beyond the sampled units is not a requirement. However if results are to be generalised to a wider population, sample sizes and methods of sampling are key issues that need consideration. See Wilson (2000) for details. Here we assume that participating respondents have been representatively chosen.

In section 7.1 below, we consider some data arising from a ranking exercise and simple methods of presentation. This is followed, in sections 7.2 and 7.3, by a brief discussion of more advanced methods of analysis applicable to ranking and scoring data, restricting attention largely to preference evaluations. Poole (1997) provides fuller coverage.

7.1 Simple methods of summary

The data structures that result from ranks/scores given to a fixed number of items, were shown using fictitious data in Tables 1 and 2. Although the numerical values shown look similar, a number of essential differences, as highlighted in section 6, must be recognised in interpreting this information. For example, mean values for columns in Table 2 are meaningful summaries and give some indication of the most serious pest problem. However, producing column totals for the ranks given in Table 1 assumes a common “distance” between any two consecutive ranks. This is particularly problematical if there are missing cells in the table, because some respondents have not ranked some items.

As illustration, suppose an overall ranking of the oils in Table 1 is required across all women sampled. A simple procedure is to give oils receiving a rank 1, 2, 3 and 4, corresponding scores of 4, 3, 2, 1, and then to average the resulting set of scores. This may result in mean scores of 3.2, 2.6, 2.4 and 1.8 respectively for oils *Covo*, *Superstar*, *Market* and *Moringa*. Of course, some would argue that this will lead to exactly the same conclusions as would be obtained if the numerical values given to each rank were averaged across the women. This is true, but the conversion to scores makes the assumptions involved in the averaging process clearer, i.e. the assumption that the degree of preference for one item over the next within an ordered list is the same irrespective of which two neighbouring ranks (or corresponding scores) are being considered, and the assumption that a missing value corresponds to a numerical zero score.

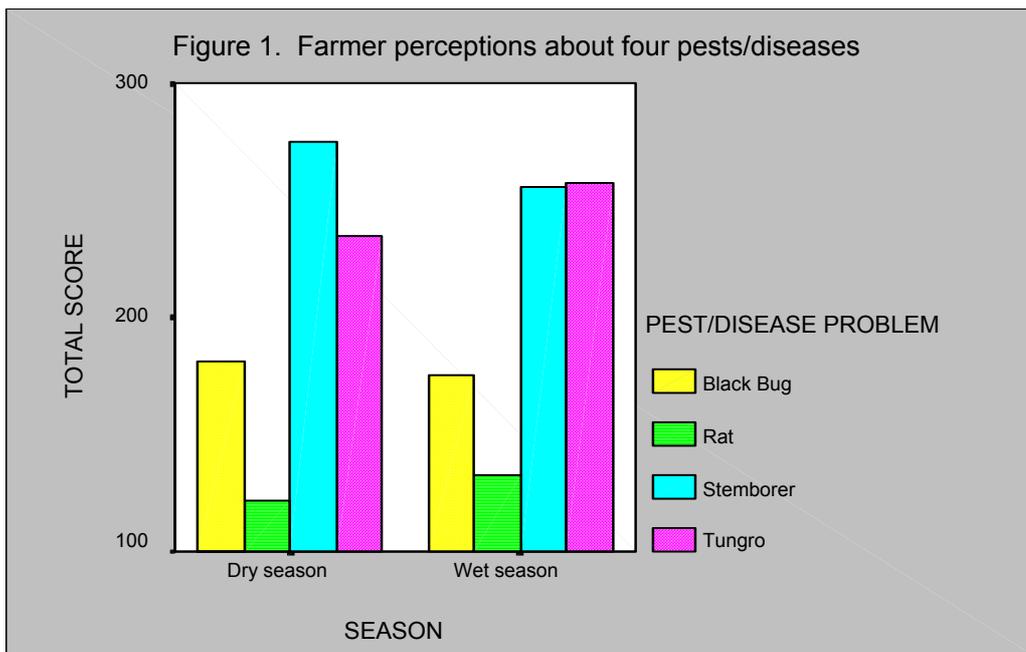
Alternative transformation scales could also be considered on the basis of comments made by those who rank the items. For example, ranks 1, 2, 3, 4 might be converted to scores 9, 5, 2, 1 respectively, if it were apparent during discussions with respondents that there is a much clearer preference for items at the top of the priority scale than those lower down.

Consider now a situation where the specific items being ranked vary from respondent to respondent. Again, a transformation from ranks to scores can be made with a zero score for items omitted during the ranking. See Abeyasekera *et al* (2000) for details. Box 1 provides a further example aimed at investigating whether a disease, known to cause high yield losses in a crop, is regarded by farmers as being the most damaging to the crop compared to damage caused by other pests and diseases. The results are shown in Figure 1. The main point to note in this example is the difficulty of summing the ranks given to pest/disease problems since many problems had been mentioned by the 226 farmers. If one problem was mentioned by only a few farmers, but ranked highest by them, its average rank, based on just those farmers, would be unrepresentative of the farming population as a whole. Scores on the other hand, include a meaningful zero, so averages work out correctly.

Simple descriptive summaries as illustrated above have a place in the analysis irrespective of whether the study is intended for making generalisable conclusions concerning a wider target population or is a case study in one village with just a few farmers. In the former case, the summary also helps to identify statistical inferential procedures relevant for demonstrating the applicability of results beyond the research setting. Such procedures are also valuable because they allow the researcher to take account of the data structure.

Box 1. An example drawn from Warburton (1998) illustrating a descriptive summary of ranks given to the three most important pests and diseases

In a study concerning farmers' practices, experiences and knowledge of rice tungro disease, 226 farmers in the Philippines were asked to name and rank the three most damaging pests or diseases affecting their rice crop. Each pest and disease named by a farmer was given a score of 3, 2 or 1 according to whether they ranked the pest or disease as being the first, second or third most damaging in their attack on the crop. A score of zero was given to any pest/disease not mentioned by the farmer. These scores were *totalled* over all farmers and are shown in Figure 1 below, for each of two seasons, for pests/diseases receiving the four highest overall scores. Tungro is clearly a recognised problem but in the dry season, stemborer is identified as being the greater problem.



7.2 Illustrating analysis of preference evaluations

Consider a situation where the raw data have been obtained by asking each respondent to order the items according to their preference so that a set of ranks results. We will assume that these ranks have been converted to scores in some meaningful way and that the researcher is interested in determining whether observed differences, in mean scores for the different items, signal real differences in preferences for these items among respondents in the target population. This is a situation where generalisability of study results is of primary concern. Assuming that the respondents have been sampled appropriately and that the sample size is adequate, such a question can be answered by applying appropriate statistical techniques.

Consider for example the ranked data shown in Table1, and suppose these have been converted to a set of scores. Suppose the main objective is to determine whether there is a greater preference for *Moringa* oil compared to the other oils. But one could question whether an overall evaluation across all women in all villages is appropriate when there is a clear structure in the data. This structure raises additional questions, e.g.

- Do preferences differ according to whether the woman is a wage earner or not?
- Do preferences differ across the different wealth categories?
- If preferences differ according to whether the woman earns or not, do such differences differ across the wealth categories?

There may be many other pieces of information available concerning the women participating in this study, e.g. size of household to which they belong. Is it possible to disentangle confounding factors such as village and household size?

The answer is “yes” and a powerful statistical procedure for this purpose is the Analysis of Variance (*anova*). This procedure involves statistical modelling[†] of the data and enables questions like the above to be answered. The procedure allows relevant comparisons to be made, after making due allowance for possible effects of other confounding factors. (See Box 2 for another example).

Box 2. An example illustrating results from a statistical modelling exercise to compare four sweet potato varieties for their tolerance to *cylas puncticollis*

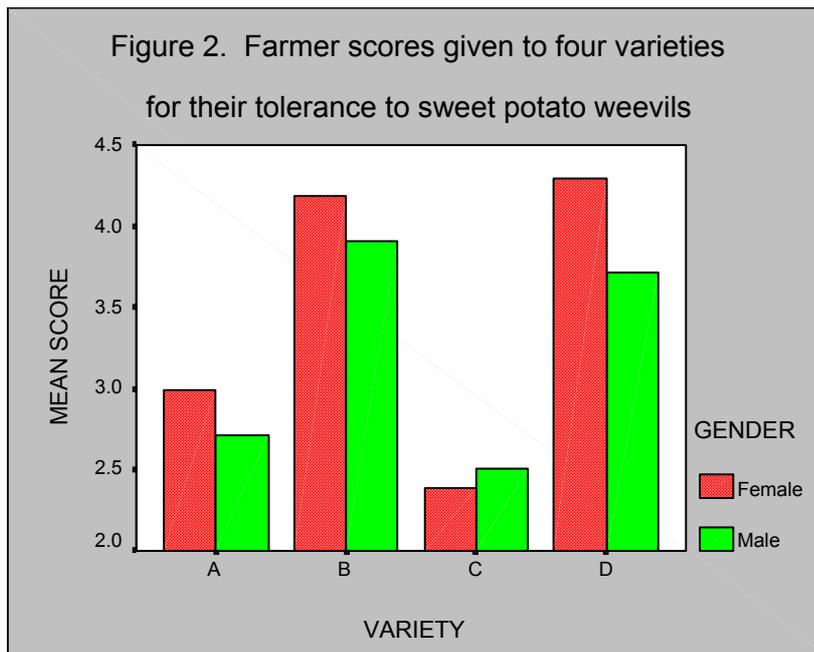
As part of an integrated pest management project in Malawi, four sweet potato varieties were studied in an on-farm trial for their tolerance to sweet potato weevils (*Ophiomyia* spp). One component of the variety evaluation involved asking twenty farmers to score separately, on a 1-5 scale (1=least tolerant; 5 = most tolerant) their assessment of each variety with respect to its tolerance to weevils. All farmers scored all four varieties resulting in 80 observations from 20 farmers. However, farmers were not divided equally by village or by gender, as is seen in the frequency table below. Modelling the data is still possible.

Village	Pindani	Chiwinja	Lidala	Maulana	Total
Female	5	1	2	2	12
Male	2	4	3	1	8

The modelling analysis demonstrated that there were no significant gender differences or differences between villages with respect to farmers’ views on pest tolerance by the different varieties. There was however strong evidence ($p < 0.001$) to indicate that two varieties, A and C, were regarded as being more susceptible than varieties B and D. Figure 2 shows some results. The modelling enables a valid comparison between male and female farmers, between villages and between varieties, despite the unequal replication of male and female farmers across the four villages. The modelling also allowed varieties to be compared “free” from possible effects due to gender and village differences, i.e. we excluded the alternative explanations that these effects might account for, or mess up, the picture of varietal differences.

* *Moringa* oil comes from pressing the seed of the tree *Moringa oleifera* to yield a high quality edible oil.

† Economists often use regression analysis, which is a special case of a general linear model where the response variate is quantitative, but the explanatory variables can be quantitative or binary or categorical qualitative variables.



Assumptions associated with the *anova* technique must also be kept in mind. The method assumes that (a) the individual evaluations are independent of each other; (b) scores given to each item come from populations of scores that have a common variance; and (c) that the data follow a normal distribution. Of these, (c) is the least problematical and can be mitigated if there are a sufficient number of respondents or respondent groups involved. See Wilson (2000) for some guidance on sample sizes.

Assumptions (b) and (c) can be checked via what statisticians refer to as a residual analysis. Data collection procedures should ensure that assumption (a) is met. For example, restricted scoring, where respondents are asked to distribute a fixed number of seeds among the items being scored, do not lead to independent observations. Responses can be regarded as being independent if ranks or scores given by one person do not directly influence the responses given by another person to the same item, or by the same person to a different item. This can be ensured during data collection, e.g. by interviewing respondents individually rather than collectively. Where it is more meaningful to collect the information in group discussions, many repeats of the process will provide the required replication.

If *anova* assumptions hold, then further analysis is possible, e.g. it would be possible to ascertain whether the oil receiving the highest mean score had a significantly higher preference than the oil receiving the next highest mean score.

The general technique used here is quite powerful and is based on an underlying general theory that can be applied even if some data are missing or if the number of items being evaluated differ across respondents. Appropriate statistics software (e.g. SPSS, Genstat, Minitab) is available to deal with the underlying data structures and other complexities.

More advanced methods of analysis do exist when large numbers of respondents are involved. For example, the frequency of farmers giving different ranks, or allocating scores, to each of a number of maize varieties, can be modelled using a proportional odds models (Agresti, 1996). The interpretation of results is then based on the odds of respondents preferring one item compared to another item. Some further advanced methods are described in SSC (2001).

7.3 Are non-parametric methods useful?

There is often a belief that non-parametric methods, i.e. methods which do not make distributional assumptions about the data, are appropriate for analysing qualitative data such as ranks. To illustrate the issues involved, we consider again the type of data shown in Table 1, assuming such data are available for a larger number of respondents or focus groups within the community targeted for research results. Firstly, a simple summary is useful. For example, in a study in India, aimed at investigating the potential for integrating aquaculture into small-scale irrigation systems managed by resource-poor farmers, 46 farmers were asked to rank four different uses of water bodies according to their importance (Felsing *et al* 2000). The data summary shown in Table 3 clearly indicates irrigation as the primary use of water bodies. Thirty-four farmers (74%) rank this as the most important use of water.

Table 3. Number of farmers giving particular ranks to different uses of water bodies

Rank	Irrigation	Livestock consumption	Household use	Clothes washing
1	34	6	5	1
2	8	16	14	8
3	1	16	14	15
4	3	8	13	22
N	46	46	46	46
R _j	65	118	127	150

The research was however intended for the larger community from which the representative sample of 46 farmers was chosen. An inferential procedure is then needed to determine whether the results of Table 3 are equally applicable to the wider population of farmers in the community. The Friedman's test (Conover, 1999) may be used to demonstrate that farmers' water uses differ significantly. The test demonstrated that the water uses ranked are not all perceived to be of equal importance. Further tests showed evidence that irrigation was indeed considered more important than other uses, and that there was insufficient evidence to distinguish the remaining three uses in terms of importance.

If ties occur in the data, an adjustment to the Friedman's test statistic is available. However, three other associated problems need consideration. The first is that the test is based on an approximation to the chi-squared distribution. If n is the number of farmers and k is the number of items being ranked, then nk must be reasonably large, say ≥ 30 , for the approximation to be reasonable. Secondly, the test makes no allowance for missing data. If missing data do occur, because (say) the number of items ranked varies from farmer to farmer, then ranks need to be converted to scores in some fashion (Abeyasekera *et al*, 2000) and then analysed using procedures described in the previous section. Thirdly, the test cannot take into account the full data structure.

Friedman's test, as well as other alternatives (e.g. Anderson, 1959; Taplin, 1997), are examples of "non-parametric" methods. Many textbooks are devoted to non-parametric procedures but many of the methods apply to measurement data that are converted to ranks before analysis on the grounds that standard methods do not apply due to failure of assumptions such as normality. Their relevance is restricted to situations where testing a hypothesis is of primary concern, and this may not always be the case with many PRA studies. The tests are also less powerful because they do not make full use of the original data. This means that differences between items being compared have to be quite substantial for statistical significance. Although many non-parametric tests exist, their use in analysing data from PRA studies is quite limited.

8. Other quantitative methods for qualitative data

Many other procedures are available for dealing with qualitative information that can be coded either as binary variables, i.e. Yes/No, presence/absence type data, or as categorical variables, e.g. high, medium or low access to regional facilities, decreasing/static/increasing dependence on forest

resources. If factors affecting qualitative features of the binary sort are to be explored, logistic regression modelling can be used. An illustration can be found in Martin and Sherington (1997).

If a categorical variable is of primary interest, e.g. in exploring factors influencing the impact of programmes introduced to improve health facilities in a region, using data in the form of peoples' perceptions of changes in their well-being (better, same, worse), then methods such as log-linear modelling or proportional odds modelling come into play. These are advanced statistical techniques and if they are relevant, some assistance from a statistician with knowledge of these techniques is advisable.

The advantage of using modelling procedures is that it enables the inter-relationships between a wide range of other factors to be taken into account simultaneously. All too often, researchers tend to look at one factor at a time, e.g. when interested in factors affecting successful co-management of forest resources. The one-factor-at-a-time approach may involve using a series of separate chi-square tests. The disadvantage in doing so is that the interactions between these factors is then ignored whereas modelling pays attention to possible interactions.

Once the results of these modelling procedures are available, the purely qualitative features of the data become important because they give breadth and depth to the formal research findings and provide the means to explain any special features that emerge. Typically, what happens is that the modelling procedure is followed by an analysis of residuals, i.e. the residual component associated with each data point after the variation in the main response variable has been adjusted for all other known factors. Some residuals may emerge as outliers. A return to the narratives is then essential for explaining and discussing these extreme cases. Identifying such cases is thus greatly facilitated by an initial quantitative treatment of the information gathered.

9. Final remarks

It is important to recognise that proceeding beyond straightforward data summaries and graphical presentations to formal statistical procedures and tests of significance has little value in helping research conclusions if sampling issues have not been appropriately addressed in the sample selection. One issue is whether the sample size gives an adequate representation of the communities being targeted for study. Here the sample size refers to the number of independent assessments obtained, either by interviewing respondents individually or by having separate discussions with a number of respondent groups. How large a sample is needed will depend on the specific objectives of the study.

In situations where the sample size is adequate and the sample has been appropriately chosen to represent the target population of interest, the application of statistical methods will provide greater validity to research conclusions. This paper provides guidance about some statistical analysis approaches for use in dealing with qualitative data, with emphasis on data in the form of ranks or scores. However many other approaches exist and some were briefly mentioned in section 8. In general, the exact approach for a particular study will be closely associated with the study objectives and other data collection activities (e.g. those related to on-farm trials). Drawing on the experiences of a survey statistician to identify the most appropriate analysis approach is likely to be beneficial.

References

Abeyasekera, S., Lawson-McDowall, J., and Wilson, I.M. (2000). Converting ranks to scores for an ad-hoc procedure to identify methods of communication available to farmers. Case-Study paper for DFID project on "Integrating qualitative and quantitative approaches in socio-economic survey work". Available at www.reading.ac.uk/ssc/

Abeyasekera, S. and Lawson-McDowall, J. (2000). Computerising and analysing qualitative information from a study concerning activity diaries. Case-Study paper for DFID project on "Integrating qualitative and quantitative approaches in socio-economic survey work". Available at www.reading.ac.uk/ssc/

- Agresti, A. (1996). *An introduction to Categorical Data Analysis*. Wiley, New York.
- Anderson, R.L. (1959). Use of contingency tables in the analysis of consumer preference studies. *Biometrics*, 15, 582-590.
- Conover, W.J. (1999). *Practical non-parametric statistics*. 3rd Edition. Wiley, New York.
- Felsing, M., Haylor, G.S., Lawrence A. and Abeyasekera S. (2000). Evaluating some statistical methods for preference testing in participatory research. DFID Aquaculture Research Programme Project.
- Fielding, W.J., Riley, J., and Oyejola, B.A (1998). Ranks are statistics: some advice for their interpretation. PLA Notes 33. IIED, London.
- Martin, A and Sherington, J. (1997) Participatory Research Methods – Implementation, Effectiveness and Institutional Context. *Agricultural Systems*, Vol. 55, NO. 2, pp. 195-216.
- Maxwell, S., and Bart, C. (1995). Beyond Ranking: Exploring Relative Preferences in P/RRA. PLA Notes 22, pp. 28-34. IIED, London.
- Moris, J. and Copestake, J. (1993). *Qualitative Enquiry for Rural Development*. IT Publications Ltd.
- Poole, J. (1997). Dealing with ranking/rating data in farmer participatory trials. MSc thesis. University of Reading.
- Pretty J.N., Gujit I., Scoones I. and Thompson J. (1995). IIED Participatory Methodology Series. Int. Institute for Environment and Development, London.
- SSC (2001). Modern Methods of Analysis. Statistical Guidelines Series supporting DFID Natural Resources Projects. Statistical Services Centre, The University of Reading, www.reading.ac.uk/ssc/
- Taplin, R.H. (1997). The Statistical Analysis of Preference Data. *Applied Statistics*, 46, No. 4, pp. 493-512.
- Warburton H., Villareal, S. and Subramaniam, P. (1998). Farmers' rice tungro management practices in India and Philippines. Proceedings of a workshop on Rice Tungro Management. IRRI, Philippines, 9-11 November 1998.
- Wilson, I.M. (2000). Sampling and Qualitative Research. Theme Paper for DFID project on "Integrating qualitative and quantitative approaches in socio-economic survey work". Available at www.reading.ac.uk/ssc/.