

Randomisation in Tools for Teaching and Testing in Statistics

Alan McLean

Monash University, Department of Econometrics and Business Statistics

900 Dandenong Road

Melbourne, Australia

Alan.McLean@Buseco.monash.edu.au

1. Introduction

It is generally accepted that in learning statistics, there is a need for sets of questions for students to use to help and to monitor their learning and to practise skills,. Such sets are widely available – any textbook is full of them - many available for use with a computer, which enables instant correction of answers and help with errors. The software for administering such a set of questions is frequently simplistic, presenting the student with the same questions every time he undertakes the topic. With repetition, he learns to answer this set of questions correctly.

In this paper I discuss certain aspects of what makes a set of questions ‘good’, and how this may be achieved. A longer version for distribution will also discuss the *Tuteman* software.

2. What are the characteristics of a good set of questions?

First, questions should show considerable **variety**, some unpredictability and freshness. A question should not be received with the reaction ‘This question again!’ Second, questions need to be **focussed** quite precisely on a particular topic, approach and level of difficulty. They should reflect the objectives of the subject. This implies that questions should be written by – or selected by – the subject coordinator.

Third, the **function** of the question set needs to be clearly identified. Is it for assessment? development of ideas? reinforcement? practice of skills? It is my firm opinion that questions designed for encouraging and developing learning are not generally appropriate for assessment. Again, this implies that the subject coordinator should be involved in the writing or selection¹.

Questions which are suitable for computer use are currently limited in type. First, a question must have an answer! The computer must be able to compare the student’s answer with what it expects to find, and recognise it as ‘correct’ or ‘incorrect’. Generally speaking, usable questions are of four types: **multiple choice** or **short answer**, in each case where the correct answer is ‘**known**’ to or can be **calculated** by the computer.

The variety we would hope to find includes **variety of stories** (the questions appear to be different but are structurally the same), **variety of data** (the calculations give different answers but are essentially the same) and **variety of structure** (the questions and the calculations are different).

To a limited extent, and with a great deal of effort, this variety can be achieved by writing a

¹ There are of course other characteristics, such as simple quality of question.

sufficiently large number of questions. This is exemplified by the use of exercises from a text book. With a computer package this variety can be introduced to a much greater extent, and with much less effort by making use of **randomisation**. This is used in two ways.

First, a relatively common facility is to utilise **random selection of questions**. This is, on its own, quite limited. It is simply too difficult to write a sufficiently large number of sufficiently good, sufficiently different questions on each relatively limited topic.

Random generation of data is the obvious way to introduce the second type of variety. This seems to be much less common than is random selection of questions, although in statistics teaching it is much used with simulations. However using random generation of data is a very efficient way of using a **small** set of questions to provide more or less unlimited variety in terms of the actual calculations. It can also be used to vary the story by selecting values of nominal variables. In the following example, the variables N, P and CONFIDENCE will provide the calculation. The other variables simply personalise the question, so that each time it appears it will look different. They can also be used to provide some humour!

[PERSON] is the [POSITION] for [COMPANY]. [SEX] is interested in estimating the proportion of the population of [CITY] who [ACTION]. In a random sample of [N] people, [P] per cent said they do [ACTION]. Obtain a [CONFIDENCE] per cent confidence interval for the required proportion.

The third type of variety – that of structure – is more difficult to introduce, and usually will involve some level of programming, although random selection of data is sufficient to allow some variation in structure. Whether such questions are easy or even possible to formulate depends on the software used. The most common approach is probably to write different questions rather than try to make individual questions too complex.

3. What are the characteristics of good questioning software?

First, and probably foremost, the software should be easy and attractive for the student to use. Second, it should provide administrative facilities such as recording of results and cumulative results. Third, if it is to be used for assessment, it must provide security facilities, such as use of a password, checking for tampering with results and encryption.

Fourth, and most important from my point of view, is that the software provides the facility to write a variety of questions, reflecting the discussion above, **easily**. Fifth, it should enable the selection of questions according to the results of previous questions. Typically, current software often enables a limited version of this, in that if a question is incorrectly answered, the same question is represented, but there is great scope for development.

RESUME

Ce papier discute l'usage de randomisation par sélection de questions et génération de données dans logiciel pour fournir pratique a etudiants. La considération est donnée à que que fait une bonne série de questions, et que que fait le bon logiciel.