TEACHING THE MEASUREMENT PROCESS IN BIOSTATISTICS

Elisabeth Svensson,
Biostatistics Branch, Mathematical Statistics
Chalmers University of Technology and Göteborg University, Sweden

The measurement process is the link between a research problem and the result. Different kinds of operational approaches to the same research problem can lead to different permissible statistical methods. As numerical assignments are common for categorical data as well as for true numerical data, it is important to discuss the meaning of response values. This paper presents a model for teaching the measurement process in biostatistics. The aim is to shed light on the importance of considering the operational process behind the numerical labels, among both statisticians and applied users of statistics. A scheme for considering different ways of operationalisation of the same variable is presented. Questions for reviewing the measurement process in a medical paper are also given.

INTRODUCTION

Numerical symbols are commonly used in the measurement of both quantitative and qualitative variables. This is often the case in medical and behavioural sciences, as questionnaires and rating scales for the evaluation of general feelings or quality of life have become important complements to traditional objective measurements of the function of body systems and processes (Bowling, 1995; McDowell and Newell, 1987). Categorical responses are often transformed into numerical scores (Merbitz and Morris, 1989; Hand, 1996; Svensson, 1993). Hence, a mixture of numerical response values with different mathematical properties can be found in the same study. Attention must therefore be paid to the measurement process that is behind the numerals before the methods for statistical analysis are chosen (Hand, 1996).

The literature concerning both applied and theoretical statistics, however, does not discuss the different theories of measurement (Hand, 1996) but focuses instead often on the statistical methods used for quantitative variables. Therefore, the responsibility falls on teachers to discuss the impact of the measurement process on the mathematical properties of numerical response values.

The aim of this paper is to describe a model for teaching the measurement process to scientists in medical and behavioural sciences as well as to graduate students in biostatistics. The purpose is to draw the attention of statisticians to the

existence of a wide variety of quantified, non-numerical data from rating scales. It is also meant to increase the awareness of medical students and scientists of the link between the properties of data and the appropriate statistical analyses.

MAIN TOPICS IN TEACHING THE MEASUREMENT PROCESS

The measurement process is the link between a research problem and the result. Concepts to be studied can often be measured in different ways that are dependent mainly on the paradigm or research tradition but also on the operational definition of the concept (Svensson, 1993). The operationalisation process is to identify observable indicators of the concept to be studied and leads to the choice of measurement instrument. There are standardized instruments for measuring physical properties, such as blood pressure and body weight, but qualitative, subjective variables can be operationalised in various ways processes (Bowling, 1995; McDowell and Newell, 1987).

Figure 1 shows a scheme for discussing the impact of operationalisation on the properties of data. Various ways of measuring the variable “functional ability” are shown. The numerical response values of “functional ability”, expressed as walking speed (m/s) and as a level on the hierarchical scale of measuring dependence in Activities of Daily Life (ADL) by means of the ADL staircase (soon, 1995) (ranging from 0 to 9), for example, have different mathematical properties (Hand, 1996; Stevens, 1946), which means that different types of statistical methods are appropriate (Figure 2).

Ordered categorical data are invariant under monotone ordered transformations, which means that categorical labels, even ordinal numbers, indicate only a rank order and not a mathematical value. Thus, the statistical methods used should not be affected by any kind of ordered relabelling of the scale categories. This means that sums of and differences between categories have no interpretable meaning. Calculation of the mean and standard deviation therefore does not make sense for ordered categorical data. Also, attention must be paid to the construction of a global score for multi-item instruments, as a sum score is not appropriate. It is important to discuss and demonstrate various rank-invariant approaches for construction of global scores for multi-item scales, such as hierarchical conditional scales, profiles and median values.

Figure 1. A scheme for discussing different operational definitions of a variable to be studied. The variable "functional ability" is used as an example.

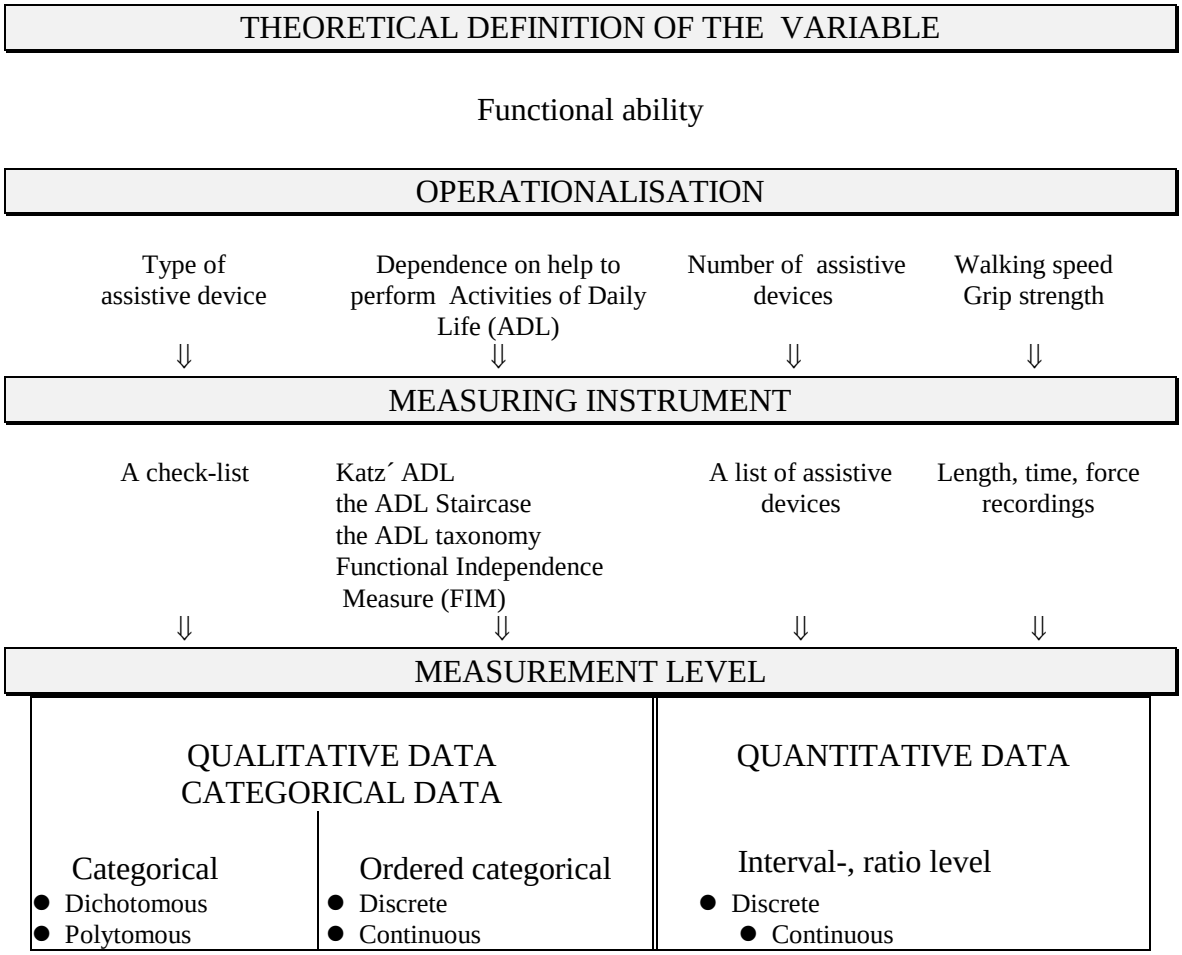

Figure 2. The properties of quantified, ordered categorical data and of true numerical data and how these properties affect the statistical treatment of data

	Ordered categorical data (ordinal level)	Quantitative data (interval-, ratio level)
Properties	Rank-invariance: the labels represent only a structure of ordering (non-numbers (Merbitz and Morris, 1989)), hence non-additivity	The values have order, equi-distance and a true mathematical meaning (true numbers)
Statistical description	Median, centiles, range, inter-quartile range	Mean, median, range, inter-quartile range, standard deviation.
Quality concepts and statistical approach	Concepts of validity and reliability; disagreement measures	Accuracy and precision; standard deviation

ARE THE NUMBERS TRUE NUMERALS?

The fact that all numbers are not true numbers is demonstrated by the example in Figure 3.

Figure 3. The basic idea in teaching the meaning of quantitative and quantified response values. Are numerical response values true numerals?

What do the numbers 1, 2, 3, ... mean in the following variables?		
Measurement level	Example of variable	Interpretation
Continuous quantitative data	Volume (ml)	2 ml stands for 1.5 to 2.5 ml 4 ml stands for 3.5 to 4.5 ml
Discrete quantitative data	Functional ability: Number of assistive devices	4 assistive devices are twice as many as 2
Ordered categorical data	Level of consciousness in acute brain disorders according to the Reaction Level Scale (RLS85) (Starmark, 1988)	2 stands for “Drowsy or confused” 4 stands for “Unconscious.” 4 indicates a higher level of consciousness than 2
Categorical data	Leisure activities	1,2,3,.. are labels for different activities: 1: badminton, 2: construction work, ..., 4: table tennis, ...

EXERCISE: THE MEASUREMENT PROCESS IN MEDICAL RESEARCH

A useful exercise for increasing awareness of the importance of identifying the conceptual framework behind measurements is to review a medical scientific paper concerning this issue (Figure 4). This exercise can be modified to describe the measurement process of a medical research project.

Figure 4. Questions for reviewing a medical scientific paper concerning the measurement process.

The purpose is to review the measurement process reported in a medical paper according to the following;

- What is the study purpose?
- Identify the main variable(s) and the measurement instrument(s).

Consider the following issues for each variable:

⇒ Is the motivation for the choice of measurement instrument clear?
⇒ What is the measurement level? Are the response values truly numerical or categorical? If categorical, how is the labelling performed?
⇒ Were the statistical methods used appropriate for the data?
⇒ Suggest an alternative operationalisation of the variable.
⇒ Review other reported studies involving the same variable. Compare and comment on the measurement processes.

DISCUSSION

The educational approach presented here is applicable to statisticians as well as to applied scientists. The teaching of the measurement process is important, as very little attention is paid to this process in the literature. In my experience, time should be made available to discuss these issues and perform the types of exercise described here.

Evaluation of the measurement process in research reports of concepts such as “Nutritional Status”, “Activities of Daily Life” (ADL) and “Quality of Life” illustrates the variety in measurement instruments and also in the interpretation of concepts. Consideration of the operational process is also valuable for a medical research project. In addition, a statistician should demand information about the operational definitions behind all types of numerical data before choosing the appropriate method of analysis.

Recently, Hand (1996) draw the attention of statisticians to the theory of measurements. The published discussions (Hand, 1996) of his paper illustrate various opinions about the importance of theories of measurements. It is, however, always

important to understand the impact of the measurement process on the choice of adequate methods of statistical analysis (Altman, 1991).

REFERENCES

Altman, D. G. (1991).Practical statistics for medical research. London: Chapman and Hall.

Bowling, A. (1995). Measuring health. A review of quality of life measurement scales. Philadelphia: Open University Press.

Hand, D. J. (1996). Statistics and the theory of measurement. (incl. discussion). *J. R. Statist. Soc. A, 159,* 445-92

McDowell, I., Newell, C. (1987). Measuring health: a guide to rating scales and questionnaires. Oxford: Oxford University Press.

Merbitz, C., and Morris J. (1989). Grip JC. Ordinal scales and foundations of misinference. *Arch Phys Med Rehabil, 70,* 308-12.

Sonn U. (1995). Longitudinal studies of dependence in daily life activities. Methodological development, use of assistive devices and relation to impairments and functional limitations (dissertation). Göteborg:Göteborg University.

Starmark, J. E. (1988). Analysing “coma scales”. The introduction of a new scale-The Reaction Level Scale (RLS85) (dissertation). Göteborg: Göteborg University.

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103,*:677-80

Svensson, E (1993). Analysis of systematic and random differences between paired ordinal categorical data (dissertation). Stockholm: Almqvist and Wiksell International.