Available online at www.sciencedirect.com

SCIENCE DIRECT®

Chemometrics and Intelligent Laboratory Systems 74 (2004) 7–24

Chemometrics and intelligent laboratory systems

www.elsevier.com/locate/chemolab

Sampling of discrete materials—a new introduction to the theory of sampling
I. Qualitative approach

Pierre Gy

Res. de Luynes, 14 Avenue Jean de Noailles, F-06400 Cannes, France

Received 22 August 2003; received in revised form 6 April 2004; accepted 28 May 2004
Available online 15 September 2004

Abstract

The purpose of a *theory of sampling* is to answer two questions: *How* should one select a sample?—*How much* material should be selected? Parts I (qualitative approach), II and III (quantitative approach) of this series propose answers to these two fundamental questions. These answers are not entirely new (answers have been formulated since 1950), but a scientific theory is a living structure that has to be kept up to date. At a course given in Brasilia in 1998, pointed questions were raised which suggested that the introduction to the qualitative approach had to be clarified. Part I represents the most updated introduction to theory of sampling (TOS). More than 200 scientific papers, books, lectures and courses on sampling theory—and practice—have been published or offered to the public by the author over a period of 50 years. A brief, chronological account of the development history of TOS is presented for the first time in part IV—with a comprehensive literature survey as part V.

Keywords: Theory of sampling; Discrete materials; Heterogeneity; Sampling errors; Accuracy; Bias; Reproducibility; Representativity; Quality control; Quality assurance; Chemometrics

1. Introduction

The accuracy of many analytical data reports is a mirage because unwitting negligence and false cost consciousness have ensured that a sample of powder taken with cursory swiftness has been examined with costly precision. *Kaye*, Illinois Institute of Technology, 1967

Chemometricians process analytical data, more often than not huge amounts of data. Are these *data* reliable? If *Kaye* is right, which fully agrees with the author's extensive experience, we are entitled to have our doubts. If the data are *biased* as a consequence of systematic sampling errors, what becomes of the chemometricians' conclusions? We have every reason to be cautious that these conclusions may be biased too. If the data are *uncertain,* for example as a consequence of high random *sampling errors* (high sampling variances), the efficiency of statistical tests will be reduced by the high residual variances. It will invariably be more difficult and/or more costly to reach safe and reliable conclusions. Few chemometricians are aware of these facts. Below it is shown that there is no such thing as a "constant sampling bias", which is the basis for many current complacent, but false sampling understandings. This tutorial is intended to highlight that a complete theory of sampling is in fact at hand—and has been for 50 years!

The heart of the matter of proper sampling is that the question of "how much?" cannot be dissociated from the question of "how?". Indeed, quantitative development of sampling theory assumes explicitly that a certain number of conditions have, by being respected, successfully suppressed the sampling bias. These conditions are presented in the qualitative approach in Part I.

E-mail addresses: kes@aue.auc.dk, kimhome@esenet.dk.

0169-7439/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemolab.2004.05.012