

# Ignorable Missingness

Michel Mouchart<sup>a</sup>

<sup>a</sup>*Institut de statistique, Université catholique de Louvain (B).*

February 6, 2008

## Abstract

The first part of this paper sketches a coherent framework for dealing with a large class of problems related to missing or incomplete data. The attention is concentrated on situations where the actually available data may be represented as a deterministic function of a latent (or, "ideal") variable and of a "signal" (or, "reporting") variable that eventually determines which part of the latent variable is actually observed. This framework provides a precise concept, and definition, of missing data. Several examples, in particular in the field of survey sampling, are shown to fit in such a framework. Within that framework, precise definitions of ignorability, in the sense of Rubin, are shown to require particular care and specific qualifications. Different versions of missingness conditions are discussed. The second part of the paper gives several theorems that display sufficient conditions to ensure that the reporting variable is ancillary or that the available data is sufficient.

*Keywords:* Specification problem, Sufficiency, Ancillarity, Minimal sufficiency, Maximal Ancillarity, Non-response, Missing-at-random, Ignorable Missingness, Data coarsening.