

Nested case-control and case-cohort designs : an overview and some future challenges

Ornulf Borgan *

borgan@math.uio.no

Cox regression and other relative risk regression models are much used in clinical and epidemiological studies to assess the influence of risk factors and other covariates on mortality or morbidity. Estimation in these models requires collection of covariate information for all individuals in a cohort also when only a small fraction of them get diseased or die. This may be very expensive, or even logistically impossible, for large cohorts. Cohort sampling techniques, where covariate information is collected for all failing individuals “cases”) but only for a sample of the non-failing ones (“controls”), then offer useful alternatives which may drastically reduce the workload of data collection and data checking. Nested case-control sampling is the most common cohort sampling design. For this design each case is compared to a small number of controls selected among those at risk at the case’s failure time. An alternative is case-cohort sampling, where a sample of control individuals (the subcohort) is selected at the outset of the study, and these individuals are used as controls throughout. In their original forms, the nested case-control and case-cohort designs use simple random sampling without replacement for the selection of controls/subcohort. Modifications using stratified sampling (and other sampling schemes) have been developed and may lead to an efficiency gain. In the presentation an overview of various cohort sampling designs will be given and the pros and cons of different designs and analysis strategies will be pointed out. Some future methodological challenges will also be discussed.

*Department of Mathematics, University of Oslo, P.O. Box 1053 Blindern, N-0316 Oslo, Norway.