Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 137 (2007) 3741 – 3752
journal of statistical planning and inference
www.elsevier.com/locate/jspi

Union–intersection principle and constrained statistical inference

Pranab Kumar Sen[a, b, *]

[a] Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27599-7420, USA
[b] Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, NC 27599-3260, USA

Available online 30 March 2007

Abstract

Most statistical models arising in real life applications as well as in interdisciplinary research are complex in their designs, sampling plans, and associated probability laws, which in turn are often constrained by inequality, order, functional, shape or other restraints. Optimality of conventional likelihood ratio based statistical inference may not be tenable here, although the use of restricted or quasi-likelihood has spurred in such environments. S.N. Roy's ingenious union–intersection principle provides an alternative avenue, often having some computational advantages, increased scope of adaptability, and flexibility beyond conventional likelihood paradigms. This scenario is appraised here with some illustrative examples, and with some interesting problems of inference on stochastic ordering (dominance) in parametric as well as beyond parametric setups.

MSC: 62A10; 62A99; 62F99; 62G99; 62H99

Keywords: Beyond parametrics; Inequality; Order and shape restraints; Likelihood; Optimality; Robustness; Stochastic ordering

1. Introduction

Statistical models usually advocated for interdisciplinary research and many real life applications are rarely very simple so as to make room for routine adaption of conventional or standard statistical inference tools. Such studies generally involve complex designs, sampling plans and the underlying stochastics relate to probability laws which, typically, not only involve a multitude of parameters but also these parameters subjected to various nonlinear restraints. Inequality, order, functional and shape constraints are commonly encountered, in probability as well as sample spaces, in such applications, and sometimes, even the large dimension and data structures may complicate the scenario considerably. For beyond parametrics (i.e., nonparametrics and semiparametrics) setups, often, there could be more complex restraints involving functional constraints. Stochastic ordering (dominance) in categorical data models, aging perspectives in life distributions (such as monotone hazards, decreasing mean remaining life (DMRL) or increasing failure rate average (IFRA)) are notable examples of this kind. In conventional statistical inference, the likelihood, sufficiency and invariance principles play a key role in finite sample methodology, and some of the finite-sample optimality properties usually transpire in the large sample case even without sufficiency or some other regularity conditions. Nevertheless, even in asymptotics, lacking support of suitable regularity assumptions, particularly in constrained environments, optimal statistical inference may encounter roadblocks of diverse types.

* Corresponding author at: Department of Biostatistics, University of North Carolina, Chapel Hill, NC 27599-7420, USA.
E-mail address: pksen@bios.unc.edu.

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2007.03.046

Generally, complex statistical models create impasses for computation of maximum likelihood estimators (MLE) and likelihood ratio tests (LRT) in closed explicit or manageable forms; often, this may become a formidable task. Even so, various algorithms have been developed for such computational convenience, the finite sample optimality properties of MLE and LRT ranging over the exponential family of densities may not automatically transpire in more complex models where the underlying probability laws are rarely bonafide members of such regular families. (Restricted) RMLE and (restricted) RLRT along with various modifications of the likelihood function have therefore been advocated for such complex models, albeit they may not universally have an established optimality property parallel to that in simple models. S.N. Roy's (1953) ingenious union–intersection principle (UIP), having its genesis in the likelihood principle (LP), has emerged as a viable alternative, often having some computational advantages, increased scope of applicability (beyond the likelihood paradigm), greater adaptability to nonstandard situations (beyond the parametrics), and good robustness perspectives.

For a general treatise of constrained statistical inference (CSI) we refer to a recent monograph by Silvapulle and Sen (2004) which has recaptured the prior developments in Barlow et al. (1972) and Robertson et al. (1988). The major emphasis in Barlow et al. (1972) has been the finite sample methodology with due consideration of the basic role of the likelihood function in such formulations. More in-depth computational aspects are additionally reported in Robertson et al. (1988). The Silvapulle–Sen (2004) treatise goes beyond that into more general setups with adequate asymptotics to simplify the methodology; in line with the Wald-type tests, in CSI such procedures are elaborated, and the UIP's basic role has also been depicted in some important problems. The present study is devoted to a display of the basic role of UIP in CSI with emphasis on stochastic ordering (dominance), clinical trials and meta analysis, the latter topic being very useful in the developing field of genomics or bioinformatics.

2. UIP: preliminary notion

The scenario of statistical inference changes drastically from the parametric to beyond parametric perspectives, and even in the parametric case from the single parameter to multiparameter setups. The more complex a model is, it is more likely that optimal statistical inference may be harder to implement. The evolving field of genomics is a pertinent citation of the enormous difficulties that conventional statistical inference tools are encountering in this high-dimensional low-sample size setups. The genesis of UIP lies in this complex. Moreover, CSI typically pertains to such complex statistical environments, and hence, it is natural to appraise the interactive role of UIP in CSI. In the early days of developments, UIP used to be motivated through the important portfolios: multiple comparisons and simultaneous statistical inference. We find it easier to illustrate UIP with a general composite hypothesis testing problem that lends itself naturally to simultaneous statistical inference as well as to CSI.

Let us consider a general hypothesis testing problem, not necessarily confined to a parametric model. Let H_0 be the null hypothesis of interest and let H_1 be the alternative one; both of them are composite so that the likelihood function is not completely specified under either of them. As it is the case with composite hypotheses testing problems, there may not be in general an optimal test for testing H_0 vs. H_1, and in many case, even finding out a similar region may restrict attention to a subclass of tests like invariant tests, conditional tests, etc.. This situation is likely to be worse in CSI where the conceived restraints may preempt the relevance of invariant or conditional tests. However, for a general class of testing problems, including in CSI, it might be possible to express

$$\mathrm{H}_0 = \bigcap_{j \in \mathscr{J}} \mathrm{H}_{0j}, \quad \mathrm{H}_1 = \bigcup_{j \in \mathscr{J}} \mathrm{H}_{1j}, \tag{2.1}$$

where $\mathscr{J}$ is a suitable index set, and for each $j \in \mathscr{J}$, there exists a suitable (and often optimal in a certain sense) test for testing H_{0j} vs H_{1j}. In a parametric framework, such a test could be the UMP (uniformly most-powerful) test whenever the latter exists, could be LMP (locally most-powerful) test in some other case, and in beyond parametrics setups, such a test can be decided on the basis of robustness, validity and efficiency considerations. Further, the index set $\mathscr{J}$ can be a finite (discrete) set, or it may even be a set in continuum. In this way, there is flexibility in the decomposition of the hypotheses and choice of appropriate test statistics. Bearing in mind the genesis of UIP in LP (Roy, 1953), we consider first the following illustrative example depicting the connection of UIP and LP.

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be n independent and identically distributed random p-vectors having a multivariate normal distribution with unknown mean vector $\boldsymbol{\mu}$ and dispersion matrix $\boldsymbol{\Sigma}$, unknown but positive definite (p.d.). Consider first the null hypothesis $\mathrm{H}_0 : \boldsymbol{\mu} = \mathbf{0}$ versus $\mathrm{H}_1 : \boldsymbol{\mu} \neq \mathbf{0}$, treating $\boldsymbol{\Sigma}$ as a nuisance parameter (matrix). There is no UMP test