Testing Semiparametric Hypotheses and Unorthodox Bootstraps

Peter Bickel

Suppose we observe a sample $X_1, \ldots, X_n$ of i.i.d. $\mathcal{X}$ valued observations distributed according to P completely unknown. In testing parametric or semiparametric hypotheses such as normality, independence, the Cox survival model and others the classical paradigm has been to either:

(I) a)   Propose a test statistic based on maximizing some type of semiparametric likelihood or something equivalent such as Pearson's χ^2 like tests and

b)   Set critical values on the basis of χ^2 distributions usually with number of degrees of freedom going to infinity,

or

(II)   Find some $n^{-1/2}$ consistent convenient measure of distance between the empirical and hypothesis distributions such as Kolmogorov–Smirnov and Cramer von Mises fit in some way and use the fitted measure as a test statistic. Here convenient usually means that there are simple asymptotic approximations to the critical value.

Tests of type (I) have approximately constant power in all directions, paid for by having asymptotic power equal to the significance level in all directions at the $n^{-1/2}$ scale. Tests of type (II) are characterized by having nontrivial power at the $n^{-1/2}$ scale in all directions, but with substantial power concentrated in a set of directions dictated by the metric used.

(III)   More recently—see Hart (1997) for a review of some of these procedures, adaptive tests of type (I) have been proposed based on what can be viewed as data determined model choice. Tests of this type implicitly order directions by considering a sieve $\{\mathcal{P}_m\}_{m \geq 0}$ of models with $\mathcal{P}_0$ being the hypothesis. One then considers a sequence of test statistics T_m for $H : P = P_0$ versus $P \in \mathcal{P}_m - \mathcal{P}_0$ with T_m usually being an approximation to the likelihood ratio test and a sequence of boundary values $\{a_m\}$ possibly depending on n. The test is then to reject if any $T_m - a_m > 0$ and accept otherwise. The tests of Bickel and Ritov (1992), and the order selection test of Eubank and Hart (1992) are of this type and do have power at the $n^{-1/2}$ scale in all directions.

In this talk we review work of ours with many collaborators, Götze, van Zwet, Ritov, Ren, Sakov and Stoker in which we propose and show how to implement the following paradigm:

[A]   Select a set of directions in which substantial power is desired and construct test statistics with this property which also retain non trivial power at the $n^{-1/2}$ scale in all directions.

[B]   Construct appropriate bootstrap methods for setting critical values for the tests statistics in question.

We show how tests of type (II) and (III) relate to our paradigm, give conditions for which our construction is valid for general semiparametric hypotheses and validate them in a number of examples.

1