Analysis of Censored Data
IMS Lecture Notes – Monograph Series (1995) Volume 27

Interval Censoring, Case 2: Alternative Hypotheses

Jon A. Wellner [1]

Department of Statistics GN-22, University of Washington, Seattle, Washington 98195

"Interval censoring case 2" involves observation times (U, V) with distribution H concentrated on the set $u \leq v$ and a time of interest X with distribution F. The goal is to estimate F based only on observation of i.i.d. copies of $(1_{[X \leq U]}, 1_{[U < X \leq V]}, U, V)$. Groeneboom (1991) initiated the study of the nonparametric maximum likelihood estimator $\widehat{F}_n$ of F; see Groeneboom and Wellner (1992), especially pages 43 - 50 and 100-108. Geskus (1992) and Geskus and Groeneboom (1994) have studied the estimation of smooth functionals (such as the mean of F) in case 2. Under hypotheses ensuring that the observations times U and V are close with (sufficiently) positive probability, Groeneboom (1991) showed that a one-step approximation $F_n^{(1)}$ to the nonparametric MLE satisfies

$$(n \log n)^{1/3}(F_n^{(1)}(t_0) - F(t_0)) \rightarrow_d 2\{\frac{3}{4} f_0^2(t_0)/h(t_0, t_0)\}^{1/3} Z$$

where Z is the last time where standard two-sided Brownian motion W minus the parabola $y(t) = t^2$ reaches its maximum. While it is conjectured in Groeneboom and Wellner (1992) that the nonparametric MLE $\widehat{F}_n$ has this same behavior, this conjecture is still unproved.

The goal of this paper is to explore alternative hypotheses under which U and V are *not* close with high probability. Under these alternative hypotheses, the one-step approximation to the nonparametric MLE will be shown to converge at rate $n^{-1/3}$ rather than $(n \log n)^{-1/3}$, much as in interval censoring case 1 (current status data). We will also briefly discuss the behavior of the one-step NPMLE with $k > 2$ observation points and estimators of smooth functionals.

Research supported in part by National Science Foundation grant DMS-9306809, NIAID grant 2R01 AI291968-04, NATO NWO Grant B 61-238.

AMS 1980 *subject classifications.* Primary: 60F17, 62E20; secondary 60B12.

Key words and phrases. asymptotic distribution, case 2, empirical processes, hypotheses, interval censored data, mean, nonparametric maximum likelihood.