Abstracts of the Ninth Workshop on Simulation (2018) 1-2

Statistical Implications of Informative Dose Allocation in Binary Regression

Nancy Flournoy [1] and Assaf Oron [2]

In many fields such acute toxicity studies, Phase I cancer trials, sensory studies and psychometric testing, binomial regression techniques are used to analyze data following informative dose allocation. We assume the simplest general case in which a univariate binary response Y has a monotone positive response probability $P(Y = 1|X = x) = F(x)$ to a stimulus or treatment X; X values are sequentially selected from a finite discrete set $\{d_1, d_2, \ldots d_M\}$ of M values to concentrate treatments in a region of interest under $F(x)$. We call a positive response a toxicity and the stimulus a dose. From first principles, we describe dependencies that are introduced by sequentially choosing informative doses.

Suppose n subjects receive treatments that were sequentially selected according to some rule using data from prior subjects. Let N_m and T_m denote the final number of subjects allocated to dose d_m and the number of toxicities observed there, respectively, $m = 1, \ldots, K$. We remind the reader that the *joint likelihood* of treatment and response data is a function of the final set of observed allocation and toxicity rates $\{N_m, T_m, m = 1, \ldots, K\}$ [1] that is similar to likelihood for binomial random variables but with random sample sizes, and without the combinatorial constants. That is, under mild conditions that typically hold:

$$\mathcal{L}_n = \prod_{m=1}^{M} F(d_m)^{T_m}[1 - F(d_m)]^{N_m - T_m}.$$

We refute the prevailing notion that T_m given N_m is a binomial random variable; and characterize, at finite sample sizes, the bias of the observed toxicity rate T_m/N_m for $F(x)$ at dose x. We show that

$$E[T_m/N_m] = F_m - Cov[T_m/N_m, N_m]/E[N_m].$$

So, the observed toxicity rate is biased for F_m when adaptive allocations, by design, induce a correlation between toxicity and allocation rates. Commonly used variance formulae are also first-order linear approximations.

Understanding this is important for small to moderately sized designs because isotonic regression methods use the toxicity rates $\{T_m/N_m, m = 1, \ldots M\}$ directly; and standard likelihood methods indirectly as first-order linear approximations. We study these biases using isotonic and likelihood-based regression methods in some commonly used (small sample size) adaptive methods including some up-and-down designs, interval designs, and the continual reassessment method.

[1]Department of Statistics, University of Missouri, Columbia, MO 65221, USA, E-mail: flournoyn@missouri.edu

[2]Institute for Disease Modeling, Bellevue, WA 98005, USA, E-mail: assaf.oron@gmail.com

References

[1] Rosenberger W., Flournoy N., Durham S.D. *Asymptotic normality of maximum likelihood estimators from multiparameter response-driven designs* // J. Stat. Plann. Inf., 1997, v. 60 p. 69-76.