Statistica Sinica Preprint No: SS-2019-0015	
Title	Entropy Learning for Dynamic Treatment Regimes
Manuscript ID	SS-2019-0015
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202019.0015
Complete List of Authors	Wenbin Lu
Corresponding Author	Wenbin Lu
E-mail	lu@stat.ncsu.edu

Statistica Sinica: Preprint
doi:10.5705/ss.202019.0015

DISCUSSION OF "ENTROPY LEARNING FOR DYNAMIC TREATMENT REGIMES" BY JIANG SONG, LI AND ZENG

Wenbin Lu

North Carolina State University

I congratulate the authors for their thoughtful article on using entropy learning to estimate optimal individualized treatment rules. Their study makes several major contributions. First, it proposes a class of smooth outcome-weighted loss functions for estimating optimal individualized treatment rules. Second, the Fisher consistency and a proper inference for the parameters of the estimated treatment rules can be established in the proposed general framework.

The proposed smooth loss function is motivated by the sign consistency of the derived optimal treatment rule. When the true optimal treatment decision rule is contained in the considered class of treatment rules, can the estimated optimal treatment rule obtained using entropy learning be shown to have sign consistency asymptotically? In addition, given a treatment rule, the proposed smooth loss function is just an approximation to the weighted classification error loss (corresponding to the value function). Is it possible to quantify the difference between the value functions under

Statistica Sinica: Preprint
doi:10.5705/ss.202019.0015

the derived optimal rule using entropy learning and the true optimal rule? Among the class of proposed smooth loss functions, is it possible to find an optimal loss function that minimizes the value difference?

To derive the asymptotic properties of the parameter estimates in the derived optimal treatment rule, the authors make a few assumptions. In particular, assumption (A3) ensures that the optimal decision is estimable. Does this assumption exclude the possibility of having a nonregular setting, that is $P(X_t^{*\prime}\beta_t^0 = 0) > 0$. Under the nonregular setting, can we establish the asymptotic distributions of the estimators in the derived optimal treatment rule and its associated estimated value function, as in Theorems 1 and 2? Finally, can the proposed entropy learning method be extended to accommodate multiple treatment options at each treatment stage? I would appreciate comments from the authors on these issues.