Art Owen's contribution to the Discussion of "Estimating means of bounded random variables by betting" by Ian Waudby-Smith and Aaditya Ramdas

Art B. Owen
Department of Statistics
Stanford University
Sequoia Hall
390 Jane Stanford Way
Stanford, CA 94305, USA
owen@stanford.edu

June 2023

I congratulate the authors on some very interesting work making connections between likelihood, betting and martingales.

What caught my eye was the connection to the empirical likelihood (EL) (Owen, 2001) and the dual likelihood (Mykland, 1995). The hindsight optimal λ solves $0 = \sum_{i=1}^{t}(x_i - m)/(1 + \lambda^\mathsf{T}(x_i - m)) = 0$ corresponding to the observation weights $w_i \propto 1/(1 + \lambda^\mathsf{T}(x_i - m))$ used in empirical likelihood. Two common alternatives use weights $w_i \propto \big(1 - \lambda^\mathsf{T}(x_i - m)\big)$ and $w_i \propto \exp(-\lambda^\mathsf{T}(x_i - m))$ arising from L_2 and entropy criteria, respectively (for different vectors λ). The entropy weights connect to exponential tilting and logistic regression; see for instance Hainmueller (2012).

The EL weights perform a kind of reciprocal tilting that gives them some special power properties. Kitamura (2003) shows that empirical likelihood tests cannot be dominated by other regular tests for moment restrictions, in a large deviations sense. His result is a nonparametric counterpart to the likelihood test optimality result of Hoeffding (1965) for multinomial distributions. Lazar and Mykland (1998) find that for true parametric models, empirical likelihood matches their power to second order and at third order either the empirical or the parametric tests could have greater power. In some overspecified moment models, empirical likelihood has such high power for detecting lack of fit that, under lack of fit, there cannot exist any pseudo-true value of the parameter for which the maximum EL estimate is root-n consistent (Schennach, 2007). We can now add that authors' hindsight optimality to this list of power properties.

A similar power optimality is achieved by the confidence bands of Berk and Jones (1979). They use the most significant binomial likelihood ratio from all n order statistics to form confidence bands with greater power than any weighted Kolmogorov-Smirnov test. It would be very interesting to see if the authors' methods could produce an always valid version of the Berk-Jones bands.

References

Berk, R. H. and Jones, D. H. (1979). Goodness-of-fit test statistics that dominate the Kolmogorov statistics. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 47(1):47–59.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political analysis*, 20(1):25–46.

Hoeffding, W. (1965). Asymptotically optimal tests for multinomial distributions. *The Annals of Mathematical Statistics*, pages 369–401.

Kitamura, Y. (2003). Asymptotic optimality of empirical likelihood for testing moment restrictions. *Econometrica*, 69(6):1661–1672.

Lazar, N. and Mykland, P. A. (1998). An evaluation of the power and conditionality properties of empirical likelihood. *Biometrika*, 85:523–534.

Mykland, P. (1995). Dual likelihood. *Annals of Statistics*, 23(2):396–421.

Owen, A. B. (2001). *Empirical Likelihood*. Chapman and Hall/CRCpress, Boca Raton, FL.

Schennach, S. (2007). Point estimation with exponentially tilted empirical likelihood. *The Annals of Statistics*, 35(2):634–672.