Other (statistical) arguments for employing betting scores (or e-values)

Aaditya Ramdas

Department of Statistics & Data Science
Machine Learning Department
Carnegie Mellon University

I personally vibe strongly with Shafer's math and philosophy,
but what if someone does not?

Some (purely) statistical utility of e-values:

P1. Dimension-agnostic inference

P2. Irregular models (composite null)

P3. Sequential inference (composite null)

P4. Multiple testing under arbitrary dependence

Batch setting

An e-value or betting score E is a nonnegative random variable with

$$\mathbb{E}_P[E] \leq 1 \text{ for any } P \in \mathscr{P}.$$

Sequential setting

A safe e-value $(E_t)_{t=0}^{\infty}$ is a nonnegative sequence of r.v. with

$$\mathbb{E}_P[E_\tau] \leq 1 \text{ for any } P \in \mathscr{P}, \text{ stopping time } \tau.$$

Shafer, Vovk, Grunwald, etc

P1. Dimension-agnostic inference

Wilk's theorem (log-LR is asymp. chi-squared) holds in certain regimes of dimension p, sample size n.

Mixture LR $\int \prod_{i=1}^{n} \frac{p_\theta(X_i)}{p_0(X_i)} dF(\theta)$ is an e-value, for any p and n.

Nonparametric generalizations exist! (Banach spaces, matrices)

Time-uniform nonparametric, nonasymptotic confidence sequences

Howard, Ramdas, McAuliffe, Sekhon, Probability Surveys 2020
+ Annals of Statistics 2020

P2. Irregular models (composite null)

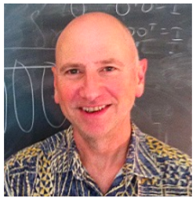
Wilk's theorem only holds for "regular models".
Irregular models: asymptotic distribution unknown.

Eg: testing whether a distribution is log-concave.

The "split-likelihood ratio" statistic is an e-value, when maximum likelihood estimation is tractable.

$$\prod_{D_0} \frac{p_{MLE(D_1;\Theta_1)}(X_i)}{p_{MLE(D_0,\Theta_0)}(X_i)}$$

Universal inference



Wasserman, Ramdas, Balakrishnan, PNAS 2020

P3. Sequential inference (composite null)

For point nulls, all admissible e-values are martingales.

Otherwise, *admissible e-values need not be martingales*.
But, they must "combine" point-null martingales!

(H_0) $X_1, X_2, \ldots$ are iid Ber(p) for some $p \in [0,1]$.

(H_1) $X_1, X_2, \ldots$ are Markov(P) for some transition $P \in [0,1]^4$.

$$E_t = \frac{\dfrac{\Gamma(n_{0|0}+\frac{1}{2})\Gamma(n_{0|1}+\frac{1}{2})\Gamma(n_{1|0}+\frac{1}{2})\Gamma(n_{1|1}+\frac{1}{2})}{2\Gamma(\frac{1}{2})^4\Gamma(n_{0|0}+n_{1|0}+1)\Gamma(n_{0|1}+n_{1|1}+1)}}{\left(\frac{n_1}{t}\right)^{n_1}\left(\frac{n_0}{t}\right)^{n_0}}$$

Jeffrey's prior over alternatives

MLE over nulls

Admissible anytime-valid sequential inference must rely on nonnegative martingales

Ramdas, Ruf, Larsson, Koolen
(arXiv:2009.03167)

P4. Multiple testing under arbitrary dependence

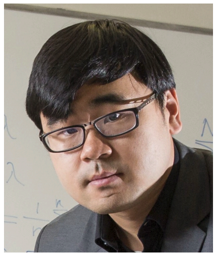
The e-BH procedure: Given $E_1, \ldots, E_K$ for K hypotheses, define

$$k^* := \max \left\{ k : E_{[k]} \geq \frac{K}{k\alpha} \right\}.$$

Reject the k^* hypotheses with largest e-values.

Theorem: The e-BH procedure controls the FDR at level α under arbitrary dependence between the e-values.

False discovery rate control with e-values



Wang, Ramdas (arXiv:2009.02824)

SAVI (Safe, Anytime Valid Inference) EURANDOM workshop (Eindhoven)

Coorganized with
Peter Grunwald

May 25-29, ~~2020~~ 2021?