Rong Jiang[1] and Keming Yu[2] contribution to the Discussion of "Estimating means of bounded random variables by betting" by Ian Waudby-Smith and Aaditya Ramdas

1 Shanghai Polytechnic University, People's Republic of China

2 Brunel University, London UB83PH, UK

We want to congratulate the authors on estimating means of bounded random variables in both with- and without-replacement settings. The authors constructed confidence intervals and time-uniform confidence sequences for the mean of a bounded random variable using test supermartingale technique. It deepens our understanding of the confidence sequence. Confidence sequence is one particular tool in sequential design that facilitates anytime-valid inference. In particular, confidence sequence is a sequence of confidence intervals that is valid at data-dependent stopping times. We offer three comments.

First, the $C_t^{PrPl-EB}$ in Theorem 2 involves $(\lambda_t)_{t=1}^{\infty}$. The authors recommend the predictable plug-in $(\lambda_t^{PrPl-EB})_{t=1}^{\infty}$ given by

$$\lambda_t^{PrPl-EB} = \sqrt{\frac{2\log(2/\alpha)}{\hat{\sigma}_{t-1}^2 t \log(1+t)}} \wedge c, \hat{\sigma}_t^2 = \frac{1/4 + \sum_{i=1}^t (X_i - \hat{\mu}_i)^2}{t+1}, \hat{\mu}_t = \frac{1/2 + \sum_{i=1}^t X_i}{t+1}.$$

Whether so many parameter estimators $\lambda_t^{PrPl-EB}$ will lead to the superposition of errors and the failure of the method. The Hoeffding process $(M_t^H(m))_{t=0}^{\infty}$ in equation (8) only need one λ. In particular, when $t < 100$, is $C_t^{PrPl-EB}$ still correct? See Figure 2, we can see the results are bed when $t < 100$. Thus, whether Theorem 2 needs to add restrictions on t. Moreover, whether $C_t^{PrPl-EB}$ is sensitive to c in $\lambda_t^{PrPl-EB}$, and if so, how to select c, although 1/2 or 3/4 is recommended

Second, the author mentioned their test supermartingle can also be inverted to get a confidence sequence for any quantile. How about mode (Chernoff, 1964), expectile (Newey and Powell, 1987) and extremile (Daouia et al., 2019)?

Third, the authors consider arbitrary distribution but bounded distribution which implies all moments exist, and require a Chernoff-type assumption on the distribution resulting in $O(\sqrt{\log t/t})$ shrinkage rates for the confidence sequences. Recently, Wang and Ramdas (2023) show that employing Catoni's estimator improves the rate to $O(\sqrt{\log\log 2t/t})$ under weaker assumptions on the distribution ($(1+\delta)$-th moment bound). We wonder if the methods and results can be generalized to unbounded observations, since the σ^2-bounded-variance assumption (Wang and Ramdas, 2023) is more realistic and easier to verify.

Reference:

[1] Chernoff, H. (1964). Estimation of the mode. Annals of Statistical Mathematics, 16: 31-41.

[2] Daouia, A., Gijbels, I., Stupfler, G., 2019. Extremiles: A new perspective on asymmetric least squares. Journal of the American Statistical Association 114, 1366-1381.

[3] Newey, W. and Powell, J. (1987). Asymmetric least squares estimation and testing. Econometrica, 55: 819-847.

[4] Wang, H. and Ramdas, A. (2023). Catoni-style confidence sequences for heavy-tailed mean estimation. arXiv:2202.01250v4.