

Robust Mean Estimation with Median-of-Means: Upper and Lower Bounds in Adversarial Settings

XABIER DE JUAN¹, SANTIAGO MAZUELAS^{1,2}

¹Basque Center of Applied Mathematics (BCAM), ²IKERBASQUE-Basque Foundation for Science

xdejuan@bcamath.org

Resumen. The Median-of-Means (MoM) is a one-dimensional robust estimator widely used in machine learning and known to be (minimax) optimal when samples are i.i.d. In more challenging scenarios, samples may be contaminated by an adversary that can inspect and modify the data. Previous work has established the suitability of MoM in some contaminated settings, but its (minimax) optimality and limitations under adversarial contamination remain unclear beyond the Gaussian case. In this talk, we present upper and lower bounds on the error of MoM under adversarial contamination for several classes of distributions. In particular, we show that MoM is (minimax) optimal in the class of distributions with finite variance, as well as in the class of distributions with infinite variance but finite absolute $(1 + r)$ -th moment. We further provide lower bounds that match the order of our upper bounds and demonstrate that MoM is sub-optimal for light-tailed distributions. This talk is based on a paper presented in NeurIPS 2025 by the same authors.

Palabras clave: estimación de la media; contaminación adversaria; estadística robusta; desigualdades de concentración

Referencias

- [1] G. Lugosi, S. Mendelson (2021). Robust multivariate mean estimation: The optimality of trimmed mean. *The Annals of Statistics*, 49(1), 393–410.
- [2] S. Bhatt, G. Fang, P. Li, G. Samorodnitsky (2022). Minimax M-estimation under Adversarial Contamination, *Proceedings of the 39th International Conference on Machine Learning*, Proceedings of Machine Learning Research, vol. 162, PMLR, 1906–1924.
- [3] P. Laforgue, G. Staerman, S. Cléménçon (2021). Generalization Bounds in the Presence of Outliers: a Median-of-Means Study, *Proceedings of the 38th International Conference on Machine Learning*, Proceedings of Machine Learning Research, vol. 139, PMLR, 5937–5947.

Indicar la preferencia (subrayar la opción elegida): póster o charla.

Indicar la preferencia (subrayar la opción elegida): Lunes/Martes o Jueves/Viernes.