Semiparametric robust mean estimations based on the orderliness of quantile averages

Tuban Lee

This manuscript was compiled on June 10, 2023

semiparametric | mean-median-mode inequality | asymptotic | unimodal | Hodges–Lehmann estimator

Furthermore, for weighted averages, separating the breakdown point into upper and lower parts is necessary.

Definition .1 (Upper/lower breakdown point). The upper breakdown point is the breakdown point generalized in Davies and Gather (2005)'s paper (**?**). The finite-sample upper breakdown point is the finite sample breakdown point defined by Donoho and Huber (1983) (1) and also detailed in (**?**). The (finite-sample) lower breakdown point is replacing the infinity symbol in these definitions with negative infinity.

Classifying Distributions by the Signs of Derivatives

Let $\mathcal{P}_{\mathbb{R}}$ denote the set of all continuous distributions over $\mathbb{R}$ and $\mathcal{P}_{\mathbb{X}}$ denote the set of all discrete distributions over a countable set $\mathbb{X}$. The default of this article will be on the class of continuous distributions, $\mathcal{P}_{\mathbb{R}}$. However, it's worth noting that most discussions and results can be extended to encompass the discrete case, $\mathcal{P}_{\mathbb{X}}$, unless explicitly specified otherwise. Besides fully and smoothly parameterizing them by a Euclidean parameter or merely assuming regularity conditions, there exist additional methods for classifying distributions based on their characteristics, such as their skewness, peakedness, modality, and supported interval. In 1956, Stein initiated the study of estimating parameters in the presence of an infinite-dimensional nuisance shape parameter (2) and proposed a necessary condition for this type of problem, a contribution later explicitly recognized as initiating the field of semiparametric statistics (3). In 1982, Bickel simplified Stein's general heuristic necessary condition (2), derived sufficient conditions, and used them in formulating adaptive estimates (3). A notable example discussed in these groundbreaking works was the adaptive estimation of the center of symmetry for an unknown symmetric distribution, which is a semiparametric model. In 1993, Bickel, Klaassen, Ritov, and Wellner published an influential semiparametrics textbook (4), which categorized most common statistical models as semiparametric models, considering parametric and nonparametric models as two special cases within this classification. Yet, there is another old and commonly encountered class of distributions that receives little attention in semiparametric literature: the unimodal distribution. It is a very unique semiparametric model because its definition is based on the signs of derivatives, i.e., $(f'(x) > 0 \text{ for } x \le M) \wedge (f'(x) < 0 \text{ for } x \ge M)$, where $f(x)$ is the probability density function (pdf) of a random variable X, M is the mode. Let $\mathcal{P}_U$ denote the set of all unimodal distributions. There was a widespread misbelief that the median of an arbitrary unimodal distribution always lies between its mean and mode until Runnenburg (1978) and van Zwet (1979) (5, 6) endeavored to determine sufficient conditions for the mean-median-mode inequality to hold, thereby implying the possibility of its violation. The class of unimodal distributions that satisfy the mean-median-mode inequality constitutes a subclass of $\mathcal{P}_U$, denoted by $\mathcal{P}_{MMM} \subsetneq \mathcal{P}_U$.

Data Availability. Data for Figure **??** are given in SI Dataset S1. All codes have been deposited in GitHub.

ACKNOWLEDGMENTS. I sincerely acknowledge the insightful comments from the editor which considerably elevated the lucidity and merit of this paper.

1. DL Donoho, PJ Huber, The notion of breakdown point. *A festschrift for Erich L. Lehmann* **157184** (1983).
2. CM Stein, Efficient nonparametric testing and estimation in *Proceedings of the third Berkeley symposium on mathematical statistics and probability.* Vol. 1, pp. 187–195 (1956).
3. PJ Bickel, On adaptive estimation. *The Annals Stat.* **10**, 647–671 (1982).
4. P Bickel, CA Klaassen, Y Ritov, JA Wellner, *Efficient and adaptive estimation for semiparametric models.* (Springer) Vol. 4, (1993).
5. JT Runnenburg, Mean, median, mode. *Stat. Neerlandica* **32**, 73–79 (1978).
6. Wv Zwet, Mean, median, mode ii. *Stat. Neerlandica* **33**, 1–5 (1979).

T.L. designed research, performed research, analyzed data, and wrote the paper.

The author declares no competing interest.

[1] To whom correspondence should be addressed. E-mail: tl@biomathematics.org