Semiparametric robust mean estimations based on the orderliness of quantile averages

Tuban Lee

This manuscript was compiled on June 8, 2023

As one of the most fundamental problems in statistics, robust location estimation has many prominent solutions, such as the symmetric trimmed mean, symmetric Winsorized mean, Hodges–Lehmann estimator, Huber M-estimator, and median of means. Recent studies suggest that their biases concerning the mean can be quite different in asymmetric distributions, but the underlying mechanisms largely remain unclear. This study exploited a semiparametric method to classify distributions by the asymptotic orderliness of location estimates with varying breakdown points, showing their interrelations and connections to parametric distributions. Further deductions explain why the Winsorized mean typically has smaller biases compared to the trimmed mean; two sequences of semiparametric robust mean estimators emerge. Building on the γ-U-orderliness, the superiority of the median Hodges–Lehmann mean is discussed.

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

In 1823, Gauss (1) proved that for any unimodal distribution, $|m-\mu| \leq \sqrt{\frac{3}{4}}\omega$ and $\sigma \leq \omega \leq 2\sigma$, where μ is the population mean, m is the population median, ω is the root mean square deviation from the mode, and σ is the population standard deviation. This pioneering work revealed that, the potential bias of the median with respect to the mean is bounded in units of a scale parameter for unimodal distributions.

1. CF Gauss, *Theoria combinationis observationum erroribus minimis obnoxiae.* (Henricus Dieterich), (1823).