Robust estimations of moments for unimodal distributions

Tuban Lee

This manuscript was compiled on June 10, 2023

A. Invariant Moments. All popular robust location estimators, such as the symmetric trimmed mean, symmetric Winsorized mean, Hodges-Lehmann estimator, Huber M-estimator, and median of means, are symmetric. As shown previously, a γ-weighted Hodges-Lehmann mean ($\mathrm{WHLM}_{k,\epsilon,\gamma}$) can achieve consistency for the population mean in any γ-symmetric distribution with a finite mean. However, it falls considerably short of consistently handling other parametric distributions that are not γ-symmetric. Shifting from semiparametrics to parametrics, consider a robust estimator with a non-sample-dependent breakdown point (defined in Subsection **??**) which is consistent simultaneously for both a semiparametric distribution and a parametric distribution that does not belong to that semiparametric distribution, it is named with the prefix 'invariant' followed by the name of the population parameter it is consistent with. Here, the recombined I-statistic is defined as

$$\mathrm{RI}_{d,\mathbf{k}_1,\mathbf{k}_2,k_1,k_2,\epsilon_1,\epsilon_2,\gamma_1,\gamma_2,n,LU_1,LU_2} := \lim_{c\to\infty}\left(\frac{\left(LU_{1\mathbf{k}_1,k_1,\epsilon_1,\gamma_1,n}+c\right)^{d+1}}{\left(LU_{2\mathbf{k}_2,k_2,\epsilon_2,\gamma_2,n}+c\right)^{d}}-c\right),$$

where d is the key factor for bias correction, $LU_{\mathbf{k},k,\epsilon,\gamma,n}$ is the LU-statistic, $\mathbf{k}$ is the degree of the U-statistic, k is the degree of the LL-statistic, ϵ is the upper asymptotic breakdown point of the LU-statistic. It is assumed in this series that in the subscript of an estimator, if $\mathbf{k}$, k and γ are omitted, $\mathbf{k}=1$, $k=1$, $\gamma=1$ are assumed, if just one γ is indicated, $\gamma_1=\gamma_2$, if n is omitted, only the asymptotic behavior is considered, in the absence of subscripts, no assumptions are made. The subsequent theorem shows the significance of a recombined I-statistic.

Theorem A.1. *Define the recombined mean as* $rm_{d,k_1,k_2,\epsilon_1,\epsilon_2,\gamma_1,\gamma_2,n,WL_1,WL_2} := RI_{d,\mathbf{k}_1=1,\mathbf{k}_2=1,k_1,k_2,\epsilon_1,\epsilon_2,\gamma_1,\gamma_2,n,LU_1=WL_1,LU_2=WL_2}$. *Assuming finite second moments,* $rm_{d=\frac{\mu-WL_{1k_1,\epsilon_1,\gamma_1}}{WL_{1k_1,\epsilon_1,\gamma_1}-WL_{2k_2,\epsilon_2,\gamma_2}},k_1,k_2,\epsilon_1,\epsilon_2,\gamma_1,\gamma_2,WL_1,WL_2}$ *is a consistent mean estimator for a location-scale distribution, where* μ, $WL_{1k_1,\epsilon_1,\gamma_1}$, *and* $WL_{2k_2,\epsilon_2,\gamma_2}$ *are different location parameters from that location-scale distribution. If* $\gamma_1=\gamma_2$, $WL = WLM$, rm *is also consistent for any* γ*-symmetric distributions.*