Near-consistent robust estimations of moments for unimodal distributions

Tuban Lee

This manuscript was compiled on June 7, 2023

Descriptive statistics for parametric models currently rely heavily on the accuracy of distributional assumptions. Here, leveraging the structures of parametric distributions and their central moment kernel distributions, a class of estimators, consistent simultanously for both a semiparametric distribution and a distinct parametric distribution, is proposed. These efficient estimators are robust to both gross errors and departures from parametric assumptions, making them ideal for estimating the mean and central moments of common unimodal distributions. This article also illuminates the understanding of the common nature of probability distributions and the measures of them.

A. Robust Estimations of the Central Moments. In 1979, Bickel and Lehmann (1), in their final paper of the landmark series *Descriptive Statistics for Nonparametric Models*, generalized a class of estimators called measures of spread, which "do not require the assumption of symmetry." From this, a popular efficient scale estimator, the Rousseeuw-Croux scale estimator (2), was derived in 1993, but the importance of tackling the symmetry assumption has been greatly underestimated. While they had already considered one version of the trimmed standard deviation, which is a measure of dispersion, in the third paper of that series (3); in the final section of that paper (1), they explored another two versions of the trimmed standard deviation based on pairwise differences, one is modified here for comparison,

$$\left[\binom{n}{2}(1-\epsilon-\gamma\epsilon)\right]^{-\frac{1}{2}}\left[\sum_{i=\binom{n}{2}\gamma\epsilon}^{\binom{n}{2}(1-\epsilon)}\left(X-X'\right)_i^2\right]^{\frac{1}{2}}, \qquad [1]$$

where $(X-X')_1 \leq \ldots \leq (X-X')_{\binom{n}{2}}$ are the order statistics of the pseudo-sample, $X_i - X_j$, $i<j$. Let $\Delta = X_i - X_j$. They showed that, when $\epsilon = 0$, [1] is $\sqrt{2}$ times the standard deviation. The paper ended with, "We do not know a fortiori which of the measures is preferable and leave these interesting questions open."

To address their open question, the nomenclature used in this paper is introduced as follows:

Nomenclature. Given a robust estimator, $\hat{\theta}$, which has an adjustable breakdown point, ϵ, that can approach zero asymptotically, the name of $\hat{\theta}$ comprises two parts: the first part denotes the type of estimator, and the second part represents the population parameter θ, such that $\hat{\theta} \to \theta$ as $\epsilon \to 0$.

Theorem A.1.

Proof. □

1. PJ Bickel, EL Lehmann, Descriptive statistics for nonparametric models iv. spread in *Selected Works of EL Lehmann.* (Springer), pp. 519–526 (2012).
2. PJ Rousseeuw, C Croux, Alternatives to the median absolute deviation. *J. Am. Stat. association* **88**, 1273–1283 (1993).
3. PJ Bickel, EL Lehmann, Descriptive statistics for nonparametric models. iii. dispersion in *Selected works of EL Lehmann.* (Springer), pp. 499–518 (2012).