Descriptive statistics for semiparametric models. ii. moments

Tuban Lee

This manuscript was compiled on June 10, 2023

A. Robust Estimations of the Central Moments. In 1976, Bickel and Lehmann (1), in their third paper of the landmark series *Descriptive Statistics for Nonparametric Models*, generalized nearly all robust scale estimators of that time as measures of the dispersion of a symmetric distribution around its center of symmetry. In 1979, the same series, they (2) proposed a class of estimators referred to as measures of spread, which consider the pairwise differences of a random variable, irrespective of its symmetry, throughout its distribution, rather than focusing on dispersion relative to a fixed point. While they had already considered one version of the trimmed standard deviation, which is essentially a trimmed second raw moment, in the third paper of that series (1); in the final section of the fourth paper (2), they explored another two versions of the trimmed standard deviation based on symmetric differences and pairwise differences, the latter 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 pairwise differences, $X_{\mathbf{i}}-X_{\mathbf{j}}$, $\mathbf{i}<\mathbf{j}$, provided that $\binom{n}{2}\gamma\epsilon \in \mathbb{N}$ and $\binom{n}{2}(1-\epsilon) \in \mathbb{N}$. They showed that, when $\epsilon=0$, the result obtained using [1] is equal to $\sqrt{2}$ times the sample standard deviation. The paper ended with, "We do not know a fortiori which of the measures is preferable and leave these interesting questions open."

Two examples of the impacts of this series are as follows. Oja (1981, 1983) (3, 4) provided a more comprehensive and generalized examination of these concepts, and integrated the measures of location, dispersion, and spread as proposed by Bickel and Lehmann (1, 2, 5), along with van Zwet's convex transformation order of skewness and kurtosis (1964) (6) for univariate and multivariate distributions, resulting a greater degree of generality and a broader perspective on these statistical constructs. Rousseeuw and Croux proposed a popular efficient scale estimator based on separate medians of pairwise differences taken over $\mathbf{i}$ and $\mathbf{j}$ (7) in 1993. However the importance of tackling the symmetry assumption has been greatly underestimated, as will be discussed later.

To address their open question (2), 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 $\epsilon \to 0$. The abbreviation of the estimator combines the initial letters of the first part and the second part. If the estimator is symmetric, the upper asymptotic breakdown point, ϵ ($\epsilon_{U_{\mathbf{k}}}$ for LU-statistic), is indicated in the subscript of the abbreviation of the estimator, with the exception of the median. For an asymmetric estimator based on quantile average, the associated γ follows ϵ.

In DSSM I, it was shown that the bias of a robust estimator with an adjustable breakdown point is often monotonic with respect to the breakdown point in a semiparametric distribution. Naturally, the estimator's name should reflect the population parameter that it approaches as $\epsilon \to 0$. If multiplying all pseudo-samples by a factor of $\frac{1}{\sqrt{2}}$, then [1] is the trimmed standard deviation adhering to this nomenclature, since $\psi_2(x_1,x_2)=\frac{1}{2}(x_1-x_2)^2$ is the kernel function of the unbiased estimation of the second central moment by using U-statistic (8). This definition should be preferable,

Theorem A.1.

Proof. □

1. PJ Bickel, EL Lehmann, Descriptive statistics for nonparametric models. iii. dispersion in *Selected works of EL Lehmann.* (Springer), pp. 499–518 (2012).
2. PJ Bickel, EL Lehmann, Descriptive statistics for nonparametric models iv. spread in *Selected Works of EL Lehmann.* (Springer), pp. 519–526 (2012).
3. H Oja, On location, scale, skewness and kurtosis of univariate distributions. *Scand. J. statistics* pp. 154–168 (1981).
4. H Oja, Descriptive statistics for multivariate distributions. *Stat. & Probab. Lett.* **1**, 327–332 (1983).
5. PJ Bickel, EL Lehmann, Descriptive statistics for nonparametric models ii. location in *selected works of EL Lehmann.* (Springer), pp. 473–497 (2012).
6. W van Zwet, Convex transformations: A new approach to skewness and kurtosis in *Selected Works of Willem van Zwet.* (Springer), pp. 3–11 (2012).
7. PJ Rousseeuw, C Croux, Alternatives to the median absolute deviation. *J. Am. Stat. association* **88**, 1273–1283 (1993).
8. PM Heffernan, Unbiased estimation of central moments by using u-statistics. *J. Royal Stat. Soc. Ser. B (Statistical Methodol.* **59**, 861–863 (1997).

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