PNAS
www.pnas.org

Supporting Information for

Semiparametric robust mean estimations based on the orderliness of quantile averages

Tuban Lee.
E-mail: tl@biomathematics.org

This PDF file includes:

Supporting text
Legend for Dataset S1

Other supporting materials for this manuscript include the following:

Dataset S1

Supporting Information Text

Then, the orderliness for parametric distributions will be discussed. For simplicity, $0 \leq \epsilon \leq \frac{1}{2}$ is assumed in the following proofs unless otherwise specified.

Theorem 0.1. *The Weibull distribution is ordered if the shape parameter* $\alpha \leqslant \frac{1}{1-\ln(2)} \approx 3.259$.

Proof. The pdf of the Weibull distribution is $f(x) = \frac{\alpha e^{-\left(\frac{x}{\lambda}\right)^{\alpha}}\left(\frac{x}{\lambda}\right)^{\alpha-1}}{\lambda}$, $x \geq 0$, the quantile function is $F^{-1}(p) = \lambda(-\ln(1-p))^{1/\alpha}$, $1 \geq p \geq 0$, $\alpha > 0, \lambda > 0$. Then, the standardized bias of a symmetric quantile average with a breakdown point ϵ, is

$$\mathrm{B}_{\mathrm{SQA}}(\epsilon,\alpha) = \frac{\mathrm{SQA}_{\epsilon} - \mu}{\sigma} = \frac{\frac{1}{2}\left(\lambda(-\ln(1-\epsilon))^{1/\alpha} + \lambda(-\ln(\epsilon))^{1/\alpha}\right) - \lambda\Gamma\left(1+\frac{1}{\alpha}\right)}{\sqrt{\lambda^2\left(\Gamma\left(1+\frac{2}{\alpha}\right) - \Gamma\left(1+\frac{1}{\alpha}\right)^2\right)}}.$$

$\frac{\partial \mathrm{B}_{\mathrm{SQA}}}{\partial \epsilon} = \frac{\frac{(-\ln(1-\epsilon))^{\frac{1}{\alpha}-1}}{\alpha-\alpha\epsilon} + \frac{(-\ln(\epsilon))^{1/\alpha}}{\alpha\epsilon\ln(\epsilon)}}{2\sqrt{\Gamma\left(\frac{\alpha+2}{\alpha}\right)-\Gamma\left(1+\frac{1}{\alpha}\right)^2}}$. Let $g(\epsilon,\alpha) = \frac{(-\ln(1-\epsilon))^{\frac{1}{\alpha}-1}}{\alpha-\alpha\epsilon} + \frac{(-\ln(\epsilon))^{1/\alpha}}{\alpha\epsilon\ln(\epsilon)} = -\left(-\ln(1-\epsilon)\right)^{\frac{1}{\alpha}}\left((1-\epsilon)\left(\ln(1-\epsilon)\right)\right)^{-1} + \left(-\ln(\epsilon)\right)^{\frac{1}{\alpha}}\left(\epsilon\ln(\epsilon)\right)^{-1}$. Arranging the equation $g(\epsilon,\alpha) = 0$, it can be shown that $\frac{\epsilon}{(1-\epsilon)} = \left(\frac{\ln(\epsilon)}{\ln(1-\epsilon)}\right)^{\frac{1}{\alpha}-1}$. Let $L(\epsilon) = \frac{\epsilon}{(1-\epsilon)}$, $R(\epsilon,\alpha) = \left(\frac{\ln(\epsilon)}{\ln(1-\epsilon)}\right)^{\frac{1}{\alpha}-1}$, $LmR(\epsilon,\alpha) = L(\epsilon,\alpha) - R(\epsilon,\alpha)$, then $\frac{\partial LmR}{\partial \alpha} = \frac{\ln\left(\frac{\ln(\epsilon)}{\ln(1-\epsilon)}\right)\left(\frac{\ln(\epsilon)}{\ln(1-\epsilon)}\right)^{\frac{1}{\alpha}-1}}{\alpha^2}$. For $0 \leq \epsilon \leq \frac{1}{2}$, $\frac{\partial LmR}{\partial \alpha} \geq 0$, so $LmR(\epsilon,\alpha)$ is monotonic with respect to α. When $\alpha = \frac{1}{1-\ln(2)}$, $g(\epsilon) = -\frac{1}{\epsilon(-\ln(\epsilon))^{\ln(2)}} + \frac{1}{(1-\epsilon)(-\ln(1-\epsilon))^{\ln(2)}}$ Let $h(\epsilon) = \epsilon\left(-\ln(\epsilon)\right)^{\ln(2)})$, $h'(\epsilon) = \frac{(-\ln(\epsilon))^{\ln(2)}\ln(2\epsilon)}{\ln(\epsilon)}$, for $0 \leq \epsilon \leq e^{-ln(2)} = \frac{1}{2}$, $h'(\epsilon) \geq 0$. As a result, $h(\epsilon)$ is monotonic increasing, $-h(1-\epsilon)$ is monotonic increasing, $h(\epsilon) - h(1-\epsilon)$ is also monotonic increasing. So, if $0 \leq \epsilon \leq \frac{1}{2}$, $h(\epsilon) - h(1-\epsilon) \leq h\left(\frac{1}{2}\right) - h\left(1-\frac{1}{2}\right) = 0$, $g(\epsilon,\alpha) \leq 0$. So, $\frac{\partial \mathrm{B}_{\mathrm{SQA}}}{\partial \epsilon} \leq 0$, $\mathrm{B}_{\mathrm{SQA}}(\epsilon,\alpha)$ is monotonic decreasing in ϵ when $\alpha \leqslant \frac{1}{1-\ln(2)}$. The assertion follows from Theorem **??**. □

SI Dataset S1 (dataset_one.xlsx)

Raw data of asymptotic biases of all estimators shown in Figure 1 in the Main Text and the standard errors of these estimators for the generalized Gaussian distribution.

References