Jensen's inequality for the Tukey median

Robert Kwiecien[1] *, Ursula Gather[2] **

[1,2] *Department of Statistics, University of Dortmund, 44221 Dortmund, Germany*

Abstract

Jensen's inequality states for a random variable X with values in $\mathbb{R}^d$ and existing expectation and for any convex function $f : \mathbb{R}^d \longrightarrow \mathbb{R}$, that $f(E(X)) \leq E(f(X))$. We prove an analogous inequality, where the expectation operator is replaced by the halfspace-median-operator (or Tukey-median-operator).

Key words: Robustness; Multivariate median; Tukey depth; Jensen's inequality

1 Introduction

Jensen's inequality is one of the most frequently used inequalities in statistics. It gives a relationship between the expectation of a random variable X and the expectation of $f(X)$, where f is a convex function. More explicitly:
For $S \subset \mathbb{R}^d$ convex, X a random variable with values in $\mathbb{R}^d$, distribution P with $supp(P) \subset S$, and existing expectation $E(X)$, Jensen's inequality reads

$$f(E(X)) \leq E(f(X)) \text{ or in terms of distributions } f(E(P)) \leq E(P^f).$$

Proving an analogue of Jensen's inequality for a median instead of the expectation causes two problems. Firstly, the expectation of a random variable (provided it exists) is a unique point in $\mathbb{R}^d$, whereas the median of a random variable may consist of a whole set. Secondly, there are several generalisations of the median in the multivariate case. Tukey (1975) introduced the halfspace depth, also known as Tukey depth. Donoho and Gasko (1992) defined the multivariate Tukey- or halfspace median using the Tukey depth. For empirical distributions Oja (1983)

* E-mail address: kwiecien@statistik.uni-dortmund.de

** E-mail address: gather@statistik.uni-dortmund.de

22 February 2007

and Liu (1990) introduced the multivariate Oja simplex median and the simplicial median, respectively. Both definitions require additional assumptions. There are many more possibilities. We will focus on the Tukey median, which coincides in the one-dimensional case with the classical median.
An analogue to Jensen's inequality for the median in the one-dimensional case was given by Merkle (2005). He considers the classical set of medians and suggests a multivariate version of a median, which is not affine linear equivariant. This leaves open a generalisation of Jensen's inequality for medians to the multivariate case. We will prove a stronger statement here and also generalize it to the multivariate case.
This paper is organized as follows. Section 2 contains some notation, two equivalent definitions of the Tukey median, and some preliminaries. In Section 3 we prove Jensen's inequality for the Tukey median and give some examples. Finally, we present some open problems in Section 4.

2 Notation and preliminaries

We start with the definition of the set of halfspace medians or Tukey medians.

Let X denote a random variable with values in $\mathbb{R}^d$ and distribution P on the Borel-σ-algebra $\mathcal{B}^d$.

Definition 2.1 *Let $\mathcal{H}^d$ denote the set of all closed halfspaces of $\mathbb{R}^d$. The halfspace depth (HD) of a point $x \in \mathbb{R}^d$ with respect to P is defined by*

$$HD(x; P) := \inf\{P(H) \; ; \; x \in H \; ; \; H \in \mathcal{H}^d\}. \tag{1}$$

The set of points at least as deep as $k \in (0, 1)$ w.r.t. P is defined by

$$TukMed_k(X) := TukMed_k(P) := \{x \in \mathbb{R}^d \; ; \; HD(x; P) \geq k\}. \tag{2}$$

Let $k' := \sup_{x \in \mathbb{R}^d} HD(x; P)$. The set of Tukey medians is defined by

$$TukMed(X) := TukMed(P) := TukMed_{k'}(P), \tag{3}$$

which can be interpreted as the set of "deepest points".

The Tukey median in the finite sample case, i.e. for an empirical distribution P_n, has been extensively investigated under computational aspects using different characterisations. For instance Donoho and Gasko (1992) use a representation of Tukey medians in the finite sample case by an intersection of certain sets. Merkle

(2005) defines a set of medians as an intersection of multivariate closed intervals. The following definition of a set of medians is equivalent to Definition 2.1; this will be shown in Section 3.

Definition 2.2 *Let*

$$\mathcal{B}^d_{cc} := \{I \in \mathcal{B}^d \; ; \; I \text{ closed and convex}\}. \tag{4}$$

With the notation of Definition 2.1 , let

$$\mathcal{I}(\widetilde{\delta}, P) := \{I \in \mathcal{B}^d_{cc}, \text{ such that } \forall \; J \supsetneq I, J \in \mathcal{B}^d_{cc} \Rightarrow P(J) > \widetilde{\delta}\}. \tag{5}$$

for some $\widetilde{\delta} \in [1/2, 1)$,

$$\delta_{min} := \delta_{min}(X) := \delta_{min}(P) := \inf\{\widetilde{\delta} \in [1/2, 1) \; ; \bigcap_{I \in \mathcal{I}(\widetilde{\delta}, P)} I \neq \emptyset\} \tag{6}$$

$$Med_{\delta'}(X) := Med_{\delta'}(P) := \bigcap_{I \in \mathcal{I}(\delta', P)} I \; , \tag{7}$$

for some $\delta' \in (0, 1)$, *and the set of medians of* X *or* P, *respectively,*

$$Med(X) := Med(P) := \bigcap_{I \in \mathcal{I}(\delta_{min}, P)} I \; . \tag{8}$$

The gravity point of $Med(P)$, *i.e. the expectation of a random variable, uniformly distributed on* $Med(P)$, *will be denoted by* $GMed(P)$.

Lemma 2.1 *The set of medians remains unchanged if we replace* $\mathcal{I}(\widetilde{\delta}, P)$ *by one of the following subsystems*

$$\widetilde{\mathcal{I}}(\widetilde{\delta}, P) := \{\widetilde{I} \in \mathcal{I}(\widetilde{\delta}, P) \; ; \; P(\widetilde{I}) > \widetilde{\delta}\} \tag{9}$$

$$\mathcal{H}(\widetilde{\delta}, P) := \{H \in \mathcal{I}(\widetilde{\delta}, P) \; ; \; H \text{ is halfspace } \} \tag{10}$$

$$\widetilde{\mathcal{H}}(\widetilde{\delta}, P) := \{\widetilde{H} \in \widetilde{\mathcal{I}}(\widetilde{\delta}, P) \; ; \; H \text{ is halfspace } \}, \tag{11}$$

i.e. for $\delta' \in [\delta_{min}, 1)$ *we have*

$$Med_{\delta'}(P) = \bigcap_{\widetilde{I} \in \widetilde{\mathcal{I}}(\delta', P)} \widetilde{I} = \bigcap_{H \in \mathcal{H}(\delta', P)} H = \bigcap_{\widetilde{H} \in \widetilde{\mathcal{H}}(\delta', P)} \widetilde{H}.$$

Proof. First, we aim at the representation $Med_{\delta'}(P) = \bigcap_{\widetilde{I} \in \widetilde{\mathcal{I}}(\delta', P)} \widetilde{I}$. For each $I \in \mathcal{I}(\delta', P)$ we define the system of sets

$$\mathcal{J}(I, \delta') := \{J \in \mathcal{B}^d_{cc} \; ; \; J \supsetneq I, \text{ such that } P(J) > \delta'\} \subset \widetilde{\mathcal{I}}(\delta', P).$$

We show that $I = \bigcap_{J\in\mathcal{J}(I,\delta')} J =: I'$ for each $I \in \mathcal{I}(\delta', P)$. Obviously, $I \subset \bigcap_{J\in\mathcal{J}(I,\delta')} J$ because by definition $I \subset J$ whenever $J \in \mathcal{J}(I, \delta')$. Clearly, if I is an element of $\mathcal{I}(\delta', P)$, I' is as well. Moreover, if $I \neq I'$, there is no set in $\mathcal{B}^d_{cc}$ "in between them", i.e.

$$\nexists K \in \mathcal{B}^d_{cc} \text{ with } I \subsetneq K \subsetneq I'. \tag{12}$$

Assume that there exists such a set K. Then K must be an element of $\mathcal{J}(I, \delta')$, and therefore $K = I' = \bigcap_{J\in\mathcal{J}(I,\delta')} J$, which implies (12). On the other hand, for each pair of sets $I, I' \in \mathcal{B}^d_{cc}$, $I \subsetneq I'$ there exists in either case a set $K \in \mathcal{B}^d_{cc}$ satisfying $I \subsetneq K \subsetneq I'$. This contradicts (12) and hence the set I cannot be a strict subset of I'. Since $I = \bigcap_{J\in\mathcal{J}(I,\delta')} J$ for each $I \in \mathcal{I}(\delta', P)$, we can represent each element of $\mathcal{I}(\delta', P)$ by intersecting elements of $\widetilde{\mathcal{I}}(\delta', P)$, which yields the first part of the proof.
We prove $Med_{\delta'}(P) = \bigcap_{H\in\mathcal{H}(\delta',P)} H$ in a similar way. For each $I \in \mathcal{I}(\delta', P)$ there exists a system of half spaces included in $\mathcal{I}(\delta', P)$ generating I by intersections. Hence, the intersection of all half spaces in $\mathcal{I}(\delta', P)$ or all elements of $\mathcal{H}(\delta', P)$, respectively, generates the set $Med_{\delta'}(P)$.
A proof of the last statement can be given in an analogous way. □

It is not immediately obvious that the medians defined in Definition 2.2 always exist. The following lemma shows that $Med(P)$ is never empty.

Lemma 2.2 *The set defined by (6) is closed from below, i.e.*

$$\begin{aligned}\delta_{min} &= \inf\{\widetilde{\delta} \in [1/2, 1) \,;\; \bigcap_{I\in\mathcal{I}(\widetilde{\delta},P)} I \neq \emptyset\} \\ &= \min\{\widetilde{\delta} \in [1/2, 1) \,;\; \bigcap_{I\in\mathcal{I}(\widetilde{\delta},P)} I \neq \emptyset\}.\end{aligned}$$

Proof. It is easy to check that there exists $\delta' \in [1/2, 1)$ such that $Med_{\delta'}(P) \neq \emptyset$. We can use the representation $Med(P) = \bigcap_{\widetilde{I}\in\widetilde{\mathcal{I}}(\delta_{min},P)} \widetilde{I}$ (Lemma 2.1). Then, for each $\delta' \in (\delta_{min}, 1)$:

$$Med_{\delta'}(P) = \bigcap_{\widetilde{I}\in\widetilde{\mathcal{I}}(\delta',P)} \widetilde{I} \neq \emptyset, \text{ and} \tag{13}$$

$$\bigcap_{\delta'\in(\delta_{min},1)} Med_{\delta'}(P) \neq \emptyset. \tag{14}$$

(13) follows immediately by (5), (6), and (9). Assume (14) does not hold. By the Heine-Borel Theorem one can find $\{\delta_1, ..., \delta_n\} \subset (\delta_{min}, 1)$ satisfying $Med_{\delta_1}(P) \cap ... \cap Med_{\delta_n}(P) = \emptyset$. Since $Med_{\delta_i}(P) \subset Med_{\delta_j}(P)$ whenever $\delta_i \leq \delta_j$, it follows that

$Med_{\min\{\delta_1,\ldots,\delta_n\}}(P) = \emptyset$, which contradicts (13). Hence, (14) holds. By Lemma 2.1 we have $\bigcap_{\delta' \in (\delta_{min},1)} Med_{\delta'}(P) = \bigcap_{\delta' \in [\delta_{min},1)} Med_{\delta'}(P) \neq \emptyset$, which completes the proof. □

Remark 2.1

a) In the one-dimensional case, Definition 2.2 yields the classical set of medians and $\delta_{min} = 1/2$. This follows immediately from Theorem 2.1 in Merkle (2005).

b) $Med(P)$ is a compact and convex set because the intersection of convex sets is convex, the intersection of closed sets is closed, and the intersection of a closed set with at least one compact set is compact.

c) Note that we claim in (6): $\widetilde{\delta} \in [1/2, 1)$. The lower bound of $1/2$ seems to be restrictive. But it is easy to check that if we allow $\widetilde{\delta}$ to be smaller, i.e.

$$\delta_{min}(P) = \inf\{\widetilde{\delta} \in (0,1) \; ; \bigcap_{I \in \mathcal{I}(\widetilde{\delta},P)} I \neq \emptyset\}, \tag{15}$$

we obtain the same set of medians. In other words, let $\delta_{min}(P)$ be defined as in (15) and let $\widetilde{\delta}$ be such that $\delta_{min}(P) \leq \widetilde{\delta} < 1/2$, then $Med_{\widetilde{\delta}}(P) = Med_{1/2}(P)$.

d) One could conjecture that $\delta_{min} = 1/2$ holds in general. But for the uniform distribution on the vertices of a regular pentagon in $\mathbb{R}^2$, we get $\delta_{min} = 3/5$.

Obviously, each $H \in \mathcal{H}(\delta_{min}, P)$ contains $Med(P)$ as a subset. For technical purposes we need the opposite inclusion.

Lemma 2.3 *If H is a half space containing $Med(P)$, i.e. $Med(P) \subset H$, then H is an element of $\mathcal{H}(\delta_{min}, P)$.*

Proof. By Lemma 2.1 $Med(P)$ can be generated by intersecting all half spaces in $\mathcal{H}(\delta_{min}, P)$. Hence, there exists $H' \in \mathcal{H}(\delta_{min}, P)$, which is a subset of H. Therefore, $H \in \mathcal{H}(\delta_{min}, P)$. □

In order to prove Jensen's inequality for the Tukey median we need that a median exists in certain subsets. The following lemma provides a criterion for this.

Lemma 2.4 *Let $K \in \mathcal{B}^d_{cc}$ and let P denote a distribution on $\mathbb{R}^d$ such that $P(K) \geq 1/2$. Then $K \cap Med(P) \neq \emptyset$, i.e. K contains at least one median of P.*

Proof. We first prove the statement for any half space H with $P(H) \geq 1/2$. Assume that $H \cap Med(P) = \emptyset$. Since H and $Med(P)$ are closed and convex, there exists a hyperplane h' separating H and $Med(P)$. Hence, there exists a closed half space H' satisfying $\partial H' = h'$ and $H' \cap H = \emptyset$ where $\partial H'$ denotes the

boundary of H'. Since H' contains $Med(P)$, by Lemma 2.3 H' is an element of $\mathcal{H}(\delta_{min}, P)$. H' is closed and convex, thus there exists a set $J \in \mathcal{B}^d_{cc}$ such that $P(J) > \delta_{min} \geq 1/2$ and $J \cap H = \emptyset$. Therefore,

$$P(H \cup J) = P(H) + P(J) > 1/2 + \delta_{min} \geq 1,$$

which is not possible. Hence,

$$H \cap Med(P) \neq \emptyset. \tag{16}$$

If we only assume that $K \in \mathcal{B}^d_{cc}$ such that $P(K) \geq 1/2$, and that the statement does not hold, there exists a half space H such that $H \cap Med(P) = \emptyset$, $P(H) \geq 1/2$, and $K \subset H$, which contradicts (16) and thus completes the proof. □

It is a natural requirement that the set of medians should not depend on the choice of the coordinate system. The following lemma states the equivariance under affine linear transformations. For technical purposes a generalisation to linear projections into lower dimensional subspaces would be desirable. But this is not possible. We therefore give the following weaker statement.

Lemma 2.5 *Let G denote the gravity-operator, i.e. G maps a compact set onto its gravity point. For each affine linear and bijective function $f : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ we have:*

$$f(Med(P)) = Med(P^f), \tag{17}$$
$$f(GMed(P)) = GMed(P^f), \tag{18}$$

where P^f denotes the distribution of $f(X)$. Moreover, for each $\delta' \geq \delta_{min}(P)$ let $Med_{\delta'}(P)$ be defined as in (13). For each linear function $L : \mathbb{R}^d \longrightarrow \mathbb{R}^{d'}$, we get

$$L(Med_{\delta'}(P)) \subset Med_{\delta'}(P^L), \tag{19}$$

and consequently, if $\delta_{min}(P) = \delta_{min}(P^L)$,

$$L(Med(P)) \subset Med(P^L). \tag{20}$$

Proof. Since f is affine linear and bijective, f maps convex sets onto convex sets. Further, f is also measurable, hence $I \in \mathcal{I}(\delta_{min}, P)$ implies $f(I) \in \mathcal{I}(\delta_{min}, P^f)$. So (17) follows from this one-to-one relation and hence also (18) because an affine linear and bijective function and the G-operator are commutative. In order to prove (19) we define for each linear function $L : \mathbb{R}^d \longrightarrow \mathbb{R}^{d'}$ an equivalence relation $\cong_L$ on the system of sets $\mathcal{I}(\delta', P)$:

$$K_1 \cong_L K_2 :\Leftrightarrow L(K_1) = L(K_2).$$

Consequently,

$$[\mathcal{I}(\delta', P)]_L := \{I_P \in \mathcal{I}(\delta', P) \; ; \; I_P = L^{-1}(L(I_P)) \; \} \tag{21}$$

is a set of representatives of the equivalence classes. Since $L(I_P) \in \mathcal{I}(\delta', P^L)$ whenever $I_P \in \mathcal{I}(\delta', P)$, it follows that

$$\mathcal{I}(\delta', P^L) = \{I = L(I_P) \; ; \; I_P \in [\mathcal{I}(\delta', P)]_L\}. \tag{22}$$

Thus,

$$\begin{array}{rcccl} L(Med_{\delta'}(P))) & = & L(\bigcap_{I_P \in \mathcal{I}(\delta', P)} I_P) & \subset & \bigcap_{I_P \in \mathcal{I}(\delta', P)} L(I_P) \\ \overset{by(21)}{=} \bigcap_{I_P \in [\mathcal{I}(\delta', P)]_L} L(I_P) & \overset{by(22)}{=} & \bigcap_{I \in \mathcal{I}(\delta', P^L)} I & = & Med_{\delta'}(P^L), \end{array}$$

which completes the proof. □

3 The main theorem and examples

With the previous results we can now prove Jensen's inequality for the median.

Theorem 3.1 *Let $f : S \longrightarrow \mathbb{R}$ be a convex function, where $S \subset \mathbb{R}^d$ is convex, and let X be a random variable and P its distribution with $supp(P) \subset S$. Then the following inequality holds:*

$$\inf\{f(Med(P))\} \leq \inf\{Med(P^f)\}. \tag{23}$$

Moreover, if P satisfies $\delta_{min}(P) = 1/2$, e.g. if $\bigcap_{H \in \mathcal{H}^d \, ; \, P(H) > 1/2} H \neq \emptyset$, and if f is continuous, then

$$\sup\{f(Med(P))\} \leq \sup\{Med(P^f)\}. \tag{24}$$

Proof. We start with the proof of (23). Without loss of generality we assume that S is not a subset of any d'-dimensional hyperplane with $d' < d$. If S is a subset of a d'-dimensional hyperplane with $d' < d$, there exists an affine linear change of coordinates such that P can be seen as a distribution on $\mathbb{R}^{d'}$ with convex support $S' \subset \mathbb{R}^{d'}$. Moreover, if f is not continuous, then the boundary ∂S contains all points of discontinuity, cf. Rockafellar (1970), Section 10. Let $\bar{f} : S \longrightarrow \mathbb{R}$ denote the continuous function satisfying $\bar{f}(x) = f(x) \; \forall x \in S \setminus \partial S$. Then we have

$$\begin{array}{l} \inf\{\bar{f}(Med(P))\} = \inf\{f(Med(P))\}, \text{ and} \\ \inf\{(Med(P^{\bar{f}}))\} \leq \inf\{(Med(P^f))\}. \end{array}$$

Hence, without loss of generality we can assume that f is continuous. Let $M \in Med(P^f)$ be arbitrary, and let $I_M \subset \mathbb{R}^d$ be defined by $I_M := f^{-1}((-\infty, M])$. I_M is closed and convex because f is continuous and convex. By Remark 2.1 a) it follows:

$$P(I_M) = P^f((-\infty, M]) \geq 1/2,$$

and therefore, by Lemma 2.4, $I_M \cap Med(P) \neq \emptyset$, i.e. there exists $m \in Med(P)$ that is also an element of I_M. The definition of I_M yields the inequality $f(m) \leq M$. Since M is an arbitrary element of $Med(P^f)$ and due to the compactness of $Med(P^f)$, we have $f(m) \leq \inf\{Med(P^f)\}$ for at least one element $m \in Med(P)$. We thus get (23). Next we prove (24). Let U be the intersection of all hyperplanes U' with $Med(P) \subset U'$, and let $IntMed(P) \subset Med(P)$ be the interior of $Med(P)$ w.r.t. the subspace topology of U. Additionally, for each $m \in IntMed(P)$ let $f'(x) := f(x) - f(m)$. We now show that

$$\sup\{Med(P^f)\} \geq f(m), \text{ i.e. } \sup\{Med(P^{f'})\} \geq 0\,. \tag{25}$$

Let $L : \mathbb{R}^d \longrightarrow \mathbb{R}$ be a linear function satisfying

$$f''(x) := L(x - m) \leq f'(x) \quad \forall\, x \in \mathbb{R}^d, \tag{26}$$

and let m' be such that

$$L(m') = \sup\{L(Med(P))\}. \tag{27}$$

Therefore,

$$\begin{aligned}
\sup\{Med(P^{f'})\} &\overset{by(26)}{\geq} \sup\{Med(P^{f''})\} \\
&\overset{by(20)\ and\ \delta_{min}(P)=1/2}{\geq} \sup\{L(Med(P))\} - \sup\{L(Med(\varepsilon_m))\} \\
&\overset{by(27)}{\geq} \sup\{L(Med(P))\} - \sup\{L(Med(\varepsilon_{m'}))\} \\
&\overset{by(27)}{=} 0\,,
\end{aligned}$$

where the "$-$" in the last two inequalities corresponds to affine movements of sets, and ε_m , $\varepsilon_{m'}$ denote the Dirac measures on m and m', respectively. Since $IntMed(P)$ is dense in $Med(P)$, and since f is continuous, replacing $f(m)$ in (25) by $\sup\{f(Med(P))\}$ completes the proof. □

A similar version for the $GMed$-operator might be desirable. The next theorem gives such a Jensen's inequality in the univariate case.

Example 3.1 *Under the assumptions of Theorem 3.1 with $S \subset \mathbb{R}$ and $f : S \longrightarrow \mathbb{R}$ continuous the following inequality holds:*

$$f(GMed(P)) \leq GMed(P^f). \tag{28}$$

Proof. Theorem 3.1 immediately implies

$$G(f(Med(P))) \leq GMed(P^f). \tag{29}$$

Observing that $GMed(P) = E(Y)$ for a random variable Y, uniformly distributed on $Med(P)$, we get

$$f(GMed(P)) \leq G(f(Med(P))) \tag{30}$$

by the standard Jensen's inequality. Combining (29) and (30) completes the proof.

□

As a simple consequence of Jensen's inequality in the multivariate case we get a kind of Hölder's inequality for medians.

Example 3.2 *Let X, Y denote real valued and non-negative random variables, and let $\delta_{min}((X,Y)) = 1/2$, then the following inequality holds for $0 \leq p \leq 1$:*

$$\inf\{Med(X^p Y^{1-p})\} \leq \inf\{Med(X)\}^p \inf\{Med(Y)\}^{1-p}.$$

Proof. Observe that for each $0 < p < 1$ the function

$$f : \mathbb{R}^2_+ \longrightarrow \mathbb{R}, \quad f(x,y) = -x^p y^{1-p}$$

is convex. By $\delta_{min}((X,Y)) = 1/2$, and by Lemma 2.5 (20) we get $Med((X,Y)) \subset Med(X) \times Med(Y)$, and hence,

$$\inf\{f(Med((X,Y)))\} \geq \inf\{f(Med(X) \times Med(Y))\}. \tag{31}$$

From Theorem 3.1 and by (31) we get the desired result. □

Example 3.3 *As mentioned in Remark 2.1 a), $\delta_{min}(P) = 1/2$ if P is a distribution on $\mathbb{R}$. Hence (23) and (24) hold in the one-dimensional case.*

Example 3.4 *We demonstrate by a counterexample that (24) does not necessarily hold if $\delta_{min}(P) \neq 1/2$. Let P be the uniform distribution on the set $\{(1,0),(0,0),(0,1)\} \subset \mathbb{R}^2$, and let $L : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be the orthogonal projection onto the x-axis, i.e. $L(x,y) = x$. Obviously, L is continuous and convex. It is easy to*

check that $\delta_{min}(P) = 2/3$ *and* $Med(P)$ *is the convex hull of* $\{(1,0),(0,0),(0,1)\}$, *but* $Med(P^L) = 0$. *Therefore, (24) does not hold.* □

Theorem 3.1 states an analogue of Jensen's inequality for medians as defined in Definition 2.2. This analogue holds also for Tukey medians if the set of Tukey medians and the set of medians defined in Definition 2.2 are equal:

Lemma 3.1 *Let* P *be a distribution on* $\mathbb{R}^d$ *and let* $k \in (0,1)$. *Then*

$$TukMed_k(P) = Med_{1-k}(P), \text{ especially } TukMed(P) = Med(P). \tag{32}$$

Proof. First note that in Lemma 3.1 we do not restrict to $1-k \geq 1/2$. As pointed out in Remark 2.1 c), $Med_{1-k}(P) = Med_{1-k\vee 1/2}(P)$, if $Med_{1-k}(P) \neq \emptyset$. In the following we make use of representations of $Med_{1-k}(P)$ as given in Lemma 2.1. We show (32) by proving that the complements of $TukMed_k(P)$ and $Med_{1-k}(P)$ are equal. Let $\mathcal{H}^d$ be defined as in Definition 2.1, let $\theta \in \mathbb{R}^d$, and assume $\theta \notin Med_{1-k}(P)$. From the definition of $Med_{1-k}(P)$ it is clear that there exists $H \in \mathcal{H}^d$ such that $\theta \notin H$ and $P(H) > 1-k$, or $1-P(H) < k$, respectively. Consequently, there exists a sequence of closed halfspaces $H'_n \in \mathcal{H}^d$, $n \in \mathbb{N}$, with the following properties:

$$H'_n \cap H = \emptyset, \text{ and } \theta \in H'_n \ \forall n \in \mathbb{N}, \text{ and}$$
$$\lim_{n\longrightarrow\infty} P(H'_n) = 1 - P(H) < k.$$

Therefore, $HD(\theta; P) < k$, and $\theta \notin TukMed_k(P)$.

Conversely, assume $\theta \notin TukMed_k(P)$, i.e. $\inf\{P(H') \ ; \ \theta \in H' \in \mathcal{H}^d\} < k$; hence there exists a sequence of closed
halfspaces H'_n satisfying

$$\theta \in H'_n \ \forall n \in \mathbb{N}, \text{ and}$$
$$\lim_{n\longrightarrow\infty} P(H'_n) < k \Leftrightarrow 1 - \lim_{n\longrightarrow\infty} P(H'_n) > 1-k,$$

and thus a sequence of closed halfspaces H_n such that

$$H'_n \cap H_n = \emptyset \text{ (and also } \theta \notin H_n), \text{ and}$$
$$\lim_{n\longrightarrow\infty} P(H_n) = 1 - \lim_{n\longrightarrow\infty} P(H'_n) > 1-k.$$

This implies $\theta \notin Med_{1-k}(P)$ which completes the proof. □

4 Discussion

We have given an analogue of the classical Jensen's inequality with the expectation replaced by the median, more precisely, by the set of Tukey medians. For technical reasons we have used an equivalent characterisation of the standard definition of Tukey medians. We have shown Jensen's inequality for the infimum of the set of Tukey medians. In the one-dimensional case Jensen's inequality even holds for the supremum and the center of gravity of the set of Tukey medians if the convex function f is continuous. Since in the one-dimensional case the set of Tukey medians coincides with the set of classical medians, Jensen's inequality also holds for the classical sample median. On the other hand, Jensen's inequality for the supremum of the set of Tukey medians requires at least one additional assumption, where the continuity of f and $\delta_{min}(P) = 1/2$, i.e. $\bigcap_{H \in \mathcal{H}^d \,;\, P(H) > 1/2} H \neq \emptyset$ is such a sufficient condition. It is an open question whether this condition is necessary. Jensen's inequality for the gravity point of the set of Tukey medians also needs additional assumptions besides the continuity of f. It is interesting whether $\delta_{min}(P) = 1/2$ is sufficient here, or not. However, other conditions that imply $\delta_{min}(P) = 1/2$ and which are easy to check also allow to apply our theorem.

Acknowledgement. The support of the Deutsche Forschungsgemeinschaft (SFB 475, "Komplexitätsreduktion in multivariaten Datenstrukturen") is gratefully acknowledged.

References

Donoho, D. and Gasko, M. (1992). Breakdown properties of location estimates based on halfspace depth and projected outlyingness. *Annals of Statistics*, 20:1803–1827.

Liu, R. (1990). On a notation of data depth based upon random simplices. *Annals of Statistics*, 18:405–414.

Merkle, M. (2005). Jensens inequality for medians. *Statistics and Probability letters*, 71:277–281.

Oja, H. (1983). Descriptive statistics for multivariate distributions. *Statistics and Probability letters*, 1:327–332.

Rockafellar, R. T. (1970). *Convex analysis*. Princeton University Press, New Jersey.

Tukey, J. W. (1975). Mathematics and picturing data. In *Proceedings of International Congress of Mathematicians, Vanvouver*, volume 2, pages 523–531.