A SHARP INEQUALITY FOR THE $p$-CENTER OF GRAVITY OF A RANDOM VARIABLE

DAVID C. COX

ABSTRACT. Let $X$ be a real-valued random variable. For $p > 1$, define the $p$-center of gravity of $X$, $C_p(X)$, as the unique number $c$ which minimizes $\|X - c\|_p$. This paper exhibits a more-or-less explicit expression for the best constant $\gamma = \gamma_p$ in the inequality $|E(X) - C_p(X)| \leq \gamma \|X - C_p(X)\|_p$, and presents asymptotic formulas for $\gamma_p$ as $p \to 1, 2$ and $+\infty$, respectively. The definition of $C_p(X)$ is extended to variables taking values in an arbitrary Hilbert space $H$, and it is shown that $\gamma_p$ is not increased by this extension.

1. Introduction. Let $X$ be a real-valued random variable. For $p \geq 1$, define the $p$-center of gravity of $X$, $C_p(X)$, as the number $c$ which minimizes $\|X - c\|_p$. One has $C_1(X) =$ any median of $X$, while $C_2(X) = E(X)$. In general, $C_p(X)$ is uniquely defined for $p > 1$, since the function $c \to \|X - c\|_p$ is then strictly convex. A question which immediately arises is: How far can $C_p(X)$ be from $E(X)$? §2 of this paper exhibits a more-or-less explicit expression for

$$\gamma_p = \sup \{ |E(X)| : C_p(X) = 0, \|X\|_p = 1 \} \quad (p > 1).$$

In general, one has

$$|E(X) - C_p(X)| \leq \gamma_p \|X - C_p(X)\|_p \leq \gamma_p \|X\|_p.$$

The best constant in the related inequality $\|X - E(X)\|_p \leq D_p \|X\|_p$ was determined by Witsenhausen [5]. His results for $D_p$ can be obtained, and in some cases extended, using methods very similar to those of the present paper; see [1] for details. Numerical values for $\gamma_p$ are provided, as well as asymptotic formulas valid for $p \to 1$, $p \to 2$, and $p \to \infty$, respectively. In §3, the definition of $C_p(X)$ is extended to random variables $X$ taking values in an arbitrary Hilbert space $H$. It is shown that the constant $\gamma_p$ is not increased by this extension.

2. Main results.

THEOREM 1. Let $p > 1$, $p \neq 2$, and $q = p/(p - 1)$. One has

$$\gamma_p = \sup_{x > 0} \frac{x - x^{q-1}}{\bigl( x^q + x \bigr)^{1/p} \bigl( 1 + x \bigr)^{1/q}} = \sup_s \frac{|\sinh(q - 2)s|}{\cosh(s)}.$$

Received by the editors October 25, 1982 and, in revised form, March 7, 1984.
1980 Mathematics Subject Classification. Primary 60E15.

©1985 American Mathematical Society
0002-9939/85$1.00 + $.25 per page

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
from which it follows that $\gamma_p = \gamma_q$. The second supremum in (1) is attained exactly for $s = \pm s_p$, where $s_p$ is the unique positive root of the equation

\[(q - 1)^2 \tanh(q - 1)s + \tanh s \{ \tanh(q - 2)s = q(q - 2) \}.
\]

**Proof.** From the obvious symmetry of the problem, one has

$$\gamma_p = \sup \left\{ E(X) : C_p(X) = 0, E\left(\left|X\right|^p\right) = 1 \right\}.$$  

The condition $C_p(X) = 0$ is equivalent to $E((\text{sgn} X)|X|^{p-1}) = 0$. Hence, from a well-known principle of the theory of moments (see, e.g., Theorem 3 of [3]), one has

$$\gamma_p = \inf \left\{ A + C : \text{there exists } B \text{ such that } x \leq A|x|^p + B(\text{sgn } x)|x|^{p-1} + C \right\}.$$  

From Theorems 4 and 5 of [3], one may confine attention, in evaluating $\gamma_p$, to random variables supported by the associated contact sets $\{ t : t = A|t|^p + B(\text{sgn } t)|t|^{p-1} + C \}$. Now, the inequality $x \leq A|x|^p + B(\text{sgn } x)|x|^{p-1} + C$ is equivalent to

\[
mx + b \leq \phi_d(x) \equiv |x|^p + d(\text{sgn } x)|x|^{p-1},
\]

where $m = A^{-1} > 0$, $b = -A^{-1}C$ and $d = A^{-1}B$. Inequality (3) holds iff the line $y = mx + b$ supports the convex hull of the graph of $\phi_d$ from below. The graph of $\phi_0$ is convex. For $d \neq 0$, the graph of $\phi_d$ is first convex, then concave and finally convex again. It follows that each contact set $\{ t : t = A|t|^p + B(\text{sgn } t)|t|^{p-1} + C \}$, mentioned above, consists of at most two points. Thus, in determining $\gamma_p$ one may assume that $X$ takes at most two values (one positive and one negative, since $E((\text{sgn} X)|X|^{p-1}) = 0$).

Let $u, v > 0$ with $P(X = v) = a$, $P(X = -u) = 1 - a$. The conditions $E((\text{sgn } X)|X|^{p-1}) = 0$ and $E(|X|^p) = 1$ lead to the equations $av^{p-1} - (1 - a)u^{p-1} = 0$ and $av^p + (1 - a)u^p = 1$, respectively. Further, $E(X) = av - (1 - a)u$. Introducing $x = a/(1 - a) = (u/v)^{p-1}$, one obtains

$$E(X) = f(x) \equiv \frac{x - x^{q-1}}{(x^q + x)^{1/p}(1 + x)^{1/q}},$$

which proves the first formula in (1). The second follows on substituting $s = (\ln x)/2$. Interchanging $p$ and $q$, and replacing $s$ by $s/(q - 1)$, one finds $\gamma_p = \gamma_q$. Setting to zero the logarithmic derivative of the second formula in (1) gives the equation

\[(q - 2) \coth(q - 2)s = \frac{q - 1}{p} \tanh(q - 1)s + \frac{1}{q} \tanh s,
\]

which is equivalent to (2). The monotonicity of the function $\tanh$ implies that (2) has a unique positive root $s_p$, at which the second supremum in (1) is attained.

In some cases, (2) can be solved explicitly. For example, if $p = 3$ one has $\gamma_3 = \gamma_3/2$, so letting $q = 3$ in (2), one finds the equation

$$4 \tanh 2s + \tanh s \right) \tanh s = 3,$$
leading to $\gamma_3 = \gamma_{3/2} = (2\sqrt{3} - 3)^{1/2}/4^{1/3}$. Likewise, $\gamma_4 = \gamma_{4/3} = 2^{1/2}3^{-3/4}$. In general, (2) is easily solved numerically. Table I provides a sample of values of $s_p$ and $\gamma_p$. For values of $p$ near 1, 2 or $+\infty$, asymptotic formulas for $s_p$ and $\gamma_p$ can be obtained.

**Table I. Selected Values of $s_p$ and $\gamma_p$**

<table>
<thead>
<tr>
<th>$p$</th>
<th>$s_p$</th>
<th>$\gamma_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>1.257757391</td>
<td>0.063081196</td>
</tr>
<tr>
<td>2.2</td>
<td>1.312362783</td>
<td>0.120239909</td>
</tr>
<tr>
<td>2.3</td>
<td>1.363924906</td>
<td>0.172125389</td>
</tr>
<tr>
<td>2.4</td>
<td>1.412794038</td>
<td>0.219321410</td>
</tr>
<tr>
<td>2.5</td>
<td>1.459259873</td>
<td>0.262346991</td>
</tr>
<tr>
<td>3.0</td>
<td>1.662885890</td>
<td>0.429160632</td>
</tr>
<tr>
<td>4.0</td>
<td>1.975436846</td>
<td>0.620403240</td>
</tr>
<tr>
<td>5.0</td>
<td>2.213433877</td>
<td>0.722673691</td>
</tr>
<tr>
<td>10.0</td>
<td>2.940014234</td>
<td>0.891257102</td>
</tr>
<tr>
<td>20.0</td>
<td>3.655311</td>
<td>0.954265</td>
</tr>
<tr>
<td>30.0</td>
<td>4.06997</td>
<td>0.97163</td>
</tr>
<tr>
<td>40.0</td>
<td>4.36279</td>
<td>0.97957</td>
</tr>
<tr>
<td>50.0</td>
<td>4.58924</td>
<td>0.98408</td>
</tr>
</tbody>
</table>

**Theorem 2.** For $p \to 2$, one has

\[
s_p = s_\ast + s_\ast (p - 2)/2 - s_\ast (p - 2)^2/6 + O((p - 2)^3),
\]

where $s_\ast$ is the unique positive root of the equation $s \tanh s = 1$, and

\[
\gamma_p = (s_\ast^2 - 1)^{1/2}|p - 2| \left(1 - \frac{1}{2}(p - 2) + \left(\frac{1}{8} + \frac{s_\ast^2}{24}\right)(p - 2)^2\right) + O((p - 2)^4).
\]

**Proof.** It is clear from (2) that $s_p$ remains bounded away from both 0 and $+\infty$ as $p \to 2$, equivalently, $q \to 2$. For example, suppose a sequence $p_i \to 2$ can be found with $s_{p_i} \to \infty$. Then, $[\tanh(q_i - 2)s_{p_i}|/(q_i - 2) \to \infty$, a contradiction. Thus, for $q \to 2$, (2) may be written $s \tanh s = 1 + O(q - 2)$. It follows that $s_p = s_\ast + O(q - 2)$. Standard resubstitution techniques (see, e.g., [4, p. 11]) lead to (4), after some calculation. Equation (5) follows routinely from (4).

**Theorem 3.** For $p \to \infty$, equivalently, $q \downarrow 1$, one has

\[
s_p = \ln(2p - 2) + \frac{\ln(2(p - 2))^2}{2(p - 1)^2} + \frac{1}{4(p - 1)^2} + O\left(\frac{\ln p}{p}\right)^3
\]

and

\[
\gamma_p = 1 - \frac{\ln 2}{p - 1} - \frac{\ln(p - 1)}{(p - 1)^2} - \frac{1 - (\ln 2)^2}{2(p - 1)^2} + O\left(\frac{\ln p}{p}\right)^3.
\]
Proof. Setting \( q = 1 \) in (2) given, \( \tanh^2 s = 1 \). Thus, \( s_p \to \infty \) as \( p \to \infty \). Let \( \varepsilon = q - 1 \). Then (2) becomes
\[
(8) \quad \varepsilon^2 \tanh \varepsilon s + \tanh s \tanh(1 - \varepsilon)s = 1 - \varepsilon^2.
\]
For \( x \to \infty \), \( \tanh x = 1 - 2e^{-2x} + O(e^{-4x}) \). I claim that \( \varepsilon^2 s_p \to 0 \) as \( p \to \infty \). For, suppose a sequence \( p_i \to \infty \) can be found with \( \varepsilon_i s_{p_i} \geq c > 0 \) for all \( i \). Since \( \tanh s_{p_i} \) and \( \tanh(1 - \varepsilon_i)s_{p_i} \) are then equal to \( 1 + o(e^2) \), and \( \tanh c > 0 \), it follows that the left side of (8) eventually exceeds 1, which is a contradiction. Equation (8) may now be written \( \tanh^2 s = 1 - e^2 + o(e^2) \), so that
\[
S = \ln\left(\frac{2}{\varepsilon}\right) + o\left(\ln\left(\frac{2}{\varepsilon}\right)\right) = \ln(2p - 2) + o(\ln(2p - 2)).
\]
Continuing in this manner one arrives at formulas (6) and (7) by routine, though tedious, calculation.

Finally, note that (6) and (7) also describe the behavior of \( s_p \) and \( \gamma_p \), respectively, as \( p \to 1 \), since \( s_q = (q - 1)s_p \) and \( \gamma_q = \gamma_p \).

3. Extensions to vector-valued random variables. First consider the complex case. Let \( Z \) be a complex-valued random variable with \( \|Z\| = E(|Z|^p)^{1/p} < \infty \), where \( p > 1 \). Because \( |z|^p \) is a strictly convex function on \( \mathbb{C} \), \( C_p(Z) \) may be uniquely defined by \( \|Z - C_p(Z)\| = \inf_{c \in \mathbb{C}} \|Z - c\| \). Let \( \gamma_p(C) = \sup\{|E(Z)| : \|Z\| = 1, C_p(Z) = 0\} \).

Theorem 4. \( \gamma_p(C) = \gamma_p \).

Proof. The condition \( C_p(Z) = 0 \) is equivalent to \( E(\text{Re} Z|Z|^{p-2}) = E(\text{Im} Z|Z|^{p-2}) = 0 \). Replacing \( Z \) by \( e^{i\theta}Z \), if necessary, one may assume \( E(Z) \) is real, i.e., \( E(\text{Im} Z) = 0 \). Hence, in determining \( \gamma_p(C) \), inequalities of the following type must be investigated [3]:
\[
\text{Re} z \leq A + B|z|^2 + C|\text{Im} z|^{p-2} + D|\text{Im} z + F||z|^p,
\]
where \( A, B, C, D, \) and \( F \) are real constants. On replacing \( z \) by \( \tilde{z} \), one may take \( C = D = 0 \) and so focus attention on inequalities of the form
\[
(9) \quad \text{Re} z \leq \tilde{A} + \tilde{B}|z|^2 + \tilde{C}\text{Re} z|z|^{p-2},
\]
valid for all \( z \in \mathbb{C} \), with \( \tilde{A}, \tilde{B}, \tilde{C} \) as real constants. Let \( z = re^{i\theta} \) and fix \( r \). Then (9) takes one of the two forms \( \cos \theta \leq d \) or \( \cos \theta \geq d \), for \( 0 < \theta < 2\pi \), where \( d \) is real. To check these inequalities only the values \( \theta = 0, \pi \) need be considered. It follows that (1) holds for all complex \( z \) if it holds for all real \( z \). Since (9) for \( z \) real reduces to \( x \leq \tilde{A} + \tilde{B}|x|^2 + \tilde{C}(\text{sgn} x)|x|^{p-1} \), the theorem is proved.

Next, let \( H \) be an arbitrary real Hilbert space and \( X \in L_p(H) \). It is well known that \( L_p(H) \) is a uniformly convex Banach space; a proof may be based on Lemma 15.4 and Theorem 15.7 of [2]. Hence, if \( S \) is a closed linear subspace of \( L_p(H) \), and \( x \not\in S \), there is a unique \( s_0 \in S \) with \( \|s_0 - x\| = \inf_{s \in S} \|s - x\| \). In particular, \( C_p(X) \) may be defined as the unique element of \( H \) which minimizes \( \|X - c\|_p \). Now let \( \gamma_p(H) = \sup\{|E(X)| : \|X\| = 1, C_p(X) = 0\} \).

Theorem 5. \( \gamma_p(H) = \gamma_p \).
Proof. Clearly, $\gamma_p(H) \geq \gamma_p$, since $\mathbb{R} \subseteq H$. Since $\gamma_p = \gamma_p(C)$, one has, from (9), that there exist real constants $\tilde{A}$, $\tilde{B}$ and $\tilde{C}$ with $\tilde{A} + \tilde{B} = \gamma_p$ and

$$x \leq \tilde{A} + \tilde{B}(x^2 + y^2)^{p/2} + \tilde{C}x(x^2 + y^2)^{(p/2)-1}$$

for all $x, y \in \mathbb{R}$. Choose an orthonormal basis $(x_a)$ for $H$ with $x_{a_0} = E(X)/\|E(X)\|$. From (10),

$$(x, x_{a_0}) \leq A + B\|x\|^p + C(x, x_{a_0})\|x\|^{p-2}$$

for all $x \in H$. Now suppose $\|X\|_p = 1$ and $C_p(X) = 0$. The latter implies, in particular, that

$$\inf_{t \in \mathbb{R}} E\left[\left(\|X\|_{a_0} - t\right)^2 + \sum_{a \neq a_0} (X, x_a)^2\right]^{p/2}$$

is assumed at $t = 0$. By differentiation, $E(X, x_{a_0})\|X\|^{p-2} = 0$. Moreover, $E(X, x_{a_0}) = \|E(X)\|$. From (11), therefore, $\|E(X)\| \leq A + B = \gamma_p$. Thus, $\gamma_p(H) \leq \gamma_p$ and so $\gamma_p(H) = \gamma_p$.

References


Battelle Memorial Institute, Washington Operations, 2030 M Street, N.W., Washington, D.C. 20036

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use