MODES AND MOMENTS OF UNIMODAL DISTRIBUTIONS

KEN-ITI SATO

(Received Sept. 10, 1985)

Summary

For a unimodal distribution relations of its mode $a$ with its absolute moment $\beta_p$ and central absolute moment $\gamma_p$ of order $p$ are considered. The best constants $A_p$ and $B_p$ are given for the inequalities $|a| \leq A_p \beta_p^p$ ($p > 0$) and $|a - m| \leq B_p \gamma_p^{1/p}$ ($p \geq 1$) where $m$ is the mean. The results follow from discussion of more general moments.

1. Introduction

Let $\mu$ be a unimodal distribution with mode $a$ and let $\beta_p$ be its absolute moment of order $p > 0$. It is shown in Sato [4] that there is a constant $A_p$ such that

\begin{equation}
|a| \leq A_p \beta_p^{1/p}.
\end{equation}

When $\mu$ has finite mean $m$, the central absolute moment of order $p \geq 1$ is denoted by $\gamma_p$. It is also shown in [4] that there is a constant $B_p$ for $p > 1$ such that

\begin{equation}
|a - m| \leq B_p \gamma_p^{1/p}.
\end{equation}

Here $A_p$ and $B_p$ are constants depending only on $p$. The latter is an extension of a result of Johnson and Rogers [3], who give (1.2) for $p = 2$ and prove that $B_2 = \sqrt{3}$ is the best constant. This result for $p = 2$ is rediscovered by Vysochanskii and Petunin [5]. By monotonicity of $\gamma_p^{1/p}$ in $p$, the existence of $B_p$ for some $p = p_0$ implies its existence for any $p \geq p_0$. We can make a similar assertion for $A_p$ by monotonicity of $\beta_p^{1/p}$ for $p > 0$. But the case of small $p$ is interesting, since there are many unimodal distributions that have absolute moments of order $p$ only for small $p$. For example stable distributions of exponent $\alpha$ ($0 < \alpha < 2$) are unimodal and have absolute moments of order $p$ only for $0 < p < \alpha$.

Key words and phrases: Unimodal distribution, mode, moment.
In this paper we will present a new proof of (1.1) and (1.2) and give the best constants. We will show that

\[(1.3) \quad B_p=(p+1)^{1/p} \]

is the best constant in (1.2) for \(p \geq 1\) (now the case \(p=1\) is included) and that the best constant in (1.1) for \(p>0\) is the unique solution of the equation

\[(1.4) \quad x^{p+1}-(p+1)x-p=0 \]

for \(x>1\). Thus \(A_p>(p+1)^{1/p}\) for \(p>0\); \(A_2=2, A_1=1+\sqrt{2}\), and, approximately, \(A_{1/2}=2.81451\).

Given a function \(g\) on the line, we call the integral \(\int g(x)\mu(dx)\) the \(g\)-moment of \(\mu\), and \(\int g(x-m)\mu(dx)\) the central \(g\)-moment of \(\mu\). We will give inequalities involving modes and \(g\)-moments of unimodal distributions. The bounds (1.1) and (1.2) are extended to more general moments.

2. Modes and \(g\)-moments

A distribution \(\mu\) is called unimodal if there is a point \(a\) such that the distribution function of \(\mu\) is convex on \((-\infty, a)\) and concave on \((a, \infty)\). The point \(a\) is called a mode of \(\mu\). If \(\mu\) is unimodal, then the set of modes of \(\mu\) is either a one point set or a closed interval. Write the restriction of \(\mu\) to an interval \(I\) as \(\mu|_I\). A distribution \(\mu\) is unimodal with mode \(a\) if and only if \(\mu|_{(-\infty,a)}\) is absolutely continuous with nondecreasing density and \(\mu|_{(a,\infty)}\) is absolutely continuous with nonincreasing density.

Let \(g(x)\) be a nonnegative continuous function on the line such that \(g(x)=g(-x)\) and \(g(x)\) is increasing for \(x>0\). The words increase and decrease are used in the strict sense.

**Theorem 2.1.** For every \(a>0\), there is a unique point \(c\) satisfying \(0<c<a\) such that, if \(\mu\) is a unimodal distribution with mode \(a\), then

\[(2.1) \quad \int g(x)\mu(dx) \geq (a+c)^{-1}\int_{-c}^{a} g(x)dx .\]

The point \(c\) is the unique point satisfying \(0<c<a\) and

\[(2.2) \quad g(c)=(a+c)^{-1}\int_{-c}^{a} g(x)dx .\]

Equality holds in (2.1) if and only if \(\mu\) is the uniform distribution on \([-c,a]\).
PROOF. Denote the Lebesgue measure on the line by \( \lambda \). Let \( a > 0 \) and let \( \mu \) be a unimodal distribution with mode \( a \). We estimate the \( g \)-moment of \( \mu \) from below in three steps.

**Step 1.** Let \( \alpha = \mu(-\infty, -a] + \mu[a, \infty) \) and let

\[
\mu_1 = \mu_{[-a, a]} + \alpha a^{-1} \lambda_{[a, \infty)} .
\]

Obviously, \( \mu_1 \) is a unimodal distribution with mode \( a \). Since \( g \) is even and increasing on the positive line, the \( g \)-moment of \( \mu_1 \) is smaller than or equal to that of \( \mu \). If \( \mu_1 \neq \mu \), then they are not equal.

**Step 2.** Let \( \beta = \mu_1(0, a) \) and let

\[
\mu_2 = \mu_1_{[-a, 0]} + \beta a^{-1} \lambda_{[a, \infty)} .
\]

Again this is unimodal with mode \( a \). Let \( f_1(x) \) be the nondecreasing density of \( \mu_1 \) on \((-a, a)\). If \( f_1 \) is flat on \((0, a)\), then \( \mu_2 = \mu_1 \). If \( f_1(0+) < f_1(a-) \), then noting \( f_1(0+) < \beta a^{-1} < f_1(a-) \) and choosing \( 0 < a' < a \) that satisfies \( f_1(a' -) \leq \beta a^{-1} \leq f_1(a' +) \), we have

\[
\int_0^a (f_1(x) - \beta a^{-1}) dx = \int_0^{a'} (f_1(x) - \beta a^{-1}) dx
\]

and

\[
\int_0^a g(x) (f_1(x) - \beta a^{-1}) dx < \int_0^a g(x) (f_1(x) - \beta a^{-1}) dx ,
\]

which implies that the \( g \)-moment of \( \mu_2 \) is smaller than that of \( \mu_1 \).

**Step 3.** Let \( \gamma = \mu_2(-a, 0) \) and \( b = \gamma \beta^{-1} a \). Then, \( 0 \leq b \leq a \). Let \( \mu_3 = \beta a^{-1} \lambda_{[-b, a]} \), the uniform distribution on \([-b, a]\). Now

\[
(2.3) \quad \int g(x) \mu_3(dx) \leq \int g(x) \mu_2(dx) ,
\]

because, letting \( f_3(x) \) be the nondecreasing density of \( \mu_3 \) on \((-a, a)\), we have

\[
\int_{-a}^a (\beta a^{-1} - f_3(x)) dx = \int_{-a}^{-b} f_3(x) dx
\]

and

\[
\int_{-b}^a g(x) (\beta a^{-1} - f_3(x)) dx \leq \int_{-a}^{-b} g(x) f_3(x) dx .
\]

Strict inequality holds in (2.3) if \( \mu_3 \neq \mu_2 \).

Now define, for fixed \( a \),
\[ \varphi(b) = (a+b)^{-1} \int_{-b}^{a} g(x) dx . \]

The above three steps show that
\[ \int g(x) \mu(dx) \geq \varphi(b) \]
for some \( b \) in \([0, a]\). Here \( b \) depends on \( \mu \). Strict inequality holds unless \( \mu \) is the uniform distribution on \([-b', a]\) for some \( b' \in [0, a] \). As \( b \) moves on \([0, a]\), the function \( \varphi(b) \) takes the minimum at a unique point \( c \). In fact, \( \varphi'(b) = (a+b)^{-2} \varphi(b) \) where
\[ \varphi(b) = (a+b)g(-b) - \int_{-b}^{a} g(x) dx = \int_{-b}^{b} (g(b) - g(x)) dx - \int_{b}^{a} (g(x) - g(b)) dx , \]
and \( \varphi(b) \) is a continuous increasing function with \( \varphi(0) < 0 \) and \( \varphi(a) > 0 \). The point \( c \) is the unique increasing function such that \( 0 < c < a \) and \( \varphi(c) = 0 \), which is equivalent to (2.2). The proof is complete.

**Remark 2.1.** Define \( k(a) \) for \( a > 0 \) by \( k(a) = c \) in Theorem 2.1 and \( k(0) = 0 \). We see that, if \( \mu \) is unimodal with mode \( a \), then
\[ \int g(x) \mu(dx) \leq g(k(|a|)) . \]

Let \( M = \sup g(x) \leq \infty \). The equation (2.2) shows that
\[ \int_{-c}^{c} (g(c) - g(x)) dx = \int_{c}^{a} (g(x) - g(c)) dx . \]
As \( a \) increases, \( c \) must increase in order to satisfy this identity. That is, \( k(x) \) is increasing in \( x \). Now it is easy to see that \( k(x) \) is a continuous increasing function from \([0, \infty)\) onto itself. Hence \( g(k(x)) \) is a continuous increasing function from \([0, \infty)\) onto \([g(0), M]\). Let \( x = h_t(y) \) be the inverse function of \( y = g(k(x)) \). If \( \mu \) is unimodal with mode \( a \), then
\[ |a| \leq h_t \left( \int g(x) \mu(dx) \right) . \]
For every \( y \geq g(0) \), \( h_t(y) \) is the supremum of modes taken over all unimodal distributions that have \( g \)-moment \( y \). In fact, for \( x = h_t(y) \), the uniform distribution on \([-k(x), x]\) has \( g \)-moment \( y \).

3. **Modes and absolute moments of order \( p > 0 \)**

The preceding theorem has the following consequence.

**Theorem 3.1.** For \( p > 0 \), let \( A_p \) be the unique solution of the equa-
tion (1.4) in (1, ∞). If \( \mu \) is unimodal with mode \( a \), then

\[
|a| \leq A_p \beta_p^{\beta_p},
\]

where \( \beta_p = \int |x|^p \mu(dx) \). Equality holds in (3.1) if and only if \( \mu \) and \( a \) satisfy one of the following:

(i) \( a = 0 \) and \( \mu \) is the \( \delta \)-distribution at 0;
(ii) \( a > 0 \) and \( \mu \) is the uniform distribution on \([-a/A_p, a] \);
(iii) \( a < 0 \) and \( \mu \) is the uniform distribution on \([a, -a/A_p] \).

**Proof.** Let \( a > 0 \). It is enough to prove the theorem in this case. Let \( g(x) = |x|^p \) in Theorem 2.1. Then

\[
\int |x|^p \mu(dx) \geq c^p,
\]

where \( c \) is the unique solution of the equation

\[
(a + c)^p - (p + 1)^{-1}(a^{p+1} + c^{p+1}) = 0
\]

for \( 0 < c < a \). We see that \( 1/A_p \) is the value of \( c \) for \( a = 1 \). The value of \( c \) for general \( a > 0 \) is \( c = a/A_p \). Hence we obtain (3.1). Equality holds in (3.1) if and only if (ii) holds, as Theorem 2.1 says. The proof is complete.

**Remark 3.1.** Let \( A \) be the unique positive solution of the equation

\[
x \log x - x - 1 = 0.
\]

The constant \( A_p \) decreases as \( p \) increases, and

\[
\lim_{p \to 0} A_p = A, \quad \lim_{p \to \infty} A_p = 1.
\]

It is easily seen from (3.2) that \( e < A < 2e \). An approximate value is \( A = 3.59112 \).

In fact, let \( a > 0 \) and let \( 0 < p < p' \). We have

\[
|a| \leq A_p \beta_p^{\beta_p} < A_p \beta_p^{\beta_p'}
\]

unless \( \mu \) is concentrated at \( a \) (see Hardy et al. [2], p. 157). Choose \( \mu \) to be the uniform distribution on \([-a/A_p, a] \). Then \( a = A_p \beta_p^{\beta_p'} \). Hence we have \( A_p < A_p \). If \( \lim_{p \to \infty} A_p > 1 \), then \( 1 - (p + 1)A_p^p - pA_p^{p-1} = 0 \) leads to a contradiction. Therefore \( A_p \) tends to 1 as \( p \to \infty \). If we fix \( x > 1 \) and let \( p \) decrease to 0, then

\[
x^{p+1} - (p + 1)x - p = p(x \log x - x - 1) + O(p^2).
\]

Hence, if \( 1 < x < A \), then \( x^{p+1} - (p + 1)x - p \) is negative for small \( p \); if
$x > A$, then it is positive for small $p$. This shows that $\lim_{p \downarrow 0} A_p = A$.

Remark 3.2. If the integral $\int \log |x| \mu(dx)$ exists, the geometric mean $g$ of $\mu$ is defined by

$$g = \exp \int \log |x| \mu(dx).$$

If $\mu$ is unimodal with mode $a$ and $\int \log |x| \mu(dx) < \infty$, then

$$|a| \leq A_g,$$

where $A$ is given in Remark 3.1. In fact, if $\mu$ has finite $\beta_p$ for some $p > 0$, then $\beta_p^{1/p}$ tends to $g$ as $p \downarrow 0$ (see [2], p. 156) and we have (3.4) from (3.1) and (3.3). If $\mu$ has infinite $\beta_p$ for every $p > 0$, then consider $\mu_n$ defined by

$$\mu_n = \mu_{[k-n, n]} + a_n \delta_n, \quad a_n = \mu(-\infty, -n] + \mu[n, \infty)$$

and note that the geometric mean of $\mu_n$ tends to $g$ as $n \to \infty$.

4. Modes and central $g$-moments

In this section let $g(x)$ be a nonnegative function such that $g(x) = g(-x)$ and $g(x)/x$ is nondecreasing in $x > 0$.

**Theorem 4.1.** If $\mu$ is unimodal with mode $a$ and has finite mean $m$ and if $m \neq a$, then

$$\int g(x - m) \mu(dx) \geq 2^{-1} |a - m|^{-1} \int |x - a|^{-1} \log |x| dx.$$

Equality holds in (4.1) if and only if $\mu$ is the uniform distribution on an interval with $a$ chosen to be an endpoint of the interval.

**Proof.** Let $\mu$ be unimodal with mode $a$ with mean $m$ and $m \neq a$. By translation and reflection, we may assume $m = 0$ and $a > 0$. We estimate the $g$-moment of $\mu$ from below by the $g$-moment of another unimodal distribution with mode $a$ and mean 0.

**Step 1.** Let $a = \mu[a, \infty)$ and let

$$\mu_1 = \mu_{(-\infty, a]} + a^{-1} \lambda_{[a, \infty)}.$$

Then $\mu_1$ is unimodal with mode $a$. If $\mu_1 \neq \mu$, then the $g$-moment of $\mu_1$ is smaller than that of $\mu$ and the mean of $\mu_1$ is negative.

**Step 2.** Let $\beta = \mu_1(0, a)$ and
\[ \mu_2 = \phi_{1, \infty, 0} + \beta a^{-1} \lambda_{0, a} \cdot \]

As in Step 2 in the proof of Theorem 2.1, we see that \( \mu_2 \) is unimodal with mode \( a \) and that, if \( \mu_2 \neq \mu_1 \), then \( \mu_2 \) has smaller \( g \)-moment and mean than \( \mu_1 \).

**Step 3.** If \( \mu_3 = \mu \), then let \( \mu_3 = \mu \). Suppose that \( \mu_3 \neq \mu \). We see that \( \mu_3(-\infty, a) \) has positive mean, since it does not have flat density. On the other hand, \( \mu_3 \) has negative mean. So we can find \( b > a \) such that

\[ \mu_3 = \mu_3(-\infty, -b) a^{-1} \lambda_{0, a} \]

has zero mean. Obviously \( \mu_3 \) is unimodal with mode \( a \) and its \( g \)-moment is smaller than that of \( \mu_2 \).

**Step 4.** We have \( \mu_4(0, a) \geq 1/2 \), since \( \mu_4 \) is unimodal with mode \( a \), concentrated on \(( -\infty, a) \) and has zero mean. The case \( \mu_4(0, a) = 1/2 \) occurs if and only if \( \mu_4 \) is the uniform distribution on \([ -a, a] \). Let \( f_4(x) \) be the density of \( \mu_4 \). We have \( f_4(x) = \gamma \) on \(( 0, a) \) for some constant \( \gamma \geq (2a)^{-1} \). We claim that

\[ \int_{-\infty}^{a} g(x) f_4(x) dx \geq \int_{-a}^{a} g(x) dx . \]

This will imply

\[ \int g(x) \mu_4(dx) \geq \gamma \int_{-a}^{a} g(x) dx \geq (2a)^{-1} \int_{-a}^{a} g(x) dx , \]

from which (4.1) follows. First note that

\[ \int_{-\infty}^{a} x f_4(x) dx = \int_{-a}^{a} x (\gamma - f_4(x)) dx . \]

Using this and increasingness of \( g(x)/x \) in \( x > 0 \), we have

\[ \int_{-\infty}^{a} g(x) f_4(x) dx \geq g(a) a^{-1} \int_{-\infty}^{a} x f_4(x) dx \]

\[ = g(a) a^{-1} \int_{-a}^{0} x |g(x) - f_4(x)| dx \geq \int_{-a}^{0} g(x) (\gamma - f_4(x)) dx . \]

Thus (4.2) follows. This proof shows that equality holds in (4.1) only if \( \mu \) is the uniform distribution on an interval with \( a \) chosen to be an endpoint of the interval. As the converse statement for the equality is obvious, proof of the theorem is complete.

**Remark 4.1.** Define \( l(x) \) by \( l(x) = (2x)^{-1} \int_{-x}^{x} g(y) dy \) for \( x > 0 \) and \( l(0) = g(0+) \). Now, if \( \mu \) is unimodal with mode \( a \) and has finite mean \( m \), then
\[ \int g(x-m)\mu(dx) \geq l(|a-m|) . \]

Noting that \( l(x) \) is a continuous increasing function from \([0, \infty)\) onto \([g(0+), \infty)\), let \( x = h_2(y) \) be the inverse function of \( y = l(x) \). Then

\[ |a - m| \leq h_2 \left( \int g(x-m)\mu(dx) \right) . \]

If \( y \geq g(0+) \), then \( h_2(y) \) is the supremum of modes taken over all unimodal distributions that have mean 0 and \( g \)-moment \( y \). In fact, for \( x = h_2(y) \), the uniform distribution on \([-x, x]\) has mean 0 and \( g \)-moment \( y \).

**Remark 4.2.** The assumption of nondecreasingness of \( g(x)/x \) in \( x > 0 \) in Theorem 4.1 cannot be replaced by nondecreasingness of \( g(x) \) in \( x > 0 \). For example, let \( g(x) = |x|^p \) with \( 0 < p < 1 \) and choose \( \mu(dx) = f(x)dx \) with \( a = 1 \) and \( m = 0 \) in the form \( f(x) = \alpha c \) on \([-b, -1/2] \), \( \alpha \) on \([-1/2, 1] \) and 0 outside of \([-b, 1] \) where \( b > 1/2 \), \( \alpha > 0 \), \( 0 < c < 1 \). Then \( \int g(x)\mu(dx) \) is smaller than \( (p+1)^{-1} \) when \( b \) is sufficiently large.

5. Modes and central absolute moment of order \( p \geq 1 \)

We apply Theorem 4.1 to central absolute moments.

**Theorem 5.1.** Let \( p \geq 1 \). If \( \mu \) is unimodal with mode \( a \) and has finite mean \( m \), then

\[ |a - m| \leq (p+1)^{1/p} \gamma_p^{1/p} , \tag{5.1} \]

where \( \gamma_p = \int |x-m|^p\mu(dx) \). Equality holds in (5.1) if and only if \( \mu \) is a \( \delta \)-distribution or a uniform distribution on an interval with a chosen to be an endpoint of the interval.

**Proof.** If \( a = m \), then (5.1) is trivial. If \( a \neq m \), then, using Theorem 4.1 for \( g(x) = |x|^p \), we get

\[ \gamma_p \geq (p+1)^{-1}|a - m|^p , \]

that is (5.1). The statement about the case of the equality also follows from the theorem.

**Remark 5.1.** The coefficient \((p+1)^{1/p}\) in (5.1) decreases from 2 to 1 as \( p \) increases from 1 to \( \infty \).
6. Modes and exponential moments

Let us consider exponential moments.

**Theorem 6.1.** Let \( g(x) = e^{\left| x \right|} - 1 \). For \( y \geq 0 \), let \( h_1(y) \) be the supremum of modes taken over all unimodal distributions that have \( g \)-moment \( y \), and let \( h_2(y) \) be the supremum of modes taken over all unimodal distributions with \( g \)-moment \( y \) and mean 0. Then,

\[
\begin{align*}
    h_1(y) &= \log y + \log \log y + \log 2 + (2^{-1} + o(1))(\log y)^{-1} \log \log y, \\
    h_2(y) &= \log y + \log \log y + (1 + o(1))(\log y)^{-1} \log \log y
\end{align*}
\]

as \( y \to \infty \).

**Proof.** By Remark 2.1, the function \( x = h_1(y) \) for \( y > 0 \) is given by \( c = \log (y+1) \) and by the equation

\[(x + c - 1)e^x - e^c + 2 = 0\]

with the condition \( x > c \). The function \( x = h_2(y) \) is, according to Remark 4.1, the inverse function of \( y = (e^x - x - 1)/x, \ x > 0 \). Hence, by the method of asymptotic expansion (see Dieudonné [1], III. 8), we can prove that \( h_1(y) \) and \( h_2(y) \) behave as in the statement of the theorem.

**NAGoya University**

**References**


