数理解析研究所講究録
916 巻 1995 年 75-89

INEQUALITIES ON THE MEAN, MEDIAN, MODE AND SKEWNESS

筑波大 数学 佐藤道一 (Michikazu Sato)

Summary

Many sufficient conditions for inequalities about the mean, median, mode and skewness have been obtained. Runnenburg gives a result and MacGillivray tries to improve it, but unfortunately, there is a mistake in it. We get a corrected and improved theorem, consider the cases of Pearson distributions, and give some counter-examples.

Key words: Mean, median, mode, skewness, Pearson distributions.

1. Introduction

Let P be a probability distribution on $\mathbf{R}$. Then we denote a random variable distributed as P by X, and the mean, median and mode by μ, m, M, respectively, and $\mu_n := E[(X-\mu)^n]$ and $\mu_n' := E(X^n)$, if each of them exists. Assume that P is asymmetric and has the density f. Many sufficient conditions for inequalities about the values above have been obtained. Van Zwet (1979) gets that if $f(\mu+x)-f(\mu-x)$ changes sign once and from negative to positive values in $x>0$, then $\mu_3>0$. (He does not

This research was supported in part by Grant-in-Aid for Science Research No. 00001777, Ministry of Education, Science and Culture, Japan.

say in this way. See the note below Theorem 2.1.) This is an improvement of Runnenburg (1978). MacGillivray (1981) says that this assumption also implies $M < \mu$ for a unimodal density, and $m < \mu$. (We can take by his paper that "$f(\mu + x) - f(\mu - x)$ changes once and from negative to positive values" is equivalent to "$\mu > M$" for a unimodal density. However, he seems to mean that under the condition that $f(\mu + x) - f(\mu - x)$ changes sign once, "from negative to positive values" is equivalent to "$\mu > M$" for a unimodal density. It is because he says that this is not necessarily so for a general unimodal density though he does not give a counter-example.) He obtains the result by using knowledge of total positivity. It is true that this assumption implies that $\mu_3 > 0$ and $m < \mu$, and we get that it can be extended to cases when a density does not exist, without knowledge of total positivity. However, it is not true that, for a unimodal density, the assumptions imply $\mu > M$, since $\mu = M$ can hold. Groneveld & Meeden (1977) also obtains sufficient conditions for $M < m$ and $m < \mu$, and van Zwet (1979) improves them. We shall give a corrected and improved theorem in Section 2, consider the cases of Pearson distributions in Section 3, and give some counter-examples in Section 4.

Generally, there are delicate problems to define the median m and the mode M. Here we adopt the following definitions:

Let F be the distribution function, that is, $F(x) = P(X \leq x)$, and denote $F(x-) = P(X < x)$. We define the median if and only if there is the unique m such that $F(m-) \leq 1/2 \leq F(m)$, and such m is called *the median*. We define the mode if and only if there is the density f (with respect to the Lebesgue measure) which has the unique M such that f (weakly) increases in $(-\infty, M]$, decreases in $[M, \infty)$, and f is either right continuous or left continuous at M. Then such M is called *the mode* and such f is said to be *unimodal*. We allow $f(M)$ to be ∞.

We define

$$\sigma := \sqrt{\mu_2}$$
$$\text{Skew}\, X := \mu_3/\sigma^{\frac{3}{2}} \qquad \text{(the usual measure of skewness)}$$
$$\text{Skew}^*\, X := (\mu - M)/\sigma \qquad \text{(the Pearson measure of skewness)}$$

if each of them exists (see Pearson, 1895, 1901, and Kendall *et al.*, 1994, pp. 108–109). Note that Pearson himself does not seem to have pay attention to their sign but considered $\beta_1 = (\text{Skew}\, X)^2$ and $|\text{Skew}^*\, X|$. Also note that by definition, we get $\text{Skew}\,(aX+b) = \text{sgn}\, a\, \text{Skew}\, X$ and $\text{Skew}^*\,(aX+b) = \text{sgn}\, a\, \text{Skew}^*\, X$ where $a (\neq 0)$ and b are constants.

2. Corrected and Improved Theorem

The following theorem gives sufficient conditions for $\mu_3 > 0$ (i.e., $\text{Skew}\, X > 0$), $M \leq \mu$ (i.e., $\text{Skew}^*\, X \geq 0$), $M < m$, $m \leq \mu$ and so on. (I) is a corrected and improved one of MacGillivray (1981). (II) and (III) are partially obtained by Groneveld & Meeden (1977) and Runnenburg (1978), and improved by van Zwet (1979).

Theorem 2.1. *Let P be an asymmetric probability distribution on* $\mathbf{R}$. *Then the following assertions hold.*

(I) [The μ-method]. *Assume that $E(|X|) < \infty$ and there exists $t > 0$ such that for all Borel sets $A \subset \mathbf{R}$,*

$$\begin{aligned} A \subset (0,t) \ \ &\text{implies}\ P(\mu + A) \leq P(\mu - A), \\ A \subset (t,\infty) \ \ &\text{implies}\ P(\mu + A) \geq P(\mu - A), \end{aligned} \tag{2.1}$$

where $\mu + A := \{\mu + a : a \in A\}$ *and* $\mu - A := \{\mu - a : a \in A\}$. *Then the* *following assertions hold.*

$$0 < \mu_3 \leq \infty \quad \text{if } \mu_3 \text{ is defined,} \tag{2.2}$$

$$P(X < \mu) > P(X > \mu), \tag{2.3}$$

$$m \leq \mu \quad \text{if } m \text{ is defined,} \tag{2.4}$$

$$m < \mu \quad \text{if } m \text{ is defined and } F \text{ is continuous,} \tag{2.5}$$

$$\mu - t \leq M \leq \mu \quad \text{if } P \text{ has a unimodal density.} \tag{2.6}$$

(II) [The m-method]. *Assume that m is defined, F is continuous, and there exists $t > 0$ such that for all Borel sets $A \subset \mathbf{R}$,*

$$A \subset (0, t) \ \text{ implies } P(m + A) \leq P(m - A),$$

$$A \subset (t, \infty) \ \text{ implies } P(m + A) \geq P(m - A),$$

where $m + A := \{m + a : a \in A\}$ and $m - A := \{m - a : a \in A\}$. *Then the following assertions hold.*

$$m < \mu \quad \text{if } \mu \text{ is defined,}$$

$$m - t \leq M \leq \mu \quad \text{if } P \text{ has a unimodal density.}$$

(III) [The M-method]. *Assume that P has a unimodal density f such that $f(M - x) \leq f(M + x)$ for $x > 0$. Then the following assertions hold.*

$$M < m,$$

$$M < \mu \quad \text{if } \mu \text{ is defined.}$$

Proof. First, for an *odd* function φ such that $E[\varphi(X)]$ is defined, we have

$$\begin{aligned}
E[\varphi(X)] &= \int_{(-\infty,\infty)} \varphi(x)P(dx) \\
&= \int_{(-\infty,0)\cup(0,\infty)} \varphi(x)P(dx) \\
&= \int_{(-\infty,0)} \varphi(x)P(dx) + \int_{(0,\infty)} \varphi(x)P(dx) \\
&= -\int_{(0,\infty)} \varphi(y)Q(dy) + \int_{(0,\infty)} \varphi(x)P(dx) \\
&\qquad [\text{where } y = -x \text{ and } Q(A) := P(-A)] \\
&= \int_{(0,\infty)} \varphi(x)(P-Q)(dx) \\
&= \int_{(0,1)} \varphi(x)(P-Q)(dx) + \varphi(1)[P(\{1\}) - Q(\{1\})] \\
&\quad + \int_{(1,\infty)} \varphi(x)(P-Q)(dx).
\end{aligned}$$

We shall show (I). We may assume that $\mu = 0$ and $t = 1$ because we may consider $(X-\mu)/t$. From the above, we have

$$\begin{aligned}
\mu_3 &= \int_{(-\infty,\infty)} x^3 P(dx) \\
&= -\int_{(0,1)} x^3\lambda(dx) + P(\{1\}) - Q(\{1\}) + \int_{(1,\infty)} x^3\nu(dx) \\
&\qquad [\text{where } \lambda = Q - P \text{ and } \nu = P - Q, \text{ which are nonnegative on} \\
&\qquad (0,1) \text{ and } (1,\infty), \text{ respectively}] \\
&\geq -\int_{(0,1)} x\lambda(dx) + P(\{1\}) - Q(\{1\}) + \int_{(1,\infty)} x\nu(dx) \\
&= \int_{(-\infty,\infty)} xP(dx) \\
&= E(X) = \mu = 0,
\end{aligned}$$

and if $\mu_3 = 0$, then $\lambda((0,1)) = \nu((1,\infty)) = 0$, so $P(|X| = 1) = 1$, and since $\mu = 0$, we have $P(X = 1) = P(X = -1) = 1/2$, hence P is symmetric. This is a contradiction. Therefore, we get (2.2). Next, we have

$$\begin{aligned}
0 = \mu = &- \int_{(0,1)} x\lambda(dx) + P(\{1\}) - Q(\{1\}) + \int_{(1,\infty)} x\nu(dx) \\
\geq &- \int_{(0,1)} \lambda(dx) + P(\{1\}) - Q(\{1\}) + \int_{(1,\infty)} \nu(dx) \\
= &\int_{(-\infty,\infty)} \operatorname{sgn} x P(dx) \\
= &P(X > 0) - P(X < 0),
\end{aligned}$$

and we get the strict inequality by a similar way to the above, hence we have (2.3). We can easily get (2.4) and (2.5) from (2.3). We shall prove (2.6). We may only show $-1 \leq M \leq 0$. Assume that $M(=: -N) < -1$. Since f decreases in $[-N,\infty)$, we have $P(-N \leq X \leq -1) \geq P(1 \leq X \leq N)$, and the reversed inequality holds by assumption, we get $P(-N \leq X \leq -1) = P(1 \leq X \leq N)$. Hence f is constant on $(-N,N)$. It contradicts the assumption that M is unique. Next, assume that $M \geq 1$. Then, by a similar way to the above, we get that f is constant on $(-1,1)$, hence $\lambda = 0$ on $(0,1)$. Therefore, $0 = \mu = \int_1^\infty x\nu(dx)$, hence $\nu = 0$ on $(1,\infty)$. Therefore, P is symmetric. This is a contradiction. Next, we assume that $0 < M < 1$. If f is left continuous at M, then, by a similar way to the above, we get that f is constant on $(-M,M]$ and have a contradiction. Therefore, f is not left continuous at M, hence right continuous at M and $f(-M) < f(M)$. So, for a sufficiently small $\varepsilon > 0$, when $-M-\varepsilon < x < -M$ and $M < y < M+\varepsilon$, we have $f(y) \leq f(-M) < f(x)$, hence $P(-M-\varepsilon < X < -M) < P(M < X < M+\varepsilon)$, which contradicts (2.1). Hence we get (2.6). We can similarly prove (II) and (III) by assuming $m = 0$ and $M = 0$, respectively.

□

Note that we can use Theorem 2.1 to get a reversed inequality. For example, if the reversed inequalities hold in (2.1), then, by considering $-X$, we get, $-\infty \le \mu_3 < 0$ if μ_3 is defined, instead of (2.2).

It is often difficult to obtain the exact value of m, but we usually need not get m concretely to use the m-method, and it is similar for the μ-method.

Van Zwet (1979) also obtains that under regularity conditions $F(m+x)+F(m-x) \le 1$ for all x implies similar results to those of the μ-method, but it is difficult to use because in applications, we usually cannot get F concretely. More abstract methods are also given in van Zwet (1964, 1979). Further, it is written in Comment of MacGillivray (1981) that one change of sign of $1-F(\mu+x)-F(\mu-x)$ in $x > 0$ (say Condition A if it is from negative to positive value) is sufficient to fix the sign of μ_3. This *seems* to be true, but he does not give its proof and states that, unfortunately, as noted by van Zwet, these (i.e., the above and other facts) are more difficult to verify. In van Zwet (1964, 1979), which MacGillivray (1981) refers, however, this is not written.

The following table shows what are implied by each assumption while what are not, for a unimodal asymmetric density such that $E(|X|) < \infty$ and μ_3 is defined. A function without $\le$ means that the function is an example in Section 4 satisfying $>$, and a function with $\le$ means that the function is an example in Section 4 satisfying equality, where ε is a sufficiently small positive number.

TABLE

	$M\&m$	$m\&\mu$	$M\&\mu$	$0\&\mu_3$
The μ-method	$f_{\sqrt{5}-1}$	$<$	$\le_{f_{\sqrt{5}-1}}$	$<$
The m-method	$\le_{f_3}$	$<$	$<$	g_ε
The M-method	$<$	h_ε	$<$	h_ε
Condition A	f_ε	$<$	f_ε	$<?$

Also note that $f_{\sqrt{5}-1}$ is a case that we can use the μ-method to show$m < \mu$ while we cannot use the m-method, and g_ε is a contrary case.

3. The Cases of Pearson Distributions

We shall consider the cases of Pearson distributions Type I to XII (Pearson, 1895, 1901, 1916). The following theorem asserts that the usually believed inequalities hold for Pearson distributions.

Theorem 3.1. *For a unimodal asymmetric Pearson distribution such that $E(|X|) < \infty$, the mean, median, mode inequality (i.e., $M < m < \mu$ or $M > m > \mu$) holds. Further if μ_3 is defined, then $M < m < \mu$ is equivalent to $\mu_3 > 0$ and $M > m > \mu$ is equivalent to $\mu_3 < 0$.*

Remark. From this theorem, the signs of Skew X and Skew* X coincide for Pearson distributions. Formal calculations in Kendall *et al.* (1994, p. 108 and pp. 215–218) seem to give a counter-example, but it is not a true counter-example. Also note that when we check Pearson's original papers, we should carefully check them because the domain of variables was not often written. For example, in Pearson (1901, pp. 446–448), if we let $3 < p < 4$, we seem to get a counter-example, but μ_3 is finite if and only if $p > 4$, hence it is not a counter-example.

Proof. We shall consider Type $\mathrm{I_L}$, that is, a density

$$f(x) = C\left(1 + \frac{x}{\alpha_1}\right)^{\nu\alpha_1}\left(1 - \frac{x}{\alpha_2}\right)^{\nu\alpha_2} \qquad (-\alpha_1 < x < \alpha_2)$$

where ν, α_1, α_2 and $C = C_{\alpha_1,\alpha_2}$ are positive and independent of x. The case $\alpha_1 = \alpha_2$ is excluded since it is symmetric. We may assume $\alpha_1 < \alpha_2$ because for the reversed case we can reduce the proof to the case $\alpha_1 < \alpha_2$

by considering $-X$. We have

$$L(x) := \log f(x) = \log C + \nu\alpha_1 \log\left(1+\frac{x}{\alpha_1}\right) + \nu\alpha_2 \log\left(1-\frac{x}{\alpha_2}\right),$$

$$L'(x) = \nu\left(\frac{1}{1+\dfrac{x}{\alpha_1}} - \frac{1}{1-\dfrac{x}{\alpha_2}}\right)$$

where they exist, hence it is unimodal and $M = 0$. Therefore, for a given θ, we have

$$\begin{aligned}
g(x) :=& \frac{d}{dx}\{\log f(\theta+x) - \log f(\theta-x)\} \\
=& L'(\theta+x) + L'(\theta-x) \\
=& \nu\left\{\left(\frac{1}{1+\dfrac{\theta+x}{\alpha_1}} + \frac{1}{1+\dfrac{\theta-x}{\alpha_1}}\right)\right. \\
&\left. - \left(\frac{1}{1+\dfrac{-\theta-x}{\alpha_2}} + \frac{1}{1+\dfrac{-\theta+x}{\alpha_2}}\right)\right\} \\
=& 2\nu\left\{\frac{1+\dfrac{\theta}{\alpha_1}}{\left(1+\dfrac{\theta}{\alpha_1}\right)^2 - \left(\dfrac{x}{\alpha_1}\right)^2} - \frac{1-\dfrac{\theta}{\alpha_2}}{\left(1-\dfrac{\theta}{\alpha_2}\right)^2 - \left(\dfrac{x}{\alpha_2}\right)^2}\right\},
\end{aligned}$$

and the sign of $g(x)$ coincides with that of

$$\frac{\left(1-\dfrac{\theta}{\alpha_2}\right)^2 - \left(\dfrac{x}{\alpha_2}\right)^2}{1-\dfrac{\theta}{\alpha_2}} - \frac{\left(1+\dfrac{\theta}{\alpha_1}\right)^2 - \left(\dfrac{x}{\alpha_1}\right)^2}{1+\dfrac{\theta}{\alpha_1}} = Ax^2 + B \quad \text{(say)},$$

where

$$A = \frac{1}{\alpha_1(\alpha_1+\theta)} - \frac{1}{\alpha_2(\alpha_2-\theta)},$$

$$B = -\theta\left(\frac{1}{\alpha_1} + \frac{1}{\alpha_2}\right).$$

Let $\theta = M(=0)$. Then we have $A > 0$ and $B = 0$, hence $L(x) - L(-x) > 0$ (i.e., $f(x) > f(-x)$) for $0 < x < \alpha_1$. For $x \geq \alpha_1$, we get $f(x) \geq f(-x)$ directly. Therefore, we have obtained $f(x) \geq f(-x)$ for all $x > 0$, so we can use the M-method. Hence we get $M < m$ and $M < \mu$. Next, let $\theta = \mu$. Since $\mu > M = 0$ from the above, we have $B < 0$. If $A \leq 0$, then we have $f(\mu + x) \leq f(\mu - x)$ for all $x > 0$, but it is a contradiction because μ is the mean. Hence $A > 0$. Therefore, there exists h such that $Ax^2 + B < 0$ for $0 < x < h$, and $Ax^2 + B > 0$ for $x > h$. Further we directly get that $f(\mu + x) - f(\mu - x)$ is continuous and nonnegative. Hence there exist $t \in (h, \alpha_1 - \mu)$ such that $f(\mu + x) < f(\mu - x)$ for $0 < x < t$ and $f(\mu + x) > f(\mu - x)$ for $x > t$. Therefore, we can use the μ-method, so $m < \mu$ and $\mu_3 > 0$. Hence we have obtained that the assertions hold for a Pearson distribution Type $\mathrm{I_L}$. (Note that we can also use the m-method.)

For any other distribution satisfying the assumptions, by a similar way to the above, we can use all of the μ-, m-, and M-methods without getting μ or m concretely, and the assertions follow. $\square$

4. Examples

We shall give some counter-examples in this section.

Let $0 < \varepsilon \leq 3$ and consider a density

$$f_\varepsilon(x) := \begin{cases} 0 & (x \leq 0), \\ Cx^\varepsilon & (0 < x \leq 1), \\ Cx^{-5} & (1 < x), \end{cases}$$

where $C = C_\varepsilon$ is independent of x. Then, apparently f_ε is continuous, unimodal, and $M = 1$. For $n = 0, 1, 2, 3$, we have

$$\begin{aligned}
\mu'_n =& \int_0^\infty x^n P(dx) \\
=& C\left(\int_0^1 x^{n+\varepsilon}dx + \int_1^\infty x^{n-5}dx\right) \\
=& C\left(\frac{1}{1+n+\varepsilon} + \frac{1}{4-n}\right) \\
=& C\frac{5+\varepsilon}{(1+n+\varepsilon)(4-n)}.
\end{aligned}$$

By letting $n = 0$, we get

$$C = \frac{4(1+\varepsilon)}{5+\varepsilon},$$

hence

$$\mu'_n = \frac{4(1+\varepsilon)}{(1+n+\varepsilon)(4-n)}.$$

We have

$$F(x) = \begin{cases} C\int_0^x y^\varepsilon dy = C\dfrac{x^{1+\varepsilon}}{1+\varepsilon} = \dfrac{4x^{1+\varepsilon}}{5+\varepsilon} & (0 \le x \le 1), \\ 1 - \int_x^\infty y^{-5}dy = 1 - \dfrac{C}{4}x^{-4} = 1 - \dfrac{1+\varepsilon}{5+\varepsilon}x^{-4} & (1 < x). \end{cases}$$

Therefore, by solving $F(x) = 1/2$, we get

$$m = \sqrt[1+\varepsilon]{\frac{5+\varepsilon}{8}},$$

hence

$$m < M \quad \text{for } 0 < \varepsilon < 3.$$

By letting $\varepsilon \downarrow 0$, we have

$$\mu'_n \to \frac{4}{(1+n)(4-n)}.$$

Hence

$$\mu = \mu_1' \to \frac{2}{3}(= 0.666\cdots), \quad \mu_2' \to \frac{2}{3}, \quad \mu_3' \to 1, \quad m \to \frac{5}{8}(= 0.625),$$

$$\mu_3 = \mu_3' - 3\mu\mu_2' + 2\mu^3 \to \frac{7}{27}(= 0.259\cdots) > 0.$$

Therefore, for a sufficiently small $\varepsilon > 0$, we have obtained that $m < \mu < M$, so Skew* $X < 0$, but $\mu_3 > 0$ that is, Skew $X > 0$. This is also a counter-example to the mean, median, mode inequality (see Theorem 3.1). Here, by calculations, for a sufficiently small $\varepsilon > 0$, we get

$$1 - F(\mu + x) - F(\mu - x)$$
$$= \begin{cases} 1 - \dfrac{4}{5+\varepsilon}\left\{(\mu + x)^{1+\varepsilon} + (\mu - x)^{1+\varepsilon}\right\} & (0 < x \le 1 - \mu), \\ \dfrac{1}{5+\varepsilon}\left\{(1+\varepsilon)(\mu + x)^{-4} - 4(\mu - x)^{1+\varepsilon}\right\} & (1 - \mu < x \le \mu), \\ 1 - F(\mu + x) & (\mu < x). \end{cases}$$

Since this is negative for $0 < x \le 1 - \mu$, positive for $\mu \le x$, and $\log(1+\varepsilon)(\mu + x)^{-4} - \log 4(\mu - x)^{1+\varepsilon}$ strictly increases for $1 - \mu \le x \le \mu$, we obtain that there exists $t \in (1 - \mu, \mu)$ such that

$$1 - F(\mu + x) - F(\mu - x) \begin{cases} < 0 & (0 < x < t), \\ > 0 & (t < x). \end{cases}$$

Therefore, Condition A is satisfied, but $M > m$ and $M > \mu$ as are announced in Table.

Next, let $\varepsilon = \sqrt{5} - 1(= 1.236\cdots)$. Here we can use the μ-method. We can get it by considering increase and decrease of $\log f_{\sqrt{5}-1}(1 + x) - \log f_{\sqrt{5}-1}(1 - x)$. So we have $m < \mu, M \le \mu$, and $\mu_3 > 0$. By direct calculations, we have

$$\mu_n' = \frac{4\sqrt{5}}{(n + \sqrt{5})(5 - \sqrt{5})} = \frac{4}{(\sqrt{5} + n)(\sqrt{5} - 1)} \quad \text{for } n = 0, 1, 2, 3,$$

hence

$$\mu = \mu_1' = 1, \quad \mu_2' = 3 - \sqrt{5}, \quad \mu_3' = \frac{\sqrt{5}-1}{2},$$

$$m = \sqrt[\sqrt{5}]{\frac{4+\sqrt{5}}{8}} (= 0.8945\cdots), \quad \mu_3 = \frac{7\sqrt{5}-15}{2} (= 0.3262\cdots) > 0.$$

Therefore, $m < M = \mu$ as is announced in Table.

Next, let $\varepsilon = 3$. Here we can use the m-method by a similar way to above. So we have $M \le m < \mu$, and $\mu_3 > 0$. By direct calculations, we have

$$\mu_n' = \frac{16}{16-n^2} \quad \text{for } n = 0, 1, 2, 3,$$

hence

$$\mu = \mu_1' = \frac{16}{15}, \quad \mu_2' = \frac{4}{3}, \quad \mu_3' = \frac{16}{7}, \quad m = 1,$$

$$\mu_3 = \frac{10544}{23625} (= 0.4463\cdots) > 0.$$

Therefore, $M = m < \mu$ as is announced in Table.

Next, let p be a unimodal density such that $M \ne m$, $E(|X|^3) < \infty$, and we can use the m-method. We may assume that $M = -1$ and $m = 0$. Let

$$q(x) = \begin{cases} \dfrac{1}{3} & (|x| \le 1), \\ \dfrac{2-x}{3} & (1 < |x| \le 2), \\ 0 & (\text{otherwise}). \end{cases}$$

This is also a density. For $0 < \varepsilon < 1$, consider a density

$$g_\varepsilon(x) = (1-\varepsilon)p(x) + \varepsilon^2 q(\varepsilon x).$$

Here $E(|X|^3) < \infty$, $M = -1$, $m = 0$, and we can use the m-method. So we have $M < m < \mu$. By letting $\varepsilon \downarrow 0$, we get

$$\mu = (1-\varepsilon)\int_{-\infty}^{\infty} xp(x)dx \to \int_{-\infty}^{\infty} xp(x)dx \in (0,\infty),$$
$$\mu_3' = (1-\varepsilon)\int_{-\infty}^{\infty} x^3p(x)dx \to \int_{-\infty}^{\infty} x^3p(x)dx \in (-\infty,\infty),$$
$$\mu_2' \geq \varepsilon^2\int_{-\infty}^{\infty} x^2q(\varepsilon x)dx = \frac{1}{\varepsilon}\int_{-\infty}^{\infty} x^2q(x)dx \to \infty,$$

hence

$$\mu_3 = \mu_3' - 3\mu\mu_2' + 2\mu^3 \to -\infty.$$

Therefore, for a sufficiently small $\varepsilon > 0$, we have $\mu_3 < 0$ as is announced in Table. This example also shows that even $M < m < \mu$ does not imply $\mu_3 > 0$.

Next, for $0 < \varepsilon < 1$, consider a density

$$h_\varepsilon(x) = \begin{cases} D_\varepsilon x & (0 < x \leq 1), \\ D_\varepsilon(2-x)^\varepsilon & (1 < x \leq 2), \\ 0 & (\text{otherwise}). \end{cases}$$

Here we have $M = 1$, and we can use the M-method. So we have $M < m$ and $M < \mu$. However, by using the μ-method (and the note below Theorem 2.1) to the case $\varepsilon = 0$, which is not unimodal in our definition, we have $m > \mu$ and $\mu_3 < 0$ for a sufficiently small $\varepsilon > 0$ as are announced in Table.

Also note that $\mu_n = 0$ for all odd n does not generally imply that P is symmetric (Ord, 1968).

References

GRONEVELD, A. & MEEDEN, G. (1977). The mode, median, and mean inequality. *Amer. Statist.* **31,** 120–121.

KENDALL, SIR M., STUART, A. & ORD, J. K. (1994). *Kendall's Advanced Theory of Statistics Volume 1: Distribution Theory (Sixth edition).* Edward Arnold, London.

MACGILLIVRAY, H. L. (1981). The mean, median, mode inequality and skewness for a class of densities. *Austral. J. Statist.* **23(2),** 247–250.

ORD, J. K. (1968). The discrete student's t distribution. *Ann. Math. Statist.* **39,** 1513–1516.

PEARSON, K. (1895). Contributions to the mathematical theory of evolution. —II. Skew variation in homogeneous material. *Phil. Trans. Roy. Soc. London.* **A186,** 343–414.

—(1901). Mathematical contributions to the theory of evolution. —X. Supplement to a memoir on skew variation. *Phil. Trans. Roy. Soc. London.* **A197,** 443–459.

—(1916). Mathematical contributions to the theory of evolution. —XIX. Second supplement to a memoir on skew variation. *Phil. Trans. Roy. Soc. London.* **A216,** 429–457.

RUNNENBURG, J. TH. (1978). Mean, median, mode. *Statistica Neerlandica.* **32,** 73–79.

ZWET, W. R. VAN (1964). Convex transformations: A new approach to skewness and kurtosis. *Statistica Neerlandica.* **18,** 433–441.

—(1979). Mean, median, mode II. *Statistica Neerlandica.* **33,** 1–5.