Nonnegativity of Odd Functional Moments of Positive Random Variables with Decreasing Density

Gerold ALSMEYER

Mathematisches Seminar, Universität Kiel, D-24098 Kiel, Germany

Abstract: In this note we give some results on the nonnegativity of odd functional moments of random variables with a decreasing density, more precisely we prove, by purely elementary arguments, $E\phi(X-EX)\geq 0$ for suitable functions ϕ that satisfy $\phi(x)=-\phi(-x)$ for all $x\geq 0$ and random variables $X\geq 0$ with a decreasing Lebesgue density on $(0,\infty)$ or counting density on $\mathbb{N}_0$. The motivation came from a problem recently published in *Statistica Neerlandica*, see Introduction below, concerning a more specialized result.

Keywords: Functional moments, skewness, decreasing density.

1. Introduction and Results

Starting point and motivation for the present article has been Problem 234 in *Statistica Neerlandica* posed by R. Gill:

> A positive, continuously distributed random variable X with finite mean μ and a decreasing density $f(x)$, $x\in(0,\infty)$, is intuitively speaking skewed to the right; hence its coefficient of skewness and more generally all its odd moments should be positive (possibly infinite).
>
> (a) Prove that this is indeed true.
>
> (b) Generalize to the lattice case, i.e. when X is positive integer-valued with a discrete counting density, or give a counterexample.

We first consider the continuous case and prove a far more general result. Denote by $\mathcal{G}_{sl}$ ($\mathcal{G}_{sl}^+$) the class of all odd, increasing functions $\phi:\mathbb{R}\to\mathbb{R}$ of (strict) *superlinear* growth, i.e. $\frac{\phi(t)}{t}$ (strictly) increases on $[0,\infty)$. Clearly, $\mathcal{G}_{sl}\supset\mathcal{G}_c$ and $\mathcal{G}_{sl}^+\supset\mathcal{G}_c^+$ where $\mathcal{G}_c$ ($\mathcal{G}_c^+$) is the class of all odd, increasing ϕ that are (strictly) convex on $[0,\infty)$. Note that $\phi(0+)=\lim_{t\downarrow 0}\phi(t)=0$ holds for every $\phi\in\mathcal{G}_{sl}$. For any such ϕ put further $\Phi(t)=\int_0^t\phi(x)\,dx$ and observe that Φ is even and convex, and that $\Phi(t^{1/2})$ is also convex (its derivative being $\phi(t^{1/2})/2t^{1/2}$). For $\phi\in\mathcal{G}_{sl}^+$, even strict convexity of $\Phi(t^{1/2})$ holds true. These facts will be essential for the proof of Theorem 1.

Theorem 1. *If X is a positive, continuously distributed random variable with finite mean μ and decreasing density $f(x)$, $x\in[0,\infty)$, then*

$$E\phi(X-\mu)\geq 0\quad \textit{for all } \phi\in\mathcal{G}_{sl}. \tag{1}$$

If X does not have a uniform distribution on $[0,2\mu]$, then

$$E\phi(X-\mu)> 0\quad \textit{for all } \phi\in\mathcal{G}_{sl}^+. \tag{2}$$

As a special case, we obtain the solution to Gill's problem in the continuous case:

Corollary 1. *For every X as given in Theorem 1,*

$$E\{sign(X-\mu)|X-\mu|^{\alpha}\} > 0 \quad \text{for all } \alpha > 1 \tag{3}$$

unless X has a uniform distribution on $[0, 2\mu]$ in which all odd moments are 0.

We now turn to the discrete case which is more difficult. The following two counterexamples show that Theorem 1 cannot be copied to this case.

Two counterexamples. (a) Let X be a random variable that satisfies $P(X=0) = P(X=1) = \frac{3}{7}$ and $P(X=2) = \frac{1}{7}$. Then $\mu = EX = \frac{5}{7}$. Consider the function $\phi(t) = sign(t)(|t| - \frac{2}{7})^+$ which clearly belongs to $\mathcal{G}_c$. Then

$$E\phi(X-\mu) \;=\; -\frac{3}{7}\cdot\frac{3}{7} \;+\; 1\cdot\frac{1}{7} \;=\; -\frac{2}{49} \;<\; 0.$$

(b) Let X be a random variable that satisfies $P(X=j) = \frac{7}{24}$ for $j \in \{0,1,2\}$ and $P(X=3) = \frac{3}{24}$. Then $\mu = \frac{5}{4}$. For the function $\phi(t) = sign(t)(|t| - \frac{3}{4})^+ \in \mathcal{G}_c$, we obtain

$$E\phi(X-\mu) \;=\; -\frac{1}{2}\cdot\frac{7}{24} \;+\; 1\cdot\frac{3}{24} \;=\; -\frac{1}{48} \;<\; 0.$$

So we see that neither $\mathcal{G}_{sl}$ nor its subclass $\mathcal{G}_c$ are suitable classes for an extension of Theorem 1 to the discrete case. As a consequence, we now introduce the classes $\mathcal{G}_{ca}$ and $\mathcal{G}_{ca}^+$ consisting of all elements ϕ of $\mathcal{G}_c$, respectively $\mathcal{G}_c^+$, that have a *convex average* $\psi(t) = \frac{\phi(t)}{t}$ on $[0,\infty)$. We will prove

Theorem 2. *Let X be a nonnegative, integer-valued random variable with finite mean μ and decreasing counting density $f(n)$, $n \in \mathbb{N}_0$. Then (1) holds true for all $\phi \in \mathcal{G}_{ca}$, and (2) for all $\phi \in \mathcal{G}_{ca}^+$ unless X is uniform on $\{0,1,...,2\mu-1\}$ (i.e. $\mu \in \mathbb{N}_0$).*

Since $\phi_\alpha(t) = sign(t)|t|^\alpha \in \mathcal{G}_{ca}^+$ for all $\alpha \geq 2$, we conclude in particular

Corollary 2. *For X as given in Theorem 2, (3) holds for all $\alpha \geq 2$ unless X is uniform on $\{0,1,...,2\mu-1\}$ in which case all odd moments are 0.*

Although we have seen that Theorem 1 cannot be right off extended to the discrete case, one may ask for an appropriate extra condition that will do it. This is the content of our final theorem. Let ν be the largest integer less than or equal to μ and $d \overset{\text{def}}{=} \mu - \nu$.

Theorem 3. *For X as given in Theorem 2, suppose that $d \in \{0, \frac{1}{2}\}$, or $d \in (0, \frac{1}{2})$ and*

$$(1-d)f(\nu+1) \;-\; df(\nu) \;\leq\; 0. \tag{4}$$

Then (1) holds for all $\phi \in \mathcal{G}_{sl}$, and (2) for all $\phi \in \mathcal{G}_{sl}^+$ unless X has a uniform distribution.

Note that condition (4) automatically holds if $d = \frac{1}{2}$ where $d = 1 - d$.

A number of earlier publications have dealt with similar problems. Hannan and Pitman(1965), Runnenburg(1978), MacGillivray(1981) and Bélisle(1991) provide sufficient conditions for nonnegative odd central moments in case where f is unimodal. Burton, Keane and O'Brien(1992) consider the discrete case for decreasing f and prove Theorem 2 with $\mathcal{G}_{ca}$ and $\mathcal{G}_{ca}^+$ replaced by the slightly smaller classes $\mathcal{G}_{cd}$ and $\mathcal{G}_{cd}^+$, respectively, consisting of all odd functions ϕ with a (strictly) convex derivative on $[0, \infty)$. Their methods, however, are completely different from ours. Finally, Rösler(1994) obtains similar results by methods based on Choquet theory and convex ordering. We give a few further comments on that approach in our final section.

2. Proofs

Proof of Theorem 1. Since $f(x)$ is decreasing, we may further assume w.l.o.g. f to be right continuous and then define a (possibly infinite) measure F on $(0, \infty)$ by

$$F(x) \;=\; F((0, x]) \;=\; f(0) - f(x) \quad \text{for all } x > 0,$$

hence $F((x, \infty)) = f(x)$. Note that F is degenerate $(= \frac{1}{2\mu}\delta_{2\mu})$ iff X has a uniform distribution on $(0, 2\mu)$ which is from now on excluded because the assertion is trivial in that case. We then have for all $k \geq 1$

$$\begin{aligned} E\phi(X-\mu) \;&=\; \int_0^\infty \phi(x-\mu) f(x)\, dx \;=\; \int_0^\infty \phi(x-\mu) \int_{(x,\infty)} F(dy)\, dx \\ &=\; \int_{(0,\infty)} \int_0^y \phi(x-\mu)\, dx\, F(dy) \;=\; \int_{(0,\infty)} \big(\Phi(y-\mu) - \Phi(\mu)\big)\, F(dy) \end{aligned} \tag{5}$$

which can be finite or $+\infty$. Let us exclude the latter and prove that then the value must be nonnegative. Two cases must be considered:

Case I. $\|F\| = f(0) \stackrel{\text{def}}{=} \lim_{x\downarrow 0} f(x) < \infty$. Then $G = F/\|F\|$ defines a nondegenerate probability distribution on $(0, \infty)$. Let Y be a random variable with that distribution. It follows from (5), $\Phi(t) = \Phi(|t|)$, the convexity of $\Phi(t^{1/2})$ and Jensen's inequality

$$E\phi(X-\mu) \;=\; \|F\| \Big(E\Phi(Y-\mu) - \Phi(\mu)\Big) \;\geq\; \|F\| \left(\Phi\Big(\Big\{E(Y-\mu)^2\Big\}^{1/2}\Big) - \Phi(\mu)\right), \tag{6}$$

and the inequality is strict if $\phi \in \mathcal{G}_{sl}^+$. But a similar calculation as in (5) leads to

$$\frac{1}{\|F\|} \;=\; \int_{(0,\infty)} y\, G(dy) \quad \text{and} \quad \frac{2\mu}{\|F\|} \;=\; \int_{(0,\infty)} y^2\, G(dy),$$

whence

$$E(Y-\mu)^2 \;=\; EY^2 - 2\mu\, EY + \mu^2 \;=\; \mu^2$$

and so the final expression in (6) equals 0.

Case II. $\|F\| = \infty$. In this case we use an approximation argument. Let $c_\infty = \sup\{t : f(t) > 0\}$ and observe that $c_\infty > 2\mu$ because f is decreasing and $f(0) = \infty > f(2\mu)$. For $n \ge 1$, let $h_n(x) = f(1/n)\mathbf{1}_{[0,1/n]}(x) + f(x)\mathbf{1}_{(1/n,\infty)}(x)$ and X_n be a random variable with density $h_n/\|h_n\|_1$, $\|\cdot\|_1$ the usual L_1-Norm. We claim that X_n is stochastically larger than X ($X_n \ge_{st} X$), i.e., $P(X_n > t) \ge P(X > t)$ for all $t \ge 0$. In particular, $EX_n \ge EX = \mu$. To see this, note that $\|h_n\|_1 < 1$ which combined with $f(0) = \infty$ and the monotonicity of f shows that there exists a unique $x_0 \in [0, 1/n]$ such that the density difference $f(x) - h_n(x)/\|h_n\|_1$ is positive on $[0, x_0)$ and nonpositive on $[x_0, \infty)$. As a consequence,

$$\Delta(t) \stackrel{\text{def}}{=} P(X > t) - P(X_n > t) = \int_t^\infty (f(x) - h_n(x)/\|h_n\|_1)\, dx$$

attains its maximum at x_0 which combined with $\Delta(\infty) = \lim_{t\to\infty} \Delta(t) = 0$ gives the asserted conclusion $\Delta(t) \ge 0$ for all $t \ge 0$.

Now consider the upper truncations $h_n^{(c)} \stackrel{\text{def}}{=} h_n\mathbf{1}_{[0,c)}$ and denote by $X_n^{(c)}$ a random variable with distribution $h_n^{(c)}/\|h_n^{(c)}\|_1$. A similar argument as before shows that, for each $n \ge 1$, $X_n^{(c_1)} \le_{st} X_n^{(c_2)} \le_{st} X_n$ whenever $0 \le c_1 \le c_2 \le \infty$. Hence $EX_n^{(c)}$ increases continuously from 0 to $EX_n \ge \mu$ which implies the existence of a maximal $c_n \le c_\infty$ such that $EX_n^{(c_n)} = \mu$. Define $f_n(x) = f(1/n)\mathbf{1}_{[0,1/n]}(x) + f(x)\mathbf{1}_{(1/n,c_n)}(x)$ so that $f_n/\|f_n\|_1$ forms a density of $X_n^{(c_n)}$. Since $X_n^{(c_n)}$ converges in distribution to X, we infer $c_n \to c_\infty$, as $n \to \infty$, and then from the dominated convergence theorem

$$\lim_{n\to\infty} E\phi(X_n^{(c_n)} - \mu) = \lim_{n\to\infty} \int_0^\infty \phi(x-\mu) f_n(x)\, dx/\|f_n\|_1 = E\phi(X-\mu),$$

provided the latter expectation is finite. But the f_n are again decreasing with $f_n(0) = f(1/n) < \infty$, so that $E\phi(X-\mu) \ge 0$ for $\phi \in \mathcal{G}_{sl}$ follows from the first part of the proof.

The strict positivity of $E\phi(X-\mu)$ for $\phi \in \mathcal{G}_{sl}^+$ yields as follows: Put $\psi(t) = \frac{\phi(t)}{t}$, $g_n = f - f_n$ and note that $\int_{(0,\infty)} (x-\mu) g_n(x)\, dx = 0$. Let n be so large that $\frac{1}{n} < \mu$ and $c_n > 2\mu$. It then follows

$$\begin{aligned} E\phi(X-\mu) &= \int_0^\infty \phi(x-\mu) f_n(x)\, dx + \int_0^\infty \phi(x-\mu) g_n(x)\, dx \\ &\ge \int_0^{1/n} \phi(x-\mu) g_n(x)\, dx + \int_{c_n}^\infty \phi(x-\mu) g_n(x)\, dx \\ &> \psi(\mu) \int_0^{1/n} (x-\mu) g_n(x)\, dx + \psi(c_n-\mu) \int_{c_n}^\infty (x-\mu) g_n(x)\, dx \\ &\ge \psi(\mu) \int_0^\infty (x-\mu) g_n(x)\, dx = 0, \end{aligned}$$

where strict monotonicity of ψ has been utilized for the second last inequality ($g_n(x) = f(x) - f(1/n) > 0$ for all $x \in [0, \varepsilon]$, $\varepsilon < \frac{1}{n}$ sufficiently small).

Proof of Theorem 2. The argument working here is different from that in the continuous case and more complicated. We henceforth exclude the trivial case $\mu = 0$ which implies $X = 0$ a.s. In view of Theorem 3, the proof of which will follow below, we may here confine ourselves to the case $d = \mu - \nu \notin \{0, \frac{1}{2}\}$. The proof is divided into three steps.

Step 1. Let Q denote the distribution of $X - \mu$ which can be written as

$$Q \;=\; \sum_{n=0}^{\nu} f(\nu - n)\delta_{-n-d} \;+\; \sum_{n=1}^{q} f(\nu + n)\delta_{n-d}$$

where $q = \sup\{k \geq 1 : f(\nu + k) > 0\}$ (possibly ∞). We will first verify that Q can be decomposed as $Q = U + V$ with U being a zero-mean measure with degenerate left tail at the leftmost mass point of Q, and V being a zero-mean measure which again has a decreasing counting density. More precisely,

$$U \;=\; f(0)\delta_{-\nu-d} + \gamma f(\nu + m)\delta_{m-d} + \sum_{k=m+1}^{q} f(\nu + k)\delta_{k-d}$$

$$\text{and} \quad V \;=\; \sum_{k=0}^{\nu-1} f(\nu - k)\delta_{-k-d} \;+\; \sum_{k=1}^{m-1} f(\nu + k)\delta_{n-d} \;+\; (1-\gamma) f(\nu + m)\delta_{m-d}$$

where $\gamma \in [0, 1)$, $m \geq \nu$ are uniquely determined by the requirement $\int x\, U(dx) = \int x\, V(dx) = 0$. In case $d \in (\frac{1}{2}, 1)$ even $m \geq \nu + 1$ holds. For a proof of this note first that there is clearly a uniquely determined finite $m \leq q$ such that

$$-f(0)(\nu + d) + \sum_{k=m}^{q} f(\nu + k)(k - d) \;>\; 0 \;\geq\; \sum_{k=m+1}^{q} f(\nu + k)(k - d)$$

which in turn implies the existence of a unique $\gamma \in [0, 1)$ such that

$$-f(0)(\nu + d) + \gamma f(\nu + m)(m - d) + \sum_{k=m+1}^{q} f(\nu + k)(k - d) \;=\; 0.$$

This settles the existence of γ, m such that U, V are both measures with mean 0. Moreover, V evidently has again a decreasing counting density on its support $\{-\nu + 1 - d, ..., -d, 1 - d, ..., m - d\}$. But for having mean zero V must possess at least equally many mass points left and right of 0 which implies $m \geq \nu$. If $d \in (\frac{1}{2}, 1)$, it must have even one more mass point right of 0, i.e., $m \geq \nu + 1$.

Finally, it is important to point out for Step 3 further below that $V/\|V\|$ forms indeed a probability measure of exactly the same type as Q (mean 0, decreasing counting density), but with a support left of 0 shrinked by one point. Consequently, a simple backward induction leads to $Q = \sum_{k=0}^{\nu} U_k$ with $U_0 = U$ and each U_k having a degenerated left tail at $-\nu + k - d$ and mean 0. We omit further details as being obvious.

Step 2. Let Y be a random variable with distribution $U/\|U\|$, hence $EY = 0$. We next consider the special element $\phi_0(t) = sign(t)t^2$ of $\mathcal{G}_{ca}$ and prove

$$E\phi_0(Y) > 0.$$

W.l.o.g. suppose $E\phi_0(Y) < \infty$.

Let first $d \in (0, \frac{1}{2})$ and $m \geq \nu + 1$, in which case $k - d \geq \nu + 1 - d > \nu + d$. Consequently,

$$\begin{aligned} \|U\| E\phi_0(Y) &= -f(0)(\nu+d)^2 \;+\; \gamma f(\nu+m)(m-d)^2 \\ &+ \sum_{k\geq m+1} f(\nu+k)(k-d)^2 \\ &> (\nu+d)\Big(-f(0)(\nu+d) \;+\; \gamma f(\nu+m)(m-d) \\ &+ \sum_{k\geq m+1} f(\nu+k)(k-d)\Big) \\ &= \|U\|(\nu+d)EY \;=\; 0. \end{aligned}$$

A similar estimation settles the proof in case $d \in (\frac{1}{2}, 1)$ and $m \geq \nu + 2$.

The remaining two cases need greater care. Consider first $d \in (0, \frac{1}{2})$ and $m = \nu$. Writing out the equation $\|U\|EY = E(X - \mu) = 0$ we obtain

$$\begin{aligned} (1-\gamma)f(2\nu)(\nu-d) &= df(\nu) \;+\; \sum_{k=1}^{\nu-1}\Big(f(\nu-k)(k+d) - f(\nu+k)(k-d)\Big) \\ &> df(\nu) \;+\; f(2\nu-1)(2\nu-2)d \;\geq\; f(2\nu-1)(2\nu-1)d, \end{aligned}$$

where the latter two inequalities use the monotonicity of f. So we infer $\gamma < 1 - \frac{(2\nu-1)d}{\nu-d}$. By using this and $\sum_{k\geq\nu+1} f(\nu+k)(k-d) = f(0)(\nu+d) - \gamma f(2\nu)(\nu-d)$, we are led to

$$\begin{aligned} \|U\|E\phi_0(Y) &\geq -f(0)(\nu+d)^2 \;+\; \gamma f(2\nu)(\nu-d)^2 \;+\; (\nu+1-d)\sum_{k\geq\nu+1} f(\nu+k)(k-d) \\ &= f(0)(\nu+d)(1-2d) \;-\; \gamma f(2\nu)(\nu-d) \\ &> f(2\nu)\Big((\nu+d)(1-2d) \;-\; \Big(1 - \frac{(2\nu-1)d}{\nu-d}\Big)(\nu-d)\Big) \\ &= f(2\nu)\Big((\nu+d)(1-2d) - (\nu-d) + (2\nu-1)d\Big) \\ &= f(2\nu)d(1-2d) \;\geq\; 0, \end{aligned}$$

which is the desired conclusion.

For $d \in (\frac{1}{2}, 1)$ and $m = \nu + 1$ we proceed similarly. Here $\|U\|EY = E(X - \mu) = 0$ gives

$$\begin{aligned} (1-\gamma)f(2\nu+1)(\nu+1-d) &= \sum_{k=1}^{\nu}\Big(f(\nu+1-k)(k-1+d) - f(\nu+k)(k-d)\Big) \\ &> f(2\nu+1)\nu(2d-1), \end{aligned}$$

the latter inequality by using $k-1+d > k-d$ for each $k = 1, ..., \nu$ and the monotonicity of f. So we infer $\gamma < 1 - \frac{\nu(2d-1)}{\nu+1-d}$. By combining this with $\sum_{k\geq\nu+2} f(\nu+k)(k-d) = f(0)(\nu+d) - \gamma f(2\nu+1)(\nu+1-d)$, we obtain

$$
\begin{aligned}
\|U\| E\phi_0(Y) \;\geq\; & -f(0)(\nu+d)^2 \;+\; \gamma f(2\nu+1)(\nu+1-d)^2 \\
& +\; (\nu+2-d) \sum_{k\geq\nu+2} f(\nu+k)(k-d) \\
=\; & f(0)(\nu+d)(2-2d) \;-\; \gamma f(2\nu+1)(\nu+1-d) \\
>\; & f(2\nu+1)\Big((\nu+d)(2-2d) \;-\; \Big(1-\frac{\nu(2d-1)}{\nu+1-d}\Big)(\nu+1-d)\Big) \\
=\; & f(2\nu+1)\Big((\nu+d)(2-2d) - (\nu+1-d) + (2d-1)\nu\Big) \\
=\; & f(2\nu+1)(1-d)(2d-1) \;\geq\; 0,
\end{aligned}
$$

which again is the desired conclusion.

Step 3. The final step uses a simple, but useful "trick". Choose an arbitrary $\phi \in \mathcal{G}_{ca}$ with $E\phi(X-\mu) < \infty$, let $\psi(t) = \frac{\phi(t)}{t}$ and note that $\psi(0) = \lim_{t\to 0} \frac{\phi(t)}{t} = 0$. In view of the decomposition of Q mentioned at the end of Step 1 it now suffices to prove that $E\phi(Y) \geq 0$ for every zero-mean random variable Y with a degenerate left tail, i.e. $p \overset{\text{def}}{=} P(Y<0) = P(Y = -y)$ for some $y > 0$, and finite, positive $E\phi_0(Y)$. Let $\beta \overset{\text{def}}{=} EY^+ = EY^-$ and notice $\beta = py$. We obtain

$$
\begin{aligned}
E\phi(Y) \;=\; & E\phi(Y^+) - E\phi(Y^-) \;=\; EY^+\psi(Y^+) - EY^-\psi(Y^-) \\
=\; & EY^+\psi(Y^+) - py\psi(y) \;=\; \beta\Big(\int \psi(Y^+)Y^+\,dP/\beta\Big) - \beta\psi(y) \\
\geq\; & \beta\psi(\beta^{-1}E(Y^+)^2) - \beta\psi(\beta^{-1}E(Y^-)^2),
\end{aligned}
$$

where the final line follows by Jensen's inequality ($Y^+(\omega)P(d\omega)/\beta$ is a probability measure and ψ convex) and the identity $y = \frac{py^2}{py} = E(Y^-)^2/\beta$. But $E\phi_0(Y) > 0$ means $E(Y^-)^2 < E(Y^+)^2$ and ψ is increasing on $[0,\infty)$. Hence $E\phi(Y) \geq 0$ and even > 0 if ψ is strictly increasing. The proof of Theorem 2 for the case $d \notin \{0, \frac{1}{2}\}$ is herewith complete.

Proof of Theorem 3. Let $\phi \in \mathcal{G}_{sl}$. We keep the notation from before and first consider the case $d = \mu - \nu \in \{0, \frac{1}{2}\}$. For $d = 0$, i.e. $\mu = \nu$, we easily infer from the monotonicity of f and of $\psi(t) = \frac{\phi(t)}{t}$

$$
\begin{aligned}
E\phi(X-\mu) \;=\; & \sum_{n=1}^{\nu} \psi(n)(nf(\nu+n) - nf(\nu-n)) \;+\; \sum_{n\geq 2\nu+1} \psi(n-\nu)(n-\nu)f(n) \\
\geq\; & \psi(\nu)\sum_{n=1}^{\nu} n(f(\nu+n) - f(\nu-n)) \;+\; \psi(\nu+1)\sum_{n\geq 2\nu+1}(n-\nu)f(n) \\
\geq\; & \psi(\nu)\sum_{n\geq 0}(n-\nu)f(n) \;=\; 0.
\end{aligned}
$$

and the inequality is obviously strict if $\phi \in \mathcal{G}_{sl}^+$ and X is not uniform on $\{0, ..., 2\mu + 1\}$. A similar symmetry argument gives the result for $d = \frac{1}{2}$. So we need not supply the details again.

Now let $d \in (0, \frac{1}{2})$ and suppose condition (4). Since f is decreasing, the latter implies

$$
\begin{aligned}
&(k+1-d)f(\nu+1+k) - (k+d)f(\nu-k) \\
&\qquad = k(f((\nu+1+k) - f(\nu-k)) \ + \ (1-d)f(\nu+1+k) - df(\nu-k) \\
&\qquad \le (1-d)f(\nu+1) - df(\nu) \ \le 0
\end{aligned}
\tag{7}
$$

for all $0 \le k \le \nu$. We then conclude

$$
\begin{aligned}
E\phi(X-\mu) \ &= \ \sum_{n=0}^{\nu} \Big(\phi(n+1-d)f(\nu+n+1) - \phi(n+d)f(\nu-n) \Big) \\
&\quad + \sum_{n>\nu} \phi(n+1-d)f(\nu+n+1) \\
&\ge \sum_{n=0}^{\nu} \psi(n+d) \Big((n+1-d)f(\nu+n+1) - (n+d)f(\nu-n) \Big) \\
&\quad + \ \psi(\nu+1-d) \sum_{n>\nu} (n+1-d)f(\nu+n+1) \\
&\ge \ \psi(\nu+d) \sum_{n\ge 0} (n-\mu)f(n) \ = \ 0,
\end{aligned}
\tag{8}
$$

where $d < 1-d$ and the monotonicity of ψ have been used for the first inequality in (8) and both these facts combined with (7) for the second one.

4. Concluding Remarks

All previous proofs are basically a combination of elementary computations with Jensen's inequality. A different approach may be based upon first reducing the problem in the following way. Each of the classes $\mathcal{G}_{sl}, \mathcal{G}_c, \mathcal{G}_{ca}$ and $\mathcal{G}_{cd}$ that have been mentioned before is easily seen to form a convex cone for which Choquet theory tells us that each element ϕ can be written as an integral of its extremal elements with respect to some measure (depending on ϕ). For the given classes these integral representations are obtained by simple partial integration. Indeed, for $\phi \in \mathcal{G}_{sl}$ we have

$$
\phi(t) \ = \ \int_{[0,\infty)} \mathbf{1}_{[x,\infty)}(t)\, t \; Q_\phi(dx), \quad t \ge 0,
$$

where Q_ϕ is the measure defined through $Q_\phi([0,t]) = \frac{\phi(t)}{t}$. Hence the extremal elements of $\mathcal{G}_{sl}$ are all odd functions ϕ_x of the form $\phi_x(t) = \mathbf{1}_{[x,\infty)}(t)t$ for $t \ge 0$. Similarly, one finds for $t \ge 0$

$$
\begin{aligned}
\phi(t) \ &= \ \int_{[0,\infty)} (t-x)^+ \; Q_\phi(dx) \quad \text{for } \phi \in \mathcal{G}_c; \\
\phi(t) \ &= \ \int_{[0,\infty)} t(t-x)^+ \; Q_\phi(dx) \quad \text{for } \phi \in \mathcal{G}_{ca}; \\
\text{and} \quad \phi(t) \ &= \ \int_{[0,\infty)} ((t-x)^+)^2 \; Q_\phi(dx) \quad \text{for } \phi \in \mathcal{G}_{cd},
\end{aligned}
$$

where Q_ϕ is a suitable measure that differs from line to line.

By using such an integral representation in $E\phi(X-\mu)$, another partial integration shows that it suffices to prove our results only for the extremal ϕ's from the respective class under consideration. E.g. for $\phi \in \mathcal{G}_{sl}$, we have

$$E\phi(X-\mu) \;=\; \int_{[0,\infty)} E\phi_x(X-\mu)\; Q_\phi(dx).$$

This reduction argument has been used by Rösler(1994) combined with a further reduction with respect to the distribution F, say, of $X-\mu$. Namely, he considers the class $\mathcal{G}_c$ for which it suffices to prove $E\phi(X-\mu) \geq 0$ only for those distributions F that are minimal in convex ordering. These distributions can be determined, in the continuous as well as in the discrete case. Note that F is less than G in convex ordering ($F \leq_c G$) if $\int \phi\, dF \leq \int \phi\, dG$ for all $\phi \in \mathcal{G}_c$.

Acknowledgement

The author wishes to thank Uwe Rösler for many stimulating discussions and also for showing him an early draft of his related article cited below.

References

Bélisle, C.(1991)
— Odd central moments of unimodal distributions. *Statist. Probab. Letters* **12**, 97-107.
Burton, R.M., Keane, M. and O'Brien, G.L.(1992)
— Positivity of central moments for lattice distributions with decreasing densities. *Preprint.*
Gill, R. (1989)
— Problem 234. *Statistica Neerlandica* **43**, 66.
Hannan, J. and Pitman, E.J.G.(1965)
— Symmetric unimodal variables conditioned to upper tails have positive central moments. *Tech. Report, Dept. of Probab. and Statist., Michigan State Univ. (East Lansing, MI).*
MacGillivray, H.L.(1981)
— The mean, median, mode inequality and skewness for a class of densities. *Austral. J. Statist.* **23**, 247-250.
Rösler, U.(1994)
— Distributions slanted to the right. To appear in *Statistica Neerlandica.*
Runnenburg, J.T.(1978)
— Mean, median, mode. *Statistica Neerlandica* **32**, 73-79.