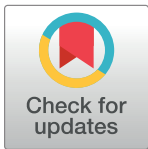RESEARCH ARTICLE

# Bounded distributions place limits on skewness and larger moments

**David J. Meer** ⓘ*, **Eric R. Weeks** ⓘ

Department of Physics, Emory University, Atlanta, Georgia, United States of America

* dmeer@emory.edu

## Abstract

Distributions of strictly positive numbers are common and can be characterized by standard statistical measures such as mean, standard deviation, and skewness. We demonstrate that for these distributions the skewness $D_3$ is bounded from below by a function of the coefficient of variation (CoV) $\delta$ as $D_3 > \delta - 1/\delta$. The results are extended to any distribution that is bounded with minimum value $x_{\min}$ and/or bounded with maximum value $x_{\max}$. We build on the results to provide bounds for kurtosis $D_4$, and conjecture analogous bounds exists for higher statistical moments.

## 1 Introduction

One often considers a probability function $P(x)$ of a random variable $X$. Distributions of $P(x)$ are characterized by quantities such as mean, median, standard deviation, and skewness [1–3]. For a continuous random variable, $X$, $P(x)$ is the probability density of finding a value $x$ in the range $(x, x + dx)$. For a discrete random variable, $P(x)$ is a discrete probability distribution which assigns a probability $p_i$ to each potential value $x_i$. The skewness is a measure of the asymmetry of a distribution [1]. While there are several possible definitions of skewness [4], a common definition depends on the third moment of the distribution compared to the second moment [5, 6]. In particular, one can define the $n$th central moment for continuous or $N$ discreet variables as [7]

$$
\begin{aligned}
m_n = \langle (x - \mu)^n \rangle &= \int_{-\infty}^{+\infty} P(x)(x - \mu)^n dx \\
&= \sum_{i=1}^{N} p_i (x_i - \mu)^n
\end{aligned}
\tag{1}
$$

where $\mu$ is the mean of the distribution and $\sqrt{m_2} = \sigma$ is the standard deviation. The standardized moments $D_n$ are defined as:

$$
D_n = \frac{m_n}{m_2^{n/2}}.
\tag{2}
$$

We define skewness as the third standardized moment, $D_3$. This definition for skewness has the advantage that it is dimensionless. It also has the useful property that distributions $P_1(x)$

**Abbreviations:** $a_+$, value of large size in a bidisperse sample; $a_-$, value of small size in a bidisperse sample; $D_n$, $n$th standardized moment; $M_n$, convenient function of $D_n$; $m_n$, $n$th central moment; $p$, population fraction of small size in a bidisperse sample; $P_r$, $P_s$, two unique distributions; $P_t$, combination of $P_r$ and $P_s$ in some proportionality; $q$, population fraction of large size in a bidisperse sample; $x_{\text{extr}}$, the extreme values of $x$, both the minimum and maximum.; $\delta$, coefficient of variation (CoV) = $\sigma/\mu$; $\Delta_+$, relative difference of large size from mean; $\Delta_-$, relative difference of small size from mean; $\Delta'_+$, difference of large size from mean; $\Delta'_-$, difference of small size from mean; $\eta$, size ratio of bidisperse sizes; $\mu$, mean; $\sigma$, standard deviation.

and $P_2(x) = cP_1(cx + d)$ have the same skewness for $c > 0$ and any $d$ [4]. Pearson [8] derived an upper boundary on the skewness:

$$D_3^2 \le m_4/m_2^2 - 1. \tag{3}$$

Alternate derivations of this result are also in the literature [9, 10]. This applies for all distributions.

Distributions of strictly positive numbers are often relevant: numbers of objects, sizes of objects such as Fig 1, ages of people, prices, barometer measurements, *etc.* Such distributions have only non-negative support; one can more broadly consider distributions with bounded support, with boundaries $x_{\text{min}}$ and/or $x_{\text{max}}$, generically $x_{\text{bound}}$. Smołalski [11] worked out upper and lower bounds on the skewness that applies for distributions with bounded support:

$$D_{3\text{min,max}} = \delta_{\text{min,max}} - \frac{1}{\delta_{\text{min,max}}} \tag{4}$$

with $\delta_{\text{min}} = \sigma/(\mu - x_{\text{min}})$ to determine $D_{3\text{min}}$ and $\delta_{\text{max}} = \sigma/(x_{\text{max}} - \mu)$ to determine $D_{3\text{max}}$.

In this paper, we present an alternative derivation for these skewness bounds. Smołalski's derivation relies on the argument that achieving the extrema of skewness requires a bidisperse distribution. We mathematically prove that this is indeed the case in Section 2. Smołalski then uses Lagrange multipliers to derive Eq (4); here, we use calculus to derive this equation and extend it to all real bounds. Our method also applies to higher order standardized moments, for which we find similar bounds in Section 3. We state the bounds, show when their behavior can be used to find the maximum or minimum standardized moment, $D_{n_{\text{extr}}}$, and conjecture that these extrema apply to all distributions, not just bidisperse.

We are treating just the value of skewness corresponding to the parent distribution, rather than the sample skewness based on a finite number of samples which has different limits, see [12]. Note also that there are other definitions of skewness, for example that use the median of the distribution as part of the calculation [1], for which other limits exist [13–15].
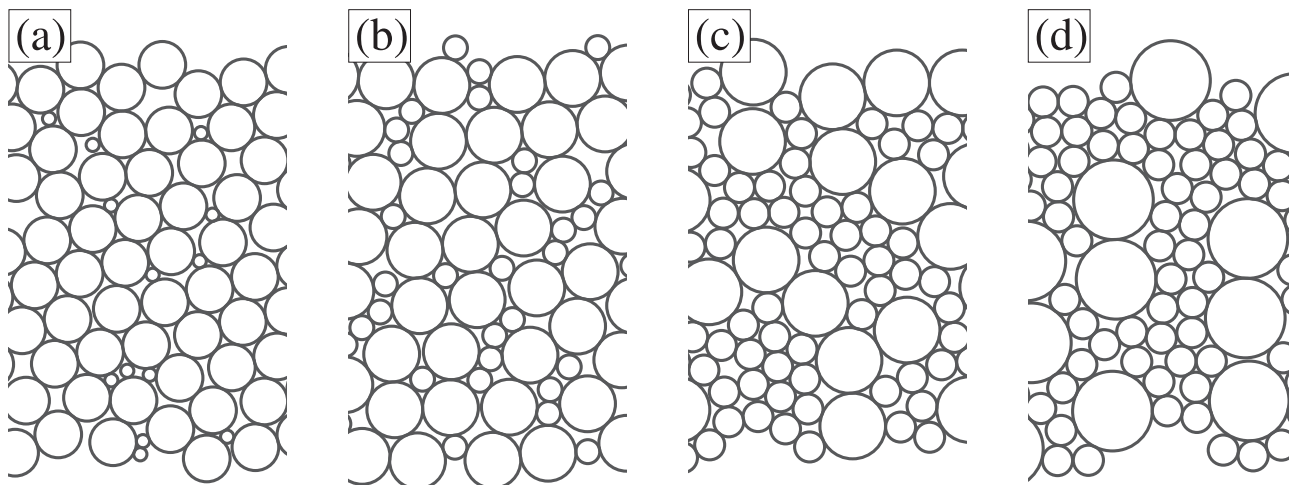


**Fig 1. Examples of circles with random bidisperse distributions of diameters with (a) CoV $\delta = 0.4$, skewness $D_3 = -1$; (b) $\delta = 0.4$, $D_3 = 0$; (c) $\delta = 0.4$, $D_3 = 1$; and (d) $\delta = 0.4$, $D_3 = +3$.** Decreasing the skewness requires the small circles becoming even smaller (compared to the mean size), as well as decreasing their frequency of occurrence. The decreasing size reaches its natural limit when the small particles have zero size at $S = -2.1$ for $\delta = 0.4$, as predicted in Eq (4).

## 2 Results for skewness

We begin in the lowest order nontrivial case $n = 3$, replicating Smołalski's skewness results. A distribution function with a low value of skewness has small values which rarely occur, for example the smallest circles seen in Fig 1a. A distribution with a high value of skewness is the opposite situation, where the large values rarely occur, for example the largest circles seen in Fig 1d. For a distribution $P(x)$ with only non-negative support, the largest possible values of $x$ are unbounded, but the lowest possible values are bounded by zero. Thus, it makes intuitive sense that the skewness will have a minimum possible value.

Our derivation will proceed by first considering bidisperse distributions with nonnegative support and showing that for a fixed $\delta$, the distribution with one value equal to zero achieves the lowest possible skewness. We then show taking two distributions obeying Eq (4) and considering a weighted sum will result in a new distribution that also obeys Eq (4). Next, we argue that any continuous distribution can be approximated by an appropriately weighted sum of bidisperse distributions. In Sec. 2.4 we will conclude by generalizing from distributions with non-negative support to distributions with arbitrary bounds, including those with $\mu \leq 0$.

### 2.1 Skewness for bidisperse distributions

We start by considering a bidisperse distribution, $P(x)$ which takes on values $a_+$, $a_-$ with probabilities $q$, $p = 1 - q$. Following [16], we define the ratio

$$\eta = a_+ / a_- \tag{5}$$

and focus on $q$ as another important variable describing the distribution. The meaning of the subscripts in $a_+$ and $a_-$ is the former is the value larger than the mean $\mu$ and the latter is smaller than $\mu$, respectively. Knowing the mean $\mu$ allows us to relate these quantities as

$$
\begin{aligned}
a_+ &= \eta\mu/(1 - q + \eta q) \\
a_- &= \mu/(1 - q + \eta q).
\end{aligned}
\tag{6}
$$

Note that a bidisperse distribution with a given $(\eta, q)$ is equivalent to a distribution with $(1/\eta, 1 - q)$ with swapped $a_+$ and $a_-$. A key concept which we will use for much of this derivation is that in addition to the mean $\mu$, in general knowing any other two quantities related to the distribution will uniquely determine the distribution. Those two quantities could be the values $a_+$ and $a_-$; they could be $\eta$ and $q$ as per Eq (6). Usefully, they can also be the standard deviation and skewness. Thus, we will show that a distribution with $a_-$ achieving the minimum possible value ($a_- = 0$) is one where the skewness $D_3$ achieves its minimum value.

Given a bidisperse distribution defined as above, the standard deviation $\sqrt{m_2} = \sigma$ and skewness $D_3$ are then expressed as

$$
\begin{aligned}
\sqrt{m_2} = \sigma &= \left((1 - q)(a_- - \mu)^2 + q(a_+ - \mu)^2\right)^{1/2} \\
D_3 &= \frac{(1 - q)(a_- - \mu)^3 + q(a_+ - \mu)^3}{m_2^{3/2}}.
\end{aligned}
\tag{7}
$$

While $P(x)$ could be a distribution of a quantity with dimensions (such as a probability distribution of weights), our goal is to understand the non-dimensional skewness $D_3$. Thus, rather than considering $\sigma$ which has dimensions of $x$, we will use the non-dimensional quantity

"coefficient of variation," (CoV) defined as:

$$\delta = \sigma/\mu. \tag{8}$$

Here we use the symbol $\delta$ and later in this manuscript we will generalize this symbol beyond the specific meaning of CoV. We can use Eq (6) to eliminate $a_+$ and $a_-$ from $m_2$ and $D_n$, resulting in

$$\frac{\sqrt{m_2}}{\mu} = \delta \quad = \quad \frac{(\eta - 1)\sqrt{q - q^2}}{1 - q + \eta q} \tag{9}$$

$$D_3 \quad = \quad \frac{2q - 1}{\sqrt{q - q^2}}. \tag{10}$$

These require $\eta > 1$. Eqs (9, 10) can be inverted to provide expressions for $q$ and $\eta$ in terms of $\delta$ and $D_3$. We include the substitution $M_3 = \sqrt{4 + D_3^2}$ which will be a reoccurring term:

$$q \quad = \quad \frac{\pm D_3 + M_3}{2M_3} \tag{11}$$

$$\eta \quad = \quad \frac{2 - \delta(\pm D_3 - M_3)}{2 - \delta(\pm D_3 + M_3)}. \tag{12}$$

These two equations give rise to two branches of solutions depending on whether the + or − is taken in each equation. Inspection shows that the negative sign in Eqs (11 and 12) arrives back at the classical definition of skewness, whereas the positive branch has no significance. For the remainder of our consideration of $D_3$, we will use the negative branch of the solutions and drop the ± symbol. We continue and calculate the two possible values according to Eq (6):

$$a_+ \quad = \quad \mu\left(1 + \frac{\delta}{2}(D_3 + M_3)\right) \tag{13}$$

$$a_- \quad = \quad \mu\left(1 + \frac{\delta}{2}(D_3 - M_3)\right). \tag{14}$$

Using Eq (14), we can do a straightforward calculation for the minimum possible skewness $D_3(\delta)$ for bidisperse distributions with $a_+$, $a_- \geq 0$. A distribution with a low skewness is one that has a small amount of small numbers: and the smallest number we can get for a distribution of strictly non-negative numbers is zero. Thus, to find the limit on skewness, we solve Eq (14) for $a_- = 0$. This also implies $a_+ = \mu/(1 - q)$. Solving for $D_3$ when $a_- = 0$ in Eq (14) lets us solve for $D_{3\min}$:

$$D_{3\min} = \delta - \frac{1}{\delta}. \tag{15}$$

For example, this gives values $D_{3\min} = -2.1$ for $\delta = 0.4$, and $D_{3\min} = 0$ for $\delta = 1$.

## 2.2 A bidisperse distribution with $a_{\min} > 0$ increases $D_3$

For a fixed value of $\delta$, if the minimum value of the distribution $a_-$ is larger than zero, then $D_3$ will increase. This is not straightforward to see from the equations above, but an alternate

formulation will work. Define:

$$
\begin{aligned}
\Delta'_+ &= a_+ - \mu > 0 \\
\Delta'_- &= \mu - a_- > 0
\end{aligned}
\tag{16}
$$

Using Eq (6) we can factor out $\mu$ and arrive at normalized definitions of $\Delta_{+,-} = \Delta'_{+,-}/\mu$ We then have the probability of $a_+$ being

$$
q = \frac{\Delta_-}{\Delta_- + \Delta_+}.
\tag{17}
$$

We can then get $\delta$ using

$$
\delta^2 = (1-q)\Delta_-^2 + q\Delta_+^2 = \Delta_- \Delta_+.
\tag{18}
$$

Given that we wish to keep $\delta$ constant, we can thus use $\Delta_+ = \delta^2/\Delta_-$ to eliminate $\Delta_+$, leading to

$$
q = \frac{1}{1 + \Delta_-^2/\delta^2}.
\tag{19}
$$

Now consider the third moment of the distribution $m_3$:

$$
\begin{aligned}
m_3 &= \mu^3((1-q)\Delta_-^3 - q\Delta_+^3) \\
&= \mu^3\delta^2(\Delta_- - \Delta_+) \\
&= \mu^3\delta^2\left(\frac{\delta^2}{\Delta_-} - \Delta_-\right).
\end{aligned}
\tag{20}
$$

The partial derivative of $m_3$ with respect to $\Delta_-$ holding $\delta$ constant is

$$
\left(\frac{\partial m_3}{\partial \Delta_-}\right)_\delta = -\mu^3\left(\frac{\delta^4}{\Delta_-^2} + \delta^2\right) < 0.
\tag{21}
$$

Increasing $\Delta_-$ always decreases $m_3$, assuming we keep $\delta$ constant and $\mu$ positive. Likewise, decreasing $\Delta_-$ (making $a_-$ larger than zero) will always increase $m_3$. Thus, making $a_-$ larger than zero must increase the skewness $D_3$. This proves that for the bidisperse distribution with a fixed $\delta$, Eq (15) is indeed the lowest possible skewness.

## 2.3 Generalizations of skewness results

Suppose we have two separate distributions $P_r(x)$ and $P_s(x)$ both with mean $\mu$ and both satisfying the bound of Eq (15). We wish to show that any combination of these two distributions, $P_t(x) = \alpha P_r(x) + (1-\alpha)P_s(x)$ (with $0 \leq \alpha \leq 1$), also satisfies Eq (15). Given that the means are identical, it is straightforward that $\delta_t^2 = \alpha\delta_r^2 + (1-\alpha)\delta_s^2$ and also $m_{3,t} = \alpha m_{3,r} + (1-\alpha)m_{3,s}$. As $D_3 = m_3/m_2^{3/2}$, we can rewrite the bound on skewness Eq (15) as

$$
m_{3,\min} \geq \mu^3(\delta^4 - \delta^2)
\tag{22}
$$

Given that both $P_r$ and $P_s$ satisfy this constraint, we have

$$
\begin{aligned}
m_{3,r} &\geq \mu^3(\delta_r^4 - \delta_r^2) \\
m_{3,s} &\geq \mu^3(\delta_s^4 - \delta_s^2),
\end{aligned}
\tag{23}
$$

and thus

$$
\begin{aligned}
m_{3,t} &= \alpha m_{3,r} + (1-\alpha) m_{3,s} \\
&\geq \mu^3(\alpha(\delta_r^4 - \delta_r^2) + (1-\alpha)(\delta_s^4 - \delta_s^2)) \\
&= \mu^3(\alpha\delta_r^4 + (1-\alpha)\delta_s^4 - \delta_t^2),
\end{aligned}
\tag{24}
$$

where the last line uses the expression for $\delta_t^2$ introduced above. Next, note that

$$
\begin{aligned}
\delta_t^4 &= (\alpha\delta_r^2 + (1-\alpha)\delta_s^2)^2 \\
&= \alpha^2\delta_r^4 + 2\alpha(1-\alpha)\delta_r^2\delta_s^2 + (1-\alpha)^2\delta_s^4.
\end{aligned}
\tag{25}
$$

On the right-hand side of Eq (24), add $\mu^3\delta_t^4$ and subtract the right-hand side of Eq (25):

$$
\begin{aligned}
m_{3,t} \geq \ \mu^3(&\alpha\delta_r^4 + (1-\alpha)\delta_s^4 - \delta_t^2 + \delta_t^4 \\
&-\alpha^2\delta_r^4 - 2\alpha(1-\alpha)\delta_r^2\delta_s^2 - (1-\alpha)^2\delta_s^4)
\end{aligned}
\tag{26}
$$

Every term without $\delta_t$ on the right-hand side can be combined as $\mu^3\alpha(1-\alpha)(\delta_r^2 - \delta_s^2)^2$ which is always non-negative, so thus

$$
m_{3,t} \geq \mu^3(\delta_t^4 - \delta_t^2),
\tag{27}
$$

proving that the combined distribution function $P_t(x)$ must satisfy Eq (4) if the two original distributions satisfy that bound.

Finally, we need to generalize from the bidisperse distribution to any distribution. Following [9], we observe that any continuous distribution with some fixed $\mu = \mu_0$ can be approximated by a discrete distribution with values $a_i$ and probabilities $p_i$ and $\mu = \mu_0$. Rohatgi and Székely then proved that any such discrete distribution can be decomposed into a sum of discrete distributions with two values and $\mu = \mu_0$, that is, the bidisperse distributions that we have been considering (see also S1 Appendix). In the previous paragraph, we have shown that sums of distributions satisfy the bound. Thus, we have proven that Eq (4) holds for any distribution $P(x)$ of strictly non-negative values of $x$.

## 2.4 Distributions bounded by $x_{\min}$ or $x_{\max}$

We have considered distributions $P(x)$ for which $x \geq 0$. By rescaling the distribution, we can enforce any value of $\mu$ we would like. However, this comes at the expense of potentially running into our bounds. For example, you cannot have some $\mu \leq 0$ without a minimum less than or equal to zero. When some values of $x$ are below 0, we cannot simply rescale by a constant multiple to enforce the bounds. Of course, an additive constant would fix a distribution and make it non-negative. As noted in the introduction, this also leaves $D_3$ unchanged: consider $P(x)$ and $P'(x) = P(x - d)$. $\mu' = \mu + d$ but as the moments are defined as $\langle (x - \mu)^n \rangle$, $m_2$ and $m_3$ are unchanged by this shift, and thus $D_3$ does not change.

Similarly, we also note that $\lim_{\mu \to 0^+}(a_+, a_-, \eta) = \lim_{\mu \to 0^-}(a_+, a_-, \eta)$. This limit can be calculated directly by multiplying by $\frac{\mu}{\mu}$ in Eq (12) and distributing the $\mu$ factor in Eqs (13 and 14), leaving us with just $\sqrt{m_2}$ where there was previously $\delta$. Therefore, we do not have to be concerned with means approaching zero.

Now consider the general case of a distribution $P(x)$ bounded by $x_{\max}$ from below and with a mean $\mu$ which might be zero. Let us assume $P(x)$ has a nontrivial domain, which is to say, it is not a distribution which is only nonzero at one value (which would thus be $\sigma = 0$, $D_3 = 0$). The transformed distribution $P'(x) = P(x + x_{\min})$ has mean $\mu' = \mu - x_{\min}$. This transformed

distribution now is nonzero only for $x \geq 0$, so is one of the distributions we considered above, and since the distribution has a nontrivial domain, $\mu' > 0$ must be true. Therefore, we have:

$$\delta = \sigma/(\mu - x_{\min}). \tag{28}$$

That is, $\delta$ depends on the standard deviation $\sigma$ and mean $\mu$ of the original distribution $P(x)$, with the additional correction of subtracting $x_{\min}$, at which point we can use Eq (15) to find $D_{3\min}$.

The other interesting case is a distribution bounded by $x_{\max}$ from above. Considering $P''(x) = P(-x)$ changes the mean to be $\mu'' = -\mu$ and the skewness to be $D_3'' = -D_3$, but does not change the standard deviation. The distribution $P''(x)$ is now bounded from below by $-x_{\max}$ so we get:

$$\delta = \sigma/(x_{\max} - \mu), \tag{29}$$

which goes into Eq (15) to calculate $D_{3\min}$. In this case, we actually have found $D_{3\max} = -D_{3\min}$. Thus, we have rederived the results of [11], that is, Eq (4).

If a distribution $P(x)$ has domain $x_{\min} \leq x \leq x_{\max}$ then the above results give both a lower and an upper bound on $D_3$. As a conceptual example, suppose that $x_{\min} = \mu - 3\sigma$ and $x_{\max} = \mu + 3\sigma$; then $-8/3 \leq D_3 \leq 8/3$. This is consistent with the empirical observation that the skewness tends to lie between -3 and +3.

As a useful check on these results, consider the bidisperse distribution again with probability $P(a_+)$ and $P(a_-)$ for sizes $a_- < a_-$. Here we have $x_{\min} = a_-$, and CoV given by Eq (9). Using Eqs (4), (6) and (28), one can solve for $D_{3\min}$ in terms of the variables $\eta$ and $q$, recovering Eq (10): that is, $D_{3\min}$ is achieved in this situation. Similarly, using $x_{\max} = a_+$ one finds again $D_{3\max} = D_3$.

If we extend Eq (15) to any arbitrary upper or lower bound $x_{\text{bound}}$, we get the following relationship for the extreme value of $D_3$, $D_{3\text{extr}}$

$$D_{3\text{extr}} = \frac{\delta}{1 - x_{\text{bound}}/\mu} - \frac{1 - x_{\text{bound}}/\mu}{\delta} \tag{30}$$

which has reprised Eq (4).

## 3 Extensions to higher order moments

### 3.1 Notes to generalize from skewness

Going forward, we note that Eq (30) is useful for more than the extreme $D_3$ of the system, when considering a bidisperse system. As noted at the start of Sec. 2, if one is given $\mu$ and two other quantities, then one can uniquely determine a bidisperse distribution. Thus knowing one size $x_{\text{bound}}$, $\delta$, and $\mu$, determines the other size and relative probabilities. By plugging in any generic size $a/\mu$, which could be $a_+/\mu$ or $a_-/\mu$ to Eq (30), this produces the $D_3$ that makes a bidisperse distribution with that size and a given CoV $\delta$. This equation can be solved for $a$ to give either of Eqs (13 and 14). In other words, if we know we have a bidisperse distribution, then Eq (30) is a formula for $D_3$ as a function of one of the sizes $a$. We will derive similar results for higher moments.

### 3.2 Kurtosis $D_4$

As noted in the introduction, previous results by Pearson [8] show that $D_4 \geq D_3^2 + 1$ for any given distribution as per Eq (3). If we now know an inequality for $D_3$ on any distribution with Eq (30), we can solve for a new limit in $D_4$. in terms of $x_{\min}$, $\mu$, and $\delta$. In particular, we have to consider two cases. Treating the situation where the distribution has only nonnegative support ($x_{\min} = 0$, then for $\delta < 1$, $D_{3,\min} < 0$. This implies that $D_3 = 0$ is also possible, and therefore we can achieve lower $D_4$ than is predicted by Eq (3) based on $D_{3,\min}$. In other words, we can

consider the bidisperse distribution with $D_3 = 0$, which can be found using Eqs ([11](#) and [12](#)), to achieve $D_{4,\min} = 1$ as per [Eq (3)](#). For $\delta \geq 1$, $D_{3,\min} \geq 0$ and the limit on $D_4$ then follows from [Eq (30)](#). Thus we have

$$
\begin{aligned}
D_4 &\geq 1 \qquad\qquad\qquad (\delta < 1)\\
D_4 &\geq \left(\delta - \frac{1}{\delta}\right)^2 + 1 \quad (\delta \geq 1)
\end{aligned}
\tag{31}
$$

for the limits on $D_4$ in the two cases.

For the more general case of a distribution bounded on one side (by either $x_{\min}$ from below, or $x_{\max}$ from above, but not both), we can define the limits on kurtosis $D_4$ in terms of the extremum bounding value $x_{\text{extr}}$. Define

$$
\delta_0 = |1 - x_{\text{extr}}/\mu|. \tag{32}
$$

That is, $\delta_0$ is the equivalent of Eqs ([28](#) and [29](#)). We then get

$$
D_4 \geq 1 \quad (\delta < \delta_0) \tag{33}
$$

$$
D_4 \geq \left(\frac{\delta}{\delta_0} - \frac{\delta_0}{\delta}\right)^2 + 1 \quad (\delta \geq \delta_0). \tag{34}
$$

In other words, whether the distribution is bounded from below or bounded from above, in both cases this sets a minimum on $D_4$—but not a maximum.

When the distribution is bounded from below by $x_{\min}$ and bounded from above by $x_{\max}$, the situation complicates further. We start by defining $\delta_{\min}$ and $\delta_{\max}$ analogously to [Eq (32)](#). While $x_{\min} < x_{\max}$, the ordering of $\delta_{\min}$ and $\delta_{\max}$ is not determined. Thus define

$$
\delta_1 = \min(\delta_{\min}, \delta_{\max}), \tag{35}
$$

$$
\delta_2 = \max(\delta_{\min}, \delta_{\max}), \tag{36}
$$

$$
D_{4,m}(\delta) = \left(\frac{\delta}{\delta_m} - \frac{\delta_m}{\delta}\right)^2 + 1, \tag{37}
$$

where $m = 1, 2$. Next define $\delta'$ using

$$
D_{4,1}(\delta') = D_{4,2}(\delta') \tag{38}
$$

which can be solved to get $\delta' = \sqrt{\delta_1 \delta_2} = \sqrt{\delta_{\min} \delta_{\max}}$. The limits on kurtosis $D_4$ are then

$$
\begin{aligned}
1 &\leq D_4 \leq D_{4,2}(\delta) \quad (\delta < \delta_1)\\
D_{4,1}(\delta) &\leq D_4 \leq D_{4,2}(\delta) \quad (\delta_1 \leq \delta < \delta')
\end{aligned}
\tag{39}
$$

and values $\delta > \delta'$ are disallowed as they would require the bidisperse distribution be composed of values that lie outside of one or both of the boundaries $(x_{\min}, x_{\max})$. At $\delta = \delta'$, the only bidisperse distribution that is valid is composed of the two values $x_{\min}, x_{\max}$ with appropriate probabilities necessary to get the value of $\delta$, and we have $D_{4,1} = D_{4,2} = D_4$.

These results are visualized in [Fig 2a](#), which illustrates a specific example with $x_{\min} = 0$, $x_{\max} = 5$, and $\mu = 1$. For this example, $\delta_1 = 1.0$ and $\delta' = 3.25$. The solid lines indicate Inequalities 39, and the symbols indicate simulated random distributions with a specified $\delta$. Specifically, we generated distributions with data lying between limits $x_{\min}, x_{\max}$, and with enforced
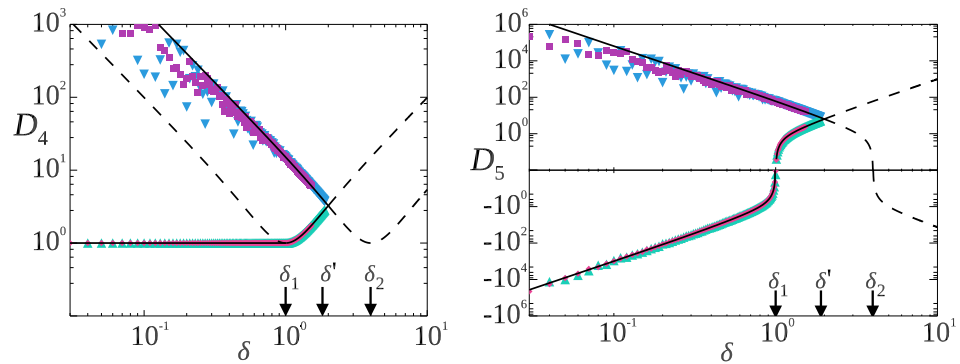
**Fig 2. Simulations with bidisperse distributions, and tri- or quad-disperse distributions, yield extrema which are plotted against the prediction (black line) given by Eq (43) for $D_4$ (left) and $D_5$ (right).** The data correspond to distributions with values limited to be greater than zero and less than $5\mu$. The bidisperse triangles are green (pointing up) for the minima and blue (pointing down) for the maxima, and the tri or quad-disperse are pink (diamonds) for minima and purple (squares) for maxima. The more extreme values from quad or tri-disperse was plotted for each polydispersity bin.

https://doi.org/10.1371/journal.pone.0297862.g002

mean $\mu$, and calculated $\delta$ and $D_4$ for all. For a given small range of $\delta$, we generated 20,000 distinct random distributions, half that are bidisperse, and the other half with three or four values. Over these 20,000 distributions, Fig 2a plots the maximum and minimum $D_4$ found for each $\delta$, all of which lie between the limits corresponding we have found (shown by the lines). While we have not proven that the bidisperse distribution sets the limits for $D_4$ for all other distributions, this is suggestive that Inequalities 39 are indeed limits for the kurtosis for any distribution.

### 3.3 Higher order generalized moments

We now proceed with an alternate derivation of Inequalities 34 which we can extend to higher moments. The generic definition of $D_n$ in the bidisperse case is:

$$D_n = \frac{(1-q)(a_- - \mu)^n + q(a_+ - \mu)^n}{m_2^{n/2}}. \tag{40}$$

If we use Eq (6) to solve for the generic definition of $D_n$ in terms of $q$ and $\eta$, we arrive at a formula of only $q$:

$$D_n = \frac{(1-q)^{n-1} + (-1)^n q^{n-1}}{(q-q^2)^{\frac{n}{2}-1}}. \tag{41}$$

Plugging in $n = 3$ arrives back at Eq (10).

For a bidisperse distribution, we can rewrite the second line of Inequality 34 as an equality in terms of $a$, one of the two bidisperse values. We then note that Eqs (30 and 34) are both

functions of $z = \delta/(1 - a/\mu)$:

$$
\begin{aligned}
D_3 &= \frac{\delta}{1 - a/\mu} - \frac{1 - a/\mu}{\delta} = z^1 - z^{-1} \\
D_4 &= \frac{\delta^4 - \delta^2(1 - a/\mu)^2 + (1 - a/\mu)^4}{\delta^2(1 - a/\mu)^2} \\
&= z^2 - z^0 + z^{-2}.
\end{aligned}
\tag{42}
$$

(In $D_4$, because only even powers of $z$ appear, the absolute value signs in Eq (32) can be dropped, allowing $z$ to have the same meaning for both $D_3$ and $D_4$.) The general pattern appears to be a finite sum of a geometric series. In fact, S2 Appendix shows how one can start from Eq (41) to derive

$$
\begin{aligned}
D_n &= \sum_{i=1}^{n-1} (-1)^{(n-i+1)} \left( \frac{\delta}{1 - a/\mu} \right)^{2i-n} \\
&= \sum_{i=1}^{n-1} (-1)^{(n-i+1)} z^{2i-n}.
\end{aligned}
\tag{43}
$$

One can immediately put in a value for $a$ of interest and get a potential limit of $D_n$. For example, for distributions bounded from below by $x_{\min}$ we conjecture

$$
D_5 \geq D_{5,\min} = z^3 - z^1 + z^{-1} - z^{-3}
\tag{44}
$$

with $z = \delta/(1 - x_{\min}/\mu)$ as above. As with $D_3$, our conjectured $D_{5,\max}$ is a similar equation using $z = \delta/(x_{\max}/\mu - 1)$. Fig 2b shows these two limits as the solid lines for the case $x_{\min} = 0$, $x_{\max} = 5$, and $\mu = 1$, along with the maximum and minimum observed $D_5$ values from numerically generated random distributions. All the random distributions lie within our conjectured analytic limits, again suggestive that they are the actual limits.

To try to show that these bounds achieve minima for any $n$, we can try a similar method as section 2.2. If we write out a more generic $m_n$:

$$
m_n = \frac{\Delta_-^2}{\Delta_-^2 + \delta^2} \left( \frac{\delta^2}{\Delta_-} \right)^n + \left( 1 - \frac{\Delta_-^2}{\Delta_-^2 + \delta^2} \right)(-\Delta_-)^n
\tag{45}
$$

we then can take its derivative with respect to $\Delta_-$, giving

$$
\begin{aligned}
\frac{\partial m_n}{\partial \Delta_-} &= \frac{(-1)^n(\delta^2 \Delta_-^{n-1})((n-2)\Delta_-^2 + n\delta^2)}{(\Delta_-^2 + \delta^2)^2} \\
&\quad - \frac{\delta^{2n}\Delta_-^{1-n}((n-2)\delta^2 + n\Delta_-^2)}{(\Delta_-^2 + \delta^2)^2}.
\end{aligned}
\tag{46}
$$

Eq (46) is negative for all odd values of $n$, implying an increase in the smallest size above zero will only increase $D_n$: thus, for odd $n$, $D_n$ is minimized for a bidisperse distribution with the smallest size set to zero. For even $n$, negative values of Eq (46) are achieved for $\Delta_-$ between 0 and $\delta$, but positive for $\Delta_- > \delta$. Thus, the minimum $m_n$ is achieved at $\Delta_- = \delta$. In fact, this recapitulates the result of Eq (31), that $D_{4,\min}$ is not a universal formula but rather depends on $\delta$. Furthermore, if we try to replicate Eqs (22–27) with $m_4$, the statements are untrue even when $\delta_r = \delta_s$. This gives credence that the boundaries of $D_n$ for even $n$ are not always given by the choice of $x_{\text{bound}}$.

Lastly, as previously noted, a bidisperse distribution can be completely described by three parameters: most directly by the values $a_-$, $a_+$, and the probability $q$ for one of these values.

Our approach has been to instead use $\mu$, $\delta$, and $a_-$ to find a constraint on $D_n$. We note that Eq (43) and the definition of $z$ is sufficient to find analogues of Eqs (11–14): thus, to use $D_n$, $\mu$, and $\delta$ to describe a bidisperse distribution. One can start with those three quantities and determine $a_-$, $a_+$, and $q$: analytically for $D_3$ as per Eqs (11–14), and numerically in other cases. This has been useful in the past for finding distributions with desired values of the moments [16]. Moreover, by then considering which values of $a_-$ and $a_+$ lie within bounds, one has a slightly alternate approach to finding bounds on $D_n$.

## 4 Conclusion

We have presented an alternative derivation of Eq (4) to that presented in [11]; this equation provides bounds on the skewness $D_3$ for a bounded distribution with a given CoV $\delta$. Equivalently, if $D_3$ is given, then this equation provides a bound for $\delta$. Returning to our starting example, if one is considering a distribution of strictly positive numbers, then for a given $D_3$, Eq (4) can be solved for the maximum possible $\delta$.

Our results for $D_3$ naturally imply limits on $D_4$ (Inequalities 34 using Pearson's formula [8], and Inequalities 39 more generally). Our general methodology is to note that bidisperse distributions are characterized by three parameters, which most naturally are the two values $a_+$ and $a_-$ as well as the probability $q$ of the value $a_+$; however, one can fruitfully choose as the three parameters the mean $\mu$, coefficient of variation $\delta$, and $a_-$. Setting $a_-$ to the lower bound of all possible distributions with a given $\mu$ and $\delta$ leads to lower bounds for $D_3$ and $D_4$. Moreover, our methodology extends to higher moments, leading to conjectures for limits on higher standardized moments as discussed in Section 3.3. One possible extension to our work would be to see if there are other relationships between general $D_n$ and $D_m$. It would also be interesting to discover a counterexample where a distribution exists that exceeds the limits of $D_n$ set by considering bidisperse distributions as in Section 3.3. We note that numerically at least, we have not found such a counterexample for $n = 5$, as seen in the data of Fig 2.

Our results have implications for a prior computational study of the packing of spheres, and how the density of such packings depend on the CoV and skewness of a particle size distribution [16]. In that prior work, the results had a varying range of skewnesses but the authors did not comment on the choice of this range. In fact, the lower bound on skewness studied in that work corresponds to result of Eq (4). This bound implies that a sphere packing composed of a distribution of radii with a given $\delta$ and lowest possible skewness is, in fact, equivalent to the packing of a distribution of equal-sized spheres; and the observed density of such packings in [16] obeyed this property, as it must. This is somewhat analogous to the circle packing shown in Fig 1a, for which the skewness has not yet reached the lower limit; nonetheless the packing is dominated by circles of the larger size.

## A Derivation of Eq (43)

We wish to show that Eqs (41 and 43) are equivalent expressions for $D_n$ for a bidisperse distribution. It is easiest to start with the end result and work backwards: Eq (43) is

$$D_n = \sum_{i=1}^{n-1} (-1)^{(n-i+1)} (z)^{(2i-n)} \tag{47}$$

$$= \sum_{i=1}^{n-1} (-1)^{(n-i+1)} \left( \frac{1 - a/\mu}{\delta} \right)^{(n-2i)} \tag{48}$$

where $a$ can represent either $a_+$ or $a_-$. We will begin by examining the term with $a$, $\mu$, and $\delta$

and work to express it in terms of $\eta$ and $q$. We will initially assume $a = a_-$ and use Eq (6) to express $a_-$ in terms of $\eta$, $\mu$, and $q$; and likewise we will use Eq (9) to express $\delta$ in terms of those same variables. This leads to

$$\frac{1 - a_-/\mu}{\delta} = \frac{\mu - a_-}{\mu\delta} \tag{49}$$

$$= \frac{\mu - [\mu/(1 - q + \eta q)]}{\mu(\eta - 1)\sqrt{q - q^2}/(1 - q + \eta q)} \tag{50}$$

$$= \frac{1 - q + \eta q - 1}{(\eta - 1)\sqrt{q - q^2}} \tag{51}$$

$$= \frac{q}{\sqrt{q(1 - q)}} = \left(\frac{q}{1 - q}\right)^{1/2}. \tag{52}$$

We can put this in to Eq (48) to give

$$D_n = \sum_{i=1}^{n-1} (-1)^{(n-i+1)} \left(\frac{q}{1 - q}\right)^{(n/2)-i} \tag{53}$$

$$= (-1)^{n+1} \left(\frac{q}{1 - q}\right)^{n/2} \left[\sum_{i=1}^{n-1} (-1)^i \left(\frac{1 - q}{q}\right)^i\right] \tag{54}$$

where now the summation is simply a finite geometric sum. The sum can be evaluated as

$$\sum_{i=1}^{n-1} \left(\frac{q - 1}{q}\right)^i = \frac{\left(\frac{q-1}{q} - \left(\frac{q-1}{q}\right)^n\right)}{1 - \frac{q-1}{q}} \tag{55}$$

$$= (q - 1) - q \left(\frac{q - 1}{q}\right)^n. \tag{56}$$

Putting this in to Eq (54), recognizing that $(q - 1)^n = (-1)^n (1 - q)^n$, and distributing the leading factor of $(-1)^{n+1}$, we get

$$D_n = \left(\frac{q}{1 - q}\right)^{n/2} [(-1)^n (1 - q) + q^{1-n}(1 - q)^n], \tag{57}$$

and this can be simplified to Eq (41).

The starting point we used above was Eq (52):

$$\frac{1 - a_-/\mu}{\delta} = \left(\frac{q}{1 - q}\right)^{1/2}.$$

If instead one focuses on $a_+$, the equivalent result is

$$\frac{1 - a_+/\mu}{\delta} = -\left(\frac{1 - q}{q}\right)^{1/2}. \tag{58}$$

Given that Eq (47) is unchanged when replacing $z \rightarrow -(1/z)$, the derivation holds whether using $a_+$ or $a_-$. Thus, the '$a$' in Eq (48) is valid for either meaning of $a$, and we have shown that Eqs (41 and 43) are equivalent.

## Supporting information

**S1 Appendix. Discreet distribution decomposition.**
(PDF)

**S2 Appendix. Derivation of Eq (43).**
(PDF)

**S1 Data. This is the data and method used to form Fig 2.**
(ZIP)

## Acknowledgments

## Author Contributions

**Conceptualization:** David J. Meer.

**Data curation:** David J. Meer.

**Formal analysis:** David J. Meer, Eric R. Weeks.

**Funding acquisition:** Eric R. Weeks.

**Investigation:** David J. Meer.

**Project administration:** Eric R. Weeks.

**Software:** David J. Meer.

**Supervision:** Eric R. Weeks.

**Writing – original draft:** David J. Meer, Eric R. Weeks.

**Writing – review & editing:** Eric R. Weeks.

## References

1. Pearson K. Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material. Phil. Trans. Royal Soc. London A. 1895; 186: 343–414. https://doi.org/10.1098/rsta.1895.0010

2. Abramowitz M and Stegun I. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables Applied Mathematics Series. Washington D.C. New York: United States Department of Commerce, National Bureau of Standards; Dover Publications; 1983; 55

3. Johnson NL, Kotz S and Balakrishnan N. Continuous Univariate Distributions 1st ed. Wiley-Interscience; 1992

4. Groeneveld RA. Sharp Inequalities for Skewness Measures. J. Royal Statistical Soc. D. 1991; 40: 387–392. https://doi.org/10.2307/2348727

5. Li X and Morris JM. On measuring asymmetry and the reliability of the skewness measure. Statistics & Probability Lett. 1991; 12: 267–271. https://doi.org/10.1016/0167-7152(91)90089-A

6. Holgersson HET. A Modified Skewness Measure for Testing Asymmetry. Communications in Statistics—Simulation and Computation. 2010; 39: 335–346. https://doi.org/10.1080/03610910903453419

7. Simpson JA and Welch BL. Table of the bounds of the probability integral when the first four moments are given. Biometrika. 1960; 47: 399–410. https://doi.org/10.1093/biomet/47.3-4.399

8.  Pearson K. Mathematical contributions to the theory of evolution.—XIX. Second supplement to a memoir on skew variation. Philos. Trans. Royal Soc. 1916; 216: 429–457. https://doi.org/10.1098/rsta.1916.0009

9.  Rohatgi VK and Székely GJ. Sharp inequalities between skewness and kurtosis. Statistics & Probability Lett. 1989; 8: 297–299. https://doi.org/10.1016/0167-7152(89)90035-7

10. Sen A. On the Interrelation Between the Sample Mean and the Sample Variance. American Statistician. 2012; 66: 112–117. https://doi.org/10.1080/00031305.2012.695960

11. Smołalski G. Identifying the asymmetry of finite support probability distributions on the basis of the first two moments. Measurement. 2020; 14: 106968. https://doi.org/10.1016/j.measurement.2019.106968

12. Cox NJ. Speaking Stata: The Limits of Sample skewness and Kurtosis. Stata J. 2010; 10: 482–495. https://doi.org/10.1177/1536867X1001000311

13. Garver R. Concerning the Limits of a Measure of Skewness. Ann. Math. Statist. 1932; 3: 358–360. https://doi.org/10.1214/aoms/1177732840

14. Majindar KN. Improved Bounds on a Measure of Skewness. Ann. Math. Statist. 1962; 33: 1192–1194. https://doi.org/10.1214/aoms/1177704482

15. Hotelling H and Solomons LM. The Limits of a Measure of Skewness. Ann. Math. Statist. 1932; 3: 141–142. https://doi.org/10.1214/aoms/1177732911

16. Desmond K and Weeks ER. Influence of particle size distribution on random close packing of spheres. Phys. Rev. E. 2014; 90: 022204. https://doi.org/10.1103/PhysRevE.90.022204 PMID: 25215730