Bounded Distributions place Limits on Skewness and Larger Moments

David J. Meer and Eric R. Weeks

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) δ as $D_3 \geq \delta - 1/\delta$. 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.

Key words and phrases: skewness, kurtosis, standardized moment, bounded distributions.

1. INTRODUCTION

Distributions of random numbers $P(x)$ are characterized by quantities such as mean, median, standard deviation, and skewness. The skewness is a measure of the asymmetry of a distribution [Pearson (1895)]. While there are several possible definitions of skewness [Groeneveld (1991)], a common definition depends on the third moment of the distribution compared to the second moment [Holgersson (2010); Li and Morris (1991)]. In particular, one can define the central moments

$$(1) \qquad m_n = \langle (x - \mu)^n \rangle = \int_{-\infty}^{+\infty} (x - \mu)^n P(x) dx$$

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

$$(2) \qquad D_n = \frac{m_n}{m_2^{n/2}}.$$

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 $P_1(x)$ and $P_2(x) = cP_1(cx + d)$ have the same skewness for $c > 0$ and any d [Groeneveld (1991)]. Pearson [Pearson (1916)] derived an upper boundary on the skewness:

$$(3) \qquad D_3^2 \leq m_4/m_2^2 - 1.$$

Alternate derivations of this result are also in the literature [Rohatgi and Székely (1989); Sen (2012)]. This applies for all distributions.

Often one considers distributions of strictly positive numbers: numbers of objects, sizes of objects (such as Figure 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 $x_{\min}$ and/or $x_{\max}$, generically x_{bound}. Smołalski [Smołalski (2020)] worked out upper and lower bounds on the skewness that applies for distributions with bounded support:

$$(4) \qquad D_{3\min,\max} = \delta_{\min,\max} - \frac{1}{\delta_{\min,\max}}$$

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

In this manuscript, 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 Equation 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 [Cox (2010)]. Note also that there are other definitions of skewness, for example that use the median of the distribution as part of the calculation [Pearson (1895)], for which other limits

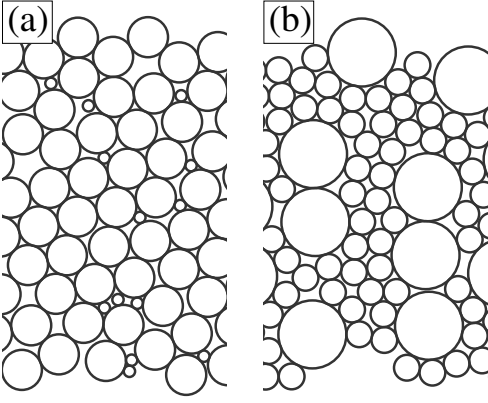*Dept. of Physics, Emory University, Atlanta, GA 30322 USA* (e-mail: *dmeer@emory.edu*).

FIG 1. *Examples of circles with random bidisperse distributions of diameters with (a) CoV $\delta = 0.4$, skewness $D_3 = -1$ and (b) $\delta = 0.4, D_3 = +3$.*

exist [Garver (1932); Hotelling and Solomons (1932); Majindar (1962)].

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 Figure 1(a). 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 Figure 1(b). 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 some minimum possible value.

Our derivation will proceed by first considering bidisperse distributions with nonnegative support and showing that for a fixed δ, the distribution with one value equal to zero achieves the lowest possible skewness. We then show taking two distributions obeying Equation 4 and considering a weighted sum will result in a new distribution that also obeys Equation 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 [Desmond and Weeks (2014)], we define the ratio

$$(5) \qquad \eta = a_+/a_-$$

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 μ and the latter is smaller than μ, respectively. Knowing the mean μm allows us to relate these quantities as

$$(6) \qquad a_+ = \eta\mu/(1 - q + \eta q)$$
$$a_- = \mu/(1 - q + \eta q).$$

Note that a bidisperse distribution with a given (η, 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 μ, 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 η and q as per Equation 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

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

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 σ which has dimensions of x, we will use the non-dimensional quantity "coefficient of variation," (CoV) defined as:

$$(8) \qquad \delta = \sigma/\mu.$$

Here we use the symbol, δ and later in this manuscript we will generalize this symbol beyond the specific meaning of CoV. We can use Equations 6 to eliminate a_+ and a_- from m_2 and D_n, resulting in

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

$$(10) \qquad D_3 = \frac{2q - 1}{\sqrt{q - q^2}}.$$

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

$$(11) \qquad q = \frac{\pm D_3 + M_3}{2M_3}$$

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

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 Equations 11,12 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 $\pm$ symbol. We continue and calculate the two possible values according to Equation 6:

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

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

Using Equation 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 Equation 14 for $a_- = 0$. This also implies $a_+ = \mu/(1-q)$. Solving for D_3 when $a_- = 0$ in Equation (14) lets us solve for $D_{3\min}$:

$$(15) \qquad D_{3\min} = \delta - \frac{1}{\delta}.$$

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 δ, 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:

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

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

$$(17) \qquad q = \frac{\Delta_-}{\Delta_- + \Delta_+}.$$

We can then get δ using

$$(18) \qquad \delta^2 = (1-q)\Delta_-^2 + q\Delta_+^2 = \Delta_-\Delta_+$$

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

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

Now consider the third moment of the distribution m_3:

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

The partial derivative of m_3 with respect to Δ_- is

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

Increasing Δ_- always decreases m_3, assuming we keep δ constant and μ positive. Likewise, decreasing Δ_- (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 δ, Equation 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 μ and both satisfying the bound of Equation 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 Equation 15. Then 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 (Equation 15) as

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

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

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

and thus

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

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

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

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

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

Every term without δ_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

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

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

Finally, we need to generalize from the bidisperse distribution to any distribution. Following [Rohatgi and Székely (1989)], 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 Appendix A). In the previous paragraph, we have shown that sums of distributions satisfy the bound. Thus, we have proven that Equation 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 μ 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 $\lim_{\mu\to0^+}(a_+, a_-, \eta) = \lim_{\mu\to0^-}(a_+, a_-, \eta)$. This limit can be calculated directly by multiplying by $\frac{\mu}{\mu}$ in Equation 12 and distributing the μ factor in Equations 13 and 14, leaving us with just $\sqrt{m_2}$ where there was previously δ. Therefore, we do not have to be concerned with means approaching zero.

Now consider the general case of a distribution $P(x)$ with bounded by $x_{\min}$ from below, and with a mean μ which might be zero. Let us assume $P(x)$ has a non-trivial 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:

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

That is, δ depends on the standard deviation σ and mean μ of the original distribution $P(x)$, with the additional correction of subtracting $x_{\min}$, at which point we can use Equation 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:

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

which goes into Equation 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 [Smołalski (2020)], that is, Equation 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 Equation 9. Using Equations 6, 28, and 4, one can solve for $D_{3\min}$ in terms of the variables η and q, recovering Equation 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 equation 15 to any arbitrary upper or lower bound x_{bound}, we get the following relationship for the extreme value of D_3, $D_{3\text{extr}}$

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

which has reprised Equation 4.

3. EXTENSIONS TO HIGHER ORDER MOMENTS

3.1 Notes to Generalize from Skewness

Going forward, we note that Equation 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 μ and two other quantities, then one can uniquely determine a bidisperse distribution. Thus knowing one size x_{bound}, δ, and μ, determines the other size and relative probabilities. By plugging in any generic size a/μ, which could be a_+/μ or a_-/μ to Equation 30, this produces the D_3 that makes a bidisperse distribution with that size and a given CoV δ. This equation can be solved for a to give either of Equations (13,14). In other words, if we know we have a bidisperse distribution, then Equation 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 [Pearson (1916)] show that $D_4 \geq D_3^2 + 1$ for any given distribution (Equation 3). If we now know an inequality for D_3 on any distribution with Equation 30, we can solve for a new limit in D_4. in terms of $x_{\min}$, μ, and δ. 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 Equation 3 based on $D_{3,\min}$. In other words, we can consider the bidisperse distribution with $D_3 = 0$ (which can be found using Equations 11 and 12) to achieve $D_{4,\min} = 1$ as per Equation 3. For $\delta \geq 1$, $D_{3,\min} \geq 0$ and the limit on D_4 then follows from Equation 30. Thus we have

$$D_4 \geq 1 \qquad\qquad (\delta < 1)$$

$$(31) \qquad D_4 \geq \left(\delta - \frac{1}{\delta}\right)^2 + 1 \qquad (\delta \geq 1)$$

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_{extr}. Define

$$(32) \qquad\qquad \delta_0 = |1 - x_{\mathrm{extr}}/\mu|.$$

That is, δ_0 is the equivalent of Equations 28 and 29. We then get

$$(33) \qquad D_4 \geq 1 \qquad\qquad (\delta < \delta_0)$$

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

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 Equation 32. While $x_{\min} < x_{\max}$, the ordering of $\delta_{\min}$ and $\delta_{\max}$ is not determined. Thus define

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

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

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

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

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

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

$$1 \leq D_4 \leq D_{4,2}(\delta) \qquad\qquad (\delta < \delta_1)$$

$$(39) \qquad D_{4,1}(\delta) \leq D_4 \leq D_{4,2}(\delta) \qquad (\delta_1 \leq \delta < \delta')$$

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 δ, and we have $D_{4,1} = D_{4,2} = D_4$.

These results are visualized in figure 2(a), 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 δ. Specifically, we generated distributions with data lying between limits $x_{\min}, x_{\max}$, and with enforced mean μ, and calculated δ and D_4 for all. For a given small range of δ, 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, Figure 2(a) plots the maximum and minimum D_4 found for each δ, 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:

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

If we use Equations 6 to solve for the generic definition of D_n in terms of q and η, we arrive at a formula of only q:

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

Plugging in $n = 3$ arrives back at Equation 10.

For a bidisperse distribution, we can rewrite Inequality 34 as an equality in terms of a, one of the two bidisperse values. We then note that Equations 30 and 34 are both functions of $z = \delta/(1 - a/\mu)$:

$$(42) \qquad D_3 = \frac{\delta}{1 - a/\mu} - \frac{1 - a/\mu}{\delta} = z^1 - z^{-1}$$

$$D_4 = \frac{\delta^4 - \delta^2(1 - a/\mu)^2 + (1 - a/\mu)^4}{\delta^2(1 - a/\mu)^2}$$
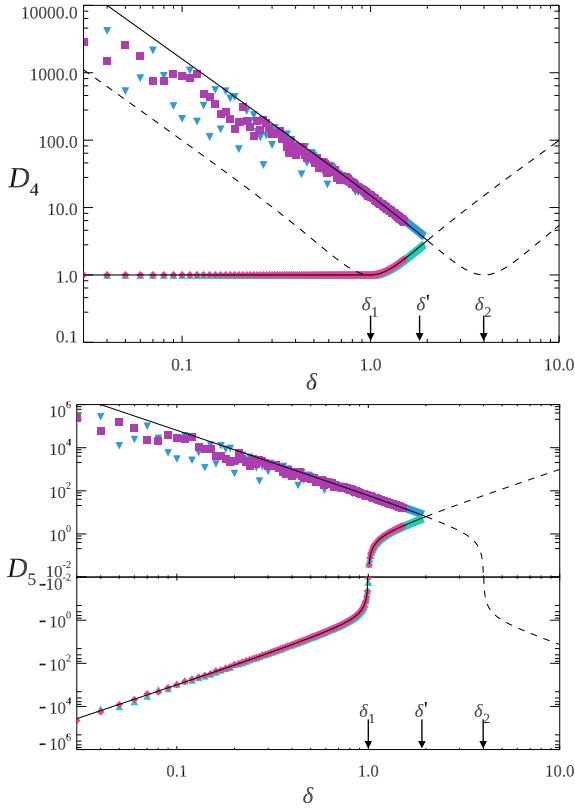
$$= z^2 - z^0 + z^{-2}.$$

FIG 2. *Simulations with bidisperse, and tri or quad-disperse extrema plotted against the prediction (black line) given by Equation (43) for D_4 (top) and D_5 (bottom) with size greater than zero and less than 5μ. 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.*

(In D_4, because only even powers of z appear, the absolute value signs in Equation 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, Appendix B shows how one can start from Equation 41 to derive

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

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

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

with $z = \delta/(1 - x_{\min}/\mu)$ as above. As with D_3, our conjectured $D_{5,\max}$ is a similar equation using $z = \delta/(x_{\max}/\mu - 1)$. Figure 2(b) 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:

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

we then can take its derivative with respect to Δ_-, giving

$$(46) \quad \frac{\partial m_n}{\partial \Delta_-} = \frac{(-1)^n \left(\delta^2 \Delta_-^{n-1} \right) \left((n-2) \Delta_-^2 + n\delta^2 \right)}{\left(\Delta_-^2 + \delta^2 \right)^2}$$
$$- \frac{\delta^{2n} \Delta_-^{1-n} \left((n-2) \delta^2 + n\Delta_-^2 \right)}{\left(\Delta_-^2 + \delta^2 \right)^2}.$$

Equation 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 Equation 46 are achieved for Δ_- between 0 and δ, but positive for $\Delta_- > \delta$. Thus, the minimum m_n is achieved at $\Delta_- = \delta$. In fact, this recapitulates the result of Equation 31, that $D_{4,\min}$ is not a universal formula but rather depends on δ. Furthermore, if we try to replicate Equations (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_{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 μ, δ, and a_- to find a constraint on D_n. We note that Equation 43 and the definition of z is sufficient to find analogues of Equations 11-14: thus, to use D_n, μ, and δ to describe a bidisperse distribution. One can start with those three quantities and determine a_-, a_+, and q – analytically for D_3 as per Equations 11 - 14, and numerically in other cases. This has been useful in the past for finding distributions with desired values of the moments Desmond and Weeks (2014). 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 Equation 4 to that presented in [Smołalski (2020)]; this equation provides bounds on the skewness D_3 for a bounded distribution with a given CoV δ. Equivalently, if D_3 is given, then this equation provides a bound for δ. Returning to our starting example, if one is considering a distribution

of strictly positive numbers, then for a given D_3, Equation 4 can be solved for the maximum possible δ.

Our results for D_3 naturally imply limits on D_4 (Inequalities 34 using Pearson's formula [Pearson (1916)], 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 μ, coefficient of variation δ, and a_-. Setting a_- to the lower bound of all possible distributions with a given μ and δ 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 Figure 2.

APPENDIX A: DISCREET DISTRIBUTION DECOMPOSITION

Rohatgi and Székely derived the result that any discrete distribution with mean μ can be decomposed into a sum of bidisperse distributions, all with mean μ [Rohatgi and Székely (1989)]. Their derivation is terse, so we rederive the result in this Appendix with a slightly lengthier presentation.

First, consider a discrete distribution $P(x)$ where x can take values a_i with probability p_i for $1 \leq i \leq n$, $\Sigma_i p_i = 1$, and with mean $\Sigma_i p_i a_i = \mu$. Replace a_n and a_{n-1} by

$$(47) \qquad a'_{n-1} = \frac{p_{n-1}}{p_{n-1} + p_n} a_{n-1} + \frac{p_n}{p_{n-1} + p_n} a_n$$

which occurs with probability $p'_{n-1} = p_{n-1} + p_n$. This is now a new distribution with mean μ and one fewer value. This can be repeated until one ends with a final distribution that takes on three discrete values, a_1, a_2, and a'_3 with probabilities p_1, p_2, and p'_3.

If we have a tridisperse distribution with three discrete values (a_1, a_2, a_3), of probabilities (p_1, p_2, p_3) and mean μ, we can decompose this into the sum of two bidisperse distributions as follows. Without loss of generality, assume $a_1 < \mu$ and $a_2 \leq \mu$. Then the first bidisperse distribution has values (a_1, a_3) with probabilities $p'_1 = \frac{a_3 - \mu}{a_3 - a_1}$ and $p'_3 = \frac{\mu - a_1}{a_3 - a_1}$, and similarly for the second distribution with values (a_2, a_3). Sampling the first distribution with probability p_1 / p'_1 and the second with probability p_2 / p'_2 recovers the original tridisperse distribution.

Now consider the distribution with four discrete values (a_1, a_2, a_3, a_4) and the related distribution (a_1, a_2, a'_3) formed using Equation 47. The latter can be decomposed

as a sum of two bidisperse distributions, as just demonstrated. This then provides a scheme to reduce the four-valued distribution to a sum of two three-valued distributions, one of which eliminates a_1 and the other which eliminates a_2. That is, the probability of finding a'_3 in each of the two bidisperse distributions is used to determine the new probabilities of finding a_3 and a_4 in the two tridisperse distributions. Proceeding by induction, each distribution with n distinct a_i values can be decomposed into two distributions of $n - 1$ distinct values, ultimately reducing down to a sum of bidisperse distributions.

APPENDIX B: DERIVATION OF EQUATION 43

We wish to show that Equations 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. Equation 43 is

$$(48) \qquad D_n = \sum_{i=1}^{n-1} (-1)^{(n-i+1)} (z)^{(2i-n)}$$

$$(49) \qquad = \sum_{i=1}^{n-1} (-1)^{(n-i+1)} \left(\frac{1 - a/\mu}{\delta} \right)^{(n-2i)}$$

where a can represent either a_+ or a_-. We will begin by examining the term with a, μ, and δ and work to express it in terms of η and q. We will initially assume $a = a_-$ and use Equation 6 to express a_- in terms of η, μ, and q; and likewise we will use Equation 9 to express δ in terms of those same variables. This leads to

$$(50) \qquad \frac{1 - a_-/\mu}{\delta} = \frac{\mu - a_-}{\mu\delta}$$

$$(51) \qquad = \frac{\mu - [\mu/(1 - q + \eta q)]}{\mu(\eta - 1)\sqrt{q - q^2}/(1 - q + \eta q)}$$

$$(52) \qquad = \frac{1 - q + \eta q - 1}{(\eta - 1)\sqrt{q - q^2}}$$

$$(53) \qquad = \frac{q}{\sqrt{q(1 - q)}} = \left(\frac{q}{1 - q} \right)^{1/2}.$$

We can put this in to Equation 49 to give

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

$$(55) \qquad = (-1)^{n+1} \left(\frac{q}{1 - q} \right)^{n/2} \left[\sum_{i=1}^{n-1} (-1)^i \left(\frac{1 - q}{q} \right)^i \right]$$

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

$$(56) \qquad \sum_{i=1}^{n-1} \left(\frac{q-1}{q} \right)^i = \frac{\left(\frac{q-1}{q} - \left(\frac{q-1}{q} \right)^n \right)}{1 - \frac{q-1}{q}}$$

$$(57) \qquad = (q-1) - q \left(\frac{q-1}{q} \right)^n .$$

Putting this in to Equation 55, recognizing that $(q-1)^n = (-1)^n (1-q)^n$, and distributing the leading factor of $(-1)^{n+1}$, we get

$$(58)$$
$$D_n = \left(\frac{q}{1-q} \right)^{n/2} \left[(-1)^n (1-q) + q^{1-n}(1-q)^n \right] ,$$

and this can be simplified to Equation 41.

The starting point we used above was Equation 53:

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

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

$$(59) \qquad \frac{1 - a_+/\mu}{\delta} = - \left(\frac{1-q}{q} \right)^{1/2} .$$

Given that Equation 48 is unchanged when replacing $z \to -(1/z)$, the derivation holds whether using a_+ or a_-. Thus, the 'a' in Equation 49 is valid for either meaning of a, and we have shown that Equations 41 and 43 are equivalent.

ACKNOWLEDGMENTS

FUNDING

REFERENCES

Cox, N. J. (2010). Speaking Stata: The Limits of Sample skewness and Kurtosis. *Stata J.* **10** 482–495. https://doi.org/10.1177/1536867X1001000311

Desmond, K. and Weeks, E. (2014). Influence of particle size distribution on random close packing of spheres. *Phys. Rev. E* **90** 022204–022204. https://doi.org/10.1103/physreve.90.022204

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

Groeneveld, R. A. (1991). Sharp Inequalities for Skewness Measures. *J. Royal Statistical Soc. D* **40** 387–392. https://doi.org/10.2307/2348727

Holgersson, H. E. T. (2010). A Modified Skewness Measure for Testing Asymmetry. *Communications in Statistics - Simulation and Computation* **39** 335–346. https://doi.org/10.1080/03610910903453419

Hotelling, H. and Solomons, L. M. (1932). The Limits of a Measure of Skewness. *Ann. Math. Statist.* **3** 141–142.

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

Majindar, K. N. (1962). Improved Bounds on a Measure of Skewness. *Ann. Math. Statist.* **33** 1192–1194. https://doi.org/10.1214/aoms/1177704482

Pearson, K. (1895). Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material. *Phil. Trans. Royal Soc. London A* **186** 343–414.

Pearson, K. (1916). Mathematical contributions to the theory of evolution.—XIX. Second supplement to a memoir on skew variation. *Philosophical Transactions of the Royal Society of London* **216** 216429–457. https://doi.org/10.1098/rsta.1916.0009

Rohatgi, V. K. and Székely, G. J. (1989). Sharp inequalities between skewness and kurtosis. *Statistics & Probability Lett.* **8** 297–299. https://doi.org/10.1016/0167-7152(89)90035-7

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

Smołalski, G. (2020). Identifying the asymmetry of finite support probability distributions on the basis of the first two moments. *Measurement* **149** 106968.

This figure "m4bounds_5.jpg" is available in "jpg" format from:

http://arxiv.org/ps/2308.05006v1

This figure "m5bounds_5.jpg" is available in "jpg" format from:

http://arxiv.org/ps/2308.05006v1