Journal of Macromolecular Science, Part A
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597274

SZOLD: Its Properties and Relationship to Other Distributions

Thomas T. Mercer[a]
[a] Department of Radiation Biology and Biophysics, University of Rochester, Rochester, New York

To cite this Article Mercer, Thomas T.(1978) 'SZOLD: Its Properties and Relationship to Other Distributions', Journal of Macromolecular Science, Part A, 12: 7, 1083 — 1088
To link to this Article: DOI: 10.1080/00222337808069406
URL: http://dx.doi.org/10.1080/00222337808069406

J. MACROMOL. SCI.–CHEM., A12(7), pp. 1083-1090 (1978)

SZOLD: Its Properties and Relationship to Other Distributions

THOMAS T. MERCER

Department of Radiation Biology and Biophysics
University of Rochester
Rochester, New York 14642

ABSTRACT

SZOLD is a form of the three-parameter lognormal distribution. As originally defined, its normalization factor permits negative values of the distributed variable for some positively skewed distributions and for all negatively skewed distributions. Correct normalization factors are given here, and the degree of truncation is shown as a function of the distribution parameters.

Rowell and Levit [1] developed the skewed zeroth-order logarithmic distribution (SZOLD) as a tool to be used for particle-size analysis of colloidal systems. They felt that SZOLD differed from a skewed normal distribution in that it did not "admit negative sizes." They defined its probability density function as

$$f(x) = N \cdot \exp - \{\ln^2 [1 + s(x - x_m)/x_m]/2\sigma^2\} \quad (1)$$

where N is a normalization factor, x_m is the mode of the distribution, s is a skew parameter, and σ is a breadth parameter. As a measure of skewness, they used the ratio, $(x_1 - x_m)/(x_2 - x_m)$, where $x_1 < x_m < x_2$ and $f(x_1) = f(x_2)$, which they called "balance." They expressed surprise that balance was independent of the skew parameter s.

Using the relationship

$$\int_0^\infty f(x)dx = 1 \tag{2}$$

and setting $y = \ln[1 + s(x - x_m)/x_m]$, they integrated the transformed equation over the range $0 \leq y \leq \infty$. Stating that the result was equal to N/2 because the integral was symmetrical about y = 0, they concluded that

$$N = (\sigma\sqrt{2\pi})^{-1}(s/x_m e^{\sigma^2/2}) \tag{3}$$

Although negative values of N, which occur when $s < 0$, appear incompatible with Eq. (2), they argued rightly that the minus sign could be ignored because "it indicates only that the direction of the coordinate has been effectively reversed by negative s."

Since y = 0 at the modal value of x, integration over the range $0 \leq y \leq \infty$ could represent N/2 only if the distribution's mode and median were identical, which is unlikely to be true of a unimodal skewed distribution. An inspection of the transformed variable y also reveals that it is symmetrical about the median only if $s \geq 1$, unless negative values of x are permitted. Since Rowell and Levit intended SZOLD to be a size distribution function for small particles they limited x to the range $0 \leq x \leq \infty$. Their definition of N [Eq. (3)] is therefore not valid when $s < 1$. Further consideration will be given to this anomaly, and to that of a skew parameter that does not appear in their quantitative definition of skew, after the distribution function has been put into a more familiar form.

Let

$$x_L = x_m [1 - (1/s)] \tag{4}$$

so that

$$x_m/s = x_m - x_L \tag{4a}$$

On combining these equations with Eqs. (1) and (3), the distribution function can be put in the form

$$- \ln f(x) = \ln(\sigma\sqrt{2\pi}) + \ln(x_m - x_L) + (\sigma^2/2)$$

$$+ [\ln(x - x_L) - \ln(x_m - x_L)]^2/2\sigma^2 \qquad (5)$$

Expanding the last term and rearranging gives

$$- \ln f(x) = \ln(\sigma\sqrt{2\pi}) + \{[\ln(x - x_L) - \ln((x_m - x_L)\, e^{\sigma^2})]^2/2\sigma^2\}$$

$$+ \ln(x - x_L) \qquad (5a)$$

The distribution function becomes

$$f(x) = [\sigma\sqrt{2\pi}\,(x - x_L)]^{-1} \exp - \{[\ln(x - x_L) - \ln((x_m - x_L)e^{\sigma^2})]^2/2\sigma^2\} \qquad (5b)$$

This is the probability density function for the three parameter log-normal distribution [2]. Since it includes the normalization factor derived by Rowell and Levit, it necessarily permits negative values of x. The nature of the distribution varies with s (or x_L) as shown in Table 1. For all positive values of s, x_L is the lower limit of the distribution, passing from x_m at $s = \infty$ to $-\infty$ at $s = +0$. For all negative values of s, x_L is the upper limit of the distribution, passing from ∞ at $s = -0$ to x_m at $s = -\infty$. The distribution function then becomes

$$f(x) = [\sigma\sqrt{2\pi}\,(x_L - x)]^{-1} \exp - \{[\ln(x_L - x) - \ln((x_L - x_m)e^{\sigma^2})]^2/2\sigma^2\} \qquad (5c)$$

Thus SZOLD is simply an awkward form of the three-parameter lognormal distribution. In its usual form, the statistical parameters of customary interest are well-known [3]:

Mean:

$$\bar{x} = x_L + (x_g - x_L)\, e^{\sigma^2/2}$$

$$= x_L + (x_m - x_L)\, e^{3\sigma^2/2}$$

TABLE 1.

s	x_L	f(x)
> 1	$0 < x_L < x_m$	Three-parameter lognormal distribution; positively skewed; $x_L \lesssim x \lesssim \infty$
1	0	Two-parameter lognormal distribution; positively skewed; $0 \lesssim x \lesssim \infty$
$0 < s < 1$	$x_L < 0$	Three-parameter lognormal distribution; positively skewed;
0	$-\infty$	$f(x) = 0$ for all x; $-\infty < x < \infty$
< 0	$> x_m$	Three-parameter lognormal distribution; negatively skewed; $-\infty < x < x_L$
$\pm \infty$	x_m	$f(x \neq x_m) = 0$; $f(x_m) = \infty$

Median:

$$x_g = x_L + (x_m - x_L)e^{\sigma^2}$$

Mode:

$$x_m = x_L + (x_g - x_L)\, e^{-\sigma^2}$$

Variance:

$$\sigma_x^2 = (\bar{x} - x_L)^2 (e^{\sigma^2} - 1) = (x_m - x_L)^2\ e^{3\sigma^2} (e^{\sigma^2} - 1) \tag{6}$$

Coefficient of variation:

$$v = (e^{\sigma^2} - 1)^{1/2}$$

Coefficient of skewness:

$$Sk = (e^{\sigma^2} + 2)\,(e^{\sigma^2} - 1)^{1/2}$$

Coefficient of kurtosis:

$$K = e^{4\sigma^2} + 2e^{3\sigma^2} + 3e^{2\sigma^2} - 6$$

The j-th moment about x_L is

$$(x_j - x_L)^j = (x_g - x_L)^j e^{j^2\sigma^2/2} = (x_m - x_L)^j e^{j^2\sigma^2/2 + j \;\; ^2} \tag{6a}$$

The ratio $(x_j - x_L)/(x_m - x_L)$ is a function only of σ^2 and j. Except for reversing the direction of skew as s changes sign, the "skew" parameter has no effect on skewness.

The normalization factor for equations 6 is given by equation 3. When the variable x represents colloid diameters, all distributions that permit negative values of x, that is, all those for which $s < 1$, must be truncated and the normalization factor adjusted accordingly. To evaluate N for these cases, let

$$z = (\sqrt{2}\sigma)^{-1} \ln [(x - x_L)/(x_m - x_L)] - \sigma/\sqrt{2} \tag{7}$$

Then

$$dx = (x_m - x_L)(\sigma\sqrt{2}) \exp [\sigma\sqrt{2}(z + \sigma/\sqrt{2})] dz \tag{7a}$$

and Eq. (2), with proper limits, becomes

$$\int_{x_1}^{x_2} f(x)dx = N(x_m - x_L)\sigma\sqrt{2}\, e^{\sigma^2/2} \int_{z_1}^{z_2} e^{-z^2} dz = 1 \tag{7b}$$

When $0 \leqq s < 1$, x_L is negative and the lower limit becomes $x = 0$. The normalization factor becomes

$$N_+ = N \left[\pi^{-1/2} \int_{z_+}^{\infty} e^{-z^2} dz\right]^{-1} = \frac{2N}{(1 - \text{erf } z_+)} \tag{8}$$

where

$$z_+ = z_1 = (\sigma\sqrt{2})^{-1} \ln (1 - s) - \sigma/\sqrt{2} \tag{8a}$$

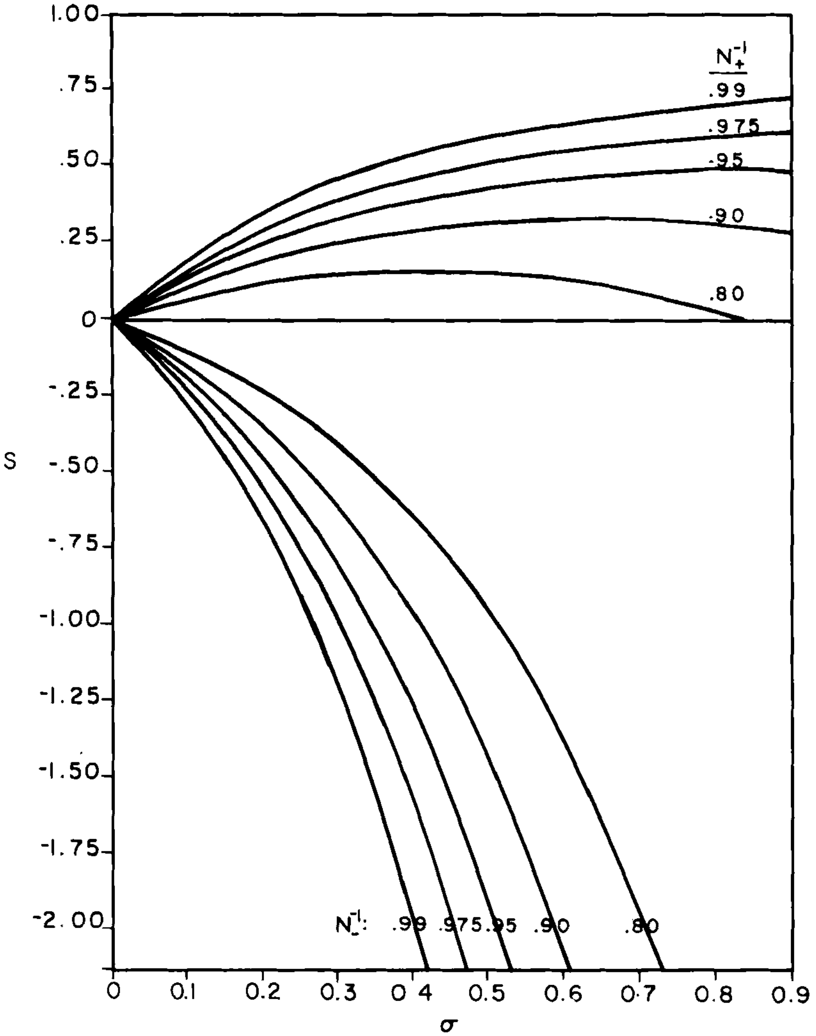

FIG. 1. Values of S and σ for several degrees of truncation.