lognormal distribution $(X \sim Lognormal(\mu, \sigma^2))$
A *continuous probability distribution* that models various natural phenomena which either depend on multiple independent (→ *independence*) *random variables* or randomly partition a finite *population*. Examples for the latter would be the distribution of the sizes of cities or the sizes of the pieces of a shattered vase. A simple example for the former would be the distribution of volumes of cuboids with heights, widths, and depths drawn randomly from a finite interval.

A random variable $X \sim Lognormal(\mu, \sigma^2)$ is lognormally distributed, if the variable $Y \sim ln(X)$ is normally distributed (→ *normal distribution*). Analogously, $X \sim e^Y$ is lognormally distributed, if Y is normally distributed. Sometimes lognormal variables are written as $X \sim e^{\mu+\sigma Z}$, where Z is the *standard normal distribution* and μ and σ are the *mean* and *standard deviation* of Z.

The relationship between normally and lognormally distributed variables is illustrated by the example of the volumes of random cuboids with *discrete* edge lengths in the range $\{1, \cdots, 10\}$. The distribution of cuboid volumes tends toward a lognormal distribution $X \sim Lognormal(ln(16.5), ln(24.75))$. The *frequency distribution* of cuboid volumes is shown in figure RVL. Because the volume of a cuboid is the product of its edge lengths, the logarithm of the volume is the sum of the edge lengths. The sum of lengths then follows a normal distribution $ln(X) \sim N(16.5, 24.75)$, also shown in figure RVL. The parameters of the lognormal distribution are shown on figure LGD, sample plots of its probability functions can be found in figure LGC.

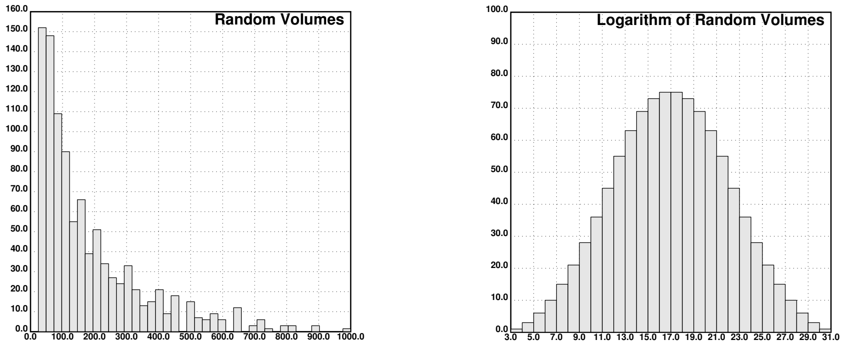

Figure **RVL:** random volumes; left: lognormal distribution of random volumes $X \sim Lognormal(ln(16.5), ln(24.75))$, right: corresponding normal distribution $Y \sim N(16.5, 24.75)$.

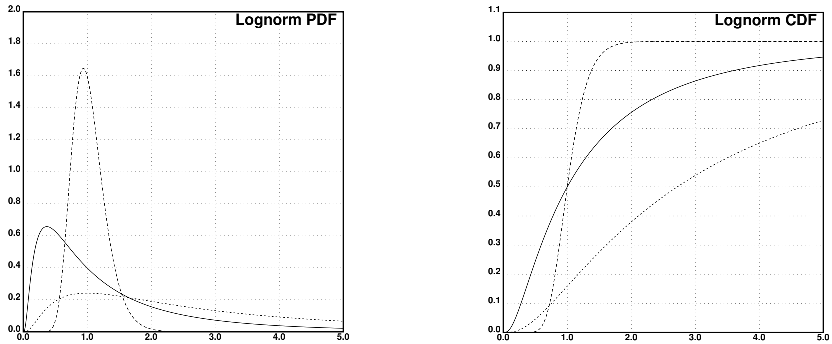

Figure **LGC:** lognormal distribution probability functions: left: PDF with $\mu = 0$, $\sigma^2 = 1$ (solid), $\mu = 0$, $\sigma^2 = 0.0625$ (dashed); $\mu = 1$, $\sigma^2 = 1$ (dotted), right: CDF with same parameters

Example: if the *average* city has $100\,000$ inhabitants and the *variance* is $10\,000^2$, then the distribution of city sizes follows a lognormal distribution $X \sim Lognormal(ln(10^5), ln(10^8))$. Therefore, the probability for a city to have between $1\,000$ and $10\,000$ inhabitants is

$$F_X(10^4) - F_X(10^3) \approx 0.154$$

and the probability for a city to have between one million and two million inhabitants is

$$F_X(2 \cdot 10^6) - F_X(10^6) \approx 0.053$$

$X \sim Lognormal(\mu, \sigma^2)$	
PDF	$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \cdot e^{-\frac{(ln(x)-\mu)^2}{2\sigma^2}}$
CDF	$F(x) = \phi(\frac{ln(x)}{\sigma})$
	$F(x) = \frac{1}{2} + \frac{1}{2} \cdot erf\left(\frac{ln(x) - \mu}{\sigma\sqrt{2}}\right)$
Statistic	$x \in \mathbf{R}_0^+$: raw score
Parameters	$\mu \in \mathbf{R}_0^+$: shape
	$\sigma^2 \in \mathbf{R}^+$: shape
μ	$e^{\mu + \frac{\sigma^2}{2}}$
σ^2	$(e^{\sigma^2} - 1) \cdot e^{2\mu + \sigma^2}$
Skewness (γ_1)	$(e^{\sigma^2} + 2) \cdot \sqrt{e^{\sigma^2} - 1}$

Figure **LGD:** lognormal distribution; ϕ is the CDF of the normal distribution; erf is the Gauss error function; the shape parameters are the variance and location (mean) of $ln(X)$; where $ln(X)$ is the normal distribution $\mu + \sigma Z$ upon which $X \sim e^{\mu + \sigma Z}$ is based