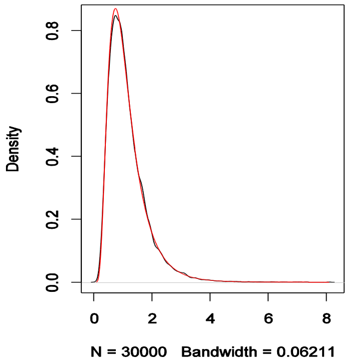
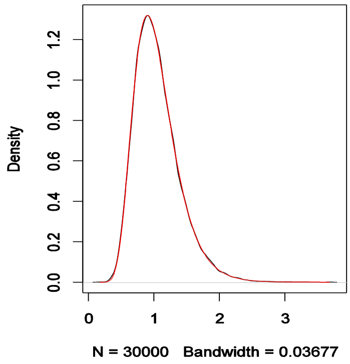
I observed the similarity between the lognormal distribution and the distribution of the sample median that comes from a lognormal population. Below is the density (in black) of the sample median (with sample size 5 and 15 respectively) coming from $LogN(0,1)$ and the fitted lognormal density (in red) using method of moments.





Since this is not a journal article, I don't write it in an academic manner. Still, it is technical and, for the sake of readiness, I use the following notations for certain special functions.

The CDF of normal distribution	$\Phi(w)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{w}\exp\left(-t^2\right)dt$
Hypergeometric function	${}_2F_1(a,b,c;w)=\sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{(c)_n}\frac{w^n}{n!};\quad (a)_n=\begin{cases}1 & n=0\\ a(a+1)\cdots(a+n-1) & n>0\end{cases}$
Error function	$erf(w)=\frac{2}{\sqrt{\pi}}\int_{0}^{w}\exp(-t^2)dt$
Beta function	$B(a,b)=\int_{0}^{1}t^{a-1}(1-t)^{b-1}dt$

Also, given the CDF of the parent population $G(x)$, the sample median distribution is given by,

$$\int_{-\infty}^{x}\frac{1}{B(k+1,k+1)}G^k(y)[1-G(y)]^k\,dG(y) \tag{1}$$

where sample size $n=2k+1$. In our case, $G(x)=\Phi\left(\frac{\ln(x)-\mu}{\sigma}\right)$. Also, having this substituted then solving this integral gives the exact sample median distribution,

$$\frac{\Phi^{k+1}\left(\frac{\ln(x)-\mu}{\sigma}\right)}{(k+1)B(k+1,k+1)}\,{}_2F_1\left(k+1,-k,k+2;\Phi\left(\frac{\ln(x)-\mu}{\sigma}\right)\right)$$

Formally I propose the lognormal approximation to the distribution of the sample median that comes from a lognormal parent,

$$\frac{\Phi^{k+1}\left(\frac{\ln(x)-\mu}{\sigma}\right)}{(k+1)B(k+1,k+1)}\,{}_2F_1\left(k+1,-k,k+2;\Phi\left(\frac{\ln(x)-\mu}{\sigma}\right)\right)\approx\Phi\left(\frac{\ln(x)-\mu}{\sigma B(k+1,k+1)4^k\sqrt{2}}\right) \quad (2)$$

That is, the sample median distribution can be approximated by a lognormal distribution, $LogN(\mu,\sigma B(k+1,k+1)4^k)$.

The proof requires the following approximation.

Proposition 1 (Approximating the error function):

$$\operatorname{erf}(w)\approx\operatorname{sgn}(w)\sqrt{1-\exp\left\{-w^2\frac{4}{\pi}\right\}} \quad (3)$$

with the maximum absolute error of 6.29×10^{-3} at $w\approx1.169$ and $\operatorname{sgn}(\cdot)$ is the sign function.

This approximation is based on the uniform approximation in Winitzki (2013) with the use of Padé approximant given by,

$$\operatorname{erf}(w)\approx\operatorname{sgn}(w)\sqrt{1-\exp\left\{-w^2\frac{P(w)}{Q(w)}\right\}}$$

where $P(w)$ and $Q(w)$ are polynomials.

Taking $P(w)$ and $Q(w)$ with the degree of two, Nealy, et al (2010) then approximates that,

$$\operatorname{erf}(w)\approx\operatorname{sgn}(w)\sqrt{1-\exp\left\{-w^2\frac{\frac{4}{\pi}+aw^2}{1+aw^2}\right\}}$$

where $a=\frac{8(\pi-3)}{3\pi(4-\pi)}$. This approximation used here further reduces the degree of the polynomials to one.

□

Alternatively, (2) can be expressed in terms of PDF.

Proposition 2 (Normal approximation to the distribution of the sample median from lognormal parent):

$$\frac{\exp\left\{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2\right\}}{x\sigma\sqrt{2\pi}B(k+1,k+1)}\Phi\left(\frac{\ln(x)-\mu}{\sigma}\right)^k\left[1-\Phi\left(\frac{\ln(x)-\mu}{\sigma}\right)\right]^k\approx\frac{\exp\left\{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma B(k+1,k+1)4^k}\right)^2\right\}}{x\sigma B(k+1,k+1)4^k\sqrt{2\pi}}$$

The proof is very straightforward. For the LHS,

$$\frac{\exp\left\{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2\right\}}{x\sigma\sqrt{2\pi}B(k+1,k+1)}\Phi\left(\frac{\ln(x)-\mu}{\sigma}\right)^k\left[1-\Phi\left(\frac{\ln(x)-\mu}{\sigma}\right)\right]^k$$

$$=\frac{\exp\left\{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2\right\}}{x\sigma\sqrt{2\pi}B(k+1,k+1)}\left[1-erf^2\left(\frac{\ln(x)-\mu}{\sigma\sqrt{2}}\right)\right]^k$$

Then Proposition 2 becomes,

$$\frac{\exp\left\{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2\right\}}{x\sigma\sqrt{2\pi}B(k+1,k+1)}\left[1-erf^2\left(\frac{\ln(x)-\mu}{\sigma\sqrt{2}}\right)\right]^k\approx\frac{\exp\left\{-\frac{\left(\ln(x)-\mu\right)^2}{2\sigma^2B^2(k+1,k+1)4^{2k}}\right\}}{xB(k+1,k+1)4^k\sigma\sqrt{2\pi}}$$

$$\Rightarrow\left[1-erf^2\left(\frac{\ln(x)-\mu}{\sigma\sqrt{2}}\right)\right]^k\approx\frac{1}{xB(k+1,k+1)4^k\sigma\sqrt{2\pi}}\exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2B^2(k+1,k+1)4^{2k}}\right\}\frac{B(k+1,k+1)4^k}{f(x;\mu,\sigma)}$$

$$\Rightarrow\left[1-erf^2\left(\frac{\ln(x)-\mu}{\sigma\sqrt{2}}\right)\right]^k\approx\frac{\exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2B^2(k+1,k+1)4^{2k}}\right\}}{\exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right\}}$$

$$\Rightarrow\left[1-erf^2\left(\frac{\ln(x)-\mu}{\sigma\sqrt{2}}\right)\right]^k\approx\left[\exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right\}\right]^{\left(\frac{1}{H}-1\right)}$$

where $H=4^{2k}B^2(k+1,k+1)$.

For x cannot be negative, use Proposition 1 on the LHS and rearrange to get,

$$\exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\frac{4k}{\pi}\right\}\approx\left[\exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right\}\right]^{\left(\frac{1}{H}-1\right)}$$

$$\Rightarrow\left[\exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right\}\right]^{\frac{4k}{\pi}}\approx\left[\exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right\}\right]^{\left(\frac{1}{H}-1\right)}$$

As noted earlier, the maximum absolute error of the error function approximation is 6.29×10^{-3}. The resulting error is not as large as it seems; $\left|\mathrm{erf}\,(w)\right|$ is strictly smaller than 1 hence the error resulted from $\left[1-\mathrm{erf}\,(w)\right]^k$ is much smaller as k increases.

We now have to show that $\frac{4k}{\pi} \sim \frac{1}{H} - 1$. Given that $H = 2^{4k} B^2(k+1, k+1)$, we first simplify the beta function.

$B(k+1, k+1) = B(k,k)\dfrac{k^2}{4k^2 + 2k}$. Then by Stirling's formula,

$$B(k+1, k+1) = B(k,k)\frac{k^2}{4k^2 + 2k} \approx \sqrt{2\pi}\,\frac{k^{2k-1}}{(2k)^{2k-0.5}}\,\frac{k^2}{4k^2 + 2k} = \frac{\sqrt{\pi k}}{2^{2k}(2k+1)}$$

Therefore, $H = 2^{4k} B^2(k+1, k+1) \approx \dfrac{\pi k}{(2k+1)^2}$ and it is straight forward that $\lim_{k\to\infty} \dfrac{\frac{1}{H} - 1}{4k/\pi} = 1$

□

A final note regarding the approximation error and the rate of convergence

- The maximum absolute error depends only on the sample size. The table below shows the value of the maximum absolute error for different sample sizes up to 53.

n	max*	n	max*
3	7.526	29	1.730
5	6.546	31	1.628
7	5.442	33	1.536
9	4.597	35	1.455
11	3.961	37	1.381
13	3.472	39	1.315
15	3.088	41	1.255
17	2.779	43	1.200
19	2.525	45	1.149
21	2.313	47	1.103
23	2.134	49	1.060
25	1.980	51	1.021
27	1.847	53	<1.000

*in 10^{-4}

- It has a faster rate of convergence than the asymptotically normal sample median distribution introduced in Cramer (1946, p.369).

- Hence, I still haven't shown how the lognormal approximation proposed here eventually converges to the one in Cramer (1946, p.369), recalling the normal approximation to lognormal distribution.