SUT Journal of Mathematics
Vol. 50, No. 2 (2014), 475–482

Tail quantile approximations for hybrid lognormal distribution

Nobumichi Shutoh

(Received July 31, 2014; Revised February 17, 2015)

Abstract. This paper provides the approximations for the tail quantile for hybrid lognormal distribution which is often applied to the distribution for some datasets such as the radiation dose. We derive the Mills ratio for the hybrid lognormal distribution, and have the main results on the basis of the Mills ratio using the idea stated in [7]. Finally, the numerical evaluation for the main results is given.

AMS 2010 *Mathematics Subject Classification.* 62E20, 41A60.

Key words and phrases. Quantile approximation, hybrid lognormal distribution, Mills ratio.

§1. Introduction

In the statistical analysis, we usually compute the quantiles for the normal or nonnormal distributions for the purpose of investigating the behavior of the distribution. However, in general, the quantiles cannot be easily expressed by an explicit form. Therefore, using iteration method depending on the initial value such as Newton-Raphson method, the quantiles are usually obtained numerically. However, it is still important work to consider the quantile approximations with analytic approach from the viewpoint of applications. Further, as mentioned in [8], the approximations reduce the computational time in the iteration method for calculating the quantiles.

For the derivation of the quantile approximations, the most of the authors derived the results via some approaches such as Cornish-Fisher expansion. Some authors approximated the distribution by adjusting the first some moments to that for another simple distribution, however, it may be the complicated forms.

In recent years, some authors considered the approximations for the tail quantiles via obtaining approximate error term between the rough approximate quantiles and true quantiles (e.g., [7, 8]). This approach is closely related to the bounds for the distribution function. For the studies on the bounds of the distribution function, see the references [1, 8]. For the estimation for the quantile from the random samples, refer to [5, 6].

In this paper, we deal with one of the nonnormal distributions extended from the normal one, named hybrid lognormal distribution, which is often applied to the datasets of the radiation dose (see [4]). For the distribution of the radiation dose, the tail quantile plays an important role in a sense of the risk management. Therefore, we primarily derive the approximations for the tail quantile which are more easily computable. In particular, focusing on the approach stated in [7, 8], we derive the tail quantile approximations for the hybrid lognormal distribution and their modified version.

The rest of this paper is organized as follows. Section 2 defines the hybrid lognormal distribution and addresses its properties. Section 3 derives the tail quantile approximations for the hybrid lognormal distribution. Section 4 evaluates the results stated in Section 3 numerically. Section 5 concludes the paper and states the direction to the further problems.

§2. Hybrid lognormal distribution and its asymptotic properties

We give the following definition of the hybrid lognormal distribution with the parameters $\rho(>0)$, μ and σ^2: $h(\rho X) \sim N(\mu, \sigma^2)$, where $h(x) = x + \ln x$ $(x > 0)$. Hereafter, we state $X \sim HLN(\rho; \mu, \sigma^2)$ if $h(\rho X)$ is distributed as $N(\mu, \sigma^2)$. By the definition, the distribution function of X can be expressed as

$$F_X(x) = \Phi\left(\frac{h(\rho x) - \mu}{\sigma}\right),$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of $N(0, 1)$. Therefore, the upper $100(1/t)$ percentile could be obtained by solving

$$F_X(x) = \Phi\left(\frac{h(\rho x) - \mu}{\sigma}\right) = 1 - 1/t, \tag{2.1}$$

i.e., $f(x) = h(\rho x) - \mu - \sigma z_{1/t} = 0$, where $z_{1/t}$ denotes the upper $100(1/t)$ percentile for the standard normal distribution. By using an implicit function, named Lambert W function (see [2]), the solution for (2.1) equals $W[\exp(\mu + \sigma z_{1/t})]/\rho$, where Lambert function $W(x)$ satisfies that $W(x)\exp[W(x)] = x$.

Using the function `ProductLog` implemented in Mathematica, the value of the Lambert W function can be obtained numerically. `ProductLog` returns the value of the function by applying the Newton-Raphson method after approximating Lambert W function by a rational function or an asymptotic expansion. Recently, [3] studied the new class of the distributions related to the Lambert W function such as the random variable $Y = U\exp(\gamma U)$ ($\gamma \in \mathbb{R}$), where U denotes a continuous random variable. If we put $Y = e^{\mu+\sigma Z}/\rho$ and $\gamma = \rho$, then it holds that $U \sim HLN(\rho; \mu, \sigma^2)$, where Z follows the standard normal distribution. As stated in [3], we can evaluate the quantiles for the distribution of Y exactly using the inverse function of $W(x)$ with an explicit form if we can use the exact quantiles for the distribution of U. However, for the hybrid lognormal distribution, we can obtain only the quantiles of Y. Our main purpose is to obtain the approximate solution for (2.1) with an explicit form under large t (henceforth, x_{tail} denotes the solution for (2.1)).

For our purpose, we have the bounds for $1-F_X(x)$ in the following theorem.

Theorem 1. *For any positive x that satisfies $h(\rho x) - \mu > 0$, it holds that*

$$\mathcal{L}(x) < 1 - F_X(x) < \mathcal{U}(x),$$

where $h(x) = x + \ln x$,

$$\begin{aligned}\mathcal{L}(x) &= \left\{\left(\frac{h(\rho x)-\mu}{\sigma}\right)^{-1} - \left(\frac{h(\rho x)-\mu}{\sigma}\right)^{-3}\right\}\phi\left(\frac{h(\rho x)-\mu}{\sigma}\right),\\ \mathcal{U}(x) &= \left(\frac{h(\rho x)-\mu}{\sigma}\right)^{-1}\phi\left(\frac{h(\rho x)-\mu}{\sigma}\right).\end{aligned}$$

Proof. We primarily show that $1 - F_X(x) < \mathcal{U}(x)$ (the lower bound for $1 - F_X(x)$ can be completed in a similar manner). It holds that

$$1 - \Phi\left(\frac{h(\rho x)-\mu}{\sigma}\right) = \frac{1}{2}\left\{1 - \mathrm{erf}\left(\frac{h(\rho x)-\mu}{\sqrt{2}\sigma}\right)\right\} = \frac{1}{\sqrt{\pi}}\int_{\frac{h(\rho x)-\mu}{\sqrt{2}\sigma}}^{\infty} \exp(-t^2)dt,$$

where $\mathrm{erf}(x)$ denotes the error function:

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x \exp(-t^2)dt.$$

Then, noting that $t^{-2}\exp(-t^2)$ is positive for all $t \in \mathbb{R}$, the proof is completed via the integration by parts. □

Applying l'Hospital's rule, we also have

$$\lim_{x\to+\infty}\frac{1-F_X(x)}{\mathcal{U}(x)} = 1, \quad \lim_{x\to+\infty}\frac{1-F_X(x)}{\mathcal{L}(x)} = 1.$$

Therefore, in the next section, we consider the approximations to the tail quantile for the hybrid lognormal distribution on the basis of $\mathcal{U}(x)$ and $\mathcal{L}(x)$.

§3. Tail quantile approximations

By replacing the left side of (2.1) with $\mathcal{U}(x)$ and taking logarithm of both sides, we have

$$\frac{1}{2}\ln 2\pi + \frac{1}{2}\left(\frac{h(\rho x)-\mu}{\sigma}\right)^2 + \ln\left(\frac{h(\rho x)-\mu}{\sigma}\right) \simeq \ln t \tag{3.1}$$

for large t if $h(\rho x) - \mu > 0$ holds. The left side of (3.1) can be expressed as $(\rho^2 x^2)/(2\sigma^2) + \mathrm{o}(x^2)$,

$$x_{tail} \simeq \frac{s}{\rho} + \varepsilon, \tag{3.2}$$

where $s = \sigma\sqrt{2\ln t}$ and $\varepsilon = \mathrm{o}(s/\rho)$. Put (3.2) into (3.1) and ignoring ε^i for $i \geq 2$, we have $\varepsilon \simeq -\mathcal{F}_0(s)/\mathcal{F}_1(s)$ under $g(s,\mu) > 0$, where $g(s,\mu) = h(s) - \mu$,

$$\begin{aligned}\mathcal{F}_0(s) &= \frac{1}{2}\ln 2\pi + \ln\frac{g(s,\mu)}{\sigma} + \frac{1}{2\sigma^2}(\ln s - \mu)(2s + \ln s - \mu),\\ \mathcal{F}_1(s) &= \frac{\rho(s+1)}{sg(s,\mu)} + \frac{\rho(s+1)}{s\sigma^2}g(s,\mu).\end{aligned}$$

Thus, for $g(s,\mu) > 0$, we have an approximate solution for (2.1) on the basis of $\mathcal{U}(x)$:

$$\widetilde{x}_u = (s/\rho) - \mathcal{F}_0(s)/\mathcal{F}_1(s). \tag{3.3}$$

Similarly to the above derivation, by replacing the left side of (2.1) with $\mathcal{L}(x)$ and taking logarithm of both sides, we have

$$\begin{aligned}&\frac{1}{2}\ln 2\pi + \frac{1}{2}\left(\frac{h(\rho x)-\mu}{\sigma}\right)^2 - \ln\left(\frac{h(\rho x)-\mu+\sigma}{\sigma}\right)\\ &-\ln\left(\frac{h(\rho x)-\mu-\sigma}{\sigma}\right) + 3\ln\left(\frac{h(\rho x)-\mu}{\sigma}\right) \simeq \ln t\end{aligned} \tag{3.4}$$

for large t if $h(\rho x) - \mu - \sigma > 0$ holds. The left side of (3.4) can be also expressed as $(\rho^2 x^2)/(2\sigma^2) + \mathrm{o}(x^2)$ and we also have (3.2) in this case. Put (3.2) into (3.4) and ignoring ε^i for $i \geq 2$, we have $\varepsilon \simeq -\mathcal{G}_0(s)/\mathcal{G}_1(s)$ under $g(s,\mu) - \sigma > 0$, where

$$\begin{aligned}\mathcal{G}_0(s) &= \frac{1}{2}\ln 2\pi + \ln\frac{g^3(s,\mu)}{\sigma\{g(s,\mu)+\sigma\}\{g(s,\mu)-\sigma\}}\\ &\quad + \frac{1}{2\sigma^2}(\ln s - \mu)(2s + \ln s - \mu),\\ \mathcal{G}_1(s) &= \frac{\rho(s+1)\{g(s,\mu)+\sqrt{3}\sigma\}\{g(s,\mu)-\sqrt{3}\sigma\}}{sg(s,\mu)\{g(s,\mu)+\sigma\}\{g(s,\mu)-\sigma\}} + \frac{\rho(s+1)}{s\sigma^2}g(s,\mu).\end{aligned}$$

Thus, for $g(s,\mu)-\sigma>0$, we have another approximate solution for (2.1):

$$\widetilde{x}_\ell \;=\; (s/\rho)-\mathcal{G}_0(s)/\mathcal{G}_1(s). \tag{3.5}$$

Furthermore, we can consider approximated function of $f(x)$ around the tail quantile approximation via the Taylor expansion. Therefore, the approximate solution up to the first and the second order for $f(x)=0$ can be obtained as

$$\begin{aligned}\alpha_1(\widetilde{x}_{tail}) &= \widetilde{x}_{tail}\left(1-\frac{f(\widetilde{x}_{tail})}{\rho\widetilde{x}_{tail}+1}\right),\\ \alpha_2(\widetilde{x}_{tail}) &= \widetilde{x}_{tail}\{\rho\widetilde{x}_{tail}+2-\sqrt{(\rho\widetilde{x}_{tail}+1)^2+2f(\widetilde{x}_{tail})}\},\end{aligned}$$

respectively, where $\widetilde{x}_{tail}$ denotes a tail quantile approximation for the distribution $F_X(x)$. Therefore, we have the modification of $\widetilde{x}_u$ and $\widetilde{x}_\ell$ if we put the quantile approximations stated in (3.3) and (3.5) into the above, respectively.

§4. Numerical results

Table 1: The results for (3.3) when $\mu=-1.0,\sigma=0.5$

t	$\widetilde{x}_u$	$\alpha_1(\widetilde{x}_u)$	$\alpha_2(\widetilde{x}_u)$	x_{tail}
10	0.542772 (0.031202)	0.440418 (0.107345)	0.446009 (0.100769)	0.446683 (0.100000)
20	0.610612 (0.012721)	0.498813 (0.054073)	0.504534 (0.050416)	0.505209 (0.050000)
40	0.674274 (0.005229)	0.553386 (0.027216)	0.559247 (0.025223)	0.559928 (0.025000)
100	0.753628 (0.001633)	0.621437 (0.010969)	0.627459 (0.010096)	0.628150 (0.010000)
200	0.810776 (0.000682)	0.670586 (0.005512)	0.676708 (0.005050)	0.677403 (0.005000)

Table 2: The results for (3.5) when $\mu=-1.0,\sigma=0.5$

t	$\widetilde{x}_\ell$	$\alpha_1(\widetilde{x}_\ell)$	$\alpha_2(\widetilde{x}_\ell)$	x_{tail}
10	0.536381 (0.033854)	0.441178 (0.106430)	0.446121 (0.100640)	0.446683 (0.100000)
20	0.606016 (0.013541)	0.499326 (0.053736)	0.504609 (0.050370)	0.505209 (0.050000)
40	0.670694 (0.005503)	0.553765 (0.027083)	0.559302 (0.025205)	0.559928 (0.025000)
100	0.750864 (0.001702)	0.621711 (0.010928)	0.627499 (0.010090)	0.628150 (0.010000)
200	0.808422 (0.000708)	0.670809 (0.005494)	0.676740 (0.005048)	0.677403 (0.005000)

Now we show the numerical results for the proposed approximations: the values $\widetilde{x}_u$, $\widetilde{x}_\ell$, $\alpha_1(\widetilde{x}_u)$, $\alpha_2(\widetilde{x}_u)$, $\alpha_1(\widetilde{x}_\ell)$, and $\alpha_2(\widetilde{x}_\ell)$ compared to the exact tail quantile calculated by applying Newton-Raphson method to $f(x)=0$. In all the tables, the values between the parentheses denote the probability $1-F_X(x)$

for the quantiles or their approximations. The parameter $\rho = 1.0$ is fixed in all the examinations because the performance of the approximations do not depend on the parameter ρ.

At first, we compare the accuracy of the approximations based on (3.3) and (3.5). Tables 1–2 list the numerical results in the both cases for $\mu = -1.0$ and $\sigma = 0.5$. In each case listed in Tables 1–2, the approximation on the basis of (3.5) performed better than that for (3.3). Further, in the other cases, we observe the similar results. Henceforth, we list the results for only the case of (3.5).

Table 3: The results for (3.5) when $\mu = -1.0, \sigma = 1.0$

t	$\widetilde{x}_\ell$	$\alpha_1(\widetilde{x}_\ell)$	$\alpha_2(\widetilde{x}_\ell)$	x_{tail}
10	0.875740 (0.040662)	0.660275 (0.106534)	0.672810 (0.100886)	0.674836 (0.100000)
20	1.061886 (0.016922)	0.816187 (0.053364)	0.828614 (0.050439)	0.830537 (0.050000)
40	1.235096 (0.007218)	0.966381 (0.026668)	0.978323 (0.025209)	0.980082 (0.025000)
100	1.448925 (0.002403)	1.157002 (0.010644)	1.168113 (0.010076)	1.169646 (0.010000)
200	1.601469 (0.001062)	1.295785 (0.005311)	1.306245 (0.005035)	1.307620 (0.005000)

Table 4: The results for (3.5) when $\mu = 1.0, \sigma = 1.0$

t	$\widetilde{x}_\ell$	$\alpha_1(\widetilde{x}_\ell)$	$\alpha_2(\widetilde{x}_\ell)$	x_{tail}
10	1.674922 (0.116888)	1.731817 (0.100099)	1.732182 (0.099998)	1.732174 (0.100000)
20	1.974275 (0.049015)	1.967888 (0.050001)	1.967891 (0.050000)	1.967891 (0.050000)
40	2.208543 (0.022703)	2.180382 (0.025005)	2.180438 (0.025000)	2.180438 (0.025000)
100	2.477799 (0.008536)	2.435891 (0.010004)	2.435992 (0.010000)	2.435993 (0.010000)
200	2.662717 (0.004120)	2.614566 (0.005002)	2.614684 (0.005000)	2.614685 (0.005000)

Next, we set the parameter $\sigma = 1.0$ in Tables 3–4. The results for Tables 2 and 3 imply that the approximations perform better for small σ. Further, it could be also observed that the results for $\mu = 1.0$ (Table 4) performed better than those for $\mu = -1.0$ (Table 3).

However, in the case of small t and large μ, our approximations cannot be calculated because our approximations require the conditions $g(s, \mu) > 0$ and $g(s, \mu) - \sigma > 0$, respectively. For example, both of the conditions does not hold for $t = 10, \mu = 3.0$, and $\sigma = 1.0$.

§5. Concluding remarks

We have the tail quantile approximations for the hybrid lognormal distribution with explicit forms. We obtain the both of the bounds for $1 - F_X(x)$ in Theorem 1 and they could be also regarded as the asymptotic approximations of $1 - F_X(x)$ for large x, respectively. In Section 3, we can obtain the rough approximations for the tail quantile on the basis of the bounds for $1 - F_X(x)$. Further, we modified the approximations by expanding $f(x)$ at $x = \widetilde{x}_u$ and $x = \widetilde{x}_\ell$, respectively. By numerical evaluations, we recommend the quantile approximation $\alpha_2(\widetilde{x}_\ell)$. It may be also applicable for one of the useful initial value for decreasing the number of replications when we apply Newton-Raphson method to (2.1).

In particular, our approximations perform better for large t and μ, small σ. However, when μ is too large, the proposed approximations could not be applied because they require the conditions $g(s, \mu) > 0$ and $g(s, \mu) - \sigma > 0$, respectively. Therefore, relaxing the condition is one of the future problems in this paper.

Further, if we can improve the rough approximations $\widetilde{x}_u$ and $\widetilde{x}_\ell$, all the approximations will perform better. Therefore, another approach may be required in order to improve $\widetilde{x}_u$ and $\widetilde{x}_\ell$ because considering the error term ε up to the second order will not perform better drastically.

Acknowledgments

The author would like to express the sincere gratitude to Professor Michiaki Kai, Oita University of Nursing and Health Sciences, who introduced the theoretical problems and the applications of the hybrid lognormal distribution to the author. Finally, the author also wish to thank the Editor and the anonymous reviewer for insightful and constructive comments, which greatly improved the paper.

References

[1] A. Capitanio, *On the approximation of the tail probability of the scalar skew-normal distribution*, Metron **68** (2010), 299–308.

[2] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, D. E. Knuth, *On the Lambert W function*, Adv. Comput. Math. **5** (1996), 329–359.

[3] G. M. Goerg, *Lambert W random variables – a new family of generalized skewed distributions with applications to risk estimation*, Ann. Appl. Stat. **5** (2011), 2197–2230.

[4] S. Kumazawa, T. Numakunai, *A new theoretical analysis of occupational dose distributions indicating the effect of dose limits*, Health Phys. **41** (1981), 465–475.

[5] N. T. Longford, *Small-sample estimators of the quantiles of the normal, log-normal and Pareto distributions*, J. Statist. Comput, Simulat. **82** (2012), 1383–1395.

[6] L. Peng, Y. Qi, *Confidence regions for high quantiles of a heavy tailed distribution*, Anal. Statist. **34** (2006), 1964–1986.

[7] R. -D. Reiss, *Approximate Distributions of Order Statistics: With Applications to Nonparametric Statistics*, Springer, New York, 1989.

[8] S. Stephan, M. Fischer, *A tail quantile approximation for the Student t distribution*, Commun. Statist. Theory Methods **41** (2012), 2617–2625.

Nobumichi Shutoh
Graduate School of Maritime Sciences, Kobe University
5-1-1, Fukae-minamimachi, Higashinada-ku Kobe, Hyogo 658-0022, Japan
E-mail: shutoh@maritime.kobe-u.ac.jp