Some Useful Properties of Log-Logistic Random Variables for Health Care Simulations

David E. Clark[1,*] and Muhammad El-Taha[2]

[1]*Department of Surgery, Center for Outcomes Research and Evaluation, Maine Medical Center, 22 Bramhall Street, Portland ME 04102, USA*

[2]*Department of Mathematics and Statistics, University of Southern Maine, 96 Falmouth Street, Portland ME 04104, USA*

Abstract: A log-logistic (LL) random variable is one whose logarithm has a logistic distribution. Since the logistic distribution is similar to the normal distribution, log-logistic random variables are similar to log-normal (LN) random variables. However, many of the important properties of LN random variables can only be described using integrals, while the corresponding properties of LL random variables can be described using simple algebra. LL random variables may therefore be a useful alternative to LN random variable for computer simulation of operating room processes or other health care applications, especially when they fit the data more closely. We review the properties of LL random variables, and derive some relationships of the mean residual time to the median residual time. We describe methods of fitting LL distributions to observed data, and discuss potential advantages of using them for simulation of operating room utilization.

Keywords: Log-logistic distribution, Log-normal distribution, Mean residual time, Median residual time, Simulation.

INTRODUCTION

The efficient use of surgical operating rooms is a major financial concern for hospitals, since an estimated 40% of their total revenue is derived from this source [1]. Patient satisfaction and staff morale can be adversely affected by delays or other inconveniences related to the operating rooms. It is even possible that patient safety could be impaired if essential personnel or facilities were unavailable due to suboptimal resource management.

Analysis of these complicated systems has been undertaken using computer simulation, with some success [2, 3]. The duration of surgical procedures has previously been modeled using log-normal distributions [4, 5]. However, there is no theoretical reason why a log-normal distribution should be preferred, and other distributional forms may have advantages for some applications.

Log-logistic distributions were first studied in detail by Fisk [6], and are sometimes referred to by his name [7]. A log-logistic random variable can be simply defined as one whose logarithm has a logistic distribution. Since a logistic distribution is similar to a normal distribution, a log-logistic (LL) distribution will be similar to a log-normal (LN) distribution, but with a somewhat narrower peak and somewhat heavier tails.

Unlike log-normal distributions, however, log-logistic distributions have hazard functions and other properties that can be specified using simple algebraic formulas, and therefore may be easier to incorporate into a computer program.

This article will present some theoretical background about LL distributions, and demonstrate some applied results to show how they may provide a useful alternative to LN distributions for the purpose of analyzing or simulating the duration of surgical procedures. We will also derive some relationships between the mean and median residual time functions for LL random variables that may add to their value in simulation modeling.

BASIC PROPERTIES OF LOG-LOGISTIC RANDOM VARIABLES

The properties of a logistic random variable Y with mean $-\infty < \mu < \infty$ and variance $\sigma^2 > 0$, which we can abbreviate $L(\mu, \sigma^2)$, have been described elsewhere [8]. This distribution is familiar in health care research because of its application to regression modeling of binary outcomes [9].

The probability density function (pdf) of a logistic random variable is

$$g_Y(y) = \frac{\pi}{\sigma\sqrt{3}} \frac{e^{-\pi(y-\mu)/\sigma\sqrt{3}}}{\left(1 + e^{-\pi(y-\mu)/\sigma\sqrt{3}}\right)^2}, \ -\infty < y < \infty.$$

It is useful at this point to make the substitutions

*Address correspondence to this author at the Department of Surgery, Center for Outcomes Research and Evaluation, Maine Medical Center, 22 Bramhall Street, Portland ME 04102, USA; Tel: 207-774-2381; Fax: 207-774-0459; E-mail: clarkd@mmc.org

$$\alpha = e^{\mu} \ , \ \ \beta = \frac{\pi}{\sigma\sqrt{3}} \ , \ (\alpha > 0, \ \beta > 0).$$

Although some previous authors have used different parameterizations for the purpose of describing LL distributions, the substitutions α and β have been most commonly used in recent years because they produce very simple formulas, especially for the LL CDF. We can rewrite the logistic pdf as

$$g_Y(y) = \frac{\beta \, (\alpha)^{\beta} \, e^{-\beta y}}{\left(1 + (\alpha)^{\beta} \, e^{-\beta y}\right)^2}$$

$$= \frac{\beta \, (1/\alpha)^{\beta} \, e^{\beta y}}{\left(1 + (1/\alpha)^{\beta} \, e^{\beta y}\right)^2} \ , \ -\infty < y < \infty.$$

Let Y be the logarithm of another random variable T (*i.e.*, $T = e^{Y}$). Then the pdf of T can be obtained using the transformation technique as

$$f_T(t) = \left| \frac{df^{-1}(t)}{dt} \right| g_Y\left(f^{-1}(t)\right)$$

$$= \left| \frac{d\ln(t)}{dt} \right| \frac{\beta \, (1/\alpha)^{\beta} \, e^{\beta(\ln(t))}}{\left(1 + (1/\alpha)^{\beta} \, e^{\beta(\ln(t))}\right)^2}$$

$$= \frac{(\beta/\alpha) \, (t/\alpha)^{\beta-1}}{\left(1 + (t/\alpha)^{\beta}\right)^2} \ , \ 0 \le t < \infty; \beta > 1.$$

We can refer to the log-logistic random variable T with the parameters α and β as LL(α,β). When $\beta > 1$, the pdf of a LL(α,β) random variable is 0 at t=0, and has a single maximum (mode) at

$$t_{pdf \max} = \alpha \left(\frac{\beta-1}{\beta+1}\right)^{\frac{1}{\beta}}.$$

From the pdf of T, we can also obtain its cumulative distribution function (CDF) as

$$F_T(t) = \int_0^t f_T(u)\,du = \frac{(t/\alpha)^{\beta}}{1 + (t/\alpha)^{\beta}} = \frac{1}{1 + (t/\alpha)^{-\beta}} \ ,$$

and its survival function $S_T(t) = 1 - F_T(t)$. The hazard function $h(t)$ can be defined as the probability that an event occurs at time t, given that it has not occurred before time t, and can be calculated for T as

$$h_T(t) = \frac{f_T(t)}{1 - F_T(t)} = \frac{(\beta/\alpha)(t/\alpha)^{\beta-1}}{1 + (t/\alpha)^{\beta}} \ .$$

The hazard function will be similar to the pdf when T is small, then become increasingly larger than the pdf as T increases, reaching a maximum at

$$t_{haz \max} = \alpha \, (\beta - 1)^{\frac{1}{\beta}} \ ,$$

and decreasing thereafter. For $\beta > 1$, it can be seen that $t_{hazmax} > t_{pdfmax}$.

The rth moments of T exist for $r < \beta$, and can be expressed using a moment-generating function [6], leading to

$$E(T^r) = \alpha^r \, \frac{r\pi/\beta}{\sin(r\pi/\beta)} \ ,$$

where $E(U)$ is the expectation operator. Thus, if $\beta > 1$ the mean of T will be

$$E(T) = \frac{\alpha(\pi/\beta)}{\sin(\pi/\beta)}$$

and if $\beta > 2$ the variance of T will be

$$Var(T) = E(T^2) - \left(E(T)\right)^2 = \frac{\alpha^2(2\pi/\beta)}{\sin(2\pi/\beta)} - \left(\frac{\alpha(\pi/\beta)}{\sin(\pi/\beta)}\right)^2 .$$

The quantile function $Q(q)$ for a log-logistic random variable can be obtained by inversion of the CDF as

$$Q(q) = F^{-1}(q) = \alpha\left(\frac{q}{1-q}\right)^{1/\beta} .$$

It can be seen that the 0.5th quantile (i.e., 50th percentile or median) of T is just α. This simple form for the quantile function also makes it easy to generate a LL(α,β) random variate: If U is a uniform (0,1) random variate, then

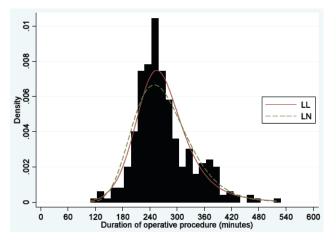$$T = \alpha\left(\frac{U}{1-U}\right)^{1/\beta} \text{ will be LL}(\alpha,\beta).$$

USING LOG-LOGISTIC RANDOM VARIABLES TO MODEL A SAMPLE OF DATA

Reasonable parameters for a log-logistic random variable (or a log-normal random variable) to approximate a sample of data can be most easily chosen using a "method of moments". A simple method of moments estimator could be obtained by setting the parameter μ equal to the mean of the log-transformed data, and setting the parameter σ^2 equal to the variance of the log-transformed data. As an example, theoretical LL and LN distributions were fitted in this way to data from 333 cases of Coronary Artery Bypass Grafting (CABG) performed at the Maine Medical Center during 2013. The use of these data for this purpose, without personal identifiers, was proposed to
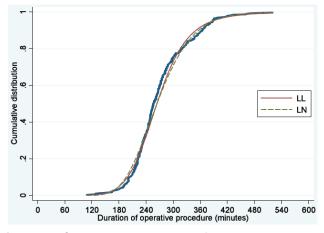
an Institutional Review Board, which ruled it exempt from further review.

Figure **1** compares a histogram of the recorded operative durations (grouped by 15-minute intervals) to the pdf of a LL(264,7.77) distribution and a LN distribution with the same mean and variance. In this example, the LL distribution is seen to have a narrower and taller peak than the LN distribution; the heavier LL "tails" are more difficult to appreciate. The shape of the histogram is somewhat closer to the LL pdf than to the LN pdf.

Figure 1: Histogram of operative procedure times for Coronary Artery Bypass Graft, with superimposed log-logistic (LL, solid line) and log-normal (LN, dotted line) probability density functions fitted using a method of moments.

Figure **2** compares the empirical CDF to the CDFs for these LL and LN distributions. A Kolmogorov-Smirnov test shows that the LL distribution fits the data somewhat better than the LN distribution (p=0.135 for LL, p=.050 for LN).

Figure 2: Cumulative distribution of operative procedure times for Coronary Artery Bypass Graft (heavy solid line), with superimposed log-logistic (LL, lighter solid line) and log-normal (LN, dotted line) cumulative distribution functions fitted using a method of moments.

It is not necessarily true that a LL distribution will fit every sample of operative times better than a LN distribution. It is also possible that a closer fit might be obtained by using other moments of the LL or LN distributions, as discussed by Ashkar and Mahdi [10]. It is also possible to fit a theoretical distribution using specific quantiles of the empirical data [11].

Many software programs are available to fit time-to-event data to parametric distributions using maximum likelihood, for example the *streg* procedure in Stata (StataCorp, College Station TX) described elsewhere [12]. LL and LN models are naturally expressed as "accelerated failure time" (AFT) models, in which the natural logarithm of the modeled duration time *t* for subject *j* is expressed as a linear function of covariates, that is

$$log(t_j) = \boldsymbol{x_j b} + \varepsilon_j = b_{0j} + b_{1j}x_{1j} + \ldots + b_{kj}x_{kj} + \varepsilon_j,$$

where $\boldsymbol{x_j} = (x_{1j} \ldots x_{kj})$ is a vector of covariates, $\boldsymbol{b} = (b_0 \ldots b_k)$ is a vector of regression coefficients, and ε_j is the error; a LL model is obtained if ε_j has a logistic distribution, and a LN model is obtained if ε_j is has a normal distribution

Exponentiating the AFT equation allows the operative duration to be estimated as

$$t_j = e^{b_{0j}} e^{b_{1j}x_{1j}} \ldots e^{b_{kj}x_{kj}}$$

Thus, an AFT model can be interpreted such that the effect of a covariate prolongs the predicted duration if the coefficient is positive and shortens the duration if the coefficient is negative. Using our example of CABG surgery, the Stata command *streg, dist*(*loglogistic*) without covariates estimates b_0=5.57. Therefore,

$$\alpha_0 = e^{b_0} = e^{5.57} = 262.4 \ .$$

The software also estimates β =7.89 (Stata actually reports γ=1/β), so the maximum likelihood estimates of α and β are similar to those obtained by a method of moments. If we add a covariate (x_1=1 if the case is an emergency, x_1=0 otherwise), Stata estimates b_0=5.57, b_1=-0.07, and β =7.92. Thus, for an emergency case we can estimate

$$\alpha_1 = e^{b_0+b_1} = e^{5.57-0.07} = 244.7 \ .$$

This can be interpreted to show that emergency surgery is associated with a shortening of the median duration by a factor of about 244.7/262.4 = 0.933.

Although AFT models may be a useful method to analyze operative duration, it is more common to analyze time-to-event variables using a proportional hazards model. Unfortunately, this is not easy with either the LL or the LN distributions. However, a similar approach can be utilized with a LL distribution and a proportional odds model. It can be seen from the formula for the LL CDF that

$$Odds(T < t) = \frac{F(t)}{1 - F(t)} = (t / \alpha)^\beta .$$

From this simple expression for the odds, it is easy to calculate an odds ratio (*OR*) measuring the effect of being an emergency case, for example

$$OR = \frac{(t / \alpha_1)^\beta}{(t / \alpha_0)^\beta} = \left(\frac{1 / \exp(b_0 + b_1)}{1 / \exp(b_0)} \right)^\beta = (\exp(-b_1))^\beta$$

$$= e^{0.07 * 7.92} = 1.74.$$

This may be interpreted that the odds of having finished the procedure for any time *t* are multiplied by a factor of 1.74 if the case was an emergency. In some settings, a proportional odds model may be preferable to an AFT model, and the LL distribution is the only one that allows both formulations [13].

MEAN AND MEDIAN RESIDUAL TIME FUNCTIONS

Any realistic simulation of operating room times should allow for reassignment of rooms for procedures scheduled to follow a case that is taking more time than expected. In "real time", information about the delayed case can be incorporated into the estimate of remaining time [14], but for simulation (or other predictions before the procedure begins) it is useful to have a method for prediction based upon the theoretical distribution of times. The most commonly used predictive functions for this purpose are the mean residual time and the median residual time [15, 16].

The mean residual time function, which we shall abbreviate as *rtm(t)*, is the expected time remaining given that a procedure has not finished by time *t,* that is,

$$rtm(t) = E(T - t \mid T > t) = \frac{\int_t^\infty S(u)\, du}{S(t)} . \tag{1}$$

Some authors have referred to this as the mean residual lifetime function [15]. For a LL distribution, the mean residual time function can be expressed as

$$rtm(t) = \left(1 + (t / \alpha)^\beta\right) \int_t^\infty \frac{1}{1 + (u / \alpha)^\beta}\, du , \tag{2}$$

but unfortunately the integral cannot be evaluated analytically. Since many statistical programs can produce values for Beta Functions, numerical values of *rtm(t)* may be most easily obtained by changing the variable of integration to *F(t)* [17]. Using *dF(u) = f(u)du*, we have the interesting result that

$$\int_t^\infty \frac{1}{1 + (u / \alpha)^\beta}\, du = \int_{F(t)}^1 \frac{1}{h(u)}\, dF(u) .$$

Expressing the integral in terms of *F(u)* leads to

$$\int_t^\infty \frac{1}{1 + (u / \alpha)^\beta}\, du = \frac{\alpha}{\beta} \int_{F(t)}^1 (F(u))^{\frac{1}{\beta} - 1} (1 - F(u))^{-\frac{1}{\beta}}\, dF(u)$$

$$= \frac{\alpha}{\beta} \left(B\left(\frac{1}{\beta}, 1 - \frac{1}{\beta}\right) - B_{F(t)}\left(\frac{1}{\beta}, 1 - \frac{1}{\beta}\right) \right),$$

where *B* denotes a complete Beta Function and $B_{F(t)}$ denotes an incomplete Beta Function evaluated at *F(t)*.

Using properties of Beta Functions [18], Equation (2) can therefore be rewritten as

$$rtm(t) = \frac{\alpha}{\beta}\left(1 + (t / \alpha)^\beta\right) \left(\frac{\pi}{\sin(\pi / \beta)} - B_{F(t)}\left(\frac{1}{\beta}, 1 - \frac{1}{\beta}\right) \right)$$

$$= \frac{\alpha(\pi / \beta)}{\sin(\pi / \beta)}\left(1 + (t / \alpha)^\beta\right) \left[1 - I\left(\frac{1}{\beta}, 1 - \frac{1}{\beta}, F(t)\right) \right], \tag{3}$$

where *I* denotes a regularized Beta Function, and the term in brackets can also be interpreted as the survival function (1-CDF) for a Beta Distribution with parameters (1/*β*, 1-1/*β*), evaluated at *F(t)*. It can be seen that *rtm(t)* will equal the mean of *T* when *t*=0, which is to be expected given the known relationship

$$\int_0^\infty S(u)du = \int_0^\infty u\, f(u)du$$

for any time-to-event distribution [19].

It is easy to find the median residual time function, $rtq_{0.5}(t)$ by solving

$$1 - F(t + rtq_{0.5}(t)) = 0.5(1 - F(t)). \tag{4}$$

For a LL random variable this results in

$$rtq_{0.5}(t) = \alpha \left(1 + 2(t/\alpha)^{\beta}\right)^{1/\beta} - t , \qquad (5)$$

and will equal the median of T when t=0. In addition to being easier to calculate, $rtq_{0.5}(t)$ has been advocated as a more robust measure of "ageing" than $rtm(t)$ [20]. Other quantile residual time functions can also be defined by substituting a fraction other than 0.5 in Equation (4).

RELATIONSHIPS OF MEAN AND MEDIAN RESIDUAL TIMES

We will derive several properties of $rtm(t)$ and $rtq_{0.5}(t)$ for LL random variables that are useful in their application to simulation studies. In the first two theorems, we will define some relationships of these functions to the hazard function $h(t)$ or its inverse

$$1/h(t) = \frac{1 + (t/a)^{\beta}}{(\beta/\alpha)(t/\alpha)^{\beta-1}} = (t/\beta)\left(1 + (t/\alpha)^{-\beta}\right). \qquad (6)$$

We will then derive additional relationships between $rtm(t)$ and $rtq_{0.5}(t)$, for the usual case where β >1.

Theorem 1: For a LL random variable, the value of t where $rtm(t)$ is minimized (t_{rtmmin}) is less than the value of t where $h(t)$ is maximized (t_{hazmax}, where also $1/h(t)$ is minimized).

Proof: We can differentiate both sides of Equation (1) to obtain

$$\frac{d\,rtm(t)}{dt} = \frac{f(t)\int_{t}^{\infty} S(u)\,du}{(S(t))^2} - \frac{S(t)}{S(t)} = h(t)\,rtm(t) - 1 . \qquad (7)$$

Thus when the derivative of $rtm(t)$ is zero, $rtm(t)$=1/$h(t)$, although the value of t at this point cannot be determined analytically. Differentiating again gives

$$\frac{d^2 rtm(t)}{dt^2} = h(t)\frac{d\,rtm(t)}{dt} + rtm(t)\frac{d\,h(t)}{dt} .$$

Gupta and Akman have shown that $rtm(t)$ for a LL distribution will have a "bathtub" (concave upward) shape [17, 21], so its second derivative must be positive. At the point where the first derivative of $rtm(t)$ is zero, the derivative of $h(t)$ must therefore be positive, demonstrating that $h(t)$ has not reached its maximum.

Corollary 1: For a LL random variable, when $t < t_{rtmmin}$, $rtm(t) < 1/h(t)$, and when $t > t_{rtmmin}$, $rtm(t) > 1/h(t)$.

Proof: This can be established by rearranging Equation (7).

Lemma 1: If a>0 and b>1, then $a^b > 1 - b + ab$.

Proof:

$$a^b = \exp(b\log a) = 1 + b\log a + \frac{(b\log a)^2}{2!} + \frac{(b\log a)^3}{3!} + \dots$$

$$= 1 + b\left(\log a + \frac{b(\log a)^2}{2!} + \frac{b^2(\log a)^3}{3!} + \dots\right)$$

$$> 1 + b(-1 + \exp(\log a)) = 1 + b(-1 + a) = 1 - b + ab .$$

Theorem 2: For a LL random variable with β >1, the value of t where $rtq_{0.5}(t)$ is minimized is less than the value of t where $h(t)$ is maximized (and $1/h(t)$ is minimized).

Proof: We can differentiate both sides of Equation (5) to obtain

$$\frac{d\,rtq_{0.5}(t)}{dt} = 2\left(1 + 2(t/\alpha)^{\beta}\right)^{\frac{1}{\beta}-1}(t/\alpha)^{\beta-1} - 1 . \qquad (8)$$

If this derivative is set equal to zero, we can find a minimum for $rtq_{0.5}(t)$ as

$$t_{rtq_{0.5}\min} = \alpha\left(2^{\frac{\beta}{\beta-1}} - 2\right)^{-\frac{1}{\beta}} .$$

Lemma 1 can be used to show that this is less than t_{hazmax}.

Corollary 2: For a LL random variable with β >1, $rtq_{0.5}(t) < \dfrac{1}{h(t)}$.

Proof: From Lemma 1, we can say that

$$\left(\frac{1}{\beta(t/\alpha)^{\beta}} + \frac{\beta+1}{\beta}\right)^{\beta} > \frac{1}{(t/\alpha)^{\beta}} + 2 .$$

After several algebraic steps, this leads to

$$(t/\beta)\left(1 + (t/\alpha)^{-\beta}\right) > \alpha\left(1 + 2(t/\alpha)^{\beta}\right)^{\frac{1}{\beta}} - t .$$

Lemma 2: If β >1, α>0, and t>0, then

$$\frac{4(\beta-1)}{\beta}(t/\alpha)^{\beta-1}\left(1 + 2(t/\alpha)^{\beta}\right)^{\frac{1}{\beta}-2} < \frac{1}{1 + (t/\alpha)^{\beta}} .$$

Proof: Consider the function

$$w(z) = z^{1-\frac{1}{\beta}}(1+2z)^{\frac{1}{\beta}-2}(1+z) \text{ , with } \beta > 1 \text{ and } z > 0. \quad (9)$$

Setting the first derivative of $w(z)$ equal to zero demonstrates a stationary point at $z = \beta$-1. The second derivative of $w(z)$ evaluated at this point is

$$-(1/\beta)(\beta-1)^{\frac{-1}{\beta}}(2\beta-1)^{\frac{1}{\beta}-3},$$

which is clearly negative, showing this point to be a maximum for $w(z)$. Evaluating $w(\beta$-1) allows us to say

$$w(z) \leq \left(\frac{\beta-1}{2\beta-1}\right)^{2-\frac{1}{\beta}}\frac{\beta}{(\beta-1)}.$$

Substituting the expression on the right-hand side of Equation (9) for $w(z)$ and rearranging terms gives

$$\left(\frac{2\beta-1}{\beta-1}\right)^{2-\frac{1}{\beta}}\frac{(\beta-1)}{\beta}z^{1-\frac{1}{\beta}}(1+2z)^{\frac{1}{\beta}-2} < \frac{1}{1+z}.$$

The first term in this expression is always greater than 4, and substituting $z = (t/\alpha)^{\beta}$ completes the proof.

Theorem 3: For a LL random variable with β >1, the derivative of $rtq_{0.5}(t)$ is less than the derivative of $rtm(t)$.

Proof: Let us refer to the expression on the left-hand side of Lemma 2 as $v(u)$, and recognize the right-hand side as $S(u)$, the survival function for a LL distribution. Now we can integrate both sides of the inequality to obtain

$$\int_t^\infty v(u)\,du < \int_t^\infty S(u)\,du,$$

and the left-hand side can be evaluated analytically to obtain

$$\left[-\frac{2\alpha}{\beta}\left(1+2(u/\alpha)^{\beta}\right)^{\frac{1}{\beta}-1}\right]_t^\infty < \int_t^\infty S(u)du$$

$$\frac{2\alpha}{\beta}\left(1+2(t/\alpha)^{\beta}\right)^{\frac{1}{\beta}-1} < \int_t^\infty S(u)du.$$

Therefore,

$$2\left(1+2(t/\alpha)^{\beta}\right)^{\frac{1}{\beta}-1}(t/\alpha)^{\beta-1}-1 <$$

$$(\beta/\alpha)(t/\alpha)^{\beta-1}\int_t^\infty S(u)du-1.$$

From Equation (2), Equation (7), and Equation (8), this means

$$\frac{d\,rtq_{0.5}(t)}{dt} < \frac{d\,rtm(t)}{dt}.$$

Corollary 3: For a LL random variable with β >1, the value of t where $rtm(t)$ is minimized is less than the value of t where $rtq_{0.5}(t)$ is minimized.

Proof: This follows immediately from Theorem 3 and the observation that when the derivative of $rtm(t)$ is zero, the derivative of $rtq_{0.5}(t)$ must be negative, so it has not reached a minimum. Combining this with Theorems 1 and 2, we have

$$t_{rtm\,\min} < t_{rtq0.5\,\min} < t_{haz\,\max}.$$

Theorem 4: For a LL random variable with β >1, $rtq_{0.5}(t) < rtm(t)$.

Proof: We have established that $rtq_{0.5}(0)$ =α and $rtm(0)$=α(π/β)/(sin(π/β)), and since (sin(π/β))/(π/β) ≤ 1 [18], $rtq_{0.5}(0) \leq rtm(0)$. With β >1, Theorem 3 means that $rtq_{0.5}(t)$ decreases more quickly (or increases more slowly) than $rtm(t)$, so $rtq_{0.5}(t) < rtm(t)$.

Corollary 4: For a LL random variable with β >1, consider the function

$$lrtm(t) = \frac{2\alpha}{\beta}\left(1+(t/\alpha)^{\beta}\right)\left(1+2(t/\alpha)^{\beta}\right)^{\frac{1}{\beta}-1}. \quad (10)$$
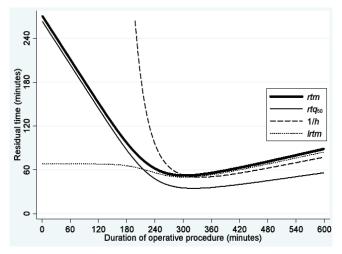
Then $lrtm(t) < rtm(t)$.

Proof: From Theorem 3 and Equation (7) we can show that

$$\left(\frac{drtq_{0.5}(t)}{dt}+1\right)/h(t) < rtm(t),$$

which leads directly to Corollary 4. Considering Corollary 4 and Theorem 4, the closest lower bound for $rtm(t)$ that we have found for a LL random variable with β >1 is the greater of $rtq_{0.5}(t)$ and $lrtm(t)$.

APPLIED EXAMPLE

The LL distribution with α=264.0 and β=7.77, which resulted from estimating the duration of Coronary Artery Bypass Graft procedures above, is used in Figure **3** to illustrate some of the relationships derived in the Theorems and Corollaries above. Equation (3) was used to calculate $rtm(t)$, Equation (5) was used to calculate $rtq_{0.5}(t)$, Equation (6) was used to calculate

$1/h(t)$, and Equation (10) was used to calculate $lrtm(t)$. It can be seen from Figure **3** that $rtq_{0.5}(0)$ is slightly less than $rtm(0)$, after which $rtq_{0.5}(t)$ and $rtm(t)$ gradually diverge. For larger values of t, $rtm(t)$ is more closely approximated by $lrtm(t)$ or even $1/h(t)$ than it is by $rtq_{0.5}(t)$.

Figure 3: For a log-logistic distribution with parameters α=264.0 and β=7.77, mean residual time function (*rtm*, bold solid line), median residual time function (*rtq₅₀*, medium solid line), inverse hazard function ($1/h$, dashed line), and the *rtm* lower bound (*lrtm,* dotted line) defined in Corollary 4.

CONCLUSIONS

The log-logistic distribution results from a simple transformation of the familiar logistic distribution, and has been found useful in survival analysis [13, 17, 22], econometrics [6, 7], and hydrology [10]. Since all these areas of study have some relationship to the flow of patients through an operating suite, it is appealing to apply this mathematical model in the health care simulation.

Log-normal distributions have been used to model operative procedure times, and the Central Limit Theorem could be invoked to argue that the product of numerous effects on these times might lead to a normal distribution for the sum of their logarithms. However, any distribution supported from zero to infinity with a unimodal hazard function could be fitted to the empirical distribution of procedure times, and the choice of one over another should be primarily determined by goodness of fit and practicality.

The mathematical simplicity of the LL hazard is a theoretical advantage compared to more complicated distributions (and even over the LN, which requires evaluation of an integral). The processing time required to compute a LL hazard function is about half of that

required to compute a LN hazard function. These computation times are in the range of a millisecond for a contemporary desktop computer using Stata, so the overall effect on a simulation program will only be noticeable if the calculation is being repeated for multiple state transitions, at frequent intervals during the period of observation, and for many patients. However, the simplicity of the functional expressions may also enable easier and more accurate programming.

In situations where simulation models require an estimate of time remaining (for example, if a surgical procedure would be moved to a different room when the originally scheduled room is "running late"), residual time functions may be useful. While $rtm(t)$ is a natural choice for this purpose, there are several reasons why $rtq_{0.5}(t)$ may be preferable [20]. However, even for skewed distributions, it cannot be simply assumed that the median is less than the mean [23]. The relationships between $rtm(t)$ and $rtq_{0.5}(t)$ we have derived for LL random variables may therefore add to their usefulness for simulation analyses.

Thus, there are several potential reasons to choose a LL distribution for modeling the duration of operative procedures. It requires a relatively large sample to determine whether or not a LL distribution fits a given set of data better than a LN with the same mean and variance [24]. Given the many other sources of randomness in simulating an operating room environment, the choice of one distributional form over another is probably not of critical importance. However, log-logistic distributions should be considered a useful option for modeling procedure durations or other health care applications when their properties are appropriate.

REFERENCES

[1]   Denton B, Viapiano J, Vogl A. Optimization of surgery sequencing and scheduling decisions under uncertainty. Health Care Manag Sci 2007; 10(1): 13-24.
http://dx.doi.org/10.1007/s10729-006-9005-4

[2]   Cardoen B, Demeulemeester E, Belien J. Operating room planning and scheduling: A literature review. Eur J Operational Res 2010; 201(3): 921-932.
http://dx.doi.org/10.1016/j.ejor.2009.04.011

[3]   Dexter F, Traub RD. How to schedule elective surgical cases into specific operating rooms to maximize the efficiency of use of operating room time. Anesth Analg 2002; 94(4): 933-942.
http://dx.doi.org/10.1097/00000539-200204000-00030

[4]   Ledolter J, Dexter F, Epstein RH. Analysis of variance of communication latencies in anesthesia: Comparing means of multiple log-normal distributions. Anesth Analg 2011; 113(4): 888-896.
http://dx.doi.org/10.1213/ANE.0b013e318227518f

[5]   Strum DP, May JH, Vargas LG. Modeling the uncertainty of surgical procedure times: Comparison of log-normal and normal models. Anesthesiology 2000; 92(4): 1160-1167.
http://dx.doi.org/10.1097/00000542-200004000-00035

[6]   Fisk PR. The graduation of income distributions. Econometrica 1961; 29(2): 171-185.
http://dx.doi.org/10.2307/1909287

[7]   Kleiber C, Kotz S. Statistical Size Distributions in Economics and Actuarial Sciences. Hoboken NJ: Wiley-Interscience; 2003.

[8]   Balakrishnan N (ed.). Handbook of the Logistic Distribution. New York: Marcel Dekker, Inc.; 1992.

[9]   Hosmer DW, Jr., Lemeshow S. Applied Logistic Regression, 2nd Edition. New York: John Wiley & Sons; 2000.

[10]   Ashkar F, Mahdi S. Fitting the log-logistic distribution by generalized moments. J Hydrol 2006; 328: 694-703.
http://dx.doi.org/10.1016/j.jhydrol.2006.01.014

[11]   Ragab A, Green J. Estimation of the parameters of the loglogistic distribution based on order statistics. Am J Mathematical and Management Sci 1987; 7: 307-323.

[12]   Cleves MA, Gould WW, Gutierrez RG, Marchenko YU. An Introduction to Survival Analysis using Stata. College Station, TX: StataCorp; 2008.

[13]   Collett D. Modelling Survival Data in Medical Research. London: Chapman & Hall; 1994.

[14]   Dexter F, Epstein RH, Lee JD, Ledolter J. Automatic updating of times remaining in surgical cases using Bayesian analysis of historical case duration data and "instant messaging" updates from anesthesia providers. Anesth Analg 2009; 108(3): 929-940.
http://dx.doi.org/10.1213/ane.0b013e3181921c37

[15]   Finkelstein MS. On the shape of the mean residual lifetime function. Appl Stochastic Models in Business and Industry 2002; 18: 135-146.
http://dx.doi.org/10.1002/asmb.461

[16]   Lillo RE. On the median residual lifetime and its aging properties: A characterization theorem and applications. Naval Res Logistics 2005; 52: 370-380.
http://dx.doi.org/10.1002/nav.20082

[17]   Gupta RC, Akman O, Lvin S. A study of log-logistic model in survival analysis. Biometrical J 1999; 41(4): 431-443.
http://dx.doi.org/10.1002/(SICI)1521-
4036(199907)41:4<431::AID-BIMJ431>3.0.CO;2-U

[18]   Abramowitz M, Stegun IA. Handbook of Mathematical Functions. New York: Dover; 1965.

[19]   Ross S. A First Course in Probability. Englewood Cliffs NJ: Prentice Hall; 1994.

[20]   Schmittlein DC, Morrison DG. The median residual lifetime: A characterization theorem and an application. Operations Res 1981; 29(2): 392-399.
http://dx.doi.org/10.1287/opre.29.2.392

[21]   Gupta RC, Akman O. Mean residual life functions for certain types of nonmonotonic ageing. Commun Statist - Stochastic Models 1995; 11: 219-225.
http://dx.doi.org/10.1080/15326349508807340

[22]   Bennett S. Log-logistic regression models for survival data. Appl Statist 1983; 32(2): 165-171.
http://dx.doi.org/10.2307/2347295

[23]   von Hippel PT. Mean, median, and skew: Correcting a textbook rule. J Stat Educ 2005; 13(2): 1-16.

[24]   Dey AK, Kundu D. Discriminating between the log-normal and log-logistic distributions. Commun Statist - Theor Meth 2010; 39: 280-292.