Proceedings of the 2016 Winter Simulation Conference
T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, and S. E. Chick, eds.

SURVIVAL DISTRIBUTIONS BASED ON THE INCOMPLETE GAMMA FUNCTION RATIO

Andrew G. Glen

Department of Mathematics and Computer Science
The Colorado College
14 E Cache La Poudre
Colorado Springs, CO 80903, USA

Lawrence M. Leemis

Department of Mathematics
The College of William & Mary
PO Box 8795
Williamsburg, VA 23187, USA

Daniel J. Luckett

Department of Biostatistics
University of North Carolina at Chapel Hill
135 Dauer Drive, 3101 McGavran-Greenberg Hall, CB #7420
Chapel Hill, NC 27599-7420, USA

ABSTRACT

A method to produce new families of probability distributions is presented based on the incomplete gamma function ratio. The distributions distributions produced also can include a number of popular univariate survival distributions, including the gamma, chi-square, exponential, and half-normal. Examples that demonstrate the generation of new distributions are provided.

1 INTRODUCTION

The gamma function

$$\Gamma(\alpha) = \int_0^\infty e^{-t} t^{\alpha-1}\, dt$$

for $\alpha > 0$ is a generalization of the factorial function that is prevalent in probability and statistics. When the lower limit of the integral is replaced by x, the resulting function is defined as the incomplete gamma function

$$\Gamma(\alpha, x) = \int_x^\infty e^{-t} t^{\alpha-1}\, dt$$

for $\alpha > 0$ and $x > 0$. The incomplete gamma ratio

$$\frac{\Gamma(\alpha, x)}{\Gamma(\alpha)}$$

for $\alpha > 0$ and $x > 0$ is bounded below at 0 and bounded above at 1, as the numerator is always smaller than the denominator. Let $g(x)$ be a monotonic and increasing function that assumes nonnegative values on the interval $(0, \infty)$. Furthermore, assume that $\lim_{x \to 0^+} g(x) = 0$ and $\lim_{x \to \infty} g(x) = \infty$. Likewise, let $r(x)$ be a monotonic and decreasing function that assumes nonnegative values on the interval $(0, \infty)$. Furthermore, let g and r be differentiable on the interval $(0, \infty)$. Also, assume that $\lim_{x \to 0^+} r(x) = \infty$ and $\lim_{x \to \infty} r(x) = 0$. A family of survival distributions for a random variable X is generated with cumulative distribution functions

$$F(x) = 1 - \frac{\Gamma(\alpha, g(x))}{\Gamma(\alpha)} \qquad x > 0 \tag{1}$$

978-1-5090-4486-3/16/$31.00 ©2016 IEEE

for any $\alpha > 0$, and

$$F(x) = \frac{\Gamma(\alpha, r(x))}{\Gamma(\alpha)} \qquad x > 0$$

for any $\alpha > 0$. The conditions on $g(x)$ and $r(x)$ ensure that $F(x)$ will be a monotonically increasing function with $F(0) = 0$ and $\lim_{x\to\infty} F(x) = 1$.

The probability density function for this family when $g(x)$ is specified is found by differentiating the cumulative distribution function:

$$f(x) = F'(x) = \frac{d}{dx}\left(1 - \int_{g(x)}^{\infty} e^{-t} t^{\alpha-1} dt\right) \Big/ \Gamma(\alpha)$$

which by the chain rule of differentiation reduces to

$$f(x) = e^{-g(x)} g(x)^{\alpha-1} g'(x) \Big/ \Gamma(\alpha) \qquad x > 0$$

for $\alpha > 0$. When $r(x)$ is specified, the probability density function is found as

$$f(x) = F'(x) = \frac{d}{dx}\left(\int_{r(x)}^{\infty} e^{-t} t^{\alpha-1} dt\right) \Big/ \Gamma(\alpha)$$

which reduces to

$$f(x) = -e^{-r(x)} r(x)^{\alpha-1} r'(x) \Big/ \Gamma(\alpha) \qquad x > 0$$

for $\alpha > 0$.

This method of creating new distribution functions from this ratio of gamma functions has a very nice graphical representation. The integrand of the gamma function $b(x) = e^{-x}x^{\alpha-1}$ can be considered the base function of the ratio. It is a positive function and has finite area underneath it on the interval $(0, \infty)$. Thus the denominator of (1) is the entire area under $b(x)$. The numerator is the right tailed area under $b(x)$ on the interval $(g(x), \infty)$. Clearly the ratio of (1) is bounded between 0 and 1, thus it represents the probability from a survivor function of a random variable. Figure 1 shows this relationship where the shaded area represents a portion of the entire area under the base, $b(x)$.

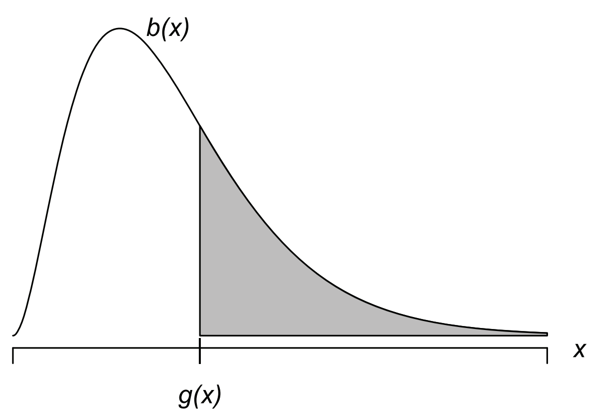

Figure 1: The base function has finite area underneath it and the shaded area represents the portion of that area to the right of $g(x)$.

A number of popular lifetime distributions can be derived using probability density functions of this form. Table 1 shows the appropriate $g(x)$ or $r(x)$ functions and α values associated with a number of popular density functions, many of which are defined in Meeker and Escobar (1998). In the table and the analysis and examples that follow, let n be a positive integer and let α, λ, κ, and σ be positive parameters.

Table 1: Parametric special cases.

Distribution	$g(x)$ or $r(x)$	α	$f(x)$
exponential	$g(x) = \lambda x$	$\alpha = 1$	$\lambda e^{-\lambda x}$
Weibull	$g(x) = (\lambda x)^\kappa$	$\alpha = 1$	$\lambda^\kappa x^{\kappa-1} \kappa e^{-(\lambda x)^\kappa}$
Rayleigh	$g(x) = (\lambda x)^2$	$\alpha = 1$	$2\lambda^2 x e^{-(\lambda x)^2}$
Lomax	$g(x) = \ln((\lambda x + 1)^\kappa)$	$\alpha = 1$	$\lambda\kappa/(1+\lambda x)^{\kappa+1}$
Muth	$g(x) = (1/\lambda)e^{\lambda x} - \lambda x - 1/\lambda$	$\alpha = 1$	$(e^{\lambda x} - \lambda)e^{-(1/\lambda)e^{\lambda x} + \lambda x + 1/\lambda}$
half-normal	$g(x) = x^2/(2\sigma^2)$	$\alpha = 1/2$	$\sqrt{2}e^{-x^2/(2\sigma^2)} \Big/ (\sigma\sqrt{\pi})$
chi-square	$g(x) = x/2$	$\alpha = n/2$	$2^{-n/2}x^{n/2-1}e^{-x/2} \Big/ \Gamma(n/2)$
Erlang	$g(x) = x/\lambda$	$\alpha = n$	$x^{n-1}e^{-x/\lambda} / (\lambda^n (n-1)!)$
gamma	$g(x) = x/\lambda$	—	$\lambda^{-\alpha}x^{\alpha-1}e^{-x/\lambda} \Big/ \Gamma(\alpha)$
generalized gamma	$g(x) = (x/\lambda)^\kappa$	—	$\kappa\lambda^{-\kappa\alpha}x^{\kappa\alpha-1}e^{-(x/\lambda)^\kappa} \Big/ \Gamma(\alpha)$
inverse gamma	$r(x) = (\lambda x)^{-1}$	—	$\lambda^{-\alpha}x^{-\alpha-1}e^{-1/(\lambda x)} \Big/ \Gamma(\alpha)$

While this paper primarily is concerned with a base function $b(x)$ that is the gamma function, any positive function with finite area underneath it would serve as a proper $b(x)$ to use this method to create new distributions. In the special case where the base function $b(x)$ is a probability distribution function of a continuous random variable B, then Figure 1 gives us a graphical depiction of the probability statement $P(B > g(X))$. Further, since $P(B > g(X)) = P(g^{-1}(B) > X)$ or $P(X < g^{-1}(B))$, one can see the relationship of B and X as the inverse of a transformation of an input random variable, i.e., $X = g^{-1}(B)$. Transformations are studied in great detail in many mathematical statistics books, see Hogg et al, (2013) and Casella and Berger (2008) for examples. Because the base functions are not limited to probability distribution functions of random variables, this method a superset of univariate transformations. The method also provides a graphical explanation of such operations.

2 ANALYSIS

The chosen function $g(x)$ or $r(x)$ is related to the hazard function of the resulting distribution. Let $f(x)$, $F(x)$, $S(x)$, and $h(x) = f(x)/S(x)$ be the population probability density function, cumulative distribution function, survivor function, and hazard function respectively. These three results follow.

Result 1. Let the random variable X have probability density function as defined above for a specified $g(x)$ and $\alpha = 1$. Then the hazard function $h(x) = g'(x)$, for $x > 0$, which implies the $g(x)$ is the cumulative hazard function for the distribution.
Proof. The hazard function of X is

$$h(x) = \frac{f(x)}{S(x)} = \frac{f(x)}{1 - F(x)} = \frac{e^{-g(x)}g'(x)}{\Gamma(1, g(x))} = \frac{e^{-g(x)}g'(x)}{\int_{g(x)}^{\infty} e^{-t}dt} = \frac{e^{-g(x)}g'(x)}{e^{-g(x)}} = g'(x)$$

for $x > 0$. □

Result 2. Let the random variable X have probability density function as defined above for a specified $r(x)$ and $\alpha = 1$. Then the hazard function $h(x) = -r'(x)/(e^{r(x)} - 1)$, for $x > 0$.

Proof The hazard function of X is

$$h(x) = \frac{f(x)}{1-F(x)} = \frac{-e^{-r(x)}r'(x)}{1-\Gamma(1,r(x))} = \frac{-e^{-r(x)}r'(x)}{1-\int_{r(x)}^{\infty} e^{-t}dt} = \frac{-e^{-r(x)}r'(x)}{1-e^{-r(x)}} = \frac{-r'(x)}{e^{r(x)}-1}$$

for $x > 0$. □

Comparing the well-known distributions in Table 1 to their hazard functions confirms these results. For example, the exponential distribution has a constant hazard rate $h(x) = \lambda$ for $x > 0$. The Weibull distribution has a hazard rate $\lambda^{\kappa}\kappa x^{\kappa-1}$ for $x > 0$. It is well-known that the Weibull distribution has an increasing failure rate when $\kappa > 1$ and a decreasing failure rate when $\kappa < 1$. This can be seen from the corresponding function $g(x) = (\lambda x)^{\kappa}$. When $\kappa > 1$, $g'(x)$ will be a monotonically increasing function.

These results tell us that a distribution specified by $g(x)$ and $\alpha = 1$ will have an increasing failure rate if $g''(x) > 0$ for all $x > 0$. The distribution will have a decreasing failure rate if $g''(x) < 0$ for all $x > 0$. Although the result in the $r(x)$ case is less tractable, similar conclusions can be made.

Next we consider taking transformations of random variables in this family. We will show that the proposed family of distributions is closed under certain 1–1 transformations. The following result holds in the case where $g(x)$ is specified.

Result 3. Let the random variable X have probability density function as defined above for a specified $g(x)$ and unspecified α. If $Y = \phi(X)$ is a 1–1 transformation from $\{x \,|\, x > 0\}$ to $\{y \,|\, y > 0\}$, then the probability density function of Y is in the same family.
Proof. Let the random variable X have probability density function

$$f_X(x) = e^{-g(x)} g(x)^{\alpha-1} g'(x) \Big/ \Gamma(\alpha) \qquad x > 0,$$

for $\alpha > 0$ and $g(x)$ for which the properties in Section 1 hold. Consider first the case of $Y = \phi(X)$ monotonically increasing. So $Y = \phi(X)$ is a 1–1 transformation from $\{x \,|\, x > 0\}$ to $\{y \,|\, y > 0\}$ with inverse

$$X = \phi^{-1}(Y)$$

and Jacobian

$$J = \frac{dX}{dY} = \frac{d}{dY}\phi^{-1}(Y).$$

By the transformation technique, the probability density function of Y is

$$\begin{aligned} f_Y(y) &= f_X\left(\phi^{-1}(y)\right) \cdot J \\ &= \frac{e^{-g(\phi^{-1}(y))} g(\phi^{-1}(y))^{\alpha-1} g'(\phi^{-1}(y))}{\Gamma(\alpha)} \cdot \frac{d}{dy}\phi^{-1}(y) \\ &= \frac{e^{-g(\phi^{-1}(y))} g(\phi^{-1}(y))^{\alpha-1}}{\Gamma(\alpha)} \cdot \frac{d}{dy} g(\phi^{-1}(y)) \qquad y > 0 \end{aligned}$$

because

$$\frac{d}{dy} g\left(\phi^{-1}(y)\right) = g'\left(\phi^{-1}(y)\right) \cdot \frac{d}{dy}\phi^{-1}(y)$$

by the chain rule of differentiation. This probability density function can be recognized as being from the proposed family and is constructed using the $g(x)$ function $g(\phi^{-1}(y))$. The case of $Y = \phi(X)$ monotonically decreasing is handled in a similar fashion. □

A similar result can be shown for the case where $r(x)$ is specified. This means that the proposed family is closed under transformations such as, for example, $Y = \phi(X) = c_0 X^{c_1}$ for positive, real constants c_0 and c_1.

3 EXAMPLES

New distributions can be developed by defining $g(x)$ or $r(x)$ functions satisfying the conditions in Section 1 that may fit a particular purpose. Here we present three examples of new distributions. Each example displays different design features that make the new distributions desirable. While each could be further developed (as in publications such as Forbes et al., 2001), the following examples demonstrate how new distributions with certain useful properties are generated.

Example 1. A new life distribution with determinable hazard function properties can be derived. Let $g(x) = (\lambda x)^{\kappa}$, for $\lambda, \kappa > 0$. Let $\alpha > 0$ be unspecified. This creates a distribution with probability density function

$$f(x) = \frac{\kappa e^{-(\lambda x)^{\kappa}} x^{\kappa\alpha-1} \lambda^{\kappa\alpha}}{\Gamma(\alpha)} \qquad x > 0.$$

When $\alpha = 1$, this reduces to the Weibull(λ, κ) distribution. When $\alpha = 2$, this creates a "semi-Weibull" distribution with probability density function of X

$$f(x) = \kappa \lambda^{2\kappa} x^{2\kappa-1} e^{-(\lambda x)^{\kappa}} \qquad x > 0.$$

This probability density function is quite similar to that of a Weibull(λ, κ) random variable. The hazard function in this case is

$$h(x) = \frac{\kappa \lambda^{2\kappa} x^{2\kappa-1}}{(\lambda x)^{\kappa} + 1} \qquad x > 0.$$

The cumulative distribution function and survivor function can also be obtained in closed form. This distribution will have a decreasing failure rate when $\kappa \le 1$ and an increasing failure rate when $\kappa \ge 1$. Another special case of this family occurs when we further assume that $\kappa = 1$. The probability density function reduces to

$$f(x) = \lambda^2 x e^{-\lambda x} \qquad x > 0,$$

which is the probability density function of an Erlang random variable with rate λ and $n = 2$ stages.

Example 2. New distributions can be created by expanding existing distributions with the α parameter. Consider the Lomax distribution with the new parameter α. Let $g(x) = \ln(x+1)$. This example introduces some likelihood calculations. The probability density function is

$$f(x) = \frac{1}{\Gamma(\alpha)(x+1)^2} \left[\ln(x+1)\right]^{\alpha-1} \qquad x > 0.$$

The log likelihood function is

$$L(\alpha) = \prod_{i=1}^{n} f(x_i) = [\Gamma(\alpha)]^{-n} \left(\prod_{i=1}^{n} \frac{1}{(1+x_i)^2} \right) \left[\prod_{i=1}^{n} \ln(1+x_i) \right]^{\alpha-1}$$

and the log likelihood function is

$$\ln L(\alpha) = -n \ln(\Gamma(\alpha)) - 2\sum_{i=1}^{n} \ln(1+x_i) + (\alpha-1)\sum_{i=1}^{n} \ln(\ln(1+x_i)).$$

The score is

$$\frac{\partial \ln L(\alpha)}{\partial \alpha} = -\frac{n\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n} \ln(\ln(1+x_i)).$$

The equation

$$\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} = \frac{1}{n}\sum_{i=1}^{n} \ln\left(\ln(1+x_i)\right)$$

must be solved numerically to compute the maximum likelihood estimator $\hat{\alpha}$ for a particular data set.

Example 3. New distributions with closed–form moments are possible. Let $g(x) = \lambda x^{1/\alpha}$. This yields the probability density function

$$f(x) = \frac{\exp(-\lambda\, x^{1/\alpha})\lambda^{\alpha}}{\Gamma(\alpha+1)} \qquad\qquad x > 0.$$

This distribution has mean

$$\mu = \frac{\lambda^{-\alpha} 4^{\alpha}\Gamma(\alpha+1/2)}{2\sqrt{\pi}}.$$

and variance

$$\sigma^2 = \frac{\lambda^{-2\alpha}\left(4\pi\Gamma(3\alpha) - 16^{\alpha}\left(\Gamma(\alpha+1/2)\right)^2\Gamma(\alpha)\right)}{4\pi\Gamma(\alpha)}.$$

Furthermore, all integer moments about the origin can be calculated with the moment function

$$E(X^n) = \int_0^{\infty} x^n f(x)\, dx = \frac{\lambda^{-n\alpha}\Gamma(\alpha+\alpha n)}{\Gamma(\alpha)}.$$

4 CONCLUSIONS

An unlimited number of survivor distributions can be generated using the incomplete gamma function ratio. Several popular survivor distributions, such as the Weibull and gamma distributions, are included in this class. The class is closed under monotonic transformations. Furthermore, any positive function with finite area can be used as a base. One can imagine ten base functions and ten $g(x)$ functions can be used to create 100 new distributions. Additionally, one is not limited to survivor distributions. Any positive function truncated to a domain of $(0,1)$ can be used to create new families of distributions with support $(0,1)$ which could possibly be an alternate to the beta distribution. Similarly, distributions with ad hoc support (a,b) or infinite support $(-\infty,\infty)$ are also possible with appropriate base functions.

ACKNOWLEDGEMENT

The authors are acknowledge the support of the Army Research Office thru its grant number 67906-MA.

REFERENCES

Casella, G., and Berger, R. 2008. *Statistical Inference*. Boston: Thomson Learning.
Forbes, C., Evans, M., Hastings, N., Peacock, B. 2011. *Statistical Distributions, Fourth Edition*. Hoboken, New Jersey: John Wiley & Sons.
Hogg, R.V., McKean, J., Craig, A.T. 2013. *Introduction to Mathematical Statistics (7th Edition)*. London: Pearson.
Meeker, W., and Escobar, L. 1998. *Statistical Methods for Reliability Data*. New York: John Wiley & Sons.

AUTHOR BIOGRAPHIES

ANDREW G. GLEN is Professor Emeritus of Operations Research at the United States Military Academy in West Point, NY. He is also a visiting professor at the Colorado College. He holds a PhD in Operations Research from The College of William & Mary. His research interests lie in computational probability.

His email address is andy.glen@coloradocollege.edu.

LAWRENCE M. LEEMIS is a Professor at the College of William & Mary. He holds a Ph.D. in Operations Research from Purdue University. His research interests include computational probability, reliability, statistics, and simulation. His email address is leemis@math.wm.edu.

DANIEL J. LUCKETT received his B.S. from the College of William & Mary. He is currently a Ph.D. candidate at the University of North Carolina at Chapel Hill, working in biostatistics. His email address is luckett@live.unc.edu.