Published in final edited form as:
Commun Stat Theory Methods. 2000 February 1; 29(4): 851–857. doi:10.1080/03610920008832519.

Properties of Doubly-Truncated Gamma Variables

Christopher S. Coffey and
A-1124 Medical Center North, Dept of Preventive Medicine, Vanderbilt Univ. School of Med., Nashville, Tennessee 37232-2637

Keith E. Muller
3105C McGavran-Greenberg, Dept. of Biostatistics, CB#7400, University of North Carolina, Chapel Hill, North Carolina 27599

Abstract

The truncated gamma distribution has been widely studied, primarily in life-testing and reliability settings. Most work has assumed an upper bound on the support of the random variable, *i.e*. the space of the distribution is $(0, u)$. We consider a doubly-truncated gamma random variable restricted by both a lower (l) and upper (u) truncation point, both of which are considered known. We provide simple forms for the density, cumulative distribution function (CDF), moment generating function, cumulant generating function, characteristic function, and moments. We extend the results to describe the density, CDF, and moments of a doubly-truncated noncentral chi-square variable.

Keywords

censoring; moments; characteristic function; noncentral chi-square; internal pilot study

1. INTRODUCTION

The truncated gamma distribution has been widely studied, primarily in life-testing and reliability settings. Most work has assumed an upper bound on the support of the random variable, *i. e*. the space of the distribution is $(0, u)$ (Johnson, Kotz, and Balakrishnan, 1994, §8.1). We consider a doubly-truncated gamma random variable restricted by both a lower (l) and upper (u) truncation point, both of which are considered known.

Coffey and Muller (1999) encountered such a distribution when computing the conditional distribution of the test statistic using an internal pilot design for sample size re-estimation as proposed by Wittes and Brittain (1990). Internal pilot studies allow researchers to use some fraction of the planned observations to re-estimate key parameters of interest and modify the final sample size if the parameters were initially mis-specified. Although internal pilot studies have been studied primarily in the clinical trials literature, these designs can be applied to a wide range of applications.

In study planning for linear models with Gaussian errors, determining the sample size required to detect a specified effect with some target power requires knowledge of the true variance, σ^2. An internal pilot study can be applied in order to protect against mis-specification of σ^2. The final sample size, N_+, then becomes a function of the variance estimate from the internal pilot sample $\hat{\sigma}_1^2$. Coffey and Muller (1999) provided an algorithm for computing the power under such a design. The power computations require first conditioning on an observed final sample size and then applying the law of total probability. Conditional upon an observed final sample size, $\hat{\sigma}_1^2$ is restricted to those values leading to that final sample size and thus follows a doubly-truncated χ^2 distribution, a special case of the doubly-truncated gamma distribution.

2. PROPERTIES OF A DOUBLY-TRUNCATED GAMMA VARIABLE

2.1 Density and CDF

Consider a gamma random variable, $X \sim \Gamma(\alpha, \beta)$, with density

$$f_{\Gamma}(x;\alpha,\beta)=\frac{x^{\alpha-1}\exp(-x/\beta)}{\Gamma(\alpha)\beta^{\alpha}}, x\in(0,\infty)\alpha>0,\beta>0, \quad (2.1)$$

where $\Gamma(\alpha)$ indicates the complete gamma function. The CDF of this distribution is simply

$$F_{\Gamma}(s;\alpha,\beta)=\int_0^s \frac{x^{\alpha-1}\exp(-x/\beta)}{\Gamma(\alpha)\beta^{\alpha}}\partial x. \quad (2.2)$$

We now consider $Y \sim \Gamma_T(\alpha, \beta, l, u)$ to be a doubly-truncated version of X with lower truncation point, l, and upper truncation point, u. Obviously,

$$f_{\Gamma}(y;\alpha,\beta,l,u)=\frac{f_{\Gamma}(y;\alpha,\beta)}{F_{\Gamma}(u;\alpha,\beta)-F_{\Gamma}(l;\alpha,\beta)} \quad ,l<y<u \quad , \quad (2.3)$$

and

$$F_{\Gamma}(s;\alpha,\beta,l,u)=\begin{cases} 0 & ,s<l \\ \frac{F_{\Gamma}(s;\alpha,\beta)-F_{\Gamma}(l;\alpha,\beta)}{F_{\Gamma}(u;\alpha,\beta)-F_{\Gamma}(l;\alpha,\beta)} & ,l<s<u \\ 1 & ,s>u. \end{cases} \quad (2.4)$$

2.2 Moment Generating Function

For a random variable X which follows a $\Gamma(\alpha, \beta)$ distribution, the moment generating function (mgf) is given by

$$M_X(t)=\mathscr{E}\{\exp(tX)\}=(1-\beta t)^{-\alpha}, t\in[0,\beta^{-1}). \quad (2.5)$$

By definition, the moment generating function of Y is

$$\begin{aligned} M_Y(t)&=\mathscr{E}\{e^{tY}\} \\ &=\frac{1}{F_{\Gamma}(u;\alpha,\beta)-F_{\Gamma}(l;\alpha,\beta)}\int_l^u \frac{y^{\alpha-1}\exp[-(1-\beta t)y/\beta]}{\Gamma(\alpha)\beta^{\alpha}}\partial y \\ &=\frac{(1-\beta t)^{1-\alpha}}{F_{\Gamma}(u;\alpha,\beta)-F_{\Gamma}(u;\alpha,\beta)}\int_l^u \frac{[(1-\beta t)y]^{\alpha-1}\exp[-(1-\beta t)y/\beta]}{\Gamma(\alpha)\beta^{\alpha}}\partial y. \end{aligned} \quad (2.6)$$

Apply the transformation $r = (1 - \beta t)y$, from which it follows $\partial r = (1 - \beta t)\partial y$. For $t \in [0, \beta^{-1})$, we use the transformation to rewrite the mgf as

$$M_Y(t)=\frac{(1-\beta t)^{-\alpha}}{F_{\Gamma}(u;\alpha,\beta)-F_{\Gamma}(l;\alpha,\beta)}\int_{l(1-\beta t)}^{u(1-\beta t)} \frac{r^{\alpha-1}\exp(-r/\beta)}{\Gamma(\alpha)\beta^{\alpha}}\partial r. \quad (2.7)$$

For the sake of brevity, define

$$P(t;\alpha,\beta,l,u)=F_{\Gamma}[u(1-\beta t);\alpha,\beta]-F_{\Gamma}[l(1-\beta t);\alpha,\beta]. \quad (2.8)$$

Since the integrand in (2.7) equals $P(t;\alpha,\beta,l,u)f_{\Gamma}[y;\alpha,\beta,l(1-\beta t),u(1-\beta t)]$, it follows that

$$\begin{aligned}M_Y(t)&=\left[\frac{P(t;\alpha,\beta,l,u)}{P(0;\alpha,\beta,l,u)}\right](1-\beta t)^{-\alpha}\\&=\left[\frac{P(t;\alpha,\beta,l,u)}{P(0;\alpha,\beta,l,u)}\right]M_X(t).\end{aligned} \quad (2.9)$$

Thus the moment generating function for a $\Gamma_T(\alpha,\beta,l,u)$ random variable equals the product of the mgf of a $\Gamma(\alpha,\beta)$ random variable and a factor which accounts for the truncation.

Using parallel reasoning, it follows that the characteristic function of the doubly-truncated gamma distribution is

$$\phi_Y(t)=\left[\frac{P(it;\alpha,\beta,l,u)}{P(0;\alpha,\beta,l,u)}\right]\phi_X(t). \quad (2.10)$$

2.3 Computing the m^{th} Moment

One can compute the moments of a $\Gamma_T(\alpha,\beta;l,u)$ using either the moment generating function or the characteristic function. However, it may be easier to compute the moments directly using the following lemma, which generalizes a χ^2 property used implicitly by Johnson, Kotz, and Balakrishnan (1994, p420).

Lemma 1— For any real number $m > -\alpha$,

$$x^m f_{\Gamma}(x;\alpha,\beta)=\frac{\beta^m\Gamma(\alpha+m)}{\Gamma(\alpha)}f_{\Gamma}(x;\alpha+m,\beta). \quad (2.11)$$

Corollary—For $Y \sim \Gamma_T(\alpha,\beta;l,u)$ and $m > -\alpha$,

$$\begin{aligned}\mathscr{E}(Y^m)&=\frac{\int_l^u y^m f_{\Gamma}(y;\alpha,\beta)}{F_{\Gamma}(u;\alpha,\beta)-F_{\Gamma}(l;\alpha,\beta)}dy\\&=\left[\frac{F_{\Gamma}(u;\alpha+m,\beta)-F_{\Gamma}(l;\alpha+m,\beta)}{F_{\Gamma}(u;\alpha,\beta)-F_{\Gamma}(l;\alpha,\beta)}\right]\frac{\beta^m\Gamma(\alpha+m)}{\Gamma(\alpha)}\\&=\left[\frac{F_{\Gamma}(u;\alpha+m,\beta)-F_{\Gamma}(l;\alpha+m,\beta)}{F_{\Gamma}(u;\alpha,\beta)-F_{\Gamma}(l;\alpha,\beta)}\right]\mathscr{E}(X^m).\end{aligned} \quad (2.12)$$

Hence the m^{th} moment of a $\Gamma_T(\alpha,\beta;l,u)$ equals the m^{th} moment of a $\Gamma(\alpha,\beta)$ times a factor depending only on the distribution functions of two gamma variables. This expression seems much easier to compute than the expression obtained directly from the mgf.

2.4 Cumulant Generating Function and Cumulants

For $t \in [0, \beta^{-1})$, let $K_X(t) = \ln M_X(t)$ represent the cumulant generating function of X. Hence, from (2.10) the cumulant generating function of the doubly-truncated random variable Y is

$$K_Y(t)=\ln P(t;\alpha,\beta,l,u)-\ln P(0;\alpha,\beta,l,u)+K_X(t). \quad (2.13)$$

In order to compute the m^{th} cumulant, we can evaluate the m^{th} derivative of $K_Y(t)$ at $t = 0$. Note that the middle term in (2.13) does not depend on t and hence can be ignored in differentiation. Furthermore, taking successive derivatives of the final term in (2.13) simply leads to κ_m, the m^{th} cumulant of a $\Gamma(\alpha,\beta)$. Thus the m^{th} cumulant of a $\Gamma_T(\alpha,\beta;l,u)$ equals

$$(\kappa_T)_m = \frac{\partial^m}{\partial t} \ln P(t;\alpha,\beta,l,u)\Big|_{t=0} + \kappa_m. \quad (2.14)$$

However, it is difficult to provide a convenient form for arbitrary derivatives of ln*P*(*t*; α, β; *l*, *u*). Alternatively, we can use the expression for computing the m^{th} moment in (2.12) to compute the cumulants using a recursion relationship such as described by Harvey (1972). Lee and Lin (1992, 1993) discussed an alternate relationship and provided an algorithm. Zheng (1998) provided an algorithm using Mathematica for computing the cumulants (analytically and numerically) in this manner.

3. EXTENSIONS TO NONCENTRAL χ^2 RANDOM VARIABLES

Clearly all of the above results apply to a central χ^2, a special case of a gamma. However, there is no direct relationship between the noncentral χ^2 and the gamma distribution. Taylor and Muller (1996) described a situation using the General Linear Model in which a power calculation was desired under censoring, *i. e*., we perform a power calculation only when the previous test was negative or positive. The setting led to a truncated *F*-distribution. A truncated noncentral χ^2 would arise under this setting when the test statistic follows a noncentral χ^2 distribution, as is often the case, at least asymptotically, for likelihood ratio test statistics. For this reason, we briefly consider some properties of a doubly-truncated, noncentral χ^2 variable, indicated $\chi^2_T(v,w,l,u)$. Let $f_{\chi^2}(u;v,w)$ and $F_{\chi^2}(u;v,w)$ represent the density and distribution function, respectively, of a noncentral chi-square variable. Clearly

$$f_{\chi^2_T}(x;v,w,l,u) = \frac{f_{\chi^2}(x;v,w)}{F_{\chi^2}(u;v,w) - F_{\chi^2}(l;v,w)} \quad, l<x<u \quad (3.1)$$

and

$$F_{\chi^2_T}(s;v,w,l,u) = \begin{cases} 0 & , s<l \\ \frac{F_{\chi^2}(s;v,w) - F_{\chi^2}(l;v,w)}{F_{\chi^2}(u;v,w) - F_{\chi^2}(l;v,w)} & , l<s<u \\ 1 & , s>u. \end{cases} \quad (3.2)$$

The noncentral χ^2 density may be written as an infinite Poisson weighted sum of central χ^2 densities (Johnson, Kotz, and Balakrishnan, 1995, p.436). For the sake of brevity, define

$$p\left(j;\frac{\omega}{2}\right) = \frac{(\omega/2)^j}{j!}\exp(-\omega/2), j=0,1,2,\ldots \quad (3.3)$$

the density of a Poisson random variable with mean $\omega/2$. We then write the density of the noncentral χ^2 as

$$f_{\chi^2}(x;v,w) = \sum_{j=0}^{\infty} p\left(j;\frac{w}{2}\right) f_{\chi^2}(x;v+2j,0), x>0. \quad (3.4)$$

For *m* any positive integer, we can use the above expression along with Lemma 1 to obtain an expression for the m^{th} moment of a $\chi^2_T(v,w,l,u)$:

$$\mathscr{E}X^m=\frac{\sum_{j=0}^{\infty}p\left(j;\frac{w}{2}\right)\int_l^u x^m f_{\chi^2}\left(x;v+2j,0\right)dx}{F_{\chi^2}(u;v,w)-F_{\chi^2}(l;v,w)}$$
$$=\frac{\sum_{j=0}^{\infty}p\left(j;\frac{w}{2}\right)\int_l^u f_{\chi^2}\left[x;v+2\left(j+m\right),0\right]dx\left\{\prod_{i=0}^{m-1}\left[v+2\left(j+i\right)\right]\right\}}{F_{\chi^2}(u;v,w)-F_{\chi^2}(l;v,w)} \tag{3.5}$$
$$=\frac{\sum_{j=0}^{\infty}p\left(j;\frac{w}{2}\right)\left\{F_{\chi^2}\left[u;v+2\left(j+m\right),0\right]-F_{\chi^2}\left[l;v+2\left(j+m\right),0\right]\right\}\prod_{i=0}^{m-1}\left[v+2\left(j+i\right)\right]}{F_{\chi^2}(u;v,w)-F_{\chi^2}(l;v,w)}.$$

The form given allows straightforward computation with contemporary computers. However, without a generalization of the lemma, expressions for the moment generating function, characteristic function, and cumulant generating function are much more complicated to develop. It is possible that recursive properties of generalized hypergeometric functions may apply.

Acknowledgments

Muller's work supported in part by NCI program project grant P01 CA47 982-04. Coffey's work supported in part by NIEHS grant 5-T32-ES07018. A portion of the work was described in a paper submitted by Coffey in partial fulfillment of the requirements for the Ph. D. degree in Biostatistics. The authors would like to thank an anonymous reviewer for pointing out a number of corrections and clarifications.

BIBLIOGRAPHY

Coffey CS, Muller KE. Exact Test Size and Power of a Gaussian Error Linear Model for an Internal Pilot Study. Statistics in Medicine. 1999; 18:1199–1214. [PubMed: 10363340]

Harvey JR. On Expressing Moments in Terms of Cumulants and Vice Versa. The American Statistician. 1972; 26:38–39.

Johnson, NL.; Kotz, S.; Balakrishnan, N. Continuous Univariate Distributions-1. New York: Wiley; 1994.

Johnson, NL.; Kotz, S.; Balakrishnan, N. Continuous Univariate Distributions-2. New York: Wiley; 1995.

Lee YS, Lin TK. High Order Cornish-Fisher Expansion. Applied Statistics. 1992; 41:233–240.

Lee YS, Lin TK. Correction to Algorithm AS 269: High Order Cornish-Fisher Expansion. Applied Statistics. 1993; 42:268–269.

Taylor DJ, Muller KE. Bias in Linear Model Power and Sample Size Calculations Due to Estimating Noncentrality. Communications in Statistics: Theory and Methods. 1996; 25:1595–1610.

Wittes J, Brittain E. The Role of Internal Pilot Studies in Increasing the Efficiency of Clinical Trials. Statistics in Medicine. 1990; 9:65–72. [PubMed: 2345839]

Zheng Q. Computing Relations between Statistical Moments and Cumulants. 1998 Proceedings of the ASA Statistical Computing Section. 1998:165–170.