Journal of Modern Applied Statistical Methods

Volume 6 | Issue 1 Article 18

5-1-2007

On the Properties of Beta-Gamma Distribution

Lingji Kong
Union College

Carl Lee
Central Michigan University

J.H. Sepanski
Central Michigan University

Journal of Modern Applied Statistical Methods
May, 2007, Vol. 6, No. 1, 187-211

1538 – 9472/07/$95.00

On the Properties of Beta-Gamma Distribution

Lingji Kong
Union College

Carl Lee J. H. Sepanski
Central Michigan University

A class of generalized gamma distribution called the beta-gamma distribution is proposed. Some of its properties are examined. Its shape can be reversed J-shaped, unimodal, or bimodal. Reliability and hazard functions are also derived, and applications are discussed.

Key words: Gamma distribution, Beta distribution, reliability function, hazard function, MLE, application.

Introduction

Let $f(\cdot)$ and $F(\cdot)$ be the probability density function and the cumulative distribution function (cdf) of a random variable, respectively. Eugene, Lee, and Famoye (2002) first introduced a generalized distribution based on the logit of the beta random variable with a cumulative distribution function given by

$$G(x) = \frac{1}{B(\alpha,\beta)} \int_0^{F(x)} t^{\alpha-1}(1-t)^{\beta-1}\,dt,\ 0<\alpha,\beta<\infty,$$

and the corresponding probability density function is

$$g(x) = \frac{1}{B(\alpha,\beta)} F(x)^{\alpha-1}\left[1-F(x)\right]^{\beta-1} f(x),\ 0<\alpha,\beta<\infty$$

Eugene, et al. (2002) studied properties of $g(x)$ when $F(\cdot)$ is the cdf of a normal distribution. Maynard (2003) examined the case when $F(\cdot)$ is the cdf of an exponential distribution.

Gamma distribution and its generalized distributions (e.g. McDonald, 1984) have been applied widely to the analyses of income distributions, life testing, and many physical and economical phenomena (e.g. Farewell, 1977, Lawless, 1980). In this article, the case when $F(\cdot)$ is the cdf of the gamma distribution is studied.

A random variable X is said to have a beta-gamma distribution, $BG(\alpha,\beta,\rho,\lambda)$, if its probability density function is given by

$$g(x) = \frac{x^{\rho-1}e^{-x/\lambda}}{B(\alpha,\beta)\Gamma(\rho)\lambda^{\rho}} F(x)^{\alpha-1}\left[1-F(x)\right]^{\beta-1},$$
$$0<\alpha,\beta,\rho,\lambda<\infty,\ x>0, \qquad (1)$$

where $F(x)$ is the cdf of the gamma distribution with parameters ρ and λ. One can also introduce a location parameter ξ in the density in (1) by replacing x with $x-\xi$ where $-\infty<\xi<\infty$. In the rest of this article, it is assumed that ξ is zero. When both α and β are integers with $\alpha+\beta$ being a bounded integer, the beta-gamma density function in (1) is the marginal probability density function of

Lingji Kong is an Assistant Professor of Mathematics. His email is lkong@unionky.edu. Carl Lee is Professor of Statistics and senior research fellow at the Center for Applied Research and Technology, Central Michigan University. His research area is in generalized distributions and data mining. His email is lee1c@cmich.edu. Jungsywan Sepanski is an Associate Professor of Statistics in the Mathematics Department at Central Michigan University. Her research area is in measurement error and estimation. Her email is sepan1jh@cmich.edu.

the α^{th} order statistic in a random sample of size $\alpha+\beta$ from the gamma distribution with parameters ρ and λ. When $\alpha=\beta=1$, the beta-gamma distribution yields the gamma distribution. When $\rho=1$, the beta-gamma distribution is beta-exponential distribution introduced in Maynard (2003).

Properties

The limit of $g(x)$ as x goes to 0 and the mode of the probability density function $g(x)$ in (1) is given in Lemma 1. The modes for cases when $\rho\alpha \le 1$ and $\rho\alpha > 1$ are studied respectively. Although some cases can be shown mathematically, plotting the function *g(x)* using Maple computer programs are employed to examine shapes and modalities for other cases. Illustrative graphs of $g(x)$ based on observations from numerous plots are presented. Numerical percentiles are presented in Table 7 to Table 9.

Limits

Lemma 1: The limit as x goes to 0 of the beta-gamma probability density function $g(x)$ in (1) is

$$\lim_{x\to 0} g(x) = \begin{cases} \infty & \text{if } \alpha\rho < 1 \\ \dfrac{1}{\Gamma^{\alpha}(\rho)B(\alpha,\beta)\rho^{\alpha-1}\lambda} & \text{if } \alpha\rho = 1 \\ 0 & \text{if } \alpha\rho > 1 \end{cases} \tag{2}$$

The proof is given in Appendix.

Modes of $g(x)$ When $\alpha\rho \le 1$

Note that the derivative $df/dx = f(\rho-1-x/\lambda)/x$. The first derivative of the logarithm of the probability density function *g(x)* is given by

$$\frac{\rho-1-x/\lambda}{x}+\frac{\alpha-1}{F}f+\frac{1-\beta}{1-F}f\,. \tag{3}$$

The mode(s) x_m of $g(x)$ if exists is the solution to the equation by setting (3) to be zero.

It is shown below that *g (x)* has a reversed-J shape when $\rho\alpha \le 1$ and $\beta \ge 1$. The derivative in (3) is equal to

$$(1-\beta)\frac{f}{1-F}+\frac{1}{xF}\left[(\alpha-1)xf+(\rho-1-x/\lambda)F\right]. \tag{4}$$

When $\beta \ge 1$, the first term in (4) is less or equal to 0. Also,

$$\begin{aligned}&\frac{d}{dx}[(\alpha-1)xf+(\rho-1-x/\lambda)F]\\ &=(\alpha-1)f+(\alpha-1)xf\frac{\rho-1-x/\lambda}{x}\\ &-F/\lambda+(\rho-1-x/\lambda)f\\ &=(\alpha\rho-1)f-\alpha xf/\lambda-F/\lambda\end{aligned}$$

which is negative when $\alpha\rho \le 1$. This implies that $(\alpha-1)xf+(\rho-1-x/\lambda)F$ is a decreasing function. Because $(\alpha-1)xf+(\rho-1-x/\lambda)F=0$ when $x=0$, the second term in (4) is therefore negative. That is, $g'(x)$ is negative. By (2) and the fact that $\lim_{x\to\infty} g(x)=0$, $g(x)$ has a reversed-J shape for the cases when $\rho\alpha \le 1$ and $\beta \ge 1$ with maximum occurring at $x=0$.

When $\alpha \le 1$ and $\rho \le 1$ regardless of β, one can see that $g(x)$ has a reversed-J shape by rewriting $g(x)$ as

$$g(x)=\frac{1}{B(\alpha,\beta)}\frac{f(x)}{1-F(x)}F(x)^{\alpha-1}[1-F(x)]^{\beta}.$$

Because the cdf F is an increasing function and the hazard function $f/(1-F)$ of the gamma distribution function is a decreasing function when $\rho \le 1$, $g(x)$ is therefore a decreasing function with $\lim_{x\to 0} g(x)=\infty$ when $\alpha \le 1$ and $\rho \le 1$.

(2.3)

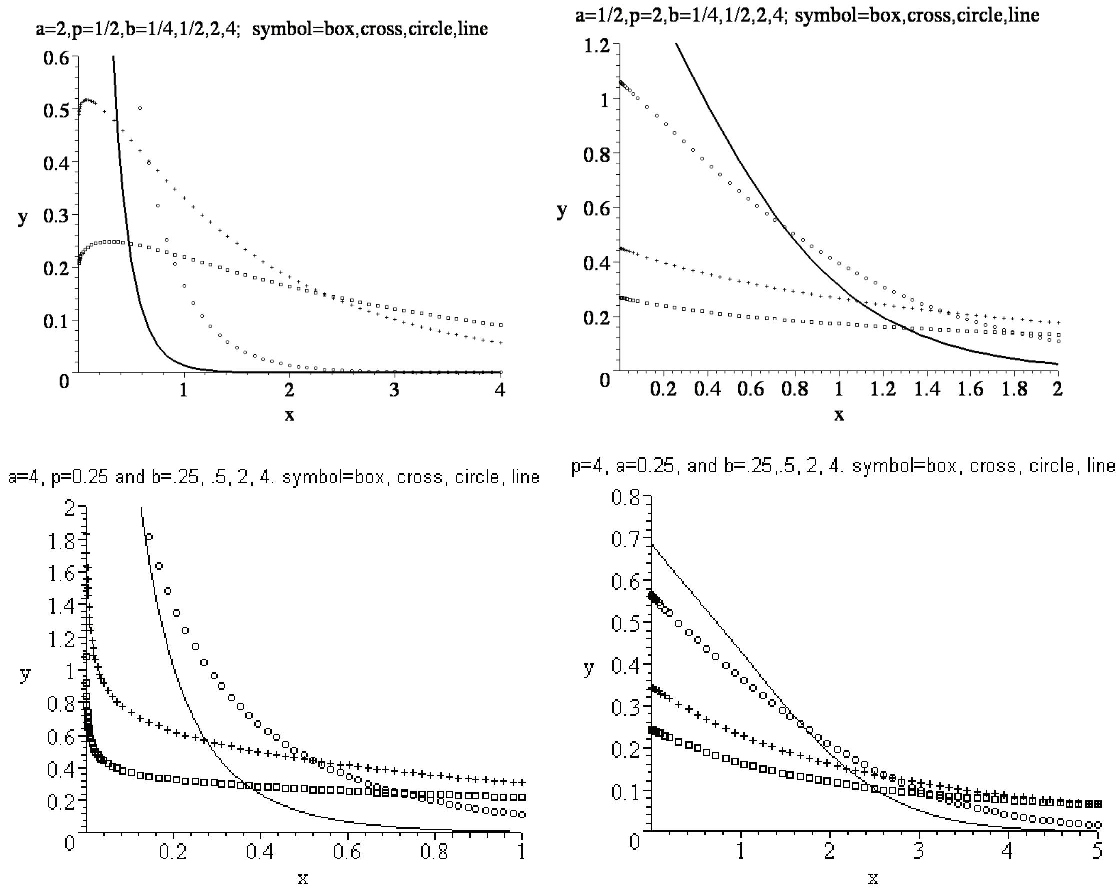

Figure 1. Plot of the density function *g(x)* when $\rho\alpha = 1$, and β=0.25, 0.5, 2, 4

Next, graphical results are shown to examine the cases when $\beta < 1$ and $\alpha\rho \leq 1$ with α or ρ greater than 1. Figure 1 represents cases when $\alpha\rho = 1$. Figure 2 contains cases when $\alpha\rho < 1$. Note that a $= \alpha$, b $= \beta$, and p $= \rho$ in all figures in this article.

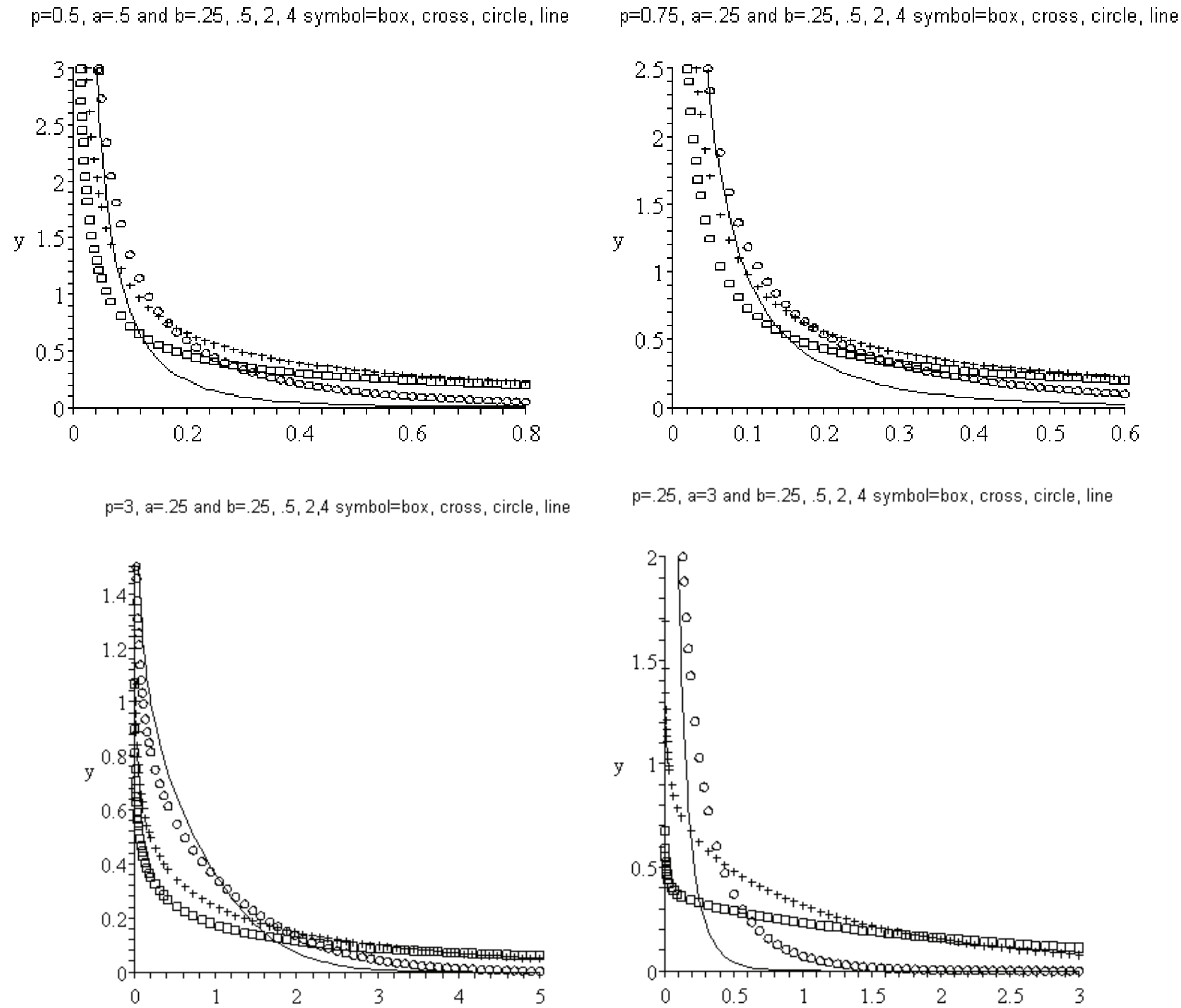

Figure 2. Plot of the density function *g(x)* when $\rho\alpha < 1$, and β=0.25, 0.5, 2, 4

When $\rho\alpha = 1$ and $\beta \geq 1$, the beta-gamma distribution appears to have a reversed-J shape. Figure 1 also shows that when $\alpha = 2$ and $\rho = 0.5$, it has a non-zero mode for β values of 0.25 and 0.5.

When $\rho\alpha < 1$ and $\beta < 1$, it is found that $g(x)$ is not necessarily a reverse J-shape, it can be bimodal (with one mode at 0). Figure 3 shows two such cases. The top two are for $\rho = 0.25,\ \alpha = 3.9,$ and $\beta = 0.5$; the bottom graph is for $\rho = 2,\ \alpha = 0.49,$ and $\beta = 0.01$. Note that the horizontal axis of the first plot ranges from 0 to 0.01 and the one of the second plot ranges from 0.01 to 2. Tables 1 – 4 give the 2nd non-zero mode in addition to the mode at $x = 0$ for some examples when $\alpha\rho < 1$ and $\rho < 1$. The empty cells are cases where $g(x)$ is reverse J-shaped and the only mode is at $x = 0$.

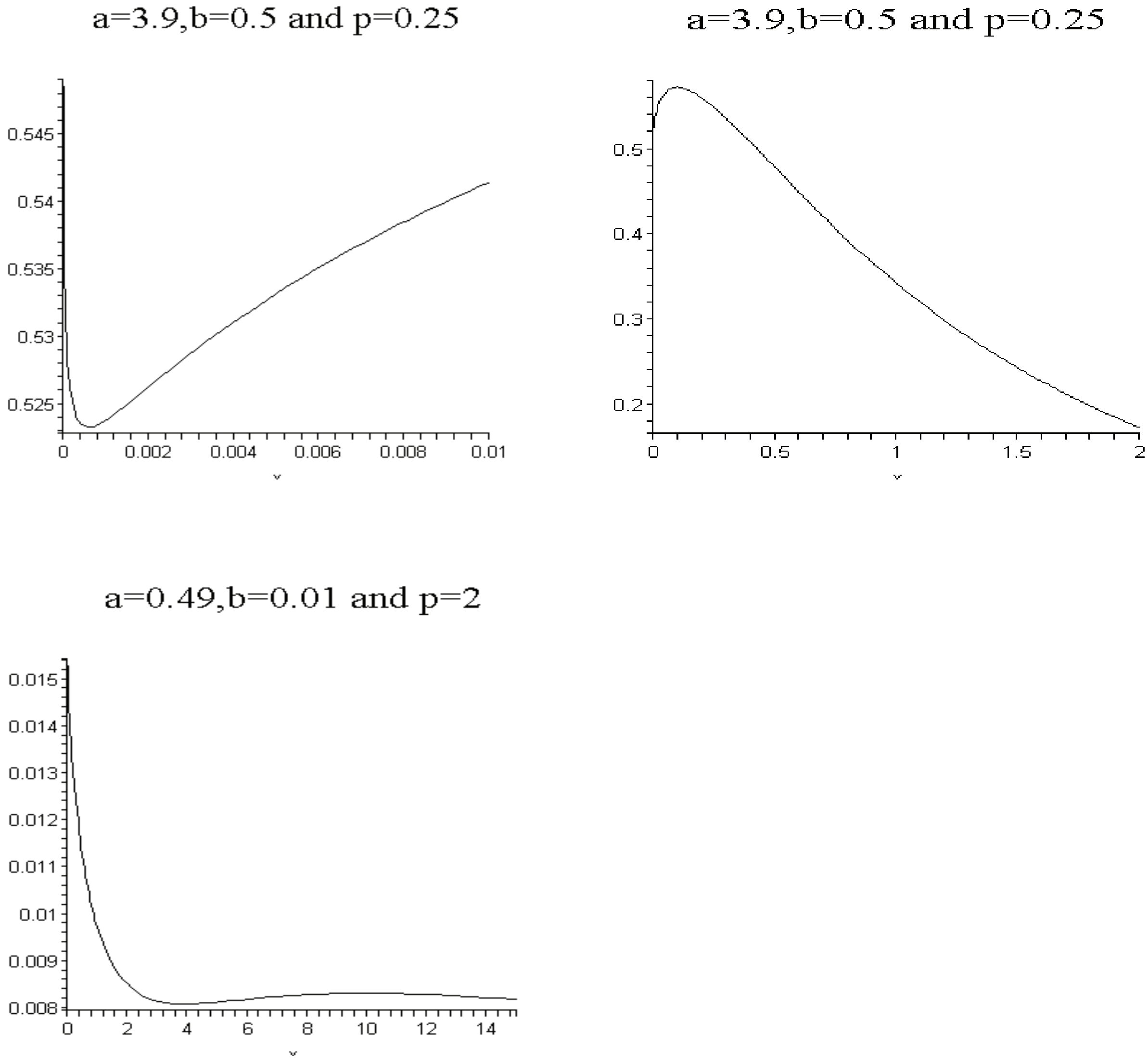

Figure 3. Graphs of $BG(\alpha, \beta, \rho, 1)$

Table 1. Nonzero 2nd mode of $BG(\alpha, \beta, \rho, 1)$ with $\rho = 1/2$.

	$\beta = 0.01$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
α =1.8	1.04									
1.9	1.31	.529	.234							
1.95	1.42	.616	.319	.167	.074					
1.99	1.50	.678	.374	.219	.013	.071	.035	.012		

Table 2(a). Nonzero 2nd mode of $BG(\alpha, \beta, \rho, 1)$ with $\rho = 1/4$.

	$\beta = 0.01$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\alpha = 3.01$										
3.02	.365									
3.1	.571									
3.2	.714									
3.3	.826	.274								
3.4	.923	.382								
3.5	1.01	.460	.193							
3.6	1.09	.526	.266	.103						
3.7	1.16	.585	.323	.173	.064					
3.8	1.22	.638	.373	.222	.124	.052				
3.9	1.29	.687	.417	.264	.165	.097	.049			

Table 2(b). Nonzero 2nd mode of $BG(\alpha, \beta, \rho, 1)$ with $\rho = 1/4$.

	$\beta = 0.91$	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99
$\alpha = 3.99$	.004	.003	.001						
3.995	.005	.004	.003	.002	.001				

Table 3. Nonzero 2nd mode of $BG(\alpha, \beta, \rho, 1)$ with $\rho = 1/6$

	$\beta = .01$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$\alpha = 5.8$	1.21	.67	.41	.27	.17	.10	.05			
5.9	1.24	.69	.44	.29	.19	.12	.07	.036	.020	
5.99	1.27	.72	.48	.31	.21	.14	.09	.052	.026	.008

Table 4. Nonzero 2nd mode of $BG(\alpha, \beta, \rho, 1)$ with $\rho = 2$

	β=0.01	0.015	0.1	0.2	0.3	0.4
$\alpha = 0.48$						
0.49	9.85					
0.499	9.86	7.656				

Note that, for example, when $\alpha = 0.48$, β =0.01, and $\rho = 2$, *g(x)* has an inversed-J shape and therefore does not have a nonzero mode. The range of β where $g(x)$ is bimodal appears to widen as α increases. When bimodality occurs, the nonzero mode increases as the parameter α increases and decreases as the parameter β increases. The bimodality property of beta-gamma distribution is not independent of the gamma parameters (α, ρ). The bimodality property also exists for beta-normal (Famoye, Lee, & Eugene, 2004).

Modes when $\alpha\rho > 1$

The second derivative of the logarithm of $g(x)$ is given by

$$\frac{\rho-1-x/\lambda}{x}\left[\frac{(\alpha-1)f}{F}+\frac{(1-\beta)f}{1-F}+\frac{\rho-1-x/\lambda}{x}\right] -\left[\frac{(\rho-1-x/\lambda)^2}{x^2}+\frac{(\alpha-1)f^2}{F^2}-\frac{(1-\beta)f^2}{(1-F)^2}+\frac{\rho-1}{x^2}\right].$$

The first term equals to 0 at the mode x_m . Hence, when $x = x_m$,

$$\frac{d^2 \ln g}{dx^2} = -\left[\begin{array}{l}\frac{(\rho-1-x)^2}{x^2}+\frac{(\alpha-1)f^2}{F^2} \\ +\frac{(\beta-1)f^2}{(1-F)^2}+\frac{\rho-1}{x^2}\end{array}\right]. \qquad (5)$$

When $\beta \geq 1, \alpha \geq 1$, and $\rho \geq 1$, (5) is less than 0 at $x = x_m$. In this case, since there must be a minimum between any two maxima and that $\lim_{x\to 0} g(x)$ =0 and $\lim_{x\to\infty} g(x)$ =0, it is concluded that $g(x)$ is unimodal with a concave shape.

When $\beta \geq 1$ and $\alpha\rho > 1$ with $\alpha < 1$ or $\rho < 1$, though not being shown mathematically, graphs of such cases indicate that beta-gamma density function $g(x)$ is also unimodal with a concave shape. Based on numerous graphs, the density functions $g(x)$ is unimodal when $\alpha\rho > 1$ regardless the value of β . The following illustrates some examples when $\alpha\rho > 1$.

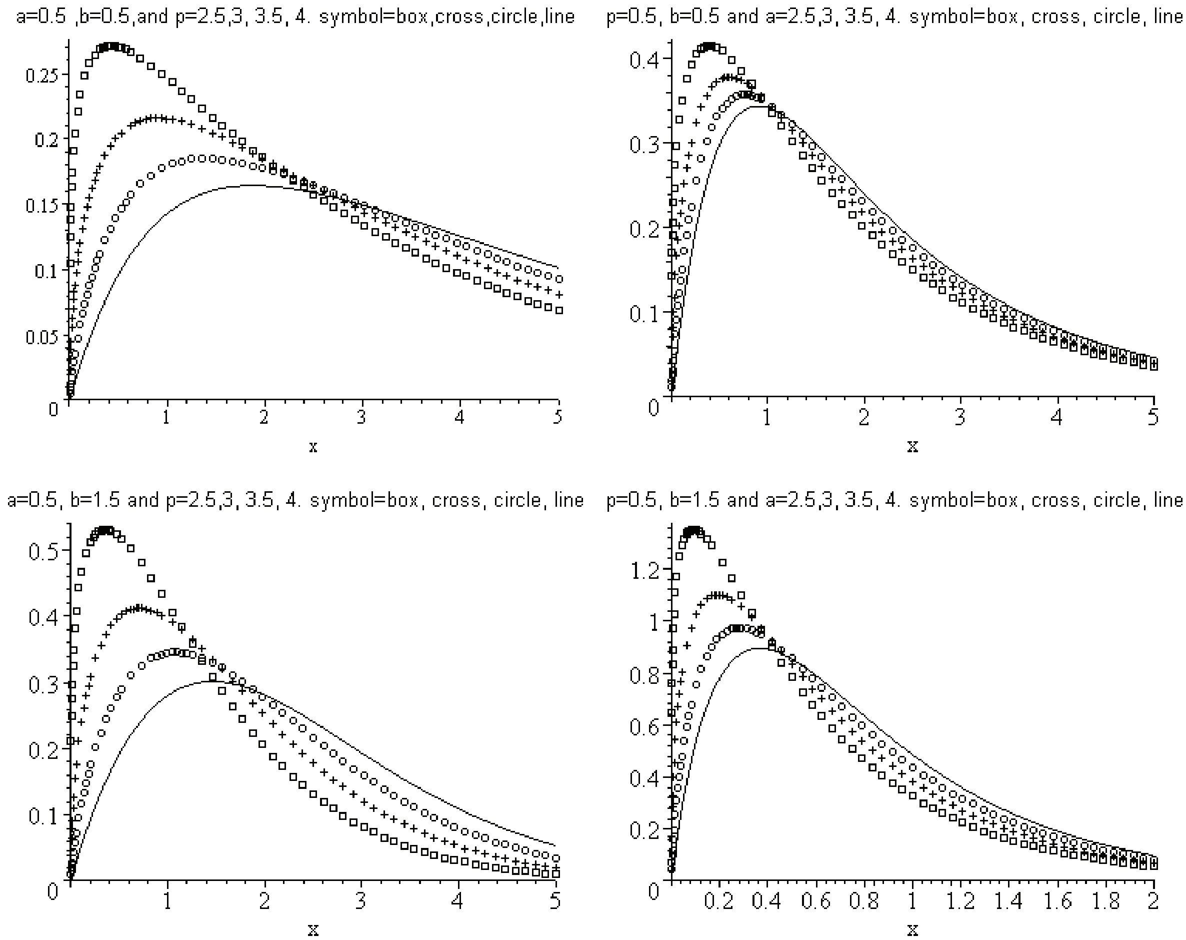

Figure 4. Plot of the density function *g(x)* when $\rho\alpha > 1$, and β=0.5, 1.5

In this case $g(x)$ is unimodal with a concave shape and the mode is nonzero. Tables 5 and 6 tabulate modes for $BG(\alpha,\beta,2,1)$ and $BG(\alpha,\beta,1/2,1)$ when $\rho\alpha > 1$.

The results indicate that when $\rho\alpha > 1$ the mode increases as α increases and that the mode decreases as β increases for both $BG(\alpha,\beta,2,1)$ and $BG(\alpha,\beta,1/2,1)$; see also Figure 4. For other cases when $\alpha\rho > 1$, this pattern holds for other values of parameters ρ and λ though the computation results are not reported here.

Percentiles of $g(x)$

The 50th, 75th, 90th, and 95th percentiles of $BG(\alpha,\beta,\rho,1)$ are computed and tabulated in the following Tables 7-9.

Table 5. Modes for $BG(\alpha,\beta,2,1)$ when $\rho\alpha > 1$

	β =0.2	0.5	1	1.5	2	2.5	5	10
α =1	2.236	1.414	1.000	.8165	.7071	.6325	.4472	.3162
α =1.5	3.206	2.162	1.555	1.270	1.097	.9775	.6815	.4744
α =2	3.729	2.628	1.938	1.598	1.386	1.238	.8644	.5994
α =2.5	4.087	2.963	2.229	1.856	1.618	1.451	1.018	.7063
α =5	5.055	3.915	3.111	2.672	2.379	2.163	1.570	1.107
α =10	5.913	4.787	3.964	3.498	3.176	3.947	2.228	1.626

Table 6. Modes for $BG(\alpha,\beta,1/2,1)$ when $\rho\alpha > 1$.

	β =0.2	0.5	1	1.5	2	2.5	5	10
α =2.5	0.832	.3919	.1692	.0903	.0545	.0359	.0291	.0154
α =5	1.788	1.129	.7150	.5137	.3923	.3114	.1337	.0735
α =10	2.653	1.798	1.286	1.016	.8411	.7155	.3925	.1812

Table 7. Percentiles of $BG(\alpha,\beta,\rho,1)$ with $\rho = 1/2$

α	β	50th	75th	90th	95th
0.25	0.25	.2275	2.011	5.260	7.855
	0.5	.0235	.4372	1.684	2.846
	1	.0031	.0831	.4479	.8765
	2	.0005	.0169	.1116	.2445
	4	.0001	.0038	.0277	.0654
0.5	0.25	.9346	3.194	6.550	9.171
	0.5	.2275	1.054	2.530	3.716
	1	.0508	.3014	.8588	1.358
	2	.0115	.0802	.2632	.4522
	4	.0027	.0207	.0752	.1370
1	0.25	1.735	4.163	7.568	10.21
	0.5	.6617	1.735	3.317	4.570
	1	.2275	.6617	1.353	1.921
	2	.0706	.2275	.5022	.7405
	4	.0202	.0706	.1678	.2576
2	0.25	2.473	4.979	8.413	11.06
	0.5	1.205	2.405	4.044	5.316
	1	.5531	1.123	1.899	2.501
	2	.2275	.4817	.8367	1.115
	4	.0816	.1824	.3306	.4504
4	0.25	3.160	5.710	9.166	11.82
	0.5	1.787	3.057	4.731	6.016
	1	.9914	1.649	2.478	3.102
	2	.5073	.8444	1.261	1.570
	4	.2275	.3872	.5869	.7351

Table 8. Percentiles of $BG(\alpha, \beta, \rho, 1)$ with $\rho = 1$

α	β	50th	75th	90th	95th
0.25	0.25	.6939	3.106	6.752	9.535
	0.5	.1882	1.050	2.710	4.070
	1	.0645	.3804	1.067	1.685
	2	.0265	.1577	.4517	.7249
	4	.0120	.0718	.2065	.3330
0.5	0.25	1.763	4.466	8.125	10.90
	0.5	.6925	1.919	3.704	5.077
	1	.2877	.8267	1.661	2.328
	2	.1285	.3729	.7592	1.075
	4	.0512	.1758	.3588	.5086
1	0.25	2.773	5.545	9.210	11.98
	0.5	1.386	2.773	4.605	5.991
	1	.6931	1.386	2.303	2.996
	2	.3466	.6931	1.151	1.498
	4	.1733	.3466	.5756	.7489
2	0.25	3.644	6.436	10.10	12.88
	0.5	2.115	3.565	5.366	6.720
	1	1.228	2.010	2.970	3.676
	2	.6931	1.120	1.631	2.000
	4	.3766	.6055	.8768	1.071
4	0.25	4.428	7.229	10.90	13.68
	0.5	2.836	4.312	6.167	7.556
	1	1.838	2.668	3.650	4.363
	2	1.159	1.641	2.187	2.571
	4	.6931	.9706	1.278	1.490

Table 9. Percentiles of $BG(\alpha,\beta,\rho,1)$ with $\rho=2$

α	β	50th	75th	90th	95th
0.25	0.25	1.678	4.874	9.049	12.09
	0.5	.7450	2.220	4.397	6.021
	1	.4035	1.142	2.244	3.094
	2	.2482	.6713	1.273	1.729
	4	.1631	.4277	.7860	1.050
0.5	0.25	3.197	6.479	10.58	13.58
	0.5	1.678	3.404	5.597	7.194
	1	.9613	1.887	3.063	3.922
	2	.5961	1.128	1.782	2.254
	4	.3893	.7157	1.102	1.374
1	0.25	4.472	7.710	11.76	14.74
	0.5	2.693	4.472	6.638	8.212
	1	1.678	2.693	3.890	4.744
	2	1.078	1.678	2.365	2.845
	4	.7095	1.078	1.487	1.767
2	0.25	5.519	8.710	12.72	15.69
	0.5	3.653	5.425	7.560	9.116
	1	2.473	3.518	4.712	5.557
	2	1.678	2.320	3.023	3.505
	4	1.135	1.536	1.963	2.250
4	0.25	6.434	9.587	13.57	16.53
	0.5	4.549	6.299	8.408	9.949
	1	3.296	4.344	5.526	6.359
	2	2.376	3.036	3.744	4.225
	4	1.678	2.103	2.543	2.834

The percentiles increase as α increases and decrease as β increases with fixed ρ, which is consistent with the results of modes. As seen in all figures, the beta-gamma distribution is skewed to the right, one would expect that the mode to be less than the median.

Moments

The closed form solutions of moments for $BG(\alpha,\beta,\rho,\lambda)$ exist only when α and β are integers. The closed form solution for the *n*th moment of $BG(\alpha,\beta,\rho,\lambda)$ is derived in Theorem 1 for the cases when α and β are integers in this section. The first four moments are also numerically computed for various parameters.

Theorem 1: When α,β are integers, the *n*th moment of the beta-gamma random variable $BG(\alpha,\beta,\rho,\lambda)$ is given by

$$\frac{1}{B(\alpha,\beta)}\sum_{j=0}^{\beta-1}(-1)^j\binom{\beta-1}{j}\left\{\sum_{k=0}^{\alpha+j-1}(-1)^k\binom{\alpha+j-1}{k}I_{n,k}\right\} \tag{6}$$

where

$$I_{n,k} = \int_0^\infty x^n f(x)(1-F)^k dx .$$

The proof is given in Appendix. The follow Corollary gives $E(X)$ and $E(X^2)$ that are used to obtain variance.

Corollary 1: When $\alpha = 2, \beta = 1$ and ρ is an integer, $E(X)$ and $E(X^2)$ are given by:

$$E(X) = 2\lambda\rho - \sum_{i=0}^{\rho-1}\frac{\lambda(\rho+i)!}{2^{\rho+i}(\rho-1)!i!}$$
$$= 2\lambda\rho - \rho\lambda\sum_{i=0}^{\rho-1}\frac{\binom{\rho+i}{\rho}}{2^{\rho+i}}; \tag{7}$$

$$E(X^2) = 2\lambda^2(\rho+1)\rho$$
$$= \rho(\rho+1)\lambda^2[2-\sum_{i=0}^{\rho-1}\frac{\binom{\rho+i+1}{\rho+1}}{2^{\rho+i+1}}] \tag{8}$$

The proof is given in Appendix.

Applying (6), the first four moments of $BG(\alpha,\beta,\rho,1)$ for a certain combinations of the parameters are evaluated and given in Tables 10 and 11.

Table 10. The mean, std, skewness and kurtosis of $BG(\alpha,\beta,\rho,1)$ with $\rho = 2$.

α	β	mean	std	skewness	kurtosis
1	1	2.000	1.414	1.415	6.005
	2	1.250	.8292	1.261	5.329
	4	.8047	.5048	1.120	4.768
	10	.4660	.2757	.9672	4.220
2	1	2.750	1.479	1.207	5.347
	2	1.824	.8975	1.010	4.588
	4	1.215	.5585	.8693	4.117
	10	.7150	.3063	.7319	3.736
4	1	3.547	1.494	1.106	5.094
	2	2.503	.9356	.8595	4.238
	4	1.747	.5987	.7038	3.779
	10	1.062	.3345	.5752	3.481
10	1	4.623	1.500	1.057	5.007
	2	3.503	.9505	.7650	4.063
	4	2.618	.6278	.5759	3.577
	10	1.705	.3654	.4363	3.302

Table 11. The mean, std, skewness and kurtosis of $BG(\alpha,\beta,\rho,1)$ with $\rho=1/2$

α	β	mean	std	skewness	kurtosis
1	1	.5000	.7071	2.829	15.00
	2	.1814	.2828	3.287	18.66
	4	.0604	.1038	3.834	26.22
	10	.0124	.0237	4.688	39.71
2	1	.8180	.8468	2.172	10.28
	2	.3523	.3830	2.290	11.11
	4	.1356	.1586	2.544	13.14
	10	.0319	.0413	3.032	17.96
4	1	1.235	.9584	1.762	8.011
	2	.6280	.4830	1.669	7.422
	4	.2868	.2289	1.718	7.614
	10	.0820	.0712	1.955	9.080
10	1	1.900	1.057	1.468	6.683
	2	1.154	.5878	1.218	5.491
	4	.6508	.3228	1.097	4.941
	10	.2505	.1290	1.119	4.968

Based on the numerical results, the mean and standard deviation appear to increase with α for a fixed β; and skewness and kurtosis appear to decrease as α increases for a fixed β in both cases when $\rho=2$ and $\rho=1/2$. Based on Figure 4, the density function has a heavier right tail as α increases. The mean and standard deviation decrease as β decreases for a fixed α. Although the skewness and kurtosis decrease with β when $\rho=2$ as shown in Table 10, the skewness and kurtosis increase with β when $\rho=1/2$ and $\alpha\rho\le 1$, as shown in Table 11. However, no clear pattern is noticed when $\alpha\rho>1$.

Reliability and Hazard Functions

The reliability and hazard functions of the beta-gamma distribution are derived in this section. The reliability function, $R(x)=1-P[X\le x]$, at time x defined to be the probability that a unit X survives beyond time x. For a beta-gamma random variable, it is given by

$$1-\frac{1}{B(\alpha,\beta)}\int_0^x F^{\alpha-1}(1-F)^{\beta-1}dF(t)$$
$$=1-\frac{1}{B(\alpha,\beta)}\int_0^{F(x)} t^{\alpha-1}(1-t)^{\beta-1}dt$$

where f and F are the density function and cdf of the gamma random variable with parameters ρ and λ, respectively. The hazard function defined to be a instantaneous measure of failure at time x given survival to time x is equal to

$$H(x) = \frac{g(x)}{R(x)} = \frac{\frac{1}{B(\alpha,\beta)} F^{\alpha-1}(1-F)^{\beta-1} f(x)}{1 - \frac{1}{B(\alpha,\beta)} \int_0^x F^{\alpha-1}(1-F)^{\beta-1} f(x)dx}.$$

Lemma 2:

(a) $\lim_{x\to 0} H(x) = \lim_{x\to 0} g(x)$

(b) $\lim_{x\to\infty} = \beta / \lambda$

The proof is given in Appendix.

The hazard functions of $BG(\alpha,\beta,\rho,1)$ are plotted. Cases with $\alpha\rho$ <1 are presented in Figure 5 and cases with $\alpha\rho \geq 1$ are given in Figure 6. The graphs in the first column represent the cases $\alpha\rho$ =1 with β =1/2, 1, and 2; and those in the second column represent the cases when $\alpha\rho$ >1 with β =1/2, 1, and 2 in Figure 6.

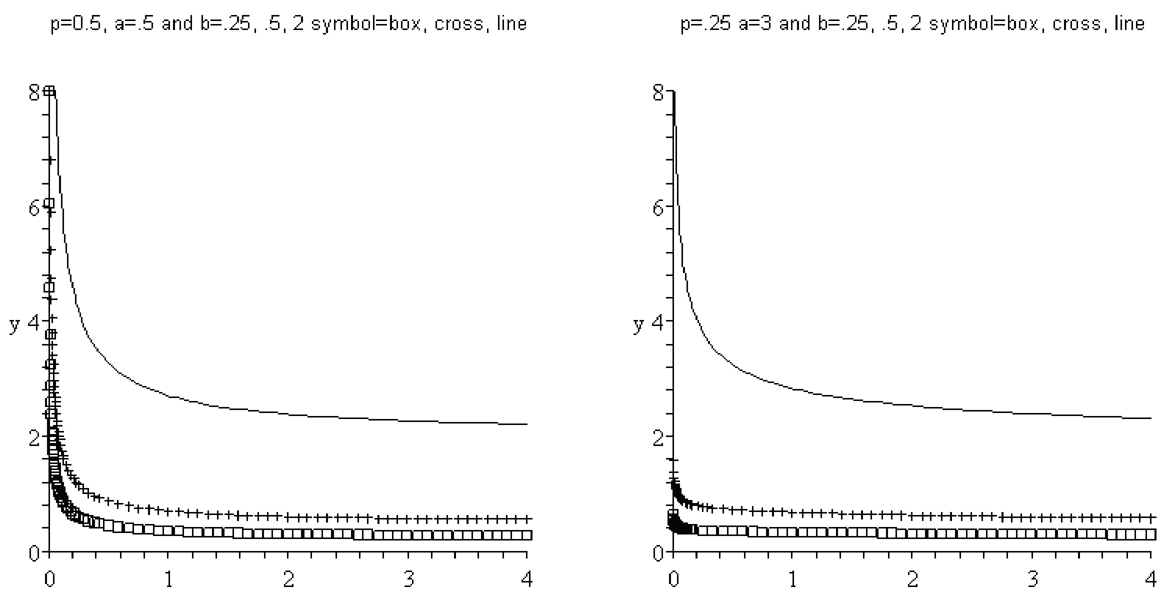

Figure 5. Hazard Function of $BG(\alpha,\beta,\rho,1)$ when $\alpha\rho < 1$

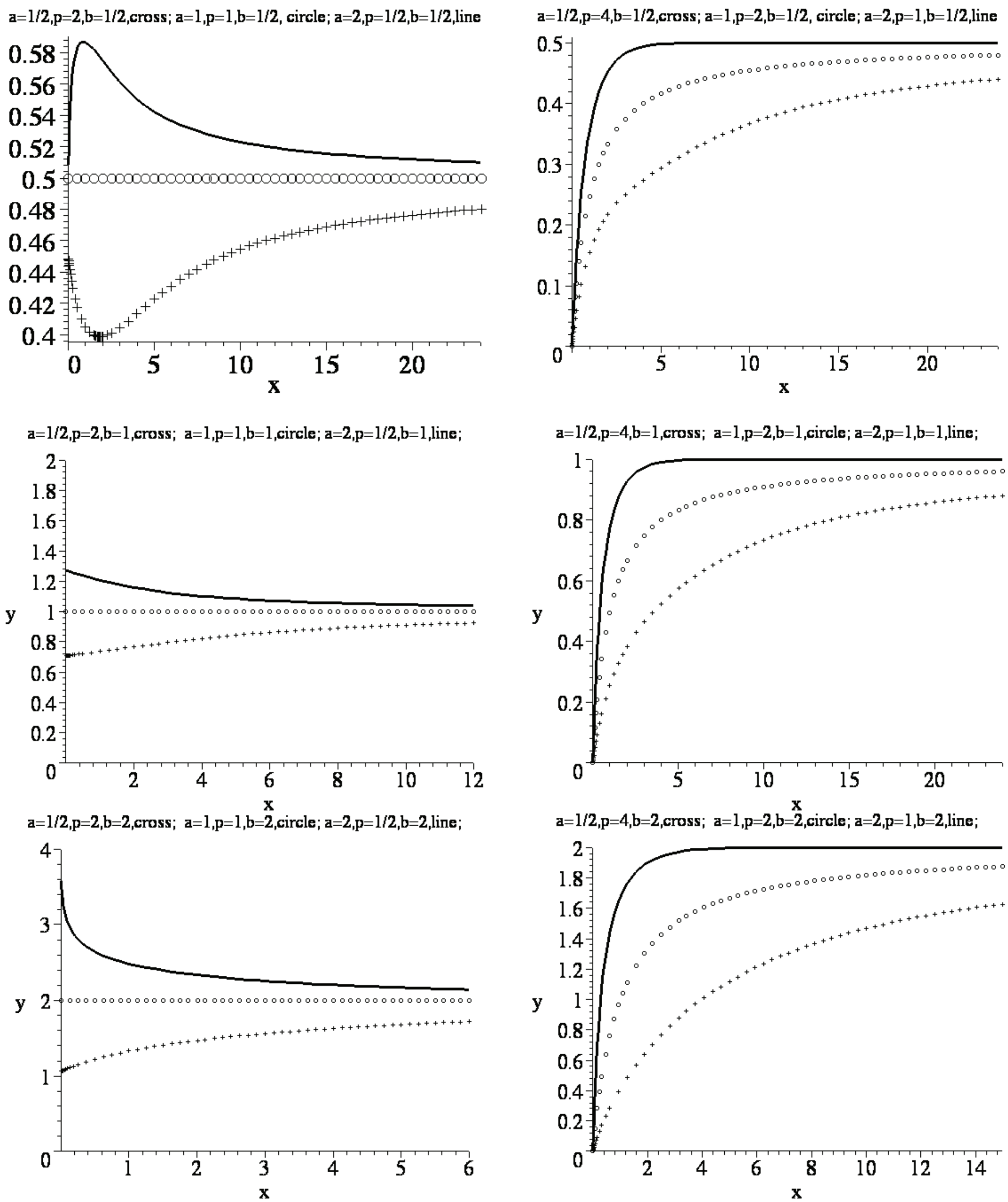

Figure 6. Hazard Function of $BG(\alpha, \beta, \rho, 1)$ when $\alpha\rho \geq 1$

As stated in Lemma 2, the curves of the hazard functions start at the values given in Lemma 1 and go to the value of β as x goes to ∞ regardless the values of other parameters. When $\alpha\rho < 1$ and $\beta \geq 1$ (see also Figure 2), $g(x)$ has a reversed J shape and the trends of hazard functions for β =1 and β =2 (both β>1) are similar (Figure 5). When $\alpha\rho = 1$ and β =1/2, the hazard function has a nonzero maximum or minimum. The hazard function is constant when $\alpha = \rho = \beta = 1$, sine *g(x)* is the exponential distribution. Within each plot, a larger α value seems to result in a larger value of the hazard function. When $\alpha\rho > 1$, $g(x)$ has a nonzero mode (see also Figure 4) and the corresponding hazard function is non-decreasing.

When $\alpha = \beta = 1$, it is Gamma function. $\lim_{x\to\infty} = 1/\lambda$, which is different from that of beta-gamma. Also, the hazard function of the beta-gamma can handle bathtub cases where gamma can not. Therefore, the beta-gamma distribution is more flexible. This is especially important when the beta parameter is not near one.

Parameter Estimation Using Maximum Likelihood Method

Let $x_1, x_2, \ldots\ldots x_n$ be a random sample of size *n* from a beta-gamma distribution defined in (1.1), the log-likelihood function $l(\alpha, \beta, \rho, \lambda)$ is then given by

$$n \log \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} + (\alpha-1)\sum_{i=1}^{n} \log F(x_i)$$

$$+ (\beta-1)\sum_{i=1}^{n} \log(1-F(x_i)) + \sum_{i=1}^{n} \log f(x_i),$$

where $f(x)$ and $F(x)$ are the pdf and cdf of the gamma distribution with parameters ρ and λ, respectively. Let $\psi(z) = d\Gamma(z)/dz$ be the digamma function. The equations for solving the maximum likelihood estimates of α, β, ρ and λ are given in Appendix.

The example in the next section, initial estimates of ρ and λ is first computed by assuming the data set follows gamma distribution with $\alpha = 1$ and $\beta = 1$, the results from MLE of (ρ, λ) along with $\alpha = 1$ and $\beta = 1$ then are used as the initial values for solving the equations (A.3) to (A.6).

Applications of the Beta-Gamma Distribution

An application of the proposed distribution is presented using the data sets given in Park, Leslie, and Mertz (1964), Park (1954), Moffa and Costantino (1977). Costantino and Desharnais (1981) established a gamma-state probability distribution for adult numbers in continuously growing populations of the flour beetle Tribolium. The hypothesis that the data set is from a beta-gamma distributed population is tested using the observed frequency distributions of adult numbers for Tribolium castaneum and Tribolium Confusum.

The beta-gamma distribution is fitted to the ten data sets discussed above, and the results are compared to those from gamma distribution and beta-normal distribution proposed by Eugene (2001) where the maximum likelihood method was used. Table 12 tabulates the resulting chi-square values form the goodness-of-fit test for the 10 data sets, and for illustration of the computations Tables 13 and 14 contains results for two of the ten data sets (Data set # 6 and #10). The expected numbers are calculated using the respective distribution with the parameters set at their maximum likelihood estimates. The chi-square goodness-of-fit test is then employed to make a comparison between the observed and expected number of observations under each distribution. Note that a class interval with an expected number less than 5 is combined with the adjacent class to avoid inflating the chi-square test statistic.

Table 12. The resulting χ^2 values (p-value, d.f.) from the goodness-of-fit tests for the 10 data sets.

Data set	Gamma	Beta-Normal	Beta-Gamma
#1	24.03 (0.0043, 9)	5.04 (0.6545, 7)	7.88 (0.3433, 7)
#2	48.16 (0, 12)	27.02 (0.0026, 10)	20.50 (0.0249, 10)
#3	129.18 (0, 17)	74.85 (0, 15)	72.63 (0, 15)
#4	78.07 (0, 11)	25.39 (0.0030, 9)	28.36 (0.0008, 9)
#5	23.62 (0.0144, 11)	19.99 (0.0180, 9)	17.89 (0.0365, 9)
#6	10.72 (0.3793, 10)	7.42 (0.4913, 8)	7.05 (0.5312, 8)
#7	21.67 (0.0169, 10)	10.56 (0.2280, 8)	12.89 (0.1157, 8)
#8	55.71 (0, 9)	25.05 (0.0007, 7)	22.28 (0.0023, 7)
#9	25.02 (0.2463, 21)	16.85 (0.6001, 19)	16.54 (0.6210, 19)
#10	17.19 (0.3076, 15)	17.07 (0.1959, 13)	15.01 (0.3067, 13)

It is of no surprise that the proposed beta-gamma distribution fits better than the gamma distribution for all the data sets. Seven of the ten data sets, the beta-gamma distribution fits better than the beta-normal distribution based on the chi-squares values. Note that, for example, the data set in Table 14 appears to have a long right tail, it is reasonable that beta-gamma distribution performed the best.

Table 13. Observed and Expected Frequencies for Tribolium Confusum Strain # 4(b)

			Expected					
$x-value$	observed		Gamma		Beta-Normal		Beta-Gamma	
37.5	5	10	2.36	8.78	2.85	10.63	4.01	12.12
42.5	5		6.42		7.78		8.11	
47.5	14		15.37		16.45		16.90	
52.5	33		28.26		27.43		28.59	
57.5	40		41.74		37.79		40.48	
62.5	49		51.32		45.59		49.06	
67.5	44		53.98		50.12		51.78	
72.5	52		49.66		50.31		48.48	
77.5	44		40.67		44.99		41.02	
82.5	28		30.07		34.99		31.64	
87.5	29		20.31		23.48		22.07	
92.5	13		12.66		13.62		13.62	
97.5	9		7.34		6.87		7.24	
102.5	1	3	3.99	7.84	3.02	5.73	3.25	5.01
107.5	1		2.05		1.16		1.23	
112.5	1		1.80		1.55		0.53	
Total	368		368		368		368	
$\hat{\alpha}$					0.45		0.17	
$\hat{\beta}$					0.23		0.69	
$\hat{\mu}$					62.79			
$\hat{\sigma}$					6.74			
$\hat{\rho}$			25.61				111.58	
$\hat{\lambda}$			2.71				0.74	
χ^2			10.72		7.42		7.05	
p-value			0.3793		0.4913		0.5312	
degree of freedom			10		8		8	

Table 14. Observed and Expected Frequencies for Tribolium Castaneum at $24^0 C$ (b)

$x-value$	observed	Expected: Gamma	Expected: Beta-Normal	Expected: Beta-Gamma
25	0	0.02	0.12	0.03
35	0 } 3	0.44 } 3.43	0.75 } 4.25	0.46 } 3.69
45	3	2.97	3.38	3.20
55	9	11.51	11.18	12.30
65	39	29.71	28.07	31.23
75	53	57.05	55.08	58.77
85	77	87.54	87.09	88.39
95	105	112.64	114.43	111.90
105	135	125.81	128.66	123.66
115	114	125.09	127.23	122.40
125	113	112.87	113.37	110.58
135	92	93.80	92.91	92.45
145	59	72.63	71.18	72.29
155	54	52.89	51.60	53.29
165	38	36.51	35.72	37.27
175	22	24.03	23.75	24.87
185	17	15.17	15.22	15.91
195	6	9.22	9.43	9.80
205	10	5.42	5.65	5.83
215	3	3.09	3.28	3.36
225	2	1.71	1.85	1.88
235	0 } 6	0.92 } 6.68	1.01 } 7.18	1.03 } 7.33
245	1	0.49	0.53	0.55
255	0	0.25	0.27	0.29
265	0	0.22	0.24	0.22
$\hat{\alpha}$			12.34	0.82
$\hat{\beta}$			0.68	0.79
$\hat{\mu}$			27.33	
$\hat{\sigma}$			47.01	
$\hat{\rho}$		13.86		17.23
$\hat{\lambda}$		8.50		6.74
χ^2		17.19	17.07	15.01
p-value		0.3076	0.1959	0.3067
degree of freedom		15	13	13

Conclusion

A beta-gamma distribution is proposed that include the gamma, exponential, and beta-exponential distributions as its special cases. When $\alpha\rho > 1$, it is unimodal with a concave shape. When $\alpha\rho \leq 1$ and $\beta \geq 1$, it has a reversed-J shape. When $\alpha < 1$ and $\rho < 1$, it also has a reversed-J shape. When $\alpha\rho = 1$ and $\beta < 1$, it can be reverse J-shaped or unimodal with a concave shape. When $\alpha\rho < 1$ and $\beta < 1$, $g(x)$ has a reversed-J shape except when $\alpha\rho$ is close to 1 with $\alpha > 1$ or $\rho > 1$ for a range of β values of less than one, in which it is bimodal with a mode of zero and a nonzero mode.

Note that the beta-normal distribution in Eugene, *et al* (2002) can be bimodal with two nonzero modes; the beta-gamma can be bimodal with a mode of zero and a nonzero mode. Closed forms of moments are derived when parameters are integers. The mean and standard deviation increase with α and decrease with β.

The hazard function of the proposed beta-gamma distribution appears to be versatile in the sense it could be constant, nondecreasing, nonincreasing, concave, and convex. This property is potentially useful in real word problems. The estimation of the parameters can be computed via maximum likelihood method. The proposed beta-gamma distribution is a generalization of the widely used gamma distribution and is at least as efficient as the beta-normal if not better.

References

Costantino, R. F. & Desharnais, R. A. (1981). Gamma distributions of adult numbers for tribolium populations in the region of their steady states. *Journal of Animal Ecology*, *50,* 667-681.

Eugene, N., Lee, C. & Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*, *31*(4), 497-512.

Famoye, F., Lee, C. & Eugene, N. (2004). Beta-normal distribution. Bimodality properties and application. *Journal of Modern Applied Statistical Methods*, *3*(1), 85-103.

Farewell, V. T. (1977). A study of distributional shape in life testing. *Technometrics*, *19*(1), 69-75.

Gupta, A.K.; Nadarajah, S. (2004). On the moments of the beta normal distribution. *Communication in Statistics-Theory and Methods*, *33*.

Lawless, J. F. (1980). Inference in the generalized gamma and log gamma distributions. *Technometrics*, *22*(3), 409-419.

Maynard, J. (2003). Unpublished doctoral dissertation. Central Michigan University.

McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica*, *52,* 647-663.

Moffa, A. M. & Costantino, R. F. (1977). Genetic analysis of a population of tribolium IV. Polymorphism and demographic equilibrium, *Genetics, 87,* 785-805.

Park, T., Leslie, P. H. & Mertz, D. B. (1964). Genetics strains and competition in population of tribolium. *Physio. Zool., 37,* 97-162.

Park, T. (1954). Experimental studies of interspecies competition. II. Temperature, humidity, and competition in two species of tribolium, *Physio. Zool., 27,* 177-238.

Appendix

Proof of Lemma 1

Using the Taylor's expansions of $e^{-x/\lambda}$, the gamma density function is

$$f(x) = \frac{x^{\rho-1}}{\lambda^{\rho}\Gamma(\rho)} \; e^{-x/\lambda}$$

$$= \frac{x^{\rho-1}}{\lambda^{\rho}\Gamma(\rho)}\left[1 - \frac{x}{\lambda} + \frac{x^2}{2!\lambda^2} + \cdots\cdots + \frac{(-1)^n x^n}{n!\lambda^n} + O(x^{n+1})\right] = \frac{x^{\rho-1}}{\lambda^{\rho}\Gamma(\rho)} + O(x^{\rho}), \tag{A.1}$$

and $F(x) = \int_0^x f(x)dx = x^{\rho} \big/ \rho\lambda^{\rho}\Gamma(\rho) + O(x^{\rho+1})$. For simplicity of presentation, let $f = f(x)$, $g = g(x),\ F = F(x)$ and $F^c = [F(x)]^c$. Using (A.1), the density function $g(x)$ in (1) becomes

$$\frac{x^{\rho-1}e^{-x/\lambda}}{\lambda^{\rho}\Gamma(\rho)B(\alpha,\beta)}\left[\frac{x^{\rho}}{\rho\lambda^{\rho}\Gamma(\rho)} + O(x^{\rho+1})\right]^{\alpha-1} (1-F)^{\beta-1}$$

$$= \frac{x^{\rho-1+\rho(\alpha-1)}e^{-x/\lambda}}{\Gamma^{\alpha}(\rho)B(\alpha,\beta)\rho^{\alpha-1}\lambda^{\rho\alpha}}[1+O(x)]^{\alpha-1}(1-F)^{\beta-1}$$

$$= \frac{x^{\rho\alpha-1}e^{-x/\lambda}}{\Gamma^{\alpha}(\rho)B(\alpha,\beta)\rho^{\alpha-1}\lambda^{\rho\alpha}}[1+O(x)]^{\alpha-1}(1-F)^{\beta-1}.$$

Lemma 1 can now be readily seen because F is a cdf and $\lim_{x\to 0} F(x) = 0$.

Proof of Theorem 1

When α and β are integers, the *n*th moment of the beta-gamma random variable with density function in (1) is

$$E(X^n) = \int_0^\infty x^n \frac{1}{B(\alpha,\beta)} F(x)^{\alpha-1}[1-F(x)]^{\beta-1} f(x)dx$$

$$= \frac{1}{B(\alpha,\beta)} \sum_{j=0}^{\beta-1} (-1)^j \binom{\beta-1}{j} \int_0^\infty x^n F^{\alpha-1} F^j f(x)dx$$

$$= \frac{1}{B(\alpha,\beta)} \sum_{j=0}^{\beta-1} (-1)^j \binom{\beta-1}{j} \int_0^\infty x^n F^{\alpha+j-1} f(x)dx .$$

$$= \frac{1}{B(\alpha,\beta)} \sum_{j=0}^{\beta-1} (-1)^j \binom{\beta-1}{j} \int_0^\infty x^n [1-(1-F)]^{\alpha+j-1} f(x)dx$$

$$= \frac{1}{B(\alpha,\beta)} \sum_{j=0}^{\beta-1} (-1)^j \binom{\beta-1}{j} \left\{ \sum_{k=0}^{\alpha+j-1} (-1)^k \binom{\alpha+j-1}{k} I_{n,k} \right\},$$

where

$$I_{n,k} = \int_0^\infty x^n f(x)(1-F)^k dx .$$

Proof: of Corollary 1

When $\alpha = 2, \beta = 1$,

$$E(X^n) = \frac{1}{B(2,1)} \sum_{k=0}^{1} (-1)^k \binom{1}{k} I_{n,k} = \frac{1}{B(2,1)}[I_{n,0} - I_{n,1}] = 2I_{n,0} - 2I_{n,1} . \tag{A.2}$$

The first term $I_{n,0}$ is given by

$$\int_0^\infty \frac{x^{n+\rho-1}}{\lambda^\rho \Gamma(\rho)} e^{-x/\lambda} dx = \int_0^\infty \frac{\lambda^{n+\rho}(\frac{x}{\lambda})^{n+\rho-1}}{\lambda^\rho \Gamma(\rho)} e^{-x/\lambda} d(\frac{x}{\lambda}) = \int_0^\infty \frac{\lambda^{n+\rho} t^{n+\rho-1}}{\lambda^\rho \Gamma(\rho)} e^{-t} d(t) = \frac{\lambda^n \Gamma(n+\rho)}{\Gamma(\rho)},$$

and $I_{n,1} = \int_0^\infty x^n f(x)(1-F)dx$ is

$$\int_0^\infty \frac{x^{\rho-1}}{\lambda^\rho \Gamma(\rho)} x^n e^{-x/\lambda} [e^{-x/\lambda} + (x/\lambda)e^{-x/\lambda} + ... + \frac{(x/\lambda)^{\rho-1} e^{-x/\lambda}}{\Gamma(\rho)}]dx$$

$$= \int_0^\infty \frac{(2x/\lambda)^{n+\rho-1} (\lambda/2)^{n+p-1}}{\lambda^\rho \Gamma(\rho)} e^{-2x/\lambda} [1 + x/\lambda + ... + \frac{(x/\lambda)^{\rho-1}}{\Gamma(\rho)}] d(2x/\lambda)(\lambda/2)$$

$$= \frac{\lambda^n}{2^{n+p}\Gamma(\rho)} \int_0^\infty t^{n+p-1} e^{-t} [1 + t/2 + ... + \frac{(t/2)^{\rho-1}}{\Gamma(\rho)}] d(t)$$

$$= \frac{\lambda^n}{2^{n+p}\Gamma(\rho)} [\int_0^\infty t^{n+p-1} e^{-t} dt + \int_0^\infty t^{n+p-1} e^{t} (t/2) dt + ... + \int_0^\infty t^{n+p-1} e^{-t} (t/2)^{\rho-1} / \Gamma(\rho) dt]$$

$$= \frac{\lambda^n \Gamma(n+\rho)}{2^{n+\rho}\Gamma(\rho)} + \frac{\lambda^n \Gamma(n+\rho+1)}{2^{n+\rho+1}\Gamma(\rho)} + \ldots + \frac{\lambda^n \Gamma(n+2\rho-1)}{2^{n+2\rho-1}\Gamma^2(\rho)}$$

$$= \sum_{i=0}^{\rho-1} \frac{\lambda^n \Gamma(n+\rho+i)}{2^{n+\rho+i}\Gamma(\rho)\Gamma(i+1)}.$$

When ρ is an integer, $I_{n,0} = \lambda^n (n+\rho-1)\cdots(\rho+1)\rho$ and $I_{n,1} = \sum_{i=0}^{\rho-1} \frac{\lambda^n (n+\rho+i-1)!}{2^{n+\rho+i}(\rho-1)!i!}$.

Substituting $I_{n,0}$ and $I_{n,1}$ into (A.2), the results of (7) and (8) are obtained.

Proof of Lemma 2:

(a) As *x* goes 0, $\lim_{x\to 0} H(x)$ is

$$\lim_{x\to 0} \frac{\frac{1}{B(\alpha,\beta)} F^{\alpha-1}(1-F)^{\beta-1} f(x)}{1-\frac{1}{B(\alpha,\beta)}\int_0^x F^{\alpha-1}(1-F)^{\beta-1} f(x)dx} = \lim_{x\to 0} g(x),$$

which is given in Lemma 1.

Proof: (b) As *x* goes to ∞, by *L'Hospital* Rule, $\lim_{x\to\infty} H(x)$ is

$$\lim_{x\to\infty} \frac{(\alpha-1)F^{\alpha-2} f(x)(1-F)^{\beta-1} f(x) + F^{\alpha-1}(\beta-1)(1-F)^{\beta-2}(-f^2) + F^{\alpha-1}(1-F)^{\beta-1} f'}{-F^{\alpha-1}(1-F)^{\beta-1} f}$$

$$= \lim_{x\to\infty}\left[\frac{(1-\alpha)f(x)}{F(x)} + \frac{(\beta-1)f(x)}{1-F(x)} - \frac{f'(x)}{f(x)}\right]$$

$$= \frac{0}{1} + (\beta-1)\lim_{x\to\infty}\frac{f'(x)}{-f(x)} - \lim_{x\to\infty}\frac{f'(x)}{f(x)}$$

$$= -\beta\lim_{x\to\infty}\frac{f'(x)}{f(x)} = -\beta\lim_{x\to\infty}\frac{\frac{-x^{\rho-1}e^{-\frac{x}{\lambda}}}{\lambda} + (\rho-1)x^{\rho-2}e^{-\frac{x}{\lambda}}}{x^{\rho-1}e^{-\frac{x}{\lambda}}}$$

$$= -\beta\lim_{x\to\infty}\left(-\frac{1}{\lambda} + \frac{\rho-1}{x}\right) = \frac{\beta}{\lambda}.$$

Note that unlike $\lim_{x\to 0} H(x)$, the limit $\lim_{x\to\infty} H(x) = \beta/\lambda$ does not depend on α and ρ. In other word, the instantaneous failure rate will not depend on α and ρ in the long run.

The Equations for Solving the Maximum Likelihood Estimates:

$$0 = \frac{\partial l}{\partial \alpha} = n\psi(\alpha+\beta) - n\psi(\alpha) + \sum_{i=1}^{n} \log F(x_i) \,. \tag{A.3}$$

$$0 = \frac{\partial l}{\partial \beta} = n\psi(\alpha+\beta) - n\psi(\beta) + \sum_{i=1}^{n} \log(1 - F(x_i)) \,. \tag{A.4}$$

$$0 = \frac{\partial l}{\partial \rho} = (\alpha-1)\sum_{i=1}^{n} \frac{1}{F(x_i)} \frac{\partial F(x_i)}{\partial \rho} - (\beta-1) \sum_{i=1}^{n} \frac{1}{1-F(x_i)} \frac{\partial F(x_i)}{\partial \rho} + \sum_{i=1}^{n} \frac{1}{f(x_i)} \frac{\partial f(x_i)}{\partial \rho}, \tag{A.5}$$

$$\frac{\partial f}{\partial \rho} = \frac{e^{-x/\lambda}[x^{\rho-1}(\log x)\lambda^{\rho}\Gamma(\rho) - x^{\rho-1}(\lambda^{\rho}\Gamma'(\rho) + \lambda^{\rho}(\log\lambda)\Gamma(\rho))]}{\lambda^{2\rho}\Gamma^{2}(\rho)}$$

$$= f(x)\frac{(\log x)\lambda^{\rho}\Gamma(\rho) - \lambda^{\rho}\Gamma'(\rho) - \lambda^{\rho}(\log\lambda)\Gamma(\rho)}{\lambda^{\rho}\Gamma(\rho)} = f(x)[\log x - \psi(p) - \log\lambda] \,,$$

$$\frac{\partial F}{\partial \rho} = \int_{0}^{x} f(x)[\log x - \psi(p) - \log\lambda]dx \,.$$

$$0 = \frac{\partial l}{\partial \lambda} = (\alpha-1)\sum_{i=1}^{n} \frac{1}{F(x_i)} \frac{\partial F(x_i)}{\partial \lambda} - (\beta-1) \sum_{i=1}^{n} \frac{1}{1-F(x_i)} \frac{\partial F(x_i)}{\partial \lambda} + \sum_{i=1}^{n} \frac{1}{f(x_i)} \frac{\partial f(x_i)}{\partial \lambda}, \tag{A.6}$$

$$\frac{\partial f}{\partial \lambda} = \frac{x^{\rho-1}}{\Gamma(\rho)} \frac{\frac{x}{\lambda^{2}} e^{-x/\lambda}\lambda^{\rho} - e^{-x/\lambda}\rho\lambda^{\rho-1}}{\lambda^{2\rho}} = f(x)(x/\lambda^{2} - \rho/\lambda) \,; \; \frac{\partial F}{\partial \lambda} = \int_{0}^{x} f(x)(x/\lambda^{2} - \rho/\lambda)dx \,.$$

The SAS IML optimization methods can be employed to solve the loglikelihood equations (A.3) – (A.6) iteratively. The lengthy and tedious second derivatives required in the algorithm are not presented here, but are available upon request.