A Note on the Means of the Beta-Geometric and Pareto of the Second Kind Distributions

Peter S. Fader
www.petefader.com

Bruce G. S. Hardie
www.brucehardie.com

Julián L. Zambrano [†]

October 2023

1 Introduction

The beta-geometric (BG) distribution and the Pareto of the second kind (P(II)) distribution are simple mixture models that can be used to characterize duration data.[1] In this note we present derivations of the mean of each distribution and explore why the mean is undefined for a subset of each distribution's parameter space.

2 Preliminaries

The Gaussian hypergeometric function is defined by the hypergeometric series

$$ {}_2F_1(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{n=0}^{\infty} \frac{\Gamma(a+n)\Gamma(b+n)}{\Gamma(c+n)} \frac{z^n}{n!} \,. \tag{1} $$

The series converges when $|z| < 1$. When $|z| = 1$, it converges if $c - a - b > 0$. When $c - a - b > 0, c \neq 0, -1, -2, \ldots$,

$$ {}_2F_1(a,b;c;1) = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)} \,. \tag{2} $$

[1] For further details, see: Fader, Peter S., Bruce G. S. Hardie, Daniel McCarthy, and Ramnath Vaidyanathan (2019), "Exploring the Equivalence of Two Common Mixture Models for Duration Data," *The American Statistician*, **73** (3), 288–295.

3 The Mean of the BG Distribution

A beta mixture of geometric distributions, also known as the beta-geometric (BG) distribution, has pmf

$$P(T = t \,|\, \gamma, \delta) = \frac{B(\gamma + 1, \delta + t - 1)}{B(\gamma, \delta)}\,, \;\; \gamma, \delta > 0,\; t = 1, 2, 3, ... \tag{3}$$

and survivor function

$$S(t \,|\, \gamma, \delta) = \frac{B(\gamma, \delta + t)}{B(\gamma, \delta)}\,, \;\; \gamma, \delta > 0,\; t = 0, 1, 2, ... \tag{4}$$

The easiest way to derive the mean of the BG is by conditioning.[2] Noting that the mean of the geometric distribution with parameter θ is $1/\theta$, we have

$$\begin{aligned} E(T \,|\, \gamma, \delta) &= \int_0^1 E(T \,|\, \theta) g(\theta \,|\, \gamma, \delta)\, d\theta \\ &= \int_0^1 \frac{1}{\theta} \frac{\theta^{\gamma-1}(1-\theta)^{\delta-1}}{B(\gamma, \delta)}\, d\theta \\ &= \frac{1}{B(\gamma, \delta)} \int_0^1 \theta^{\gamma-2}(1-\theta)^{\delta-1} d\theta \,. \end{aligned} \tag{5}$$

We know from Euler's integral of the first kind that for $\gamma > 1$ the integral is the beta function with parameters $\gamma - 1$ and δ. Therefore,

$$\begin{aligned} E(T \,|\, \gamma, \delta) &= \frac{B(\gamma - 1, \delta)}{B(\gamma, \delta)} \\ &= \frac{\gamma + \delta - 1}{\gamma - 1}\,. \end{aligned} \tag{6}$$

But what happens when $\gamma \leq 1$? Can we derive an expression for $E(T \,|\, \gamma, \delta)$? Is the mean defined?

3.1 An Alternative Derivation

We know from the definition of a geometric series that

$$\frac{1}{\theta} = \sum_{t=0}^{\infty} (1-\theta)^t\,, \text{ for } |\theta| < 1.$$

[2] See the appendix for a derivation using the standard formula for the mean of a discrete distribution.

(This also follows from a Taylor series expansion of $1/\theta$.) We can therefore rewrite (5) as

$$
\begin{aligned}
E(T \mid \gamma, \delta) &= \int_0^1 \left\{ \sum_{t=0}^{\infty} (1-\theta)^t \right\} \frac{\theta^{\gamma-1}(1-\theta)^{\delta-1}}{B(\gamma,\delta)} d\theta \\
&= \sum_{t=0}^{\infty} \int_0^1 (1-\theta)^t \frac{\theta^{\gamma-1}(1-\theta)^{\delta-1}}{B(\gamma,\delta)} d\theta \\
&= \sum_{t=0}^{\infty} \frac{B(\gamma,\delta+t)}{B(\gamma,\delta)} \qquad (7) \\
&= \frac{\Gamma(\gamma+\delta)}{\Gamma(\delta)} \sum_{t=0}^{\infty} \frac{\Gamma(\delta+t)}{\Gamma(\gamma+\delta+t)} \, . \qquad (8)
\end{aligned}
$$

Note that (7) also follows from the so-called alternative expectation formula:

$$E(X) = \sum_{x=0}^{\infty} S(x) \, .$$

Looking closely at (1) and (8), we see that $E(T \mid \gamma, \delta) = {}_2F_1(1, \delta; \gamma+\delta; 1)$. It follows from (2) that for $\gamma - 1 > 0$ (i.e., $\gamma > 1$),

$$
\begin{aligned}
E(T \mid \gamma, \delta) &= \frac{\Gamma(\gamma+\delta)\Gamma(\gamma+\delta-1-\delta)}{\Gamma(\gamma+\delta-1)\Gamma(\gamma+\delta-\delta)} \\
&= \frac{\gamma+\delta-1}{\gamma-1} \, .
\end{aligned}
$$

So the solution to (8) is (6) when $\gamma > 1$. We know from the definition of the Gaussian hypergeometric function that the series does not converge when $\gamma \leq 1$, and therefore the mean is undefined. We could stop here, but let us explore this convergence (or lack thereof) in greater depth.

3.2 Exploring Convergence

Let

$$S_n = \frac{\Gamma(\gamma+\delta)}{\Gamma(\delta)} \sum_{t=0}^{n} \frac{\Gamma(\delta+t)}{\Gamma(\gamma+\delta+t)} \, .$$

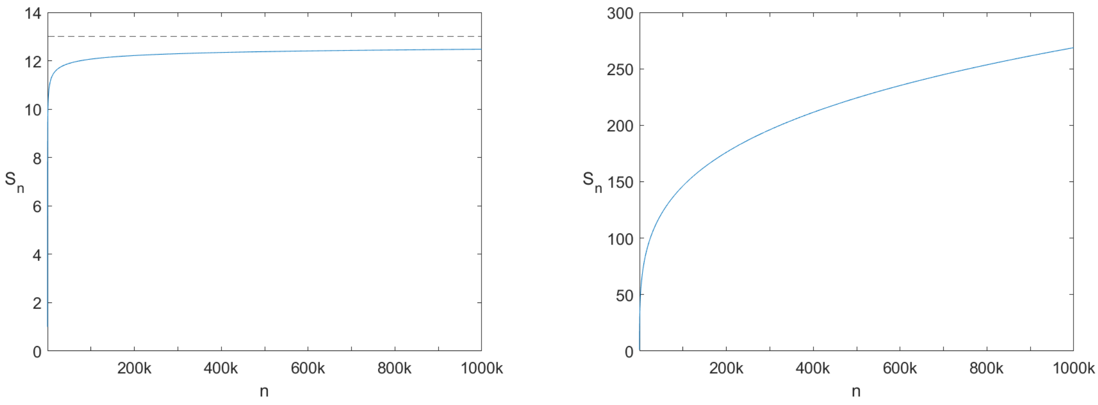
(Note that $E(T) = \lim_{n\to\infty} S_n$.)

Consider the parameter values $\gamma = 1.25, \delta = 3.00$. Given these parameters, $E(T) = 13$. We plot in Figure 1a S_n for $n = 0, 1, 2, \ldots, 1{,}000{,}000$. We see that S_n is converging to a limit.

Now consider the parameter values $\gamma = 0.75, \delta = 3.00$. Given our analysis above, $E(T)$ is undefined for these parameter values. We plot in Figure 1b S_n for $n = 0, 1, 2, \ldots, 1{,}000{,}000$. It appears that S_n is not converging to a limit.

3.3 Proving Divergence

While we can appeal to the properties of the Gaussian hypergeometric function, let us formally prove that (8) does not converge when $\gamma \leq 1$.



(a) (b)

Figure 1: Plot of S_n for (a) $\gamma = 1.25, \delta = 3.00$ and (b) $\gamma = 0.75, \delta = 3.00$.

We will make use of the integral test for convergence, which can be expressed as follows:

> Let $f(x)$ be a continuous, positive, decreasing function on $[k, \infty)$ and $f(n) = a_n$. The series
> $$\sum_{t=k}^{\infty} a_n$$
> converges if and only if
> $$\int_k^{\infty} f(x)dx$$
> is finite (i.e., convergent). If the integral is infinite (i.e., divergent), the series diverges.

The integral test is often expressed setting k to 1. However it also holds for $k = 2, 3, \ldots$.

Let us first consider the case of $\gamma = 1$, in which case

$$f(t) = \frac{\Gamma(\delta + t)}{\Gamma(\delta + t + 1)} = \frac{1}{\delta + t} .$$

Now

$$\begin{aligned} \int_1^{\infty} \frac{1}{\delta + t} dt &= \lim_{s \to \infty} \int_1^{s} \frac{1}{\delta + t} dt \\ &= \lim_{s \to \infty} \Big\{ \ln(\delta + s) - \ln(\delta + 1) \Big\}, \end{aligned}$$

which is divergent. Therefore

$$\sum_{t=1}^{\infty} \frac{1}{\delta + t}$$

diverges. It follows that

$$\sum_{t=0}^{\infty} \frac{1}{\delta + t}$$

diverges, and therefore (8) diverges when $\gamma = 1$ and the mean is undefined.

Turning to the case of $\gamma < 1$, consider

$$A = \sum_{t=2}^{\infty} \frac{\Gamma(\delta + t)}{\Gamma(\gamma + \delta + t)}$$

and

$$\begin{aligned} B &= \sum_{t=2}^{\infty} \frac{\Gamma(\delta + t)}{\Gamma(\delta + t + 1)} \\ &= \sum_{t=2}^{\infty} \frac{1}{\delta + t} . \end{aligned}$$

We know from the integral test for convergence that B diverges. If we can show that

$$\frac{\Gamma(\delta + t)}{\Gamma(\gamma + \delta + t)} \geq \frac{\Gamma(\delta + t)}{\Gamma(\delta + t + 1)} \; \forall \, t \in \{2, 3, \ldots\}$$

then we can conclude that A diverges. Now,

$$\frac{\Gamma(\delta + t)}{\Gamma(\gamma + \delta + t)} \geq \frac{\Gamma(\delta + t)}{\Gamma(\delta + t + 1)}$$

implies $\Gamma(\delta + t + 1) \geq \Gamma(\gamma + \delta + t)$. We see from Figure 2 that $\Gamma(x)$ is monotonically increasing for $x > 2$ (strictly speaking, for $x > 1.46163\ldots$). Therefore, when $\gamma < 1$, $\Gamma(\delta + t + 1) > \Gamma(\gamma + \delta + t)$ for $t \geq 2$. This means A diverges and, by extension, (8) diverges and the mean is undefined.

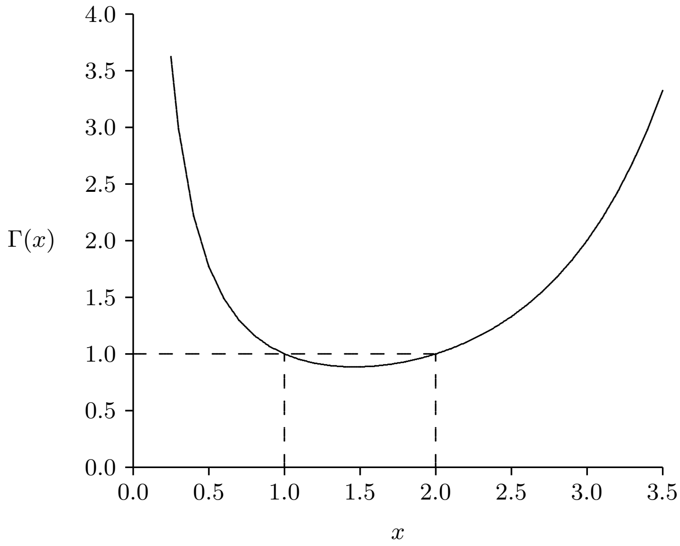


Figure 2: Graph of $\Gamma(x)$ for $x > 0$

4 The Mean of the P(II) Distribution

A gamma mixture of exponential distributions, which is one chance mechanism that yields the Pareto distribution of the second kind, has cdf

$$F(t) = 1 - \left(\frac{\alpha}{\alpha + t}\right)^r, \quad r, \alpha > 0,\ t \geq 0\,,$$

where r is the shape parameter and α is the inverse-scale parameter of the gamma mixing distribution.[3]

The easiest way to derive the mean of the P(II) is by conditioning. Noting that the mean of the exponential distribution with parameter λ is $1/\lambda$, we have

$$\begin{aligned} E(T \mid r, \alpha) &= \int_0^\infty E(T \mid \lambda) g(\lambda \mid r, \alpha) d\lambda \\ &= \int_0^\infty \frac{1}{\lambda} \frac{\alpha^r \lambda^{r-1} e^{-\alpha\lambda}}{\Gamma(r)} d\lambda \\ &= \frac{\alpha}{r-1} \int_0^\infty \frac{\alpha^{r-1} \lambda^{r-2} e^{-\alpha\lambda}}{\Gamma(r-1)} d\lambda\,. \end{aligned}$$

When $r > 1$, the integrand is a gamma pdf with shape parameter $r - 1$ and inverse scale parameter α, which means the integral, by definition equals one, and

$$E(T \mid r, \alpha) = \frac{\alpha}{r-1}\,. \tag{9}$$

But what happens when $r \leq 1$? Can we derive an expression for $E(T \mid r, \alpha)$? Is the mean defined?

4.1 An Alternative Derivation

Using the so-called alternative expectation formula, which for a continuous random variable is

$$E(X) = \int_0^\infty S(x)\, dx\,,$$

we have

$$\begin{aligned} E(T \mid r, \alpha) &= \int_0^\infty \left(\frac{\alpha}{\alpha + t}\right)^r dt \\ &= \lim_{s\to\infty} \int_0^s \left(\frac{\alpha}{\alpha + t}\right)^r dt \\ &= \lim_{s\to\infty} \frac{\alpha}{r-1} \left\{1 - \left(\frac{\alpha}{\alpha + s}\right)^{r-1}\right\}. \end{aligned}$$

When $r > 1$, this converges to $\alpha/(r - 1)$, i.e., (9). When $r = 1$, we are dividing by 0 and the solution is undefined. When $r < 1$, this diverges (to ∞). Therefore the mean of the P(II) is undefined when $r \leq 1$.

[3] In many references, α is called a scale parameter; in R, it is called a rate parameter.

Appendix

We can also derive of the mean of the BG using the standard formula:

$$\begin{aligned}
E(T\,|\,\gamma,\delta) &= \sum_{t=1}^{\infty} t\,P(T=t\,|\,\gamma,\delta)\\
&= \sum_{t=1}^{\infty} t\,\frac{B(\gamma+1,\delta+t-1)}{B(\gamma,\delta)}\\
&= \frac{\Gamma(\gamma+1)\Gamma(\gamma+\delta)}{\Gamma(\gamma)\Gamma(\delta)}\sum_{t=1}^{\infty} t\,\frac{\Gamma(\delta+t-1)}{\Gamma(\gamma+\delta+t)}
\end{aligned}$$

which, letting $y = t-1$

$$\begin{aligned}
&= \frac{\Gamma(\gamma+1)\Gamma(\gamma+\delta)}{\Gamma(\gamma)\Gamma(\delta)}\sum_{y=0}^{\infty}(y+1)\,\frac{\Gamma(\delta+y)}{\Gamma(\gamma+\delta+y+1)}\\
&= \frac{\Gamma(\gamma+1)\Gamma(\gamma+\delta)}{\Gamma(\gamma)\Gamma(\delta)}\sum_{y=0}^{\infty}\frac{\Gamma(y+2)}{\Gamma(y+1)}\frac{\Gamma(\delta+y)}{\Gamma(\gamma+\delta+y+1)}
\end{aligned}$$

which, recalling (1)

$$\begin{aligned}
&= \frac{\Gamma(\gamma+1)\Gamma(\gamma+\delta)}{\Gamma(\gamma)\Gamma(\delta)}\frac{\Gamma(2)\Gamma(\delta)}{\Gamma(\gamma+\delta+1)}\,{}_2F_1(2,\delta;\gamma+\delta+1;1)\\
&= \frac{\Gamma(\gamma+1)\Gamma(\gamma+\delta)}{\Gamma(\gamma)\Gamma(\gamma+\delta+1)}\,{}_2F_1(2,\delta;\gamma+\delta+1;1)
\end{aligned}$$

which, recalling (2)

$$\begin{aligned}
&= \frac{\Gamma(\gamma+1)\Gamma(\gamma+\delta)}{\Gamma(\gamma)\Gamma(\gamma+\delta+1)}\frac{\Gamma(\gamma+\delta+1)\Gamma(\gamma-1)}{\Gamma(\gamma+\delta-1)\Gamma(\gamma+1)} \quad \text{when } \gamma>1\\
&= \frac{\Gamma(\gamma-1)}{\Gamma(\gamma)}\frac{\Gamma(\gamma+\delta)}{\Gamma(\gamma+\delta-1)}\\
&= \frac{\gamma+\delta-1}{\gamma-1}\,.
\end{aligned}$$