The Pareto-Gamma Mixture

Greg McNulty, FCAS

Abstract: In this paper we will review some established properties and derive some new properties of a Pareto distribution with fixed scale whose unknown shape parameter is Gamma distributed. Namely:

- that Gamma is a conjugate prior to the Pareto distribution
- the formula for the posterior parameters of the Gamma given observed data
- a closed form for the CDF of the Pareto-Gamma mixture
- that the mean and all higher moments of the distribution are infinite
- a formula for the moments of the limited expected value of a random variable following this distribution
- the intractability of the closed form

Keywords: Pareto distribution, Gamma distribution, Conjugate prior, Bayesian statistics, Reinsurance pricing

1. INTRODUCTION

Our company recently performed a study comparing the exposure based large loss models for a long-tailed line of business across a number of reinsurance cedents. The industry curves indicated losses excess of a common threshold followed a Pareto distribution with the scale parameter varying by cedent. Fitting a distribution to the sample of shape parameter values showed that they approximately followed a Gamma probability distribution.

Hoping to benefit from the simple formulas of a conjugate prior relationship for a Bayesian model (and avoid more a difficult programming exercise using R) we found in the *Wikipedia* entry for "conjugate prior" [1] that Gamma was indeed conjugate prior to the Pareto. However, unlike for almost every other such pair of distributions, the closed form of the posterior predictive distribution was not given. We wondered if this was because it was not available or did not exist.

It turns out the closed form does exist, and this paper will show a derivation relying at times on other well-established facts rather than providing full mathematical proofs. We will also explain why the mean and all higher moments of the distribution are infinite. Finally, we will derive the formula for powers of the limited expected value of x, although we will find them to be of limited usefulness.

Please note, different texts and tools will use different parameterizations of the gamma distribution. This document attempts to be self-consistent using the definition in Section 2.2.

2. CONJUGATE PRIOR RELATIONSHIP

Proofs will be supplied for each step, but the reasoning is as follows: the Pareto distribution is "log-Exponential"; Gamma is conjugate prior to the Exponential distribution; the conjugate prior relationship is preserved under the log transformation; therefore Gamma is conjugate Prior to "log-Exponential", aka the Pareto distribution.

2.1 Pareto is "Log-Exponential"

To see this, let's just write the PDF of a random variable whose log is exponentially distributed:

$$f(\ln(x)) \sim e^{-\lambda * \ln(x)} = x^{-\lambda}$$

To be a little more precise, we could start with the CDF of the Pareto:

$$F(x) = 1 - \left(\frac{x}{\theta}\right)^{-\lambda}; x > \theta$$

Now we can transform x into a random variable supported on $[0, \infty)$ by letting $y = \ln(x/\theta)$. Then we would have:

$$F(y) = 1 - (e^{y})^{-\alpha} = 1 - e^{-\alpha y}$$

This is the CDF of the exponential distribution, meaning a Pareto distributed random variable's scaled natural log is exponentially distributed. We also see here the relationship between the parameters of the distributions, namely being equal under the correct parameterizations.

2.2 Gamma is Conjugate Prior to Exponential

For this section we will use Bayes' Theorem:

$$p(\theta|x) \sim p(x|\theta) * p(\theta)$$

where x represents the data and θ represents the model parameters. We use the following parameterizations of the Exponential and Gamma distributions:

$$f(x; \lambda) = \lambda e^{-\lambda x}$$

$$\Gamma(x; \alpha, \beta) \sim x^{\alpha-1} e^{-\beta x}$$

In this case conjugate prior would mean that if *x* is Exponentially distributed, and $p(\lambda)$ is Gamma distributed, then $p(\lambda|x)$ is also Gamma distributed. Note that by *x* we mean a set of *n* observations of the random variable, which could also be denoted $\{x_i\}$. Then we have:

$$p(x|\theta) = \prod_i \lambda e^{-\lambda x_i}$$

Using Bayes' Theorem:

$$p(\lambda|x) \sim \left(\prod_i \lambda e^{-\lambda x_i} \right) * \lambda^{\alpha-1} e^{-\beta\lambda}$$

$$= \lambda^{\alpha+n-1} * e^{-(\beta+\sum x_i)\lambda}$$

$$\sim \Gamma(\lambda; \alpha + n, \beta + \sum x_i)$$

So, the posterior distribution of the Exponential parameter is again Gamma distributed, and we also have expressions for the posterior parameters of the Gamma distribution.

2.3 Conjugate Prior Relationship Preserved Under Logarithm

Now we can show that Gamma is a conjugate prior to the Pareto distribution. Suppose *x* is Pareto distributed:

$$F(x) = 1 - \left(\frac{x}{\theta}\right)^{-\lambda} ; x > \theta$$

Further suppose the Pareto α is Gamma distributed:

$$\Gamma(\lambda; \alpha, \beta) \sim \lambda^{\alpha-1} e^{-\beta\lambda}$$

Typically, the Pareto parameter shape parameter is called "alpha", but in order to avoid confusion with the Gamma parameter we choose a different letter. Let's again apply the transformation $y = \ln{(x/\theta)}$. Then we see *y* is Exponentially distributed according to:

$$G(y) = 1 - e^{-\lambda y}$$

Now suppose we observe a set $\{x_i\}$. This is equivalent to a set $\{y_i\} = \{\ln{(x_i/\theta)}\}$. Because Gamma is conjugate prior to the Exponential, we know the posterior distribution is:

$$p(\lambda|\{y_i\}) \sim \Gamma(\lambda; \alpha + n, \beta + \sum y_i)$$

But the λ of the Exponential is the same as the λ of the Pareto. We know that the Bayesian posterior distribution is Gamma and we can rewrite the above equation as:

$$p(\lambda|\{x_i\}) \sim \Gamma(\lambda; \alpha + n, \beta + \sum \ln(x_i/\theta))$$

Which means that Gamma is conjugate prior to the Pareto.

3. CLOSED FORM OF THE MIXTURE

We use the same logarithm transformation of a Pareto distributed variable into an Exponentially distributed variable to access the simpler derivation of the closed form for the Gamma-Exponential mixture.

3.1 Gamma-Exponential Mixture

Suppose we have that:

$$f(x; \lambda) = \lambda e^{-\lambda x}$$

$$\mathrm{g}(\lambda; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda}$$

In other words, *x* is Exponentially distributed, with the rate parameter itself a variable which is Gamma distributed. Can we obtain a closed form for this mixture, i.e. a formula for *f(x)* unconditional on λ that does not contain integrals?

The derivation below was posted on StackExchange [2] courtesy of user "heropup":

$$f(x) = \int_0^\infty f(x|\lambda) * g(\lambda) d\lambda$$

$$= \int_0^\infty \lambda e^{-\lambda x} \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} d\lambda$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \lambda^\alpha e^{-(\beta+x)\lambda} d\lambda$$

The integral now looks similar to the Gamma function, so we use the substitution $u = (\beta + x) * \lambda$. Note that the limits of the integral do not change since this is just a scalar multiple and both *x* and β are positive. Continuing from above:

$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \left(\frac{u}{\beta + x}\right)^\alpha e^{-u} \frac{du}{\beta + x}$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\beta+x}\right)^{\alpha+1}\int_0^\infty u^\alpha e^{-u}du$$

$$= \frac{\beta^\alpha}{\Gamma(\alpha)}\left(\frac{1}{\beta+x}\right)^{\alpha+1}\Gamma(\alpha+1)$$

Using the identity $\Gamma(\alpha+1) = \alpha * \Gamma(\alpha)$ we find:

$$f(x) = \frac{\alpha\beta^\alpha}{(\beta+x)^{\alpha+1}}$$

$$= \frac{\alpha}{\beta}\left(1+\frac{x}{\beta}\right)^{-(\alpha+1)}$$

This is the PDF of the Lomax distribution. It is essentially a standard Pareto distribution except the values are the amount in excess of the scale or threshold, which in this case is β. If the reader wishes to align these distributions with the Pareto types [4], then the standard Pareto refenced above is Type I, and the Lomax is a special case of Type II having $\mu = 0$.

3.2 Gamma-Pareto Mixture

We showed before that the Pareto is "Log-Exponential". Suppose we have a Pareto distributed random variable x with a fixed scale or threshold, but the Pareto shape or λ parameter is Gamma distributed. We have:

$$F(x) = 1 - \left(\frac{x}{\theta}\right)^{-\lambda}; x > \theta$$

$$\text{g}(\lambda; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}$$

Using the substitution $y = \ln(x/\theta)$, from the previous section we know that y, unconditional on λ, is Lomax distributed with parameters α, β. Referencing Wikipedia for the CDF of the Lomax [3], we have:

$$F(x) = 1 - \left(1 + \frac{\ln\left(\frac{x}{\theta}\right)}{\beta}\right)^{-\alpha}$$

Note that the PDF would not just be plugging the substitution into the PDF of the Lomax for *y*. The PDF is the derivative of the CDF, and so must contain an *x* somewhere due to the logarithm term. For completion, taking the derivative we obtain:

$$f(x) = \alpha \left(1 + \frac{\ln\left(\frac{x}{\theta}\right)}{\beta} \right)^{-(\alpha+1)} * \frac{\theta}{\beta x}$$

4. MOMENTS AND LIMITED EXPECTED VALUES

The introduction mentions that the mean and all moments of the "Log-Lomax" distribution are infinite. We can see this in two ways. First, we recall that the Pareto distribution only has a finite mean for shape parameter $\lambda > 1$. For this mixture, we have a certain probability of all λ's from zero to infinity. Since there is a positive probability of having $\lambda \leq 1$, then the overall mean of the mixture must be at least that probability times infinity, hence it is infinite. The argument applies for all higher moments since the k^{th} moment of a Pareto only exists for $\lambda > k$.

The other way to see (which is actually the same reason mathematically) is that the PDF decays like:

$$f(x) \sim \frac{1}{x \ln(x)^{\alpha+1}}$$

Obviously if we multiply by x^k for integer $k > 1$, then this quantity actually increases, giving an infinite integral. But even for $k = 1$, i.e. the mean, the quantity $\ln(x)^{\alpha+1}$ grows more slowly than *x* for any positive α, hence we have for large enough *R*:

$$\int_R^\infty \frac{1}{\ln(x)^{\alpha+1}} dx \geq \int_R^\infty \frac{1}{x} dx = \infty$$

4.1 Limited Expected Value

Even though the mean and moments are infinite, the limited expected values must obviously be finite. Especially if we are interested in reinsurance pricing applications, typical quantities of interest would be the expected loss, average severity, and standard deviation of losses in a layer. Those can all be derived from the moments of the limited expected value of *x*.

Again suppose:

$$F(x) = 1 - \left(1 + \frac{\ln\left(\frac{x}{\theta}\right)}{\beta}\right)^{-\alpha}$$

We then calculate the limited expected value:

$$E((x \wedge L)^k) = \int_\theta^\infty \min(x, L)^k * f(x)dx$$

$$= \int_\theta^L x^k \frac{\alpha\theta}{\beta x}\left(1 + \frac{\ln\left(\frac{x}{\theta}\right)}{\beta}\right)^{-(\alpha+1)} dx + L^k * S(L)$$

For the second term we use the CDF formula from above:

$$L^k * S(L) = L^k\left(1 + \frac{\ln\left(\frac{L}{\theta}\right)}{\beta}\right)^{-\alpha}$$

The first term can be expressed as:

$$\frac{\alpha\theta}{\beta}\int_\theta^L x^{k-1}\left(1 + \frac{\ln\left(\frac{x}{\theta}\right)}{\beta}\right)^{-(\alpha+1)} dx$$

$$= \frac{\alpha\theta}{\beta}\int_\theta^L x^{k-1}\beta^{\alpha+1}\left(\beta + \ln\left(\frac{x}{\theta}\right)\right)^{-(\alpha+1)} dx$$

$$= \alpha\theta\beta^\alpha\int_\theta^L x^{k-1}\left(\beta + \ln\left(\frac{x}{\theta}\right)\right)^{-(\alpha+1)} dx$$

To simplify the integral, we make the substitution:

$$u = k(\beta + ln(x/\theta))$$

This gives us:

$$du = \frac{k\theta}{x}dx \Rightarrow dx = \frac{x}{k\theta}du$$

$$x = \theta e^{u/k-\beta}$$

$$x = \theta \Rightarrow u = \theta e^{\theta/k-\beta}$$

$$x = L \Rightarrow u = \theta e^{L/k-\beta}$$

Substituting those into the integral we get:

$$\alpha\theta\beta^{\alpha}\int_{\theta e^{\theta/k-\beta}}^{\theta e^{L/k-\beta}} x^{k-1}(u/k)^{-(\alpha+1)}\frac{x}{k\theta}du$$

$$= \frac{\alpha\beta^{\alpha}}{k^{\alpha+2}}\int_{\theta e^{\theta/k-\beta}}^{\theta e^{L/k-\beta}} x^{k}(u)^{-(\alpha+1)}du$$

$$= \alpha\beta^{\alpha}\int_{\theta e^{\theta/k-\beta}}^{\theta e^{L/k-\beta}} (\theta e^{u/k-\beta})^{k}(u)^{-(\alpha+1)}du$$

$$= \alpha\beta^{\alpha}\theta^{k}e^{-k\beta}\int_{\theta e^{\theta/k-\beta}}^{\theta e^{L/k-\beta}} e^{u}u^{-(\alpha+1)}du$$

Unfortunately, the integral part of the expression above does not have a closed form. It can be approximated numerically, or as we show below it can be expressed as a difference of values in the Incomplete Gamma Function (whose values themselves are numerically approximated).

Plugging the results back into the original expression for the limited expected value we obtain:

$$E((x \wedge L)^{k}) = \alpha\beta^{\alpha}\theta^{k}e^{-k\beta}\int_{\theta e^{\theta/k-\beta}}^{\theta e^{L/k-\beta}} e^{u}u^{-(\alpha+1)}du + L^{k}\left(1+\frac{\ln\left(\frac{L}{\theta}\right)}{\beta}\right)^{-\alpha}$$

4.2 Incomplete Gamma Function

The unsolvable integral from above is of the form:

$$\int_a^b e^x x^{-c}\, dx$$

This looks very similar to the Gamma function. One could imagine substituting $u = -x$ and getting:

$$\int_{-b}^{-a} e^{-u}(-u)^{-c}\, du$$

Suspending disbelief momentarily, let's assume all the integrals and quantities involved exist and are finite. If the lower incomplete gamma function, defined by:

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

existed for negative values of *x*, then we could "simplify" the unsolvable integral above as:

$$\int_{-b}^{-a} e^{-u}(-u)^{-c}\, du = (-1)^{-c} \int_{-b}^{-a} e^{-u} u^{-c}\, du$$
$$= (-1)^{-c}(\Gamma(-c+1, -b) - \Gamma(-c+1, -a))$$

The two questions we need to ask are: is this true, and is this helpful? The problem with this being true is that we would need to know that $\Gamma(a, x)$ exists for negative *x*. Gautschi [3, pp. 3-4] references earlier works (that were unavailable to this author) on the incomplete gamma function which define:

$$\gamma^*(a, x) = \frac{x^{-a}}{\Gamma(a)} \gamma(a, x)$$

where:

$$\gamma(a, x) = \int_0^x e^{-t} t^{a-1} dt = \Gamma(a) - \Gamma(a, x)$$

and show that $\gamma^*$is real valued for real *a* and *x* and exists for *x*<0. It's well established that gamma functions can be extended to negative *a* values (excluding negative integers) using recurrence relations arising from integration by parts.

Putting these results together, we can see how the $(-1)^{-c}$ term of our expression will be cancelled out by the x^{-a} term in the definition of $\gamma^*(a,x)$ to show that the representation of the integral by the incomplete gamma function will exist and be real.

But is this useful? Unfortunately, no for a couple of reasons. First, not all tools, programs and packages return values for negative arguments of the incomplete gamma function. Sage (available through cocalc.org) and some online calculators do, but standard packages in R and Python do not. Secondly, the size of some of the terms becomes large enough that the accuracy of the program used becomes questionable. Even in a simple example tested, the $\alpha\beta^\alpha\theta^k e^{-k\beta}$ term became something on the order of 10^{70}, and this was after attempting to use multiple tools, some of which returned an overflow error. Should we really trust if a program tells us $1.234 \ldots * 10^{70} - 1.234 \ldots * 10^{70} = 42$?

However useful or not, the closed form formula for limited moments of the Pareto-Gamma mixture is given by the following:

$$E((x \wedge L)^k) = \alpha\beta^\alpha\theta^k e^{-k\beta}(-1)^{-(\alpha+1)}\left(\Gamma\left(-\alpha, -\theta e^{\frac{L}{k}-\beta}\right) - \Gamma\left(-\alpha, -\theta e^{\frac{\theta}{k}-\beta}\right)\right)$$

$$+L^k\left(1+\frac{\ln\left(\frac{L}{\theta}\right)}{\beta}\right)^{-\alpha}$$

Due to these constraints, the author recommends approximating quantities of interest by simulating random draws and taking a sample average. Take care however not to do this with quantities whose true value is infinite, e.g. the uncapped mean of the mixture.

5. CONCLUSION

The Pareto-Gamma mixture can arise in the context of Bayesian models for large losses. We have shown some properties of the distribution, and derived closed form formulas for some quantities of interest, although unfortunately some of them are not of practical use. Hopefully the availability of this information can aid and encourage the recognition of parameter risk in large loss modeling, or perhaps the discovery of a similar model matching empirical data, but with more tractable formulas for quantities of interest.

6. REFERENCES

[1] heropup, "compound of gamma and exponential distribution", StackExchange, https://math.stackexchange.com/questions/646852/compound-of-gamma-and-exponential-distribution/, 2014.
[2] Conjugate Prior. (January 14, 2021). In *Wikipedia*. https://en.wikipedia.org/wiki/Conjugate_prior
[3] Lomax Distribution. (May 21, 2020). In *Wikipedia*. https://en.wikipedia.org/wiki/Lomax_distribution
[4] Pareto Distribution. (January 13, 2021). In *Wikipedia*. https://en.wikipedia.org/wiki/Pareto_distribution
[5] Gautsche, Walter. "Incomplete Gamma Function Since Tricomi". In Tricomi's Ideas and Contemporary Applied Mathematics, Atti dei Convegni Lincei, n. 147, Accademia Nazionale dei Lincei, 1998. http://www.cs.purdue.edu/homes/wxg/Tricomi.ps

Abbreviations and notations

CDF, cumulative distribution function
PDF, probability distribution function

About the Author

Greg McNulty is Senior Vice President, Head of Americas Actuarial Pricing for SCOR Reinsurance and is based in Chicago. He has a BS in Mathematics from UCLA and an MS in Mathematics from the University of Michigan. He is a Fellow of the CAS. His prior works have been awarded the Ratemaking Prize and the Variance Prize.