The Single Parameter Pareto Revisited

Rajesh Sahasrabuddhe, FCAS, MAAA

November 19, 2016

Abstract

In one of his seminal works, Stephen W. Philbrick proposed an elegant solution to the complex problem of modeling claims amounts in excess layers.

1 Introduction

In one of his seminal works, Stephen W. Philbrick[1] proposed an elegant solution to the complex problem of modeling claims amounts in excess layers. That solution involved the use a Pareto Type I distribution for claims above an excess threshold.

This paper expands on that work in several ways by providing the following:

1. A consistent notational framework
2. A review of the various formulæ including supporting derivation in the Appendices
3. A discussion of parameter values less than or equal to 1
4. A summary of the formulæto calcualte policy frequency and seveirty. This section also included errata related to certain issues in Philbrick.
5. An approach to allow consideration of maximum probable loss using order statistics
6. Quantification of parameter uncertainty using Bayesian Markov chain Monte Carlo (MCMC) methods (expanding on the work of Meyers and Reichle & Yonkunis)
7. Frequency and severity trends
8. Dealing with claims less than the threshold (expanding on the work of Fackler)
9. Provides an errata to address certain issues in *Philbrick*

2 The *Single Parameter* Pareto

Philbrick's Single Parameter Pareto (SPP) is a special case of the Pareto Type I distribution which has the following cumulative distribution function:

$$F(x) = 1 - \left(\frac{k}{x}\right)^a \tag{1}$$

Many readers are often confused by the reference to the *single parameter*. After all, *Philbrick*, in Section III initially presents the Pareto with two parameters (Equation 1), k and a, and then later adds that claims should be "normalized" by dividing by the "selected lower bound"

This presentation leaves many readers not understanding how the lower bound, k, lost its parameter status. Philbrick explains that this is because:

> Although there may be situations where this value must be estimated, in virtually all insurance applications this value will be selected in advance. (Section III)

This explanation often leaves readers less than satisfied.

We offer the alternative that users of the SPP model should consider the process of normalizing the claims to be a transformation of the data rather than the application of a parameter. An analogous transformation occurs when we take logs. When we do that, we do not consider the base of the logarithm to be a parameter. Similarly, we will not consider the lower bound to be a parameter.

To improve clarity of this concept, we present the following:

Random Variable	Observed Claim Amount	Normalized Claims Amount
Symbol	$\mathbf{Y}$	$\mathbf{X}$
Transformation		$\mathbf{X} = g(\mathbf{Y})$ $g(\mathbf{Y}) = \mathbf{Y}/\text{lower bound}$
Domain	[lower bound, ∞]	[1, ∞]
Density	$a\dfrac{k^a}{y^{a+1}}$	$qx^{-(q+1)}$
Parameters	$k > 0$ (scale); $a > 0$ (shape)	$q > 0$ (shape)

Table 1: The Pareto Type I and the SPP

We can now work with model forms in the space of $\mathbf{X}$ and then use g^{-1} to transform back into the space of $\mathbf{Y}$. We can also now present the density and distribution functions.

$$f(x) = qx^{-(q+1)} \tag{2}$$

$$F(x) = 1 - x^{-q} \tag{3}$$

In Appendix B.1, we provide the derivation of the distribution function (Equation (3)).

2.1 Other Pareto Forms

In Appendix A, we provide an inventory of Pareto distributions including the SPP presented in Equations (2) and (3).

3 Actuarial Formulæ

In general, we leave formula derivation to the Appendices of this paper. In Appendix B.2, we present the derivation of the expected value:

$$\mathbb{E}[X] = \begin{cases} \dfrac{q}{q-1}, & \text{if } q > 1 \\ \text{undefined}, & \text{if } q \leq 1 \end{cases} \tag{4}$$

We can use Equation (14), to calculate the limited expected value through b as presented in Appendix B.3 as:[1]

$$\mathbb{E}[X; b] = \frac{q - b^{1-q}}{q-1} \tag{5}$$

In that derivation, there is no restriction that $q > 1$. That is, we recognize that, although the limited expected value is undefined, expected values are defined when we have an upper limit (such as a policy limit). As such, users of the SPP need not "fear" q values less than 1 for most insurance applications.

So for the expected claim amount for the layer between AP and L, we have:

$$\begin{aligned} \mathbb{E}[X; AP, L] &= \frac{q - L^{1-q}}{q-1} - \frac{q - AP^{1-q}}{q-1} \\ &= \frac{AP^{1-q} - L^{1-q}}{q-1} \end{aligned} \tag{6}$$

[1] Philbrick used b to refer to both the "lower bound" and the policy limit. We will not do that in this paper primarily for clarity as using a variable to represent the lower bound implied at least the possibility that the lower bound was a parameter. Conveniently, it also allows us to use the traditional policy notation as attaching at AP through limit L with the resulting layer width equal to $L - AP$.

This is, of course, likely the most common use of the SPP, estimating claims for an excess policy. This was also a focus of Section III of *Philbrick*.

3.1 Policy Claims Estimate

The purpose of the Philbrick calculation was likely demonstrate that the average claim size in the layer between AP and L was equal to the expected value of claims limited to L/AP net of the lower bound but multiplied by AP. The latter is calculated as Equation (5) -1 which simplifies to:

$$\frac{1-b^{1-q}}{q-1} \times AP \tag{7}$$

We can demonstrate that as follows using Equation (6) and the survival function as follows:

$$\begin{aligned}
\frac{\frac{AP^{1-q}-L^{1-q}}{q-1}}{S(AP)} &= \frac{\frac{AP^{1-q}-L^{1-q}}{q-1}}{AP^{-q}} \\
&= \frac{1}{q-1} \times \frac{AP}{AP} \times \frac{AP^{1-q}-L^{1-q}}{AP^{-q}} \\
&= \frac{AP}{q-1} \times \frac{AP^{1-q}-L^{1-q}}{AP^{1-q}} \\
&= \frac{AP}{q-1} \times \left(1-\left(\frac{L}{AP}\right)^{1-q}\right) \\
&= \frac{1-(L/AP)^{1-q}}{q-1} \times AP
\end{aligned} \tag{8}$$

As mentioned, the most common actuarial application of the SPP is to estimate the number of claims, their average value and the resulting aggregate claim amount to a policy. We therefore summarize those formulæ in the Table 3.

Number of Claims	$S(AP) = AP^{-q}$
Average Value of Individual Claims	$\frac{1-(L/AP)^{1-q}}{q-1} \times AP$
Aggregate Claim Amount	$\frac{AP^{1-q}-L^{1-q}}{q-1}$

Table 2: Policy Analysis

4 The Shortcomings of the SPP

In Section IV., Philbrick indicates that:

> ... but most actual data suggests that the tail of the Pareto is still somewhat too thick at extremely high loss amounts. In other words, the theoretical density at high loss amounts is larger than empirical experience tends to indicate. Rather than discard the Pareto, it is easier to postulate that the distribution is censored or truncated at some high, but finite, value. As we have seen earlier, any upper limitation (either censorship point or truncation point) will produce formulæ for the mean claim size that are finite for all possible values of q.

We noted in Section 2, that the analyst defines the domain of claims to which the SPP is fit. In *Philbrick*, the reference is to the domain is limited to the lower bound only. However, we may specify the domain using both lower and upper bounds. The upper bound may be defined using an estimate of the maximum probable loss for casualty exposures or property values for those exposures. To overcome this shortcoming, we create the "Truncated Single Parameter Pareto" (TSPP). We discuss this distribution is Section 5. Before leaving the SPP however, we provide an errata to the *Philbrick*.

4.1 Errata

I believe that that section includes one minor typographical error and a second calculation error. These are discussed below.

4.1.1 Philbrick Errata #1

In the application of formula (5), we should understand that there is a minor typographical error in *Philbrick*. The second paragraph following Equation (6) appears on Page 56 and includes the following:

$$\begin{gathered} b = 20 \times (500,000/25,000) \\ \text{which should be} \\ b = 20 = 500,000/25,000 \end{gathered}$$

4.1.2 Philbrick Errata #2

Starting at the bottom on Page 58 and extending to Page 59, Philbrick presents an example with a q parameter of 1.5 and expected claim count of 7 that results in the following (where $S(x)$ represents the survival function):

$$\begin{aligned}
F(4) &= 1 - 4^{-1.5} \\
F(4) &= 7/8 \\
S(4) &= 1 - F(4) = 1/8
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}[n] &= 7 \\
\mathbb{E}[n; x > 4] &= 7 \times S(4) = 7/8
\end{aligned} \tag{9}$$

(It is unfortunate that, in this example both $\mathbb{E}[n; x > 4]$ and $S(4)$ both equal 7/8.)

$$\begin{aligned}
\mathbb{E}[X] &= \frac{1.5}{1.5 - 1} \\
\mathbb{E}[X] &= 3
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}[X; 4] &= \frac{1.5 - 4^{1-1.5}}{1.5 - 1} \\
\mathbb{E}[X; 4] &= \frac{1.5 - 4^{-0.5}}{0.5} \\
\mathbb{E}[X; 4] &= \frac{1.5 - .5}{0.5} \\
\mathbb{E}[X; 4] &= 2
\end{aligned}$$

The average severity of claims in the layer is $(\mathbb{E}[X] - \mathbb{E}[X; 4])/S(4) = 8$. Using the frequency calculated in Equation (9), we estimate claims in the layer to be $8 \times 7/8 = 7$ which agrees with Philbrick's calculation.

The error occurs when the example is extended to calculate claims in the layer from $AP = 3$ to $L = 7.5$. Using the approach above, we have the following:

$$\begin{aligned}
\mathbb{E}[X; 3] &= 1.845299 \\
\mathbb{E}[X; 7.5] &= 2.269703 \\
F(3) &= 0.8075499 \\
<<<<<<< HEADS(3) &= 0.1924501.
\end{aligned}$$

The resulting the average claim amount in the layer is

$$\frac{\mathbb{E}[X; 7.5] - \mathbb{E}[X; 3]}{1 - F(3)} = 2.205267$$

which agrees with Philbrick's calculation of "net average claim size". However, the corresponding frequency should be $7 \times S(3) = 1.347151$ and a resulting expected claims in the layer of 2.970827. The purpose of the $F(2.5)$ term in the frequency calculation is not entirely clear to me.

$$S(3) = 0.1924501.$$

With that we have average claim amounts in the layer at

$$\frac{\mathbb{E}[X; 7.5] - \mathbb{E}[X; 3]}{1 - F(3)} = 2.205267$$

which agrees with Philbrick's calculation of "net average claim size" on Page 59. However, the corresponding frequency should be $7 \times S(3) = 1.347151$ and a resulting expected claims in the layer of 2.970827. The purpose of the $F(87,500/75,000) = F(2.5)$ term in the frequency calculation is not entirely clear to this author. ¿¿¿¿¿¿¿ origin/master

4.1.3 Philbrick Errata #3

Equation (11) indicates that "nth moment of the Pareto distribution with no upper limit is" $\dfrac{q}{q+n}$. Then, in Equation (12) the second moment is represented in the calculation of variance by $\dfrac{q}{q-n}$ and of course we have the calculation of mean as $\dfrac{q}{q-1}$. We can clearly see the error in Equation (11).

5 The Truncated Single Parameter Pareto

Using V to represent the upper bound, we have:

$$F(V) = 1 - V^{-q} \tag{10}$$

$$f(x) = \frac{qx^{-(q+1)}}{1 - V^{-q}} \tag{11}$$

$$F(x) = \frac{1 - x^{-q}}{1 - V^{-q}} \tag{12}$$

¡¡¡¡¡¡¡ HEAD The presentation above differs from the truncated SPP in *Philbrick*. In the latter, the SPP is an unlimited distribution the is truncated in application. That is, there is no change to the underlying density function. As noted in Equation 11, the presentation above is actually a new distribution with, importantly, a different maximum likelihood estimator for q.

======= As mentioned, the most common actuarial application of the SPP is to estimate the number of claims, their average value and the resulting aggregate claim amount to a policy. We therefore summarize those formulæ in the Table 3.

¿¿¿¿¿¿¿ origin/master

Number of Claims	$S(AP) = AP^{-q}$
Average Value of Individual Claims	$\frac{1-(L/AP)^{1-q}}{q-1} \times AP$
Aggregate Claim Amount	$\frac{AP^{1-q} - L^{1-q}}{q-1}$

Table 3: Policy Analysis

6 The Parameter

7 Trend

One of the more “controversial” issues in *Philbrick*relates to the application of trend. Specifically, Section V. states that “However, the parameter of the Pareto distribution in this paper is unchanged due to trend.”

Appendices

A An Inventory of Pareto Distributions

Appendix first content.

B Derivation of Forumulæ

B.1 SPP Cumulative Distribution Function

$$
\begin{aligned}
F(x) &= \int_1^x f(x)\,\mathrm{d}x \\
&= \int_1^x q x^{-(q+1)}\,\mathrm{d}x \\
&= q \int_1^x x^{-(q+1)}\,\mathrm{d}x \\
&= q \frac{1}{-(q+1)+1} x^{-q}\Big|_1^x \\
&= q \frac{1}{-q} x^{-q}\Big|_1^x \\
&= -x^{-q}\Big|_1^x \\
&= -x^{-q} - (-1^{-q}) \\
F(x) &= 1 - x^{-q} \qquad (13)
\end{aligned}
$$

B.2 SPP Expected Value

We can derive the expected values as:

$$
\begin{aligned}
\mathbb{E}[X] &= \int_1^{\infty} x f(x)\,\mathrm{d}x \\
&= \int_1^{\infty} x q x^{-(q+1)}\,\mathrm{d}x \\
&= q \int_1^{\infty} x^{-q}\,\mathrm{d}x \\
&= q \frac{1}{-q+1} x^{-q+1} \Big|_1^{\infty} \\
&= \frac{q}{1-q} x^{-q+1} \Big|_1^{\infty} \\
\mathbb{E}[X] &= \frac{q}{1-q} \frac{1}{x^{q-1}} \Big|_1^{\infty}
\end{aligned} \tag{14}
$$

We can see that for $x = 1$ (the lower limit of integration) equation (14) evaluates to $\frac{q}{1-q}$. However for $x = \infty$, we have the following[2]:

$$
\frac{q}{1-q} \frac{1}{x^{q-1}} = \begin{cases} 0, & \text{if } q > 1 \\ \text{undefined}, & \text{if } q = 1 \\ \infty, & \text{if } q < 1 \end{cases}
$$

and therefore we have:

$$
\mathbb{E}[X] = \begin{cases} 0 - \frac{q}{1-q}, & \text{if } q > 1 \\ \text{undefined}, & \text{if } q = 1 \\ \infty, & \text{if } q < 1 \end{cases}
$$

or more simply:

$$
\mathbb{E}[X] = \begin{cases} \frac{q}{q-1}, & \text{if } q > 1 \\ \text{undefined}, & \text{if } q \leq 1 \end{cases} \tag{15}
$$

[2] In the limit as $x \to \infty$, the expression evaluates to $-\frac{q}{q+1}$. However evaluated at ∞, the expression is undefined.

B.3 SPP Limited Expected Value

The limited expected value is calculated as:

$$
\begin{aligned}
\mathbb{E}[X;b] &= \frac{q}{1-q}\frac{1}{x^{q-1}}\Big|_1^b + b(1-F(b)) \\
&= \frac{q}{1-q}\left[\frac{1}{b^{q-1}} - \frac{1}{1^{q-1}}\right] + b\left[1-(1-b^{-q})\right] \\
&= \frac{q}{1-q}\left[\frac{1}{b^{q-1}} - 1\right] + b\left[b^{-q}\right] \\
&= \frac{q}{q-1}\left[1 - \frac{1}{b^{q-1}}\right] + b^{1-q} \\
&= \frac{q}{q-1}\left[1 - b^{1-q}\right] + b^{1-q} && (16) \\
&= \frac{1}{q-1}\left[q - qb^{1-q} + (q-1)b^{1-q}\right] \\
&= \frac{1}{q-1}\left[q - b^{1-q}\right] \\
&= \frac{q-b^{1-q}}{q-1} && (17)
\end{aligned}
$$

B.4 Maximum Likelihood Estimator for Parameter

The negative log-likelihood (NLL) function given data $D = x_1 \ldots x_n$ is defined as:

$$
\begin{aligned}
L(q) &= \prod_{i=1}^{n} f x_i \\
\text{NLL} &= -\sum_{i=1}^{n} \ln(f x_i) \\
\text{NLL} &= -\sum_{i=1}^{n} \ln(q x_i^{-(q+1)}) \\
\text{NLL} &= -\sum_{i=1}^{n} \left[\ln q + \ln x_i^{-(q+1)}\right] \\
\text{NLL} &= -\sum_{i=1}^{n} \left[\ln q - (q+1)\ln x_i\right] \\
\text{NLL} &= -n \ln q + \sum_{i=1}^{n} (q+1)\ln x_i \\
\text{NLL} &= -n \ln q + (q+1)\sum_{i=1}^{n} \ln x_i
\end{aligned}
$$

We can calculate the MLE of q by taking partial derivatives and setting equal to 0.

$$
\begin{aligned}
0 &= \frac{\partial}{\partial q}\left[-n \ln q + (q+1)\sum_{i=1}^{n} \ln x_i\right] \\
0 &= \frac{\partial}{\partial q}\left[\sum_{i=1}^{n} \ln(q) + \sum_{i=1}^{n} -(q+1)\ln x_i\right] \\
0 &= n - \frac{\partial}{\partial q}\sum_{i=1}^{n} q \ln x - \frac{\partial}{\partial q}\sum_{i=1}^{n} \ln x
\end{aligned}
$$

References

[1] Stephen W. Philbrick. A Practical Guide to Single Parameter Pareto Distribution. *Proceedings of the Casualty Actuarial Society*, LXXII:44–84, 1985.