# Classifying the Tails of Loss Distributions

Leigh J. Halliwell, FCAS, MAAA

**Abstract.** Of the several classifications which actuaries have proposed for the heaviness of loss-distribution tails, none has been generally accepted. Here we will show that the ultimate settlement rate, or asymptotic failure rate, provides a natural tripartite division into light, medium, and heavy tails. We prove that all the positive moments of light- and medium-tailed distributions are finite. Within the heavy-tailed distributions, we will define very heavy-tailed and super heavy-tailed, and we will explain how the power and exponential transformations are the basis for these subdivisions. An appendix relates extreme value theory to our findings.

**Keywords:** loss distribution, ultimate settlement rate, power transform, exponential transform, extreme value theory

## 1. INTRODUCTION

Many actuaries are as fascinated with the "heaviness" of loss-distribution tails as chemists are with heavy elements and as physicists are with heavy particles. However, unlike those scientists, with their periodic table and standard particle theory, "actuarial scientists" have no generally accepted standard of tail comparison. In this paper we will propose one that gives every indication of being natural, comprehensive, and insightful. To outline our progress, after briefly defining in Section 2 what constitutes a loss distribution, in Section 3 we will introduce the ultimate settlement rate and derive the settlement rates of several familiar distributions. Then in Section 4 we will show how the most basic transformation, a change of a distribution's scale parameter, provides the basis for a division of loss distributions into light-, medium-, and heavy-tailed. An immediate benefit from this is a proof in Section 5 that all the positive moments of light- and medium-tailed distributions are finite. Infinite moments are a sufficient, but not a necessary condition, for being heavy-tailed. Section 6 takes up the next logical transformation, the power transformation, and will show its effect on the tail class of a distribution. A symmetric "multiplication" table there, showing the medium-tailed distribution to be like an identity element among distributions, will be crucial to the following sections. Section 7 contains an abstract examination into the results so far, finishing with a diagram that will make memorable the classification schema. In Section 8 we will treat the next logical transformation, the exponential, which is the key to loss-tail heaviness. Then in Section 9 we will treat the moments of exponentially transformed random variables, vindicating the power of this classification by the results. Section 10 is a brief treatment of two other transformations, inverting and mixing. Finally, before concluding, in Section 11 we will show that the classification is indefinitely expandable, encompassing ever more distant realms of heavy and light tails. An appendix will fit extreme value theory into the classification schema.

## 2. LOSS DISTRIBUTIONS DEFINED

For the purposes of this paper $X$ is a loss distribution[1] if its survival function $S_X(x) = Prob[X > x]$ has the following properties:

(i) $S_X(0) = 1$

(ii) $S_X(x) > 0$

(iii) For all $x_1 < x_2$, $S_X(x_1) \geq S_X(x_2)$. And there exists some ξ such that for all $\xi < x_1 < x_2$, $S_X(x_1) > S_X(x_2)$

(iv) $\lim_{x \to \infty} S_X(x) = 0$

(v) For all $x$ greater than some ξ, $S''_X(x) > 0$

Although these properties are standard, some commentary will be helpful. Property (i) implies that $X$ must be positive; in particular, there is no probability mass at zero. So this definition disqualifies the Tweedie distribution (Meyers [5]; cf. Footnote 15). The property provides for $1/X$ to be a loss distribution, which we deem desirable.[2] Property (ii) requires the tail of a loss distribution to be infinite. We are not interested in classifying tails of finite distributions; they might as well be "no-tailed." Property (iii) requires for the survival function never to increase, and beyond some point for it strictly to decrease. Property (iv) precludes any probability that $X$ might be infinite. Though we will often encounter limits to infinity, infinity is not a real number. This property is allied with the first, for if somehow $Prob[X = \infty] > 0$, then the inverse $1/X$ would have a probability mass at zero. And property (v) demands beyond some point for the survival function to be concave upward. Of course, for the second derivative to exist the first derivative must also exist. By implication, the left and right derivatives must be equal. This property ensures that at some point the survival function "settles down." Thereafter there will be no more discrete jumps or probability masses, no more vertices or corners, and no more undulations or inflections. Having defined a loss distribution, we name the set of all loss distributions Ξ.

## 3. THE ULTIMATE SETTLEMENT RATE

[1] More accurately, $X$ is a "loss random variable," whose probability obeys a "loss distribution." But since "loss random variable" sounds odd, we will use 'random variable' and 'distribution' interchangeably.

[2] We also desire $E[X^0]$ to equal unity, which would fail if there were any probability of the indeterminate $0^0$. Cf. Section 5, esp. Footnote 6.

Of the several ways described by Klugman [4, pp. 86-92] and Corro [1] by which to compare the tails of loss distributions, we believe the best to be the "asymptotic failure rate" [4, p. 87] or "ultimate settlement rate" [1, p. 451]. Since $S_X(x) = Prob[X > x]$ is the probability for *X* to "survive" at least until *x*, we may think of *X* as being subject to a force of mortality $\lambda_X(x) = -\frac{d\,ln\,S_X(x)}{dx} = -\frac{1}{S_X(x)}\frac{dS_X(x)}{dx} = \frac{f_X(x)}{S_X(x)}$. Corro's ultimate settlement rate is $\tau_X = \lim_{x\to\infty}\lambda_X(x)$. Just as an account compounding at a higher interest rate will eventually overtake an account compounding at a lower rate, regardless of their current positive balances, so too if $\tau_X < \tau_Y$, $\frac{S_X(x)}{S_Y(x)}$ will grow infinitely large with *x*, i.e., $\lim_{x\to\infty}\frac{S_X(x)}{S_Y(x)} = \infty$ or $\lim_{x\to\infty}\frac{S_Y(x)}{S_X(x)} = 0$.[3] But, following Klugman [4, p. 88], from L'Hôpital's rule we may express $\tau_X$ in terms of the probability density function:

$$\begin{aligned}
\tau_X &= \lim_{x\to\infty}\lambda_X(x) \\
&= \lim_{x\to\infty}\frac{f_X(x)}{S_X(x)} \\
&= \lim_{x\to\infty}\frac{f_X(x)}{\int\limits_{u=x}^{\infty} f_X(u)} \qquad \text{a } \frac{0}{0} \text{ form} \\
&= \lim_{x\to\infty}\frac{f_X'(x)}{(-f_X(x))} \\
&= -\lim_{x\to\infty}\frac{d\,ln\,f_X(x)}{dx}
\end{aligned}$$

This does not mean that $\lambda_X(x) = -\frac{d\,ln\,f_X(x)}{dx}$; it is only true in the limit as $x \to \infty$. Of course, for $0 \le \xi \le x$, where ξ is the "settling down" point required by property (v):

$$S_X(x) = S_X(\xi)e^{-\int\limits_{u=\xi}^{x}\lambda_X(u)du}$$

Let us look at the settlement rates of some well known distributions. If $X \sim Gamma(\alpha,\theta)$, or equivalently $X/\theta \sim Gamma(\alpha,1)$, then $f_X(x) = \frac{1}{\Gamma(\alpha)}e^{-\frac{x}{\theta}}\left(\frac{x}{\theta}\right)^{\alpha-1}\frac{1}{\theta}$, for positive α. If $-\alpha < k$, $E[X^k] = \theta^k\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}$. Now:

---

[3] This implies an infinite right tail for *X* (property ii). Although in order for $S_X(x)$ to reach zero the force of mortality $\lambda_X(x)$ must become infinite, once $S_X(x)$ "flatlines" at zero, $\lambda_X(x) = 0/0$. It is meaningless to speak of the growth (or mortality) rate of something whose quantity is zero; a zero balance in a bank account remains zero at any interest rate. Property (v) guarantees the existence of $\lambda_X(x)$ far enough out, as well as for $\lim_{x\to\infty}\lambda_X(x)$ either to converge to a non-negative real number or to diverge to positive infinity.

$$\tau_{Gamma(\alpha,\theta)} = -\lim_{x\to\infty}\frac{d\,ln\,f_X(x)}{dx} = -\lim_{x\to\infty}\frac{d\left(-\frac{x}{\theta}+(\alpha-1)ln\,x\right)}{dx} = -\lim_{x\to\infty}\left(-\frac{1}{\theta}+\frac{\alpha-1}{x}\right) = \frac{1}{\theta}$$

The ultimate settlement rate of a Gamma-distributed random variable depends only on its scale parameter θ. But the force of mortality of the exponential random variable is $\lambda_{Gamma(1,\theta)}(x)=\frac{1}{\theta}$. For this reason it is legitimate to say that *far enough out in the tail, every gamma distribution looks like an exponential distribution*. Compare this with the right tail of a normal distribution:[4]

$$\tau_{N(\mu,\sigma^2)} = -\lim_{x\to\infty}\frac{d\,ln\,f_X(x)}{dx} = -\lim_{x\to\infty}\frac{d\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right)}{dx} = \lim_{x\to\infty}\left(\frac{x-\mu}{\sigma^2}\right) = \infty$$

So the right tail of the normal distribution is "lighter" than that of the gamma.

For the inverse-gamma random variable, $(X/\theta)^{-1} \sim Gamma(\beta,1)$, or $X/\theta \sim 1/Gamma(\beta,1)$, where $\beta > 0$. Its density function is $f_X(x)=\frac{1}{\Gamma(\beta)}e^{-\frac{\theta}{x}}\left(\frac{\theta}{x}\right)^{\beta+1}\frac{1}{\theta}$, and $E\left[X^k\right]=\theta^k\frac{\Gamma(\beta-k)}{\Gamma(\beta)}$, for $k<\beta$. As for its ultimate settlement rate:

$$\tau_{InvGamma(\beta,\theta)} = -\lim_{x\to\infty}\frac{d\,ln\,f_X(x)}{dx} = -\lim_{x\to\infty}\frac{d\left(-\frac{\theta}{x}-(\beta+1)ln\,x\right)}{dx} = -\lim_{x\to\infty}\left(\frac{\theta}{x^2}-\frac{\beta+1}{x}\right) = 0$$

So the inverse-gamma is "heavier-tailed" than the gamma distribution, since $0<1/\theta$. One might have surmised this from the non-existence of its positive moments greater than or equal to β.

If $X$ is a lognormal random variable, then $ln\,X \sim N(\mu,\sigma^2)$ and $f_X(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{1}{2}\frac{(ln\,x-\mu)^2}{\sigma^2}}\frac{1}{x}$. For all real *k*, $E\left[X^k\right]=e^{k\mu+k^2\sigma^2/2}$. All the moments of the lognormal random variable exist, even the negative ones. Its ultimate settlement rate is:

$$\tau_{LogNorm(\mu,\sigma^2)} = -\lim_{x\to\infty}\frac{d\,ln\,f_X(x)}{dx} = -\lim_{x\to\infty}\frac{d\left(-\frac{1}{2}\frac{(ln\,x-\mu)^2}{\sigma^2}-ln\,x\right)}{dx} = \lim_{x\to\infty}\left(\frac{ln\,x-\mu}{\sigma^2}\cdot\frac{1}{x}+\frac{1}{x}\right) = 0$$

Hence, the lognormal distribution is heavier-tailed than the gamma. Although its settlement rate equals the inverse-gamma's, the existence of all its moments implies that it is not as heavy-tailed as

[4] The normal distribution with its infinite left tail is not a loss distribution. But we may still calculate the ultimate settlement rate of its right tail. Alternatively, we could also consider the right tail of the absolute value of the standard normal distribution (i.e., $X/\theta \sim |N(0,1)|$) and arrive at the same result (cf. Footnote 10).

the inverse gamma.

Last, let $X$ be a generalized-Pareto random variable. In our parameterization this will mean that $\frac{X}{\theta} \sim \frac{Gamma(\alpha,1)}{Gamma(\beta,1)}$, where the two gamma random variables are independent. The distribution function for $X$ is $f_X(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\left(\frac{x}{x+\theta}\right)^{\alpha-1}\left(\frac{\theta}{x+\theta}\right)^{\beta-1}\frac{\theta}{(x+\theta)^2}$, and its moments are $E\left[X^k\right] = \theta^k \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)}\frac{\Gamma(\beta-k)}{\Gamma(\beta)}$, for $-\alpha < k < \beta = (-\alpha < k) \cap (k < \beta)$. Division by the $Gamma(\beta,1)$, or multiplication by the $InvGamma(\beta,1)$ random variable, places a positive limit on *k*. Its ultimate settlement rate is:

$$\tau_{GenPareto(\alpha,\beta,\theta)} = -\lim_{x\to\infty}\frac{d\,ln\,f_X(x)}{dx} = -\lim_{x\to\infty}\frac{d((\alpha-1)ln\,x-(\alpha+\beta)ln(x+\theta))}{dx} = \lim_{x\to\infty}\left(\frac{\alpha-1}{x}-\frac{\alpha+\beta}{x+\theta}\right) = 0$$

Again, this is the same rate as the inverse-gamma's and the lognormal's. But the domain of its positive moments makes its tail like the inverse-gamma's. The non-existence of negative moments is relevant only to the tail of the inverse distribution.

To conclude this section, the tail of the normal distribution is lighter than the tail of the gamma distribution, which is lighter than the tails of the lognormal, inverse-gamma, and generalized-Pareto distributions, even as $\infty > \theta > 0$. The non-existence, or infinitude, of positive moments hints at secondary orderings within the last three distributions.

## 4. THE ULTIMATE SETTLEMENT RATE UNDER A SCALE TRANSFORMATION

In the previous section we saw that the ultimate settlement rate of a gamma random variable is the inverse of its scale parameter. Here we will generalize, and form the basis for classifying the (right) tails of loss distributions.

If $X$ is a random variable and $\theta$ a positive constant, the scale transformation of $X$ is the random variable $Y/\theta = X$ . Accordingly:

$$S_Y(u) = Prob[Y > u] = Prob[Y/\theta > u/\theta] = Prob[X > u/\theta] = S_X(u/\theta)$$

Hence:

$$\tau_Y = -\lim_{u\to\infty}\frac{d\,ln\,S_Y(u)}{du} = -\lim_{u\to\infty}\frac{d\,ln\,S_X(u/\theta)}{du} = -\lim_{u/\theta\to\infty}\frac{d\,ln\,S_X(u/\theta)}{d(u/\theta)}\frac{1}{\theta} = \frac{\tau_X}{\theta}$$

A scale transformation should not be the basis for tail class; in fact, most loss distributions are parameterized to include one "scale" parameter along with one or more "shape" parameters.[5] Because $0/\theta = 0$, $+/\theta = +$, and $\infty/\theta = \infty$, there are only three essential values for ultimate settlement rates, zero (0), positive (+), and infinity (∞), with the ordering, $0 < + < \infty$. Since smaller $\tau_X$ means heavier tail, we will classify a loss distribution as light-, medium, or heavy-tailed according as $\tau_X$ is ∞, +, or 0. The meaning of the symbols in the partition $\Xi = \Xi_0 \cup \Xi_+ \cup \Xi_\infty$ should be obvious. This is the gist of our classification; the rest of the paper merely draws out its implications.

[5] The scale parameter bears the unit of the random variable $Y$, so $Y/\theta$ is unitless, or a pure number. Most accurate is to divide each random variable by a parameter and to relate them by a unitless factor (called a "scalar"), as in $Y/\theta_1 = \eta(X/\theta_2)$. This is a safeguard in the physical sciences, but here it would be stilted. Cf. Footnote 13.

## 5. POSITIVE MOMENTS AND THE ULTIMATE SETTLEMENT RATE

In Section 3 we found loss distributions whose ultimate settlement rates spanned the range of zero, positive, and infinity. The normal distribution is light-tailed; the gamma distribution is medium-tailed; and the inverse-gamma, lognormal, and generalized-Pareto distributions are heavy-tailed. We also noted that some of the moments of the inverse-gamma and generalized-Pareto random variables were infinite. In this section we will prove that all the positive moments of light- and medium-tailed random variables are finite. But before that we will prove a partitioning lemma about non-negative moments, viz., that if $E[X^l]$ is finite for $0 < l$, then $E[X^k]$ is finite for $0 < k < l$.

If $X > 0$, then $X^0 = 1$.[6] Because according to property (i) $Prob[X > 0] = S_X(0) = 1$, $Prob[X^0 = 1] = 1$ and so $E[X^0] = E[1] = 1$. In words, the zeroth moment of a loss distribution exists and equals unity. Next consider $E[X^k]$ for $k > 0$. Because over this range of integration $x^k$ is non-negative:

$$E[X^k] = \int_{x=0}^{\infty} x^k dF_X(x) = \int_{x=0}^{1} x^k dF_X(x) + \int_{x=1}^{\infty} x^k dF_X(x) \le 1 + \int_{x=1}^{\infty} x^k dF_X(x)$$

So whether or not $E[X^k]$ is finite depends on $\int_{x=1}^{\infty} x^k dF_X(x)$. But for $x \ge 1$ and $k < l$, $1 \le x^k \le x^l$. So, if $\int_{x=1}^{\infty} x^l dF_X(x)$ converges, then so to does $\int_{x=1}^{\infty} x^k dF_X(x)$. Likewise, if $\int_{x=1}^{\infty} x^k dF_X(x)$ diverges, so too does $\int_{x=1}^{\infty} x^l dF_X(x)$. Therefore, for $0 < k < l$, if $E[X^l]$ is finite, so too is $E[X^k]$. And if $E[X^k]$ is infinite, so too is $E[X^l]$. The existence or non-existence of moments partitions the non-negative real numbers into two subsets. The lower partition is not empty, since it includes zero. The upper partition is empty when all the positive moments converge.

To return to the theorem of this section, let the distribution of $X$ be light- or medium-tailed. So $\tau_X = \lim_{x\to\infty} \lambda_X(x) > 0$. And let $\rho = \tau_X/2$, if $\tau_X$ is finite; let $\rho = 1$, if it is infinite. In either case, $\rho > 0$ and there exists a $\xi > 0$ such that for all $x \ge \xi$, $\rho < \lambda_X(x)$. So for all $x \ge \xi$:

$$S_X(x) = S_X(\xi) e^{-\int_{u=\xi}^{x} \lambda_X(u)du} \le S_X(\xi) e^{-\int_{u=\xi}^{x} \rho du} = S_X(\xi) e^{-\rho(x-\xi)}$$

---

[6] The form $0^0$ is undefined, even as $0^0 = e^{ln 0 \cdot 0} = e^{-\infty \cdot 0}$. Corro's "convention" that $0^0 = 1$ [1, p. 453] is equivalent to the convention that $\infty \cdot 0 = 0/0 = 0$. This convention is wired into the arithmetic of some programming languages (e.g., APL and J. R is inconsistent: $0^0 = 1$, but $0/0$ is undefined). However, such conventions should not be placed on undefined, or indeterminate, forms, since in limiting cases they may assume different values.

By integration by parts one can show that $E[h(X)] = h(0) + \int_{x=0}^{\infty} S_X(x)dh(x)$.[7] So, for positive *k*:

$$E[X^k] = 0^k + \int_{x=0}^{\infty} S_X(x)dx^k = \int_{x=0}^{\infty} S_X(x)dx^k = \int_{x=0}^{\xi} S_X(x)dx^k + \int_{x=\xi}^{\infty} S_X(x)dx^k$$

$$\leq \int_{x=0}^{\xi} 1 \cdot dx^k + \int_{x=\xi}^{\infty} S_X(\xi)e^{-\rho(x-\xi)}dx^k$$

Finally, we simplify the inequality:

$$\begin{aligned}
E[X^k] &\leq \int_{x=0}^{\xi} 1 \cdot dx^k + \int_{x=\xi}^{\infty} S_X(\xi)e^{-\rho(x-\xi)}dx^k \\
&= \xi^k + kS_X(\xi)e^{\rho\xi} \int_{x=\xi}^{\infty} e^{-\rho x}(\rho x)^{k-1} d(\rho x) \cdot \rho^{-k} \\
&\leq \xi^k + kS_X(\xi)e^{\rho\xi} \int_{\rho x=0}^{\infty} e^{-\rho x}(\rho x)^{k-1} d(\rho x) \cdot \rho^{-k} \\
&= \xi^k + kS_X(\xi)e^{\rho\xi}\, \Gamma(k)/\rho^k
\end{aligned}$$

Thus we prove that $E[X^k]$ is not infinite. As a result, we know that all the positive moments of light- and medium-tailed distributions exist.

This converse ("Not all the positive moments of heavy-tailed distributions exist.") is not true, for the lognormal is heavy-tailed, yet all its moments exist. But a random variable that lacks even one positive moment is heavy-tailed. This suggests a subclass of the heavy-tailed distributions $\Xi_0$. Those lacking in positive moments are heavier-tailed than those not lacking. And the heaviest of many heavier-tailed distributions is the one with fewest positive moments (or the one with the most infinite moments).[8] But the next section will provide a better subclassification.

## 6. THE ULTIMATE SETTLEMENT RATE UNDER A POWER TRANSFORMATION

In Section 4 we found the tail classification of a random variable to be invariant to a scale transformation; more accurately, we devised that classification for it to be invariant. But just as

[7] For details cf. Halliwell [3, Appendix A]

[8] Are there distributions so heavy-tailed that they have no positive moments? In a Section 9 we will prove that there are such distributions. However, it seems that their worth is purely theoretical.

Klugman [4, pp. 92-93] advances from scale transformations to power transformations, so too will we in this section.

Our form of the power transformation is $\dfrac{Y}{\theta} = X^{\gamma}$, for positive γ.[9] The equation $E\left[Y^k\right] = \theta^k E\left[\left(\dfrac{Y}{\theta}\right)^k\right] = \theta^k E\left[\left(X^{\gamma}\right)^k\right] = \theta^k E\left[X^{k\gamma}\right]$ puts the moments of $X$ and $Y$ into a one-to-one correspondence. Thus, distributions with infinite positive moments remain heavy-tailed under a power transformation. But how other distributions power-transform requires the following analysis.

Since $x^{\gamma}$ strictly increases:

$$S_Y(u) = Prob[Y > u] = Prob\left[X^{\gamma} > u/\theta\right] = Prob\left[X > (u/\theta)^{1/\gamma}\right] = S_X\left((u/\theta)^{1/\gamma}\right)$$

Therefore:

$$\begin{aligned}
\tau_Y &= -\lim_{u\to\infty} \frac{d\, ln\, S_Y(u)}{du} \\
&= -\lim_{u\to\infty} \left\{ \frac{d\, ln\, S_X\left((u/\theta)^{1/\gamma}\right)}{d(u/\theta)^{1/\gamma}} \frac{d(u/\theta)^{1/\gamma}}{du} \right\} \\
&= \lim_{\substack{u\to\infty \\ v(u)\to\infty}} \left\{ -\frac{d\, ln\, S_X(v)}{dv} \cdot \frac{1}{\gamma} (u/\theta)^{1/\gamma - 1} \Big/ \theta \right\} \\
&= \frac{1}{\gamma\theta} \lim_{v\to\infty} \left\{ \lambda_X(v) \cdot \left(v^{\gamma}\right)^{1/\gamma - 1} \right\} \\
&= \frac{1}{\gamma\theta} \lim_{v\to\infty} \left\{ \lambda_X(v) \cdot v^{1-\gamma} \right\}
\end{aligned}$$

Now $\lim_{v\to\infty} \lambda_X(v) = \tau_X$. And $\lim_{v\to\infty} v^{1-\gamma} = \begin{cases} \infty & 0 < \gamma < 1 \\ 1 & \gamma = 1 \\ 0 & \gamma > 1 \end{cases}$. By the product rule we can express $\tau_Y$ in the following three-valued multiplication table (so $1/\gamma\theta$ may be ignored):

[9] Klugman [4, p. 93] uses 'τ' for the exponent ($Y = X^{1/\tau}$); we use 'γ' to avoid confusion with Corro's ultimate settlement rate τ. We also invert the exponent, because we believe it easier to see that $\gamma < 1$ thins the tail (taking a root) and $\gamma > 1$ thickens it (raising to a power).

| $\tau_Y \quad \left(\dfrac{Y}{\theta} = X^{\gamma}\right)$ | *Thinner* $0 < \gamma < 1$ | $\gamma = 1$ | *Thicker* $\gamma > 1$ |
|---|---|---|---|
| *Light* : $\tau_X = \infty$ | $\infty$ | $\infty$ | $\infty \cdot 0$ |
| *Medium* : $\tau_X = +$ | $\infty$ | $+$ | $0$ |
| *Heavy* : $\tau_X = 0$ | $0 \cdot \infty$ | $0$ | $0$ |

Most obvious is the sensitivity of the medium-tailed random variable: the slightest exponent $\gamma = 1 \pm \varepsilon$ knocks it off the medium ridge into light or heavy valleys, from which the inverse exponent can restore it. For example, if $X$ is medium-tailed, then $Y = X^2$ is heavy-tailed. And if $Y = X^{0.5}$, then $Y$ is light-tailed.[10] So by power transformation, a medium-tailed distribution can become either heavy or light. But because $\left(Y^{\gamma}\right)^{1/\gamma} = Y = \left(Y^{1/\gamma}\right)^{\gamma}$, power transformations are invertible. By repeated transformations and inversions, one can cycle a medium-tailed distribution through all three types; e.g., $X \to \sqrt{X} \to \left(\sqrt{X}\right)^4 = X^2 \to \sqrt{X^2} = X$ is a three-stop roundtrip from medium to light to heavy and back to medium.

## 7. SET-THEORETIC PRESENTATION AND DIAGRAM OF RESULTS SO FAR

Define $PT[X; \gamma, \theta]$, the power transformation of random variable $X$ with positive parameters $\gamma$ and $\theta$, as the distribution $Y$ such that $\dfrac{Y}{\theta} = X^{\gamma}$. Therefore, $PT[X; \gamma, \theta] = \theta X^{\gamma}$. A compound power transformation reduces to a simple one:

$$\begin{aligned}
PT\left[PT\left[X; \gamma_1, \theta_1\right]; \gamma_2, \theta_2\right] &= \theta_2 \left(PT\left[X; \gamma_1, \theta_1\right]\right)^{\gamma_2} \\
&= \theta_2 \left(\theta_1 X^{\gamma_1}\right)^{\gamma_2} \\
&= \theta_2 \theta_1^{\gamma_2} X^{\gamma_1 \gamma_2} \\
&= PT\left[X; \gamma = \gamma_1 \gamma_2, \theta = \theta_2 \theta_1^{\gamma_2}\right]
\end{aligned}$$

Because the four original parameters are positive, the two reduced parameters are defined and

[10] Consider the transformation $Y = X^{0.5}$, where $X^{0.5} \sim Gamma(½, 2)$. Because the gamma distribution is medium-tailed and $\gamma = ½$, $Y$ is light-tailed: $medium \times (\gamma < 1) = \infty = light$. Moreover:

$$f_Y(x) = \frac{1}{\Gamma(½)} e^{-\frac{x^2}{2}} \left(\frac{x^2}{2}\right)^{½-1} \frac{1}{2} \frac{dx^2}{dx} = \frac{1}{\sqrt{\pi}} e^{-\frac{x^2}{2}} \frac{\sqrt{2}}{x} x = \frac{2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} = f_{|N(0,1)|}(x)$$

Hence, $Y \sim |N(0,1)|$; the negative support of the standard normal distribution has been reflected onto the positive support. Thus, the right tail of the normal distribution is light, in confirmation of what we derived in Section 3.

positive. By repetition, an *n*-step power transformation is always equivalent to a direct one. Due to the asymmetry of the scale formula, power transformation is not commutative. But since $\theta_3\left(\theta_2\theta_1^{\gamma_2}\right)^{\gamma_3} = \left(\theta_3\theta_2^{\gamma_3}\right)\theta_1^{\gamma_2\gamma_3}$, it is associative. And power transformation can always be inverted: $PT\left[PT\left[X;\gamma,\theta\right];1/\gamma,\theta^{-1/\gamma}\right] = PT\left[X;1,1\right] = X$.

So if $X$ can power-transform into $Y$, it can do so in one step. And from $Y$ it can return to $X$ in one step. Therefore the range within which $X$ can power-transform is a closed network.[11] The *power-transformation network* of $X$ is the set $ptn(X) = \{Y \in \Xi : \exists\gamma,\theta > 0 : Y = PT[X;\gamma,\theta]\}$. And the *power-transformation range* of $\Phi \subseteq \Xi$, where $\Phi$ is a set of random variables (or of their distributions), can be defined as $PTR(\Phi) = \{Y \in \Xi : \exists X \in \Phi, \exists\gamma,\theta > 0 : Y = PT[X;\gamma,\theta]\}$ or as $PTR(\Phi) = \bigcup\{ptn(X) : X \in \Phi\}$. Unlike $ptn(X)$, there is no guarantee of a power-transformation connection between any two elements of $PTR(\Phi)$. Because 'network' connotes interconnectedness, we changed the noun here to 'range'.

Obviously, $PTR(\varnothing) = \varnothing$ and $PTR(\Xi) = \Xi$. But of interest here is $PTR(\Xi_+)$, the set of all distributions that can be formed by power-transforming medium-tailed distributions. Above we saw that the power transformation "knocks distributions off the medium ridge into light or heavy valleys." Therefore, this set is larger than $\Xi_+$, i.e., $\Xi_+ \subset PTR(\Xi_+)$. It spills into $\Xi_0$ and $\Xi_\infty$, or in symbols $\Xi_0 \cap PTR(\Xi_+) \neq \varnothing$ and $\Xi_\infty \cap PTR(\Xi_+) \neq \varnothing$. But there are distributions in $\Xi_0$ and $\Xi_\infty$ that are unattainable from $\Xi_+$ by power transformation. We found above that power transformation cannot unseat distributions that are so heavy as to have infinite moments; hence, a trip to them from $\Xi_+$ to them is precluded. But even the lognormal, whose moments are all finite, power-transforms back to lognormal.

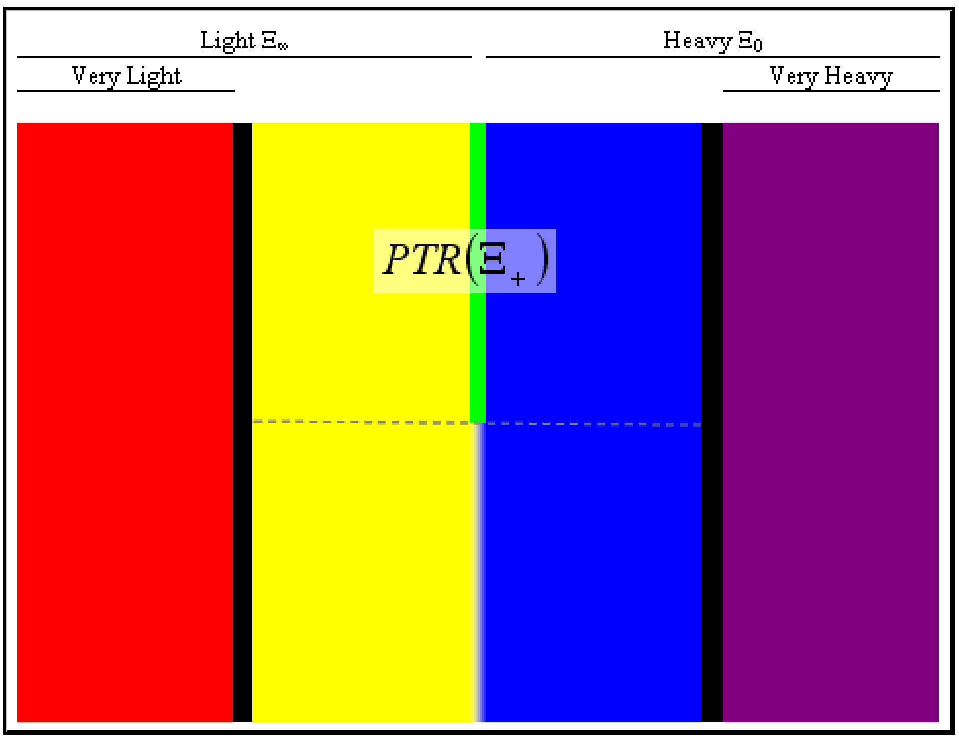
On the other hand, the symmetry of the multiplication table hints that some light-tailed distributions might be too light to power-transform elsewhere. Indeed, the survival function of one such distribution is $S_Q(x) = e^{-\left(e^x - 1\right)}$. It is light-tailed, since $\tau_Q = -\lim_{x\to\infty}\dfrac{d\,ln\,S_Q(x)}{dx} = \lim_{x\to\infty}\dfrac{d\left(e^x - 1\right)}{dx} = \infty$. But if $\dfrac{Y}{\theta} = Q^\gamma$, then:

---

[11] Technically, it is an algebraic group, whose set $G$ is $\{\langle\gamma,\theta\rangle \in \Re^+ \times \Re^+\}$ and whose function $f : G \times G \to G$ is $f\left(\langle\gamma_1,\theta_1\rangle,\langle\gamma_2,\theta_2\rangle\right) = \left\langle\gamma_1\gamma_2,\theta_2\theta_1^{\gamma_2}\right\rangle$. The function is associative; $\langle 1,1\rangle$ is the identity element; and every element has an inverse: $f\left(\langle\gamma,\theta\rangle,\left\langle 1/\gamma,\theta^{-1/\gamma}\right\rangle\right) = f\left(\left\langle 1/\gamma,\theta^{-1/\gamma}\right\rangle,\langle\gamma,\theta\rangle\right) = \langle 1,1\rangle$.

$$\tau_Y = \frac{\gamma}{\theta}\lim_{v\to\infty}\left\{-\frac{d\,ln\,S_Q(v)}{dv}\cdot v^{\frac{\gamma-1}{\gamma}}\right\} = \frac{\gamma}{\theta}\lim_{v\to\infty}\left\{\frac{d\left(e^v-1\right)}{dv}\cdot v^{\frac{\gamma-1}{\gamma}}\right\} = \frac{\gamma}{\theta}\lim_{v\to\infty}\left\{e^v\cdot v^{\frac{\gamma-1}{\gamma}}\right\} = \infty\,.$$

Hence, this distribution remains light under power transformation. So within the $0\cdot\infty$ and $\infty\cdot 0$ cells of the table are distributions so heavy-tailed and so light-tailed as to be unmoved by power transformation. So these cannot belong to $PTR(\Xi_+)$. Because of this duality, we deem the power transformation to be a better basis for subclassification than the divergence of positive moments. Tail-class immutability to power transformation merits the adverb 'very'. Thus we will now speak of "very light-tailed" and "very heavy-tailed" distributions and random variables. The lognormal is very heavy-tailed, though not as heavy-tailed as something with missing moments. Quite appropriately, nothing is "very" medium-tailed; medium is just medium. The following diagram will conclude this section:




The diagram, which looks like a painted tennis court with half a net, represents a tripartite form of Ξ. The black regions are boundaries; Ξ is the union of the colored regions. The middle partition is the set of all loss distributions whose tail classes change under power transformation. All the medium-tailed distributions, the green area, must belong to this set. The red and violet regions contain all the loss distributions whose tail classes do not change. These unchangeable distributions are either light-tailed and in the red region or heavy-tailed and in the violet region. The yellow region contains the changeable light-tailed distributions, the blue the changeable heavy-tailed. By definition, power transformation cannot "jump" from the red or violet regions. But if perchance, it could jump from the middle, it could not jump back. So since power transformation is reversible, the black regions are barriers to power transformation.

Now consider all the changeable distributions as organized into horizontal slices of power-transformation networks. Whatever might be the position of distribution $X$ in its network, $Y = X^{\gamma}$ transforms $X$ to the right if $\gamma > 1$ and to the left if $\gamma < 1$. The movement approaches the black boundaries as $\gamma$ approaches infinity or zero. If the movement passes through a medium-tailed distribution, then it is within the power-transform range $PTR(\Xi_+)$. Since a power-transformation range can contain at most one medium-tailed distribution,[12] the set of medium-tailed distributions $\Xi_+$ is merely an interface between $\Xi_0$ and $\Xi_\infty$. It is not intended for the green region to appear thick.

But one might think that the transition between light and heavy implies that some $X^{\gamma}$ is medium-tailed. A counterexample disproves this: let R be the random variable $S_R(x) = (1+x)^{-x}$. Its hazard rate is $\lambda_R(x) = ln(1+x) + x/(1+x)$. Hence:

$$\begin{aligned}
\tau_{\frac{Y}{\theta}=R^{\gamma}} &= \frac{1}{\gamma\theta}\lim_{v\to\infty}\left\{\lambda_R(v)\cdot v^{1-\gamma}\right\} \\
&= \frac{1}{\gamma\theta}\lim_{v\to\infty}\left\{\left(ln(1+v)+v/(1+v)\right)\cdot v^{1-\gamma}\right\} \\
&= \frac{1}{\gamma\theta}\lim_{v\to\infty}\left(\frac{ln(1+v)+v/(1+v)}{ln(v)}\cdot ln(v)\cdot v^{1-\gamma}\right) \\
&= \frac{1}{\gamma\theta}\lim_{v\to\infty}\left(ln(v)\cdot v^{1-\gamma}\right) \\
&= \begin{cases} \infty & 0 < \gamma \leq 1 \\ 0 & \gamma > 1 \end{cases}
\end{aligned}$$

So there are power transformations back and forth between $\Xi_\infty$ and $\Xi_0$ which avoid $\Xi_+$. For this reason, the area underneath the green is porous; it shades from yellow to blue

[12] More accurately, it contains at most one medium-tailed distribution per $\theta$. The diagram cannot represent $\Xi$ as a metric space; only $\tau_X$ and $X^{\gamma}$ are represented.

## 8. THE ULTIMATE SETTLEMENT RATE UNDER EXPONENTIAL AND LOGARITHMIC TRANSFORMATIONS

Our third transformation is the exponential, which Klugman defines as $Y = e^X$ [4, p. 95]. However, for two reasons we prefer the form $\frac{Y}{\theta} = \frac{e^{\eta X} - 1}{\eta}$, for $\eta > 0$.[13] First, although $Y = e^X$ works for such random variables as the normal, with support over $\Re$, we are transforming loss distributions, whose support is positive. We wish all our transformations $y = h(x)$ to be strictly increasing functions from $[0, \infty)$ *onto*, not just *into*, $[0, \infty)$. Therefore, $0 = h(0)$, as it does in the above forms. Second, the parameter $\eta$, though not strictly necessary, standardizes the transformation; it sets the derivative at zero to unity, or $1 = h'(0)$. The standardized transformation looks like $y = x$ in the neighborhood of the origin. In fact, the limit of the standardized transformation as $\eta \to 0^+$ is the identity function $h(x) = x$. The appeal of this limiting case is the second reason for our form.

As for the ultimate settlement rate under the exponential transformation:

$$S_Y(u) = Prob\left[\frac{Y}{\theta} > \frac{u}{\theta}\right] = Prob\left[\frac{e^{\eta X} - 1}{\eta} > u/\theta\right] = Prob\left[X > \frac{1}{\eta} ln\left(1 + \eta\frac{u}{\theta}\right)\right] = S_X\left(\frac{1}{\eta} ln\left(1 + \eta\frac{u}{\theta}\right)\right)$$

Therefore:

[13] Actually, for precision and to ensure unitless parameters in transcendental functions we should include two scale parameters: $\frac{Y}{\theta} = \frac{e^{\eta\frac{X}{\phi}} - 1}{\eta}$. But, again, as in Footnote 5, this overparameterizes, for $\frac{Y}{(\theta/\phi)} = \frac{e^{(\eta/\phi)X} - 1}{(\eta/\phi)}$. This transformation is valid and meaningful even for $\eta < 0$, as explained in the appendix.

$$\begin{aligned}
\tau_Y &= -\lim_{u\to\infty}\frac{d\,ln\,S_Y(u)}{du} \\
&= -\lim_{u\to\infty}\left\{\frac{d\,ln\,S_X\left(\frac{1}{\eta}ln\left(1+\eta\frac{u}{\theta}\right)\right)}{d\left(\frac{1}{\eta}ln\left(1+\eta\frac{u}{\theta}\right)\right)}\frac{d\left(\frac{1}{\eta}ln\left(1+\eta\frac{u}{\theta}\right)\right)}{du}\right\} \\
&= \lim_{\substack{u\to\infty \\ v(u)\to\infty}}\left\{-\frac{d\,ln\,S_X(v)}{dv}\cdot\frac{1}{1+\eta\frac{u}{\theta}}\cdot\frac{1}{\theta}\right\} \\
&= \lim_{v\to\infty}\left\{\lambda_X(v)\cdot\frac{1}{e^{\eta v}}\cdot\frac{1}{\theta}\right\} \\
&= \frac{1}{\theta}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot e^{-\eta v}\right\} \\
&= \frac{1}{\theta}\tau_X\cdot 0
\end{aligned}$$

Since $\eta > 0$, medium- and heavy-tailed random variables exponentially transform into heavy-tailed ones; light-tailed random variables are indeterminate.

But let $X$ be light-tailed ($\tau_X = \infty$), but not very light-tailed. This puts $X$ in the yellow region of the diagram. Using "simply" for "not very," we can say that $X$ is simply light-tailed. Then, it becomes heavy-tailed ($\tau_Z = 0$) under some $\gamma > 1$ power transformation $Z = X^{\gamma}$. Hence:

$$0 = \gamma\cdot 0 = \gamma\cdot\tau_Z = \gamma\cdot\frac{1}{\gamma}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot v^{1-\gamma}\right\} = \lim_{v\to\infty}\left\{\lambda_X(v)\cdot v^{1-\gamma}\right\}$$

This information resolves the indeterminacy of the exponential transformation. The following proof makes use of the truth that $\lim_{v\to\infty}\left\{e^{-\eta v}v^{\gamma-1}\right\} = 0$ for $\eta > 0$:

$$\begin{aligned}
0 &= \frac{1}{\theta} \cdot 0 \cdot 0 \\
&= \frac{1}{\theta} \cdot \lim_{v\to\infty}\left\{\lambda_X(v) \cdot v^{1-\gamma}\right\} \cdot \lim_{v\to\infty}\left\{e^{-\eta v} v^{\gamma-1}\right\} \\
&= \frac{1}{\theta} \cdot \lim_{v\to\infty}\left\{\lambda_X(v) \cdot v^{1-\gamma} e^{-\eta v} v^{\gamma-1}\right\} \\
&= \frac{1}{\theta} \cdot \lim_{v\to\infty}\left\{\lambda_X(v) \cdot e^{-\eta v}\right\} \\
&= \tau_Y
\end{aligned}$$

So, in the "simply light" case, the indeterminacy of $\tau_Y = \frac{1}{\theta}\tau_X \cdot 0 = \frac{1}{\theta} \cdot \infty \cdot 0$ resolves to heavy.

However, we do not yet know whether $Y$ is simply heavy or very heavy (blue or violet). So now let $Z$ now be a power transformation of $\frac{Y}{\theta}$, i.e., $Z = \left(\frac{Y}{\theta}\right)^{\gamma} = \left(\frac{e^{\eta X} - 1}{\eta}\right)^{\gamma}$. And so:

$$S_Z(u) = Prob\left[\left(\frac{e^{\eta X} - 1}{\eta}\right)^{\gamma} > u\right] = Prob\left[X > \frac{1}{\eta} ln\left(1 + \eta u^{\frac{1}{\gamma}}\right)\right] = S_X\left(\frac{1}{\eta} ln\left(1 + \eta u^{\frac{1}{\gamma}}\right)\right)$$

We have seen just above that because $X$ is simply light-tailed, $\lim_{v\to\infty}\left\{\lambda_X(v) \cdot e^{-\eta v}\right\} = 0$. But this holds true any $\eta > 0$. And since $\gamma > 0$, it will hold true also for $\gamma\eta > 0$. Therefore, knowing that $\lim_{v\to\infty}\left\{\lambda_X(v) \cdot e^{-\gamma\eta v}\right\} = 0$, we can determine the value of $\tau_Z$:

$$\tau_Z = -\lim_{u\to\infty}\left\{\frac{d\,ln\,S_X\left(\frac{1}{\eta}ln\left(1+\eta u^{\frac{1}{\gamma}}\right)\right)}{d\left(\frac{1}{\eta}ln\left(1+\eta u^{\frac{1}{\gamma}}\right)\right)}\frac{d\left(\frac{1}{\eta}ln\left(1+\eta u^{\frac{1}{\gamma}}\right)\right)}{du}\right\}$$

$$= \lim_{\substack{u\to\infty \\ v(u)\to\infty}}\left\{-\frac{d\,ln\,S_X(v)}{dv}\cdot\frac{\frac{1}{\gamma}u^{\frac{1}{\gamma}-1}}{1+\eta u^{\frac{1}{\gamma}}}\right\}$$

$$= \lim_{v\to\infty}\left\{\lambda_X(v)\cdot\frac{\frac{1}{\gamma}\left(\frac{e^{\eta v}-1}{\eta}\right)^{1-\gamma}}{e^{\eta v}}\right\}$$

$$= \frac{1}{\gamma\cdot\eta^{1-\gamma}}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot\left(\frac{e^{\eta v}-1}{e^{\eta v}}\right)^{1-\gamma}\cdot\frac{1}{e^{\gamma\eta v}}\right\}$$

$$= \frac{1}{\gamma\cdot\eta^{1-\gamma}}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot e^{-\gamma\eta v}\right\}$$

$$= \frac{1}{\gamma\cdot\eta^{1-\gamma}}\cdot 0$$

$$= 0$$

Consequently, $Y$ is a heavy-tailed random variable whose tail class is invariant to power transformation; it is very heavy-tailed. This proves that an exponential transformation of a simply light-tailed random variable is a very heavy-tailed random variable. In the diagram exponential transformation moves from the yellow region to the violet; unlike power transformation, it is capable of jumping a least the right barrier.

In the exponential-transformation of medium- and heavy-tailed random variables, there is no indeterminacy to $\tau_Y = \frac{1}{\theta}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot e^{-\eta v}\right\} = 0$. But again, the ultimate settlement rate of a subsequent power transformation is $\tau_Z = \frac{1}{\gamma\cdot\eta^{1-\gamma}}\lim_{v\to\infty}\left\{\lambda_X(v)\cdot e^{-\gamma\eta v}\right\} = 0$. Hence, exponential transformation turns medium and simply heavy tails into very heavy tails. In sum, it transforms yellow, green, and blue into violet. But its differing effect on moments will be treated in the next section.

We can be brief about the logarithmic transformation $\frac{Y}{\theta}=\frac{1}{\eta}ln(1+\eta X)$. It is the inverse of the exponential transformation, but less obviously than in the case of power transformation. Define two operators: the exponential transformation $ET[X;\eta,\theta]=\theta\frac{e^{\eta X}-1}{\eta}$ and the logarithmic transformation $LT[X;\eta,\theta]=\theta\cdot\frac{1}{\eta}ln(1+\eta X)$. Just as $ET[X;\eta,\theta]\to\theta X$ as $\eta\to 0^{+}$, so too $LT[X;\eta,\theta]\to\theta X$ as $\eta\to 0^{+}$. One who performs the algebra will find that $ET\left[LT[X;\eta,\theta];\frac{\eta}{\theta},\frac{1}{\theta}\right]=LT\left[ET[X;\eta,\theta];\frac{\eta}{\theta},\frac{1}{\theta}\right]=X$. Since $\eta,\theta>0$, the inverting parameters exist. Now if $\frac{Y}{\theta}=\frac{1}{\eta}ln(1+\eta X)$, then $S_Y(u)=S_X\left(\frac{e^{\eta\frac{u}{\theta}}-1}{\eta}\right)$. The reader should be able to prove by now that $\tau_Y=\frac{1}{\theta}\lim_{v\to\infty}\{\lambda_X(v)\cdot(1+\eta v)\}=\frac{1}{\theta}\tau_X\cdot\infty$. So $\eta>0$ logarithmic transformation turns light and medium into light; the heavy-tailed random variables are indeterminate. It is not necessary to repeat the power-on-top-of-exponential-transformation argument. Because of exponential-logarithmic inversion, the question "Into what do yellow, green, and blue logarithmically transform?" is equivalent to "What exponentially transforms into yellow, green, and blue?" Whatever it is, it can't be undone by power transformation. Therefore, what exponentially transforms into the middle region of the diagram must be very light. So exponential and logarithmic transformations from the middle jump the barriers.

## 9. POSITIVE MOMENTS AND THE EXPONENTIAL TRANSFORMATION

Section 5 proved that all the positive moments of light- and medium-tailed random variables are finite. An infinite moment is a sufficient, but not a necessary, condition for a heavy tail. Here we will examine the positive moments of the exponentially transformed $Y=\frac{e^{\eta X}-1}{\eta}$. But $Y^{k}=\left(\frac{e^{\eta X}-1}{\eta}\right)^{k}$. Although this is on the order of $e^{k\eta X}$, it is not the same. Since our findings depend on the behavior of $e^{k\eta X}$, we must first prove that $E[Y^{k}]$ is finite if and only if $E[e^{k\eta X}]$ is finite.

To begin, $E[Y^{k}]=E\left[\left(\frac{e^{\eta X}-1}{\eta}\right)^{k}\right]=\frac{E\left[\left(e^{\eta X}-1\right)^{k}\right]}{\eta^{k}}$. Since $\eta$ and *k* are positive, $\eta^{k}$ is positive. So $E[Y^{k}]$ is finite if and only $E\left[\left(e^{\eta X}-1\right)^{k}\right]$ is finite. And since $0\le e^{\eta x}-1<e^{\eta x}$ over the support of *X*, $0^{k}=0\le\left(e^{\eta x}-1\right)^{k}<e^{k\eta x}$. So $Prob\left[0\le\left(e^{\eta X}-1\right)^{k}<e^{k\eta X}\right]=1$ and $0\le E\left[\left(e^{\eta X}-1\right)^{k}\right]<E\left[e^{k\eta X}\right]$. Therefore, if $E\left[e^{k\eta X}\right]$ is finite, then so too is $E\left[\left(e^{\eta X}-1\right)^{k}\right]$. As for the converse:

$$
\begin{aligned}
E\left[\left(e^{\eta X}-1\right)^k\right] &= \int_{x=0}^{\infty}\left(e^{\eta x}-1\right)^k dF_X(x) \\
&= \int_{\eta x=0}^{\infty}\left(e^{\eta x}-1\right)^k dF_X(x) \\
&= \int_{\eta x=0}^{ln\,2}\left(e^{\eta x}-1\right)^k dF_X(x)+\int_{\eta x=ln\,2}^{\infty}\left(e^{\eta x}-1\right)^k dF_X(x) \\
&= \int_{\eta x=0}^{ln\,2}\left(e^{\eta x}-1\right)^k dF_X(x)+\int_{\eta x=ln\,2}^{\infty}\left(\frac{e^{\eta x}-1}{e^{\eta x}}\right)^k e^{k\eta x}dF_X(x) \\
&\geq \int_{\eta x=0}^{ln\,2}(0)^k dF_X(x)+\int_{\eta x=ln\,2}^{\infty}\left(\frac{1}{2}\right)^k e^{k\eta x}dF_X(x) \\
&\geq \left(\frac{1}{2}\right)^k \int_{\eta x=ln\,2}^{\infty} e^{k\eta x}dF_X(x)
\end{aligned}
$$

Consequently, $\int_{\eta x=ln\,2}^{\infty} e^{k\eta x}dF_X(x)\leq 2^k\,E\left[\left(e^{\eta X}-1\right)^k\right]$. Furthermore:

$$
\begin{aligned}
E\left[e^{k\eta X}\right] &= \int_{\eta x=0}^{\infty} e^{k\eta x}dF_X(x) \\
&= \int_{\eta x=0}^{ln\,2} e^{k\eta x}dF_X(x)+\int_{\eta x=ln\,2}^{\infty} e^{k\eta x}dF_X(x) \\
&\leq \int_{\eta x=0}^{ln\,2} 2^k\,dF_X(x)+\int_{\eta x=ln\,2}^{\infty} e^{k\eta x}dF_X(x) \\
&\leq 2^k+\int_{\eta x=ln\,2}^{\infty} e^{k\eta x}dF_X(x) \\
&\leq 2^k+2^k\,E\left[\left(e^{\eta X}-1\right)^k\right]
\end{aligned}
$$

So $0\leq E\left[e^{k\eta X}\right]\leq 2^k+2^k\,E\left[\left(e^{\eta X}-1\right)^k\right]$. Therefore, if $E\left[\left(e^{\eta X}-1\right)^k\right]$ is finite, so too is $E\left[e^{k\eta X}\right]$. Thus have we shown that $E\left[Y^k\right]$ is finite if and only if $E\left[e^{k\eta X}\right]$ is finite.

Now we continue with the simpler problem of examining the moments of $E\left[e^{k\eta X}\right]$. Using again the theorem from Section 5 that $E[h(X)]=h(0)+\int_{x=0}^{\infty}S_X(x)dh(x)$, we have:

$$
E\left[e^{k\eta X}\right]=e^{k\eta 0}+\int_{x=0}^{\infty}S_X(x)de^{k\eta x}=1+k\eta\int_{x=0}^{\infty}S_X(x)e^{k\eta x}dx
$$

Therefore, $E\left[Y^k\right]$ is finite if and only if $\int_{x=0}^{\infty} S_X(x)e^{k\eta x}dx$ is finite. Dispensing with mathematical rigor, we know that $\int_{x=0}^{\infty} S_X(x)e^{k\eta x}dx$ is finite if and only there exists a $\xi > 0$ such that for all $x \geq \xi$, $S_X(x) \leq S_X(\xi)e^{-k\eta(x-\xi)}$. In words, in order for the integral to converge, at some point the survival function must decay at a rate greater than $k\eta$. But the limit of this decay is the ultimate settlement rate $\tau_X$. Hence, $E\left[Y^k\right]$ is finite if and only if $k\eta < \tau_X$, or $k < \tau_X/\eta$. Therefore, all the positive moments of exponential transformations of light-tailed ($\tau_X = \infty$) distributions are finite. The positive moments of exponential transformations of medium-tailed ($0 < \tau_X < \infty$) distributions are finite for $k < \tau_X/\eta$ and infinite for $k \geq \tau_X/\eta$. And all the positive moments of exponential transformations of heavy-tailed ($\tau_X = 0$) distributions are infinite.[14]

## 10. INVERTING AND MIXING LOSS DISTRIBUTIONS

The commentary on property (i) in Section 2 stated the desirability for a loss distribution to be invertible. But our only use of an inverse distribution was to derive the ultimate settlement rate of the inverse gamma in Section 3. Moreover, Klugman lists a fourth transformation, viz., mixing [4, pp. 97-99]. Both inverting and mixing are involved in the generalized Pareto, because:

$$\begin{aligned} GenPareto(\alpha,\beta,\theta) &\sim \frac{\theta}{Gamma(\beta,1)} \cdot Gamma(\alpha,1) \\ &\sim InvGamma(\beta,\theta) \cdot Gamma(\alpha,1) \\ &\sim Gamma(\alpha, InvGamma(\beta,\theta)) \end{aligned}$$

So the generalized Pareto can be formed as a gamma distribution whose scale parameter is an inverse-gamma distribution. In this section we will explain why inverting and mixing tend to produce heavy-tailed distributions.

[14] Thus indirectly we verify what we know about the lognormal distribution, an exponential transformation of the light-tailed normal, viz., that all its moments are finite: $E\left[e^{kN(\mu,\sigma^2)}\right] = e^{k\mu + k^2\sigma^2/2}$. $E\left[Y^k\right] \approx E\left[e^{k\eta X}\right] = M_X(k\eta)$. If $X \sim Gamma(\alpha,\theta)$, $M_X(k\eta) = (1-\theta k\eta)^{-\alpha}$, which diverges for $k\eta \geq 1/\theta = \tau_X$. The missing moments explain the intractability, or even the nonexistence, of the moment generating functions of all but the simplest distributions. However, the zeroth moment is finite: $E\left[Y^0\right] = E\left[e^{0X}\right] = 1$. So by virtue of absolute convergence in the complex numbers, viz., $\left|\varphi_X(t)\right| = \left|E\left[e^{itX}\right]\right| \leq E\left[\left|e^{itX}\right|\right] = E[1] = 1$, the characteristic function is more successful. All the imaginary moments of all real random variables exist as complex numbers.

First, as to inverting, if $Y = 1/X$, then:

$$S_Y(u) = Prob\left[\frac{1}{X} > u\right] = Prob\left[X < \frac{1}{u}\right] = 1 - Prob\left[X \geq \frac{1}{u}\right] = 1 - Prob\left[X = \frac{1}{u}\right] - S_X\left(\frac{1}{u}\right)$$

Here we will assume that $X$ has no probability mass, at least not in the neighborhood of zero. So $S_Y(u) = 1 - S_X\left(\frac{1}{u}\right)$. Therefore:

$$\lambda_Y(u) = -\frac{d\,ln\,S_X(x)}{dx} = -\frac{d\,ln\left(1 - S_X\left(\frac{1}{u}\right)\right)}{dx} = \frac{-S'_X\left(\frac{1}{u}\right)}{1 - S_X\left(\frac{1}{u}\right)}\frac{1}{u^2} = \frac{f_X\left(\frac{1}{u}\right)}{1 - S_X\left(\frac{1}{u}\right)}\frac{1}{u^2}$$

And so the ultimate settlement rate of $Y$ is:

$$\tau_Y = \lim_{u\to\infty}\lambda_Y(u) = \lim_{u\to\infty}\frac{f_X\left(\frac{1}{u}\right)}{1 - S_X\left(\frac{1}{u}\right)}\frac{1}{u^2} = \lim_{v\to 0^+}\frac{f_X(v)}{1 - S_X(v)}v^2 = \lim_{v\to 0^+}\frac{f_X(v)}{F_X(v)}v^2 = \lim_{v\to 0^+}\frac{f_X(v)}{h(v)}v,$$

where $h(v) = \frac{F_X(v)}{v} = \frac{1}{v}\int_{x=0}^{v} f_X(x)$. It is the average height of $f_X$ over the interval $(0, v]$. If $f_X$ increases as $x \to 0^+$, $h(v) > f_X(v) > 0$. Then $0 < \frac{f_X(v)}{h(v)} < 1$, and $\tau_Y = 0$. This holds true even if $f_X$ approaches infinity. If $f_X$ is bounded within two positive numbers, then again, $\tau_Y = 0$. The remaining possibility is that $f_X$ decreases to zero as $x \to 0^+$, in which case $1 < \frac{f_X(v)}{h(v)}$.

Now if $f_X$ is zero in some interval $[0, \varepsilon]$, then $Prob[X \leq \varepsilon] = 0$; so $S_Y(1/\varepsilon) = Prob[Y > 1/\varepsilon] = 0$ and property (ii) would disqualify $Y$ as a loss distribution. So $f_X$ decreases to zero, but equals zero only at the origin. The obvious choice is a power-function approach into the origin, i.e., $f_X(v) \propto v^{\gamma-1}$ for $\gamma > 1$. But then $F(v) \propto v^\gamma/\gamma$ and $\tau_Y = \lim_{v\to 0^+}\frac{v^{\gamma-1}}{(v^\gamma/\gamma)}v^2 = 0$. So even power-function approaches are not slow enough. For $\tau_Y$ to be positive, near the origin, $\frac{f_X(v)}{F_X(v)}$ must be on the order of $v^{-2}$. So $\frac{d\,ln\,F_X(v)}{dv} = \frac{f_X(v)}{F_X(v)} = \frac{k}{v^2}$, and $ln\,F_X(v) - ln\,F_X(\varepsilon) = \int_{x=\varepsilon}^{v}\frac{k}{x^2}dx = \frac{k}{\varepsilon} - \frac{k}{v}$. Thus, $F_X(v) = F_X(\varepsilon)e^{\frac{k}{\varepsilon}}e^{-\frac{k}{v}}$. The solution which satisfies $F_X(0) = 0$ is $F_X(x) = \begin{cases} e^{-\frac{k}{x}} & x > 0 \\ 0 & x = 0 \end{cases}$. But this is the inverse exponential cumulative density function. So, the only likely way of obtaining anything other than a heavy-tailed distribution by inversion is to invert an already inverted distribution. One may expect inversion to produce heavy-tailed distributions.

As for mixing, let $S_X(x) = \int_\theta S_{X|\theta}(x)dh(\theta)$. The survival function of the mixed distribution is the weighting according to $dh(\theta)$ of the distributions indexed by θ. Hence:

$$\lambda_X(x) = -\frac{S'_X(x)}{S_X(x)}$$

$$= \frac{\int_\theta S'_{X|\theta}(x)dh(\theta)}{\int_\theta S_{X|\theta}(x)dh(\theta)}$$

$$= \frac{\int_\theta \lambda_{X|\theta}(x)\cdot S_{X|\theta}(x)dh(\theta)}{\int_\theta S_{X|\theta}(x)dh(\theta)}$$

$$= \int_\theta \lambda_{X|\theta}(x)\cdot dw(x,\theta)$$

In the last equation $dw(x,\theta) = S_{X|\theta}(x)dh(\theta) \Big/ \int_\theta S_{X|\theta}(x)dh(\theta)$. The weights vary by *x*; but for all *x*, $\int_\theta dw(x,\theta) = 1$. As $x \to \infty$, the weighting will shift more and more toward the "surviving" distributions, i.e., in favor of the distributions whose τ is least. Consequently, $\tau_X = \inf\{\tau_{X|\theta}\}$. For the mixed exponential distribution $S_{MX}(x) = \sum_{i=1}^{n} p_i e^{-\frac{x}{\theta_i}}$, $\tau_X = min(1/\theta_i) = 1/max(\theta_i)$. This is medium-tailed; but $S_{MX}(x) = \sum_{i=1}^{\infty} p_i e^{-\frac{x}{\theta_i}}$ is heavy-tailed, if $\lim_{i\to\infty}\theta_i = \infty$.[15]

## 11. SUPER LIGHT AND SUPER HEAVY

The region within the barriers of the diagram is like the everyday world. Its span is that of the power transformation. But of course, in the long run $e^x$ overwhelms $x^n$. To what others mean loosely by "in the long run"[16] mathematicians have given precision, viz., $\lim_{x\to\infty}$.[17] Just over the right

---

[15] The Tweedie distribution is $T = X_1 + \ldots + X_N$, for $X \sim Gamma(\alpha, \theta)$ and $N \sim Poisson(\lambda)$. Therefore, $T|N \sim Gamma(N\alpha, \theta)$ and $\tau_{T|N} = 1/\theta$. So, $\tau_T = \inf\{\tau_{T|N}\} = 1/\theta$, and $T$ is medium-tailed.

[16] Such loose speech harbors specious arguments, for which Keynes expressed disdain in his famous quip, "In the long run we're all dead." Many use the adverb 'exponentially', as in "Something is growing exponentially," to express alarm, as if dealing with that thing were a critical matter. The sober truth is that almost all growth is exponential, but of limited duration. Mathematically, for $x \approx 0$, $e^{\gamma x} \approx 1 + \gamma x$. Moreover, the obverse is never considered: no one ever expresses alarm by saying that something is decaying exponentially.

[17] Still amazing even after 150 years are the accomplishments of such mathematicians as Cauchy, Weierstrass, and Dedekind concerning the nature of the real numbers, which finally put to rest the 2300-year-old paradoxes of Zeno. One who might try to resurrect them on the basis of today's quantum theory would ignore the fact that the paradoxes themselves presuppose continuity.

barrier are some familiar enough distributions, the very heavy-tailed ones. Hopping over it, first we'll find exponential transformations of simply light tails with all their moments. Second we'll find ETs of middle tails with moments up to a point. Third we'll find ETs of simply heavy tails with no moments at all. The familiar enough distributions are of the first two types; the distributions with no moments we will call "super heavy-tailed." We could refine our diagram's color scheme in conformity with the spectrum: first indigo, second violet, and third ultraviolet. The span of these distributions is on the order of $e^{\eta\left(x^{\gamma}\right)}$. But because $e^{\left(x^{\gamma}\right)} >> \left(e^{x}\right)^{\gamma}$, their span is greater than that of the power transformation. In fact, their span is "power-on-top-of-exponential." But at the end of that span is another barrier, over which another exponential transformation jumps, and so forth. The same applies in the other direction, into the microworld, with the logarithmic transformation. In descending order of heaviness are logarithmic transforms of simply heavy tails, which we could color orange. Second are LTs of middle tails, which remain red. And third are LTs of simply light tails, "super light-tailed" distributions whose color is infrared. And then we find a barrier to be surmounted by another LT.[18] So the classification is indefinitely extendable; but current needs remain within one transformation of the center.

## 11. CONCLUSION

Good classifications are not arbitrary; they are not set by convention or decree. Natural classifications should actually help those who study a subject to understand it and eventually to make deeper discoveries. Work is made easier with the right tools, and the essential tool for intellectual work is clear definition and classification. In this paper we entered the house of loss distributions through the door of the medium-tail distribution. We explored the first floor with the help of the power transformation, and then found exponential and logarithmic staircases to the second floor and the basement. Some the mathematics was formidable; but it all reduces to the interaction between power and exponential functions. The classification scheme yielded new and beautiful insights. Surely there is much more to be discovered; but the classification of distributions into light, medium, and heavy, as well as the subclassifications "very" and "super," almost as surely will play an important role therein. Though it might be hard for now to put this theory to practical use (we've given no list of "which distribution for which purpose"), actuaries have a right to appreciate the beauty of their subject – its æsthetic value. And many, perhaps most, practical benefits have arisen from what once had been considered "mere theory."[19]

[18] But however rarefied these tails may become, they are still infinite.

[19] That good theory aids discovery and technological progress (and conversely, that bad theory impedes them) is illustrated in modern physics. On the basis of quantum theory in 1928 Wolfgang Pauli predicted the existence of antimatter, in particular, the anti-electron or positron, which was discovered in 1932 and whose discovery now benefits mankind in positron emission tomography – commonly performed in hospitals as PET scans. Since the 1960s the standard model of particle physics has predicted the one still missing particle, the Higgs boson, whose existence many

## REFERENCES

[1] Corro, Dan, "Grouping Loss Distributions by Tail Behavior," *CAS E-Forum* (Fall 2008), 448-541, www.casact.org/pubs/forum/08fforum.

[2] Embrechts, Paul, Resnick, Sidney I., and Samorodnitsky, Gennady, "Extreme Value Theory as a Risk Management Tool," *North American Actuarial Journal*, 3:2, 1999, 30-41, www.casact.org/library/studynotes/Embrechts_ExtremeValue.pdf.

[3] Halliwell, Leigh J., "The Mathematics of Excess Losses," *Variance*, 6:1, 2012, 32-47, www.variancejournal.org/issues/06-01/32.pdf.

[4] Klugman, Stuart A., and Panjer, Harry H., and Willmot, Gordon E., *Loss Models: From Data to Decisions*, 1998, John Wiley & Sons, New York.

[5] Meyers, Glenn, "Pure Premium Regression with the Tweedie Model," The Actuarial Review (May 2009), 12-13, www.casact.org/newsletter/pdfUpload/ar/Brainstorms_corrected_from AR_May2009.pdf.

[6] Wikipedia contributors, "Generalized Pareto distribution," *Wikipedia, The Free Encyclopedia*, http://en.wikipedia.org/wiki/Generalized_Pareto_distribution.

---

physicists believe is about to be verified (as of January 2013) in CERN's Large Hadron Collider. Although science is inductive, it is not haphazard. On the basis of theory scientists make hunches. Theory is essential for posing the right questions, and helps in determining how and where to look for their answers. The present author discovered much about loss distributions while developing this classification, the following three things in particular. First, it led to his discovery of super heavy-tailed distributions; before he would have insisted that every distribution must have some positive moments. Second, the symmetry of the "multiplication table" of Section 6 beckoned the existence of super light-tailed distributions, one of which he then readily found, the *Q* distribution of Section 7. Third, again in Section 7, he spent considerable time trying to prove that every power-transformation range must include a medium-tailed distribution, i.e., that the green line of the diagram cuts all the way through. Twice he thought to have proven it only on checking to be disappointed. This opened him to try to falsify the proposition, during which he discovered a new expression for the ultimate settlement rate that specified a sufficient characteristic for a counterexample, which again he readily found, the *R* distribution. Thus, he is convinced that this classification is not arbitrary, no mere convention or convenience. Rather, it is fertile of discovery.

# APPENDIX

## Extreme Value Theory

Most casualty actuaries have studied the forms of loss distributions that are given in Klugman [3, Appendix A]. However, in the field of extreme-value theory, there is a generalized-Pareto distribution that differs from our $GenPareto(\alpha, \beta, \theta)$. In this appendix we will translate it into forms more familiar to actuaries. The survival function of this generalized-Pareto is:

$$S_X(x \geq 0; \xi, \theta) = Prob[X > x] = \begin{cases} \left(1 + \xi \dfrac{x}{\theta}\right)^{-\frac{1}{\xi}} & \text{for } \xi \neq 0 \\ e^{-\frac{x}{\theta}} & \text{for } \xi = 0 \end{cases}$$

This is the definition given in [2, p. 33] and [5], except that we have zeroed a location parameter and used 'θ' instead of 'σ' for the scale parameter. The shape parameter ξ may be any real number, but the scale parameter θ, as always, must be positive. The function is defined for $\xi = 0$ as $S_X(x; 0, \theta) = \lim_{\xi \to 0} S_X(x; \xi, \theta)$, which pertains to the $Gamma(1, \theta)$ or $Exponential(\theta)$ distribution.

For $\xi \neq 0$, the probability density function is $f_X(x) = -\dfrac{dS_X(x)}{dx} = \dfrac{1}{\xi}\left(1 + \xi \dfrac{x}{\theta}\right)^{-\frac{1}{\xi} - 1} \dfrac{\xi}{\theta}$. If $\xi > 0$, the exponent is negative. In this case, the function translates as follows:

$$
\begin{aligned}
f_X(x) &= \frac{1}{\xi}\left(1+\xi\frac{x}{\theta}\right)^{-\frac{1}{\xi}-1}\frac{\xi}{\theta} \\
&= \frac{1}{\xi}\left(\frac{1}{x/(\theta/\xi)+1}\right)^{\frac{1}{\xi}+1}\frac{1}{(\theta/\xi)} \\
&= \frac{1}{\xi}\left(\frac{x/(\theta/\xi)}{x/(\theta/\xi)+1}\right)^{1-1}\left(\frac{1}{x/(\theta/\xi)+1}\right)^{\frac{1}{\xi}-1}\frac{1/(\theta/\xi)}{(x/(\theta/\xi)+1)^2} \\
&= \frac{\Gamma\left(1+\frac{1}{\xi}\right)}{\Gamma(1)\Gamma\left(\frac{1}{\xi}\right)}\left(\frac{x/(\theta/\xi)}{x/(\theta/\xi)+1}\right)^{1-1}\left(\frac{1}{x/(\theta/\xi)+1}\right)^{\frac{1}{\xi}-1}\frac{1/(\theta/\xi)}{(x/(\theta/\xi)+1)^2} \\
&= f_{GenPareto\left(1,\frac{1}{\xi},\frac{\theta}{\xi}\right)}(x) \quad\Rightarrow\quad \frac{X}{(\theta/\xi)} \sim \frac{Gamma(1,1)}{Gamma\left(\frac{1}{\xi},1\right)}
\end{aligned}
$$

If $\xi < 0$, $-\frac{1}{\xi}$ is positive, and the translation is:

$$
\begin{aligned}
f_X(x) &= \frac{1}{-\xi}\left(1-\xi\frac{x}{\theta}\right)^{\frac{1}{-\xi}-1}\frac{-\xi}{\theta} \\
&= \frac{1}{-\xi}\left(1-x/(\theta/-\xi)\right)^{\frac{1}{-\xi}-1}\left(1/(\theta/-\xi)\right) \\
&= \frac{1}{-\xi}\left(x/(\theta/-\xi)\right)^{1-1}\left(1-x/(\theta/-\xi)\right)^{\frac{1}{-\xi}-1}\left(1/(\theta/-\xi)\right) \\
&= \frac{\Gamma(1+1/-\xi)}{\Gamma(1)\Gamma(1/-\xi)}\left(x/(\theta/-\xi)\right)^{1-1}\left(1-x/(\theta/-\xi)\right)^{\frac{1}{-\xi}-1}\left(1/(\theta/-\xi)\right) \\
&= f_{(\theta/-\xi)\cdot Beta\left(1,\frac{1}{-\xi}\right)}(x)
\end{aligned}
$$

$$
\Rightarrow\quad \frac{X}{(\theta/-\xi)} \sim Beta\left(1,\frac{1}{-\xi}\right) \sim \frac{Gamma(1,1)}{Gamma(1,1)+Gamma\left(1,\frac{1}{-\xi}\right)}
$$

Therefore, depending on the shape parameter, the tail of this distribution can be finite ("no-tailed" $\xi < 0$), medium-tailed ($\xi = 0$), or very heavy-tailed ($\xi > 0$).

What is most relevant to our analysis of tail characteristics is that this distribution is an exponential transformation of the exponential distribution:

$$\frac{X}{\theta} \sim \frac{e^{\xi \cdot Exponential(1)} - 1}{\xi}$$

Under this transformation the light-tailed exponential distribution becomes very heavy-tailed.