MATH5633 Loss Models I Autumn 2024

Chapter 2: Loss Severity Models

Lecturer: Kenneth Ng

Preview

This chapter introduces models for loss severity. We shall first look at some parametric distributions that are commonly used in modelling loss variables. We then introduce more sophisticated models by transformations and mixing. We will also discuss various coverage modifications on loss variables.

Key topics in this chapter:

1. Common parametric distributions for modelling severity – Exponential, Gamma, Weibull, Pareto, Beta;
2. Coverage Modifications: Deductibles, Policy Limits, Coinsurance and Inflation;
3. Tails of distributions;
4. Transformation of random variables;
5. Mixing.

1 Parametric Continuous Distributions

In what follows, we will let X be a *loss variable*, i.e., a random variable that describes the loss of a peril or the size of claim of an insurance policy. In this section, we will look at some simple continuous, parametric distributions that could be used to model X.

1.1 Exponential Distribution

X follows an exponential distribution with *mean parameter* $\theta > 0$, denoted by $X \sim \text{Exp}(\theta)$, if it has the following pdf:

$$\boxed{f_X(x) = \frac{1}{\theta}e^{-\frac{x}{\theta}},\ x > 0.}$$

- Some important distributional quantities of $X \sim \text{Exp}(\theta)$ on the next page:

1. **Mean and Variance:**
$$\mathbb{E}[X] = \theta, \quad \text{Var}[X] = \theta^2.$$
2. **Higher Moments:**
$$\mathbb{E}[X^k] = n!\theta^k.$$
3. **CDF:**
$$F_X(x) = 1 - e^{-\frac{x}{\theta}}, \quad S_X(x) = e^{-\frac{x}{\theta}}, \ x > 0.$$
4. **Generating Functions:**
$$M_X(t) = \frac{1}{1-\theta t}, \ t < \frac{1}{\theta},$$
$$P_X(t) = \frac{1}{1-\theta \ln t}, \ t < e^{\frac{1}{\theta}}.$$

- $X \sim \text{Exp}(\theta)$ is ***memoryless***: for any $x, a > 0$,
$$\mathbb{P}(X > x + a | X > a) = \mathbb{P}(X > x).$$
- The pdf of X is decreasing in x. This may not be an optimal model for severity, since the distribution of losses from insurance claims usually has a peak (mode) around small/intermediate values.
- When θ increases, more weight is placed on large values of x. The shape of the pdf appears to be flatter, more spread out with a *heavier tail*; see Figure 1.

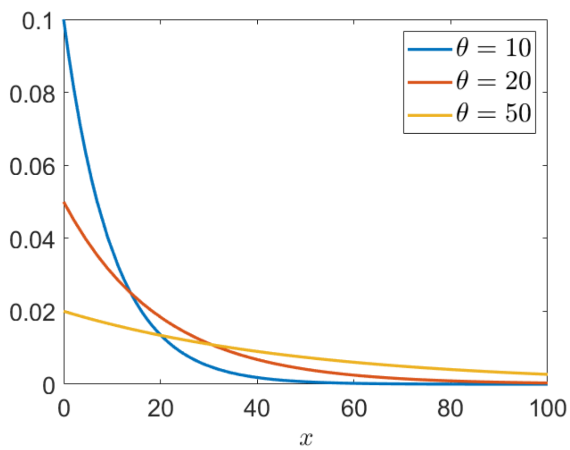


Figure 1: The pdfs of exponential distributions with mean parameter $\theta = 10, 20$ and 50

1.2 Gamma Distribution

X follows a Gamma distribution with *shape parameter* $\alpha > 0$ and *scale parameter* $\theta > 0$, denoted by $X \sim \text{Gamma}(\alpha, \theta)$, if it has the following pdf:

$$\boxed{f_X(x) = \frac{1}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\theta}},\ x > 0,}$$

where $\Gamma(\alpha)$ is the *Gamma function*:

$$\Gamma(\alpha) := \int_0^\infty t^{\alpha-1} e^{-t} dt,\ \alpha > 0.$$

Some properties about the gamma function:

1. for $n \in \mathbb{N}$, $\Gamma(n) = (n-1)!$;
2. for any $\alpha > 0$, $\Gamma(1+\alpha) = \alpha\Gamma(\alpha)$.

- Some important distributional quantities of $X \sim \text{Gamma}(\alpha, \theta)$:

1. **Mean and Variance:**

$$\mathbb{E}[X] = \alpha\theta, \quad \text{Var}[X] = \alpha\theta^2.$$

2. **Higher Moments:**

$$\mathbb{E}[X^k] = \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} \theta^k.$$

3. **Generating Functions:**

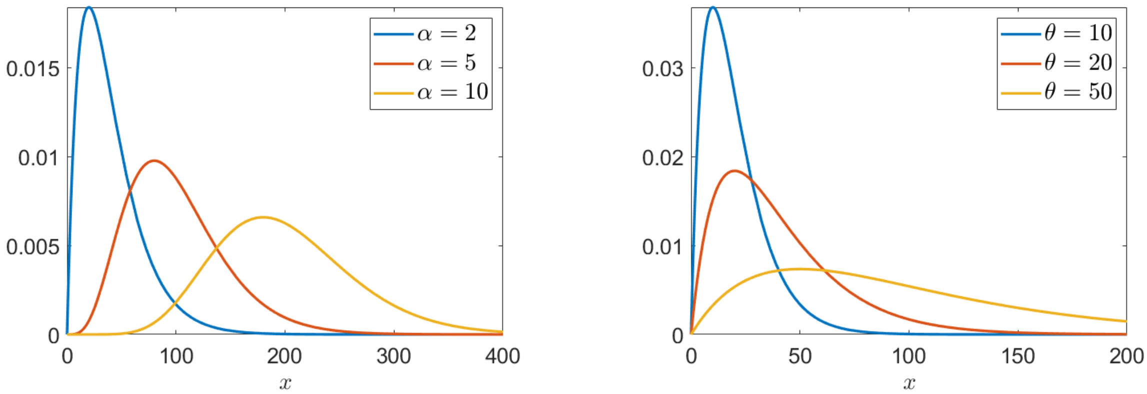
$$M_X(t) = (1-\theta t)^{-\alpha},\ t < \frac{1}{\theta},$$
$$P_X(t) = (1-\theta \ln t)^{-\alpha},\ t < e^{\frac{1}{\theta}}.$$

- $\text{Gamma}(1, \theta) = \text{Exp}(\theta)$. In general, for any positive integer n, $X \sim \text{Gamma}(n, \theta)$ is the sum of n ***independent and identically distributed (i.i.d.)*** exponentially distributed random variables with mean parameter θ, i.e.,

$$X = X_1 + X_2 + \cdots + X_n,$$

where $\{X_i\}_{i=1}^n$ is i.i.d. with $X_i \sim \text{Exp}(\theta)$.

- The shape parameter α controls the *skewness of the distribution.* As α increases, the mode shifts towards right and becomes less distinct. The distribution also becomes more symmetric; see Figure 2a. If $\alpha \in (0, 1)$, the pdf is decreasing in x with mode 0.
- Similar to the exponential distribution, the scale parameter θ controls *how spread out and tail of the distribution*: the higher the θ, the flatter the distribution; see Figure 2b.



(a) $\alpha = 2, 5, 10$; $\theta = 20$

(b) $\theta = 10, 20, 50$; $\alpha = 2$

Figure 2: The pdfs of Gamma distributions with different α and θ

1.3 Weibull Distribution

The Weibull distribution is named after the Swedish physicist Waloddi Weibull. It is often used to model the time to failure. In insurances, it is also used to model the size of excess of loss reinsurance claims.

X follows a Weibull distribution with *shape parameter* $\alpha > 0$ and *scale parameter* $\theta > 0$, denoted by $X \sim \text{Weibull}(\alpha, \theta)$, if it has the following pdf:

$$\boxed{f_X(x) = \frac{\alpha}{\theta^\alpha} x^{\alpha-1} e^{-\left(\frac{x}{\theta}\right)^\alpha}, \ x > 0.}$$

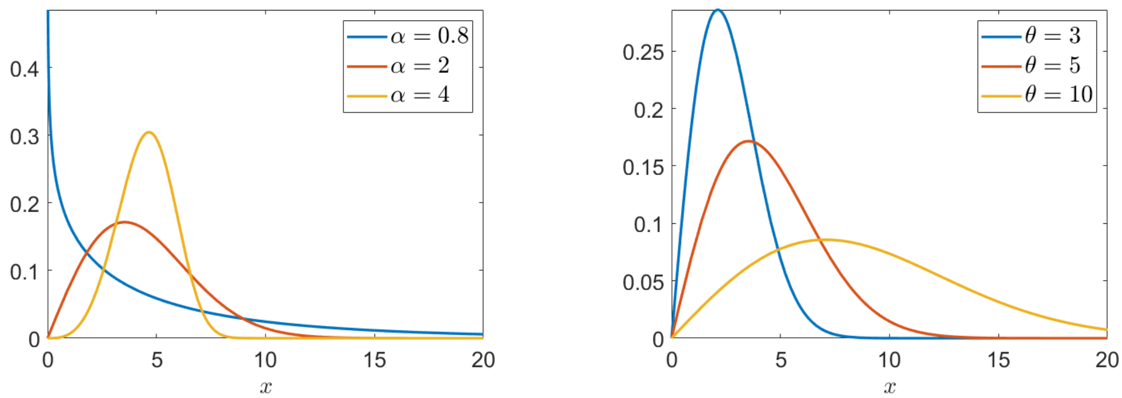
- Some important distributional quantities of $X \sim \text{Weibull}(\alpha, \theta)$:

1. **Mean and Variance:**
$$\mathbb{E}[X] = \theta\Gamma\left(1 + \frac{1}{\alpha}\right), \quad \text{Var}[X] = \theta^2 \left[\Gamma\left(1 + \frac{2}{\alpha}\right) - \left(\Gamma\left(1 + \frac{2}{\alpha}\right)\right)^2\right].$$
2. **Higher Moments:**
$$\mathbb{E}[X^k] = \theta^k \Gamma\left(1 + \frac{k}{\alpha}\right).$$
3. **CDF:**
$$F_X(x) = 1 - e^{-\left(\frac{x}{\theta}\right)^\alpha}, \quad S_X(x) = e^{-\left(\frac{x}{\theta}\right)^\alpha}, \ x > 0.$$

- $\text{Weibull}(1, \theta) = \text{Exp}(\theta)$. In general, if $Y \sim \text{Exp}(\theta^\alpha)$, then $X := Y^{1/\alpha} \sim \text{Weibull}(\alpha, \theta)$.
- The shape parameter α controls the rate of decay, and thus of tail of the distribution. The

distribution has a *heavy tail* if $\alpha < 1$; and a light tail if $\alpha > 1$; see Figure 3a,

- If $\alpha > 1$, the peak of the pdf is positive; otherwise the pdf is a decreasing function, and thus the peak is located at 0.
- As before, the higher the value of the scale parameter θ, the more spread out the distribution is; see Figure 3b.



(a) $\alpha = 0.8, 2, 4$; $\theta = 5$

(b) $\theta = 3, 5, 10$; $\alpha = 2$

Figure 3: The pdfs of Weibull distributions with different α and θ

1.4 Pareto Distribution

The Pareto distribution is named after Italian economist Vilfredo Pareto. It has two formulations, *two-parameter* or *single parameter*. In the sequel, we shall refer a Pareto distribution to as the two-parameter formulation.

1.4.1 Two-Parameter Formulation

X follows a Pareto distribution with *shape parameter* $\alpha > 0$ and *scale parameter* $\theta > 0$, denoted by $X \sim \text{Pareto}(\alpha, \theta)$, if it has the following pdf:

$$\boxed{f_X(x) = \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}, \quad x > 0.}$$

- Some important distributional quantities of $X \sim \text{Pareto}(\alpha, \theta)$:

1. **Mean and Variance:**

$$\mathbb{E}[X] = \begin{cases} \dfrac{\theta}{\alpha - 1}, & \text{if } \alpha > 1; \\ \infty, & \text{otherwise,} \end{cases} \quad , \quad \text{Var}[X] = \begin{cases} \dfrac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}, & \text{if } \alpha > 2; \\ \infty, & \text{otherwise.} \end{cases}$$

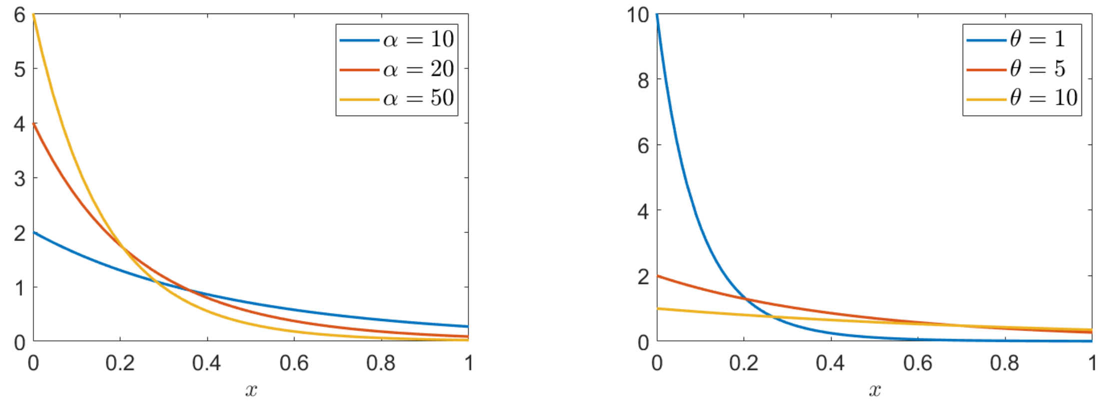
2. **Higher Moments:**

$$\mathbb{E}[X^k] = \begin{cases} \dfrac{\theta^k k!}{(\alpha-1)\cdots(\alpha-k)}, & \text{if } \alpha > k; \\ \infty, & \text{otherwise.} \end{cases}$$

3. **CDF:**

$$F_X(x) = 1 - \left(\frac{\theta}{x+\theta}\right)^{\alpha}, \quad S_X(x) = \left(\frac{\theta}{x+\theta}\right)^{\alpha}, \ x > 0.$$

- Unlike the former distributions (Exponential, Gamma, and Weibull) whose pdfs decay exponentially, the pdf of a Pareto distribution only has a *power rate of decay*. Hence, it has a much heavier tail, and is often used to model loss with large or catastrophic losses.
- The shape parameter $\alpha > 0$ controls the rate of decay of the pdf. The smaller the value of α, the lower the decay rate and thus the *heavier* the tail it has. The higher the value of α, the higher the pdf starts off around $x = 0$; see Figure 4a.
- Given a fixed α, the smaller the scale parameter $\theta > 0$, the higher the pdf starts off around $x = 0$; see Figure 4b.
- For $\mathbb{E}[X^k]$ to exist, the pdf should decay fast enough ($\alpha > k$). Since the k-th moment exists only for $k < \alpha$, the mgf of a Pareto distribution *does not exist*.



(a) $\alpha = 10, 20, 30$; $\theta = 5$ (b) $\theta = 1, 5, 10$; $\alpha = 10$

Figure 4: The pdfs of Pareto distributions with different α and θ

1.4.2 Single-Parameter Formulation

If $Y \sim \text{Pareto}(\alpha, \theta)$, the random variable $X := Y + \theta$ follows a single-parameter Pareto distribution, whose pdf is given by

$$f_X(x) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}}, \; x > \theta.$$

$\theta > 0$ determines the support of X, and is not considered as a parameter of the distribution.

Some important distributional quantities of the single-parameter Pareto distribution:

1. **Mean and Variance:**

$$\mathbb{E}[X] = \begin{cases} \frac{\alpha\theta}{\alpha - 1}, & \text{if } \alpha > 1; \\ \infty, & \text{otherwise,} \end{cases} \quad , \quad \text{Var}[X] = \begin{cases} \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}, & \text{if } \alpha > 2; \\ \infty, & \text{otherwise.} \end{cases}$$

2. **Higher Moments:**

$$\mathbb{E}[X^k] = \begin{cases} \frac{\alpha\theta^k}{\alpha - k}, & \text{if } \alpha > k; \\ \infty, & \text{otherwise.} \end{cases}$$

3. **CDF:**

$$F_X(x) = 1 - \left(\frac{\theta}{x}\right)^\alpha, \quad S_X(x) = \left(\frac{\theta}{x}\right)^\alpha, \; x > \theta.$$

1.5 Beta Distribution

Beta distribution has a finite support. X follows a beta distribution with shape parameters $\alpha, \beta > 0$, denoted by $X \sim \text{Beta}(\alpha, \beta)$, if it has the following pdf:

$$f_X(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1}(1-x)^{\beta-1}, \; 0 < x < 1,$$

where $B(\cdot, \cdot)$ is the beta function defined by

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1} dt = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \; \alpha, \beta > 0.$$

- The mean and variance of $X \sim \text{Beta}(\alpha, \theta)$ is given by :

$$\mathbb{E}[X] = \frac{\alpha}{\alpha + \beta}, \quad \text{Var}[X] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

- When $\alpha = \beta = 1$, the distribution is reduced to the uniform $[0, 1]$ distribution.
- The beta distribution is the ***conjugate prior*** of the binomial, negative binomial, and geometric distributions. We will go through these distributions when we study frequency models.
- When $\alpha = \beta$, the pdf is symmetric. When $\alpha, \beta > 1$, the mode of the distribution increases with α, and decreases with β. Otherwise, the pdf explodes at $x = 0$ if $\alpha < 1$, and at $x = 1$ if $\beta < 1$.

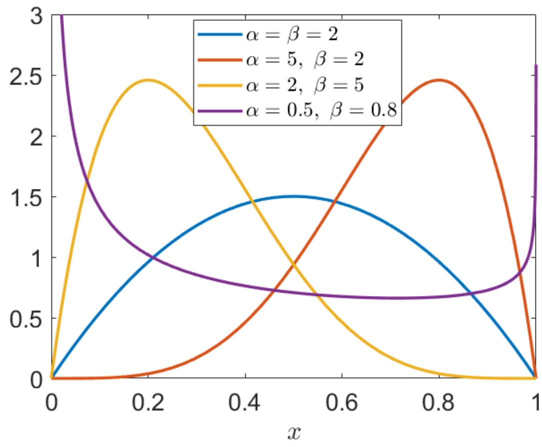


Figure 5: The pdfs of beta distributions with different α and β

2 Coverage Modifications

In most situations, an insurance policy does not fully cover the loss an insured suffered. Henceforth, to model the insurance payment accurately, relevant coverage modifications should be incorporated into the loss variable X. In this section, we shall introduce a couple of commonly seen coverage modifications – deductibles, policy limit, and coinsurance.

2.1 Deductibles

An ordinary deductible d is the amount that the policyholder agrees to pay before the insurer starts to cover the loss. If $X > d$, the insurer will pay any amount exceeding d, i.e., $X - d$; otherwise, the policyholder pays the loss and the insurer does not have to pay anything.

2.1.1 Stop Loss Variable

The *stop loss variable* models the payment by the insurer under a deductible d, which is defined below:

Definition 2.1 For a loss variable X, the ***stop loss variable*** (a.k.a. ***payment per loss variable***) with deductible d is defined as

$$\boxed{Y^L := (X-d)_+ = \begin{cases} X-d, & \text{if } X > d; \\ 0, & \text{if } X \le d, \end{cases}}$$

where $\boxed{x_+ := \max\{x, 0\}}$ for any $x \in \mathbb{R}$.

The expected value of $Y^L = (X-d)_+$ is known as the ***expected stop loss*** or ***expected payment per loss***, which can be computed as follows:

$$\boxed{\mathbb{E}[(X-d)_+] = \begin{cases} \displaystyle\sum_{x>d}(x-d)p_X(x), & \text{if } X \text{ is discrete;} \\ \displaystyle\int_d^\infty (x-d)f_X(x)dx, & \text{if } X \text{ is continuous.} \end{cases}} \tag{1}$$

Higher moments of the stop loss variable can be computed by replacing $(x-d)$ with $(x-d)^k$ in (1).

The survival function and the cdf of the variable Y^L are given by the following:

Proposition 2.1 For the stop loss variable Y^L of loss X and deductible d, we have

$$\boxed{S_{Y^L}(y) = \begin{cases} 1, & \text{if } y < 0; \\ S_X(y+d), & \text{if } y \ge 0 \end{cases}, \quad F_{Y^L}(y) = \begin{cases} 0, & \text{if } y < 0; \\ F_X(y+d), & \text{if } y \ge 0. \end{cases}} \tag{2}$$

Proof. We only derive $S_{Y^L}(y)$, as $F_{Y^L}(y) = 1 - S_{Y^L}(y)$.

1. If $y < 0$, $S_{Y^L}(y) = \mathbb{P}(Y^L > y) = 1$ since $Y^L \ge 0$.
2. If $y \ge 0$, $Y^L > y \iff (X-d)_+ > y \iff X > y+d$. Hence,
$$S_{Y^L}(y) = \mathbb{P}(X > y+d) = F_X(y+d).$$

□

Using Equation (2), we can derive an alternative expression of the expected stop loss:

Proposition 2.2 Let X be an integrable random variable. For any $d \in \mathbb{R}$,

$$\boxed{\mathbb{E}[(X-d)_+] = \int_d^\infty S_X(x)dx.}$$

Proof. Since $Y^L = (X-d)_+ \geq 0$, using Equation (4) of Chapter 1 and Equation (2), we have

$$\mathbb{E}[(X-d)_+] = \mathbb{E}[Y^L] = \int_0^\infty S_{Y^L}(y)dy = \int_0^\infty S_X(y+d)dy = \int_d^\infty S_X(x)dx$$

where the last line follows by the change of variable $x = y + d$. □

Example 2.1 Calculate the expected stop loss with a deductible d, if the loss variable X has the following distribution:

(a) $X \sim \text{Exp}(\theta)$;

(b) $X \sim \text{Pareto}(\alpha, \theta)$ with $\alpha > 1$.

<u>*Solution:*</u>

(a) Using the survival function of $X \sim \text{Exp}(\theta)$,

$$\mathbb{E}[(X-d)_+] = \int_d^\infty S_X(x)dx = \int_d^\infty e^{-\frac{x}{\theta}}dx = \theta e^{-\frac{d}{\theta}}.$$

(b) Using the survival function of $X \sim \text{Pareto}(\alpha, \theta)$,

$$\mathbb{E}[(X-d)_+] = \int_d^\infty S_X(x)dx = \int_d^\infty \left(\frac{\theta}{x+\theta}\right)^\alpha dx = \frac{\theta^\alpha}{(\alpha-1)(d+\theta)^{\alpha-1}}.$$

2.1.2 Excess Loss/Payment Per Payment Variable

The *excess loss variable* is another random variable that evaluates the loss of an insurance claim with a deductible d.

Definition 2.2 Let X be a loss variable. Given a fixed number $d > 0$ such that $\mathbb{P}(X > d) > 0$, the ***excess loss variable*** (a.k.a. ***payment per payment variable***) is defined by

$$\boxed{Y^P := X - d | X > d.}$$

Remark 2.3. If $\mathbb{P}(X > d) = 0$, the excess loss variable is not undefined.

The excess loss variable models the insurer's payment in excess of a *deductible* d, given that $X > d$. The difference between Y^P and the stop loss variable Y^L is as follows:

- $Y^L = (X-d)_+$ evaluates the payment the insurance company needs to pay for each policy/loss, including those do not exceed the deductible, thus the name *payment per loss*. Y^L is ***left-censored***

- For $Y^P = X - d|X > d$, only claims that exceed the deductible are reported and evaluated, and thus the name *payment per payment*. Y^P is ***left-truncated***.

- *Left-censored data* vs *Left-truncated data*:
  - Left-censored: Claims with losses less than or equal to d are still recorded, but are recorded as 0 regardless of the actual loss amount;
  - Left-truncated: Claims with losses less than or equal to d are NOT recorded.

In life contingencies where we interpret X as the lifetime of an insured, the excess loss variable is called the *residual lifetime*, which is the remaining life expectancy beyond the age d. The expected value of the excess loss variable is called the *mean excess loss function*.

Definition 2.3 The expected value of the excess loss variable is called the ***mean excess loss*** (a.k.a. ***mean residual lifetime (MRL)***) is defined as

$$\boxed{e_X(d) := \mathbb{E}[X - d | X > d] = \frac{\mathbb{E}[(X-d)\mathbb{1}_{\{X>d\}}]}{\mathbb{P}(X > d)} = \frac{\mathbb{E}[(X-d)_+]}{S_X(d)}.}$$

Remark 2.4. $e_X(d)$ is always greater than or equal to the expected stop loss, $\mathbb{E}[(X-d)_+]$.

Depending on whether the distribution of X is discrete or continuous, we can compute $e_X(d)$ by the following:

$$\boxed{e_X(d) := \mathbb{E}[X - d | X > d] = \begin{cases} \dfrac{\sum_{x>d}(x-d)p_X(x)}{S_X(d)}, & \text{if } X \text{ is discrete;} \\ \dfrac{\int_d^\infty (x-d) f_X(x)dx}{S_X(d)}, & \text{if } X \text{ is continuous.} \end{cases}}$$

Alternatively, regardless of the support of X, discrete or continuous, we can compute $e_X(d)$ based on Proposition 2.2:

$$\boxed{e_X(d) = \frac{\mathbb{E}[(X-d)_+]}{S_X(d)} = \frac{\int_d^\infty S_X(x)dx}{S_X(d)}.}$$

Proposition 2.5 For the payment per payment variable Y^P of loss X and deductible d, we have

$$\boxed{S_{Y^P}(y) = \begin{cases} 1, & \text{if } y < 0; \\ \dfrac{S_X(y+d)}{S_X(d)}, & \text{if } y \geq 0 \end{cases}, \quad F_{Y^P}(y) = \begin{cases} 0, & \text{if } y < 0; \\ \dfrac{F_X(y+d) - F_X(d)}{1 - F_X(d)}, & \text{if } y \geq 0. \end{cases}} \tag{3}$$

Proof. Again, we only derive $S_{Y^P}(y)$. Notice that $Y^P = X - d | X > d = X - d | X - d > 0 = Y^L | Y^L > 0$, since $X - d$ implies $Y^L = X - d$. Using this, we have:

1. If $y < 0$, $S_{Y^P}(y) = \mathbb{P}(Y^L > y | Y^L > 0) = 1$.
2. If $y \geq 0$, $Y^L > y \iff (X-d)_+ > y \iff X > y + d$. Hence, using Equation (6),

$$S_{Y^P}(y) = \mathbb{P}(Y^L > y | Y^L > 0) = \frac{\mathbb{P}(Y^L > y, Y^L > 0)}{\mathbb{P}(Y^L > 0)}$$
$$= \frac{\mathbb{P}(Y^L > y)}{\mathbb{P}(Y^L > 0)} = \frac{S_{Y^L}(y)}{S_{Y^L}(0)} = \frac{S_X(y+d)}{S_X(d)}.$$

□

Example 2.2 Calculate the mean excess loss with a deductible d, if the loss variable X has the following distribution:
(a) $X \sim \text{Exp}(\theta)$;
(b) $X \sim \text{Pareto}(\alpha, \theta)$ with $\alpha > 1$.

Solution:
(a) Using Example 2.1, we can compute the mean excess loss by

$$e_X(d) = \frac{\mathbb{E}[(X-d)_+]}{S_X(d)} = \frac{\theta e^{-\frac{d}{\theta}}}{e^{-\frac{d}{\theta}}} = \theta.$$

(b) Similarly, the mean excess loss for $X \sim \text{Pareto}(\alpha, \theta)$ is given by

$$e_X(d) = \frac{\mathbb{E}[(X-d)_+]}{S_X(d)} = \frac{\frac{\theta^\alpha}{(\alpha-1)(d+\theta)^{\alpha-1}}}{\left(\frac{\theta}{d+\theta}\right)^\alpha} = \frac{d+\theta}{\alpha-1}.$$

2.1.3 Franchise Deductible

When a policy has a *franchise deductible* of amount $d > 0$, the policyholder has to pay the loss if $X \leq d$; otherwise, if $X > d$, the insurer will cover the entire loss X (instead of $X - d$ for ordinary deductibles). We can define the *payment per loss* and *payment per payment* variables as follows.

Definition 2.4 Let d be the amount of a franchise deductible. The ***payment per loss variable*** is defined as

$$Y^L := \begin{cases} 0, & \text{if } X \leq d; \\ X, & \text{if } X > d \end{cases} = X\mathbb{1}_{\{X>d\}} = (X-d)_+ + d\mathbb{1}_{\{X>d\}}.$$

The ***payment per payment*** **variable** is defined as

$$\boxed{Y^P := X|X > d.}$$

The expected payment per loss, and the expected payment per payment with a franchise deductible can be computed by

$$\boxed{\begin{aligned} \mathbb{E}[Y^L] &= \begin{cases} \displaystyle\sum_{x>d} x p_X(x), & \text{if } X \text{ is discrete;} \\ \displaystyle\int_d^\infty x f_X(x) dx, & \text{if } X \text{ is continuous,} \end{cases} \\ \mathbb{E}[Y^P] &= \frac{\mathbb{E}[Y^L]}{S_X(d)}. \end{aligned}}$$

Alternatively, one can relate the expected payments of franchise deductible with an ordinary deductible as follows:

$$\boxed{\begin{aligned} \mathbb{E}[Y^L] &= \mathbb{E}[(X-d)_+] + dS_X(d), \\ \mathbb{E}[Y^P] &= \frac{\mathbb{E}[(X-d)_+] + dS_X(d)}{S_X(d)} = e_X(d) + d, \end{aligned}}$$

where $e_X(d)$ is the mean excess loss with an ordinary deductible d.

We can also compute the cdf and the survival function of the payment per loss variable under a franchise deductible:

Proposition 2.6 For the payment per loss variable Y^L of loss X and deductible d, we have

$$\boxed{S_{Y^L}(y) = \begin{cases} 1, & \text{if } y < 0; \\ S_X(d), & \text{if } 0 \le y < d; \\ S_X(y), & \text{if } y \ge d \end{cases}, \quad F_{Y^L}(y) = \begin{cases} 0, & \text{if } y < 0; \\ F_X(d), & \text{if } 0 \le y < d; \\ F_X(y), & \text{if } y \ge d. \end{cases}}$$

Proof. We derive the cdf of Y^L.

1. If $y < 0$, $F_{Y^L}(y) = 0$ since $Y^L \ge 0$.
2. If $0 \le y < d$, $Y^L \le y \iff X \le d$. Hence, $F_{Y^L}(y) = \mathbb{P}(X \le d) = F_X(d)$.
3. If $y > d$, $Y^L > y \iff X > y$. Hence, $S_{Y^L}(y) = \mathbb{P}(X > y) = S_X(y)$, and thus $F_{Y^L}(y) = F_X(y)$.

□

Example 2.3 A loss variable follows a continuous distribution with the following pdf:

$$f_X(x) = \frac{4(100-x)^3}{100^4}, \; 0 < x < 100.$$

Find the expected payment per loss, and the expected payment per payment with a franchise deductible $d = 20$.

Solution: The expected payment per loss is given by

$$\begin{aligned}
\mathbb{E}[Y^L] &= \mathbb{E}[X\mathbb{1}_{\{X>20\}}] \\
&= \int_{20}^{100} \frac{4x(100-x)^3}{100^4}dx \\
&= \int_{20}^{100} \frac{-4(100-x)^4 + 400(100-x)^3}{100^4}dx \\
&= \frac{1}{100^4}\left[\frac{4}{5}(100-x)^5 - 100(100-x)^4\right]_{20}^{100} \\
&= 36(0.8)^4 = 14.7456.
\end{aligned}$$

To calculate the expected payment per payment, we consider

$$S_X(20) = \int_{20}^{100} \frac{4(100-x)^3}{100^4}dx = -\frac{(100-x)^4}{100^4}\Bigg|_{20}^{100} = 0.8^4.$$

Therefore, the expected payment per payment is given by

$$\mathbb{E}[Y^P] = \frac{\mathbb{E}[Y^L]}{S_X(20)} = \frac{36(0.8)^4}{0.8^4} = 36.$$

2.2 Policy Limits

When a policy limit of amount u is imposed, the insurer is only responsible to cover the loss up to u, and any exceeding amount $X - u$ would be retained by the policyholder. The *limited loss variable* defined below models the payment the insurer needs to make under a policy limit u:

Definition 2.5 The ***limited loss variable*** with a policy limit u is defined by

$$\boxed{Y := X \wedge u = \min\{X, u\},}$$

where $\boxed{a \wedge b := \min\{a, b\}}$ for any $a, b \in \mathbb{R}$.

The limited loss variable is ***right-censored***: losses above u are recorded, but are recorded as u. Depending on the support of the loss variable X, we can compute the k-th moment of the limited loss variable as follows.

$$\mathbb{E}[(X \wedge u)^k] = \begin{cases} \displaystyle\int_{-\infty}^{u} x^k f_X(x)dx + u^k S_X(u), & \text{if } X \text{ is continuous;} \\ \displaystyle\sum_{x \leq u} x^k p_X(x) + u^k S_X(u), & \text{if } X \text{ is discrete.} \end{cases} \tag{4}$$

Proof. If X is continuous,

$$\begin{aligned} \mathbb{E}[(X \wedge u)^k] &= \int_{-\infty}^{\infty} (x \wedge u)^k f_X(x)dx \\ &= \int_{-\infty}^{u} x^k f_X(x)dx + \int_{u}^{\infty} u f_X(x)dx \\ &= \int_{-\infty}^{u} x^k f_X(x)dx + uS_X(u). \end{aligned}$$

Similarly, if X is discrete,

$$\begin{aligned} \mathbb{E}[(X \wedge u)^k] &= \sum_{x} (x \wedge u)^k p_X(x) \\ &= \sum_{x \leq u} x^k p_X(x) + \sum_{x > u} u^k p_X(x) \\ &= \sum_{x \leq u} x^k p_X(x) + u^k S_X(u). \end{aligned}$$

□

Example 2.4 Let $X \sim \text{Pareto}(\alpha, \theta)$, where $\alpha \neq 1$, compute $\mathbb{E}[X \wedge u]$.
<u>*Solution:*</u>
By Proposition 2.7, we have

$$\begin{aligned} \mathbb{E}[X \wedge u] &= \int_{0}^{u} x f_X(x)dx + uS_x(u) \\ &= \int_{0}^{u} x \left(\frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}} \right) dx + u \left(\frac{\theta}{u+\theta} \right)^\alpha \\ &= \alpha\theta^\alpha \int_{0}^{u} \frac{x \ dx}{(x+\theta)^{\alpha+1}} + u \left(\frac{\theta}{u+\theta} \right)^\alpha \end{aligned}$$

$$
\begin{aligned}
&= \alpha\theta^\alpha \int_0^u \left(\frac{1}{(x+\theta)^\alpha} - \frac{\theta}{(x+\theta)^{\alpha+1}} \right) dx + u \left(\frac{\theta}{u+\theta} \right)^\alpha \\
&= \alpha\theta^\alpha \left(\frac{1}{\alpha-1} \left(\frac{1}{\theta^{\alpha-1}} - \frac{1}{(u+\theta)^{\alpha-1}} \right) - \frac{\theta}{\alpha} \left(\frac{1}{\theta^\alpha} - \frac{1}{(u+\theta)^\alpha} \right) \right) + u \left(\frac{\theta}{u+\theta} \right)^\alpha \\
&= \frac{\theta}{\alpha-1} - \frac{\theta^\alpha(\alpha u+\theta)}{(\alpha-1)(u+\theta)^\alpha} + u \left(\frac{\theta}{u+\theta} \right)^\alpha \\
&= \frac{\theta}{\alpha-1} \left(1 - \left(\frac{\theta}{u+\theta} \right)^{\alpha-1} \right).
\end{aligned}
$$

The stop loss variable and the limited loss variable are related by the following:

$$
\boxed{(X-d)_+ + X \wedge d = X.} \tag{5}
$$

Equation (5) can be shown as follows: if $X < d$, we have $(X-d)_+ = 0$ and $X \wedge d = X$, and thus $(X-d)_+ + X \wedge d = X$; if $X \geq d$, $(X-d)_+ = X - d$ and $X \wedge d = d$, which again sums to X. Intutively, Equation (5) means that, a policy with a deductible d, along with a policy with a policy limit d, renders a full-coverage policy.

Proposition 2.7 If X is a non-negative random variable. Then, for any $u \geq 0$,

$$
\boxed{\mathbb{E}[X \wedge u] = \int_0^u S_X(x)dx.}
$$

Proof. Using Equation (5), Proposition 2.2, and Equation (4) of Chapter 1, we have

$$
\begin{aligned}
\mathbb{E}[X \wedge u] &= \mathbb{E}[X] - \mathbb{E}[(X-u)_+] \\
&= \int_0^\infty S_X(x)dx - \int_u^\infty S_X(x)dx = \int_0^u S_X(x)dx.
\end{aligned}
$$

□

We can also find the cdf and the survival function of Y:

Proposition 2.8 For the limited loss variable Y for the loss X with policy limit u, we have

$$
\boxed{S_Y(y) = \begin{cases} S_X(y), & \text{if } y < u; \\ 0, & \text{if } y \geq u \end{cases}, \quad F_Y(y) = \begin{cases} F_X(y), & \text{if } y < u; \\ 1, & \text{if } y \geq u. \end{cases}} \tag{6}
$$

Proof. We derive the cdf of Y:

1. If $y < u$, we have $Y \le y \iff X \wedge u \le y \iff X \le y$, which implies $F_Y(y) = \mathbb{P}(X \le y) = F_X(y)$.
2. If $y \ge u$, since $Y = X \wedge u \le u < y$, we have $F_Y(y) = 1$.

□

Using Equation (6), we also have the following formula to compute higher moments of the limited loss variable:

Proposition 2.9 If X is a non-negative random variable. Then, for any $u \ge 0$,

$$\boxed{\mathbb{E}[(X \wedge u)^k] = \int_0^u kx^{k-1} S_X(x)dx.}$$

Proof. By Theorem 4.1 of Chapter 1, and also Equation (6), we have

$$\begin{aligned}
\mathbb{E}[(X \wedge u)^k] &= \int_0^\infty ky^{k-1} S_Y(y)dy \\
&= \int_0^u ky^{k-1} S_X(y)dy + \int_0^u ky^{k-1} \times 0dy \\
&= \int_0^u ky^{k-1} S_X(y)dy.
\end{aligned}$$

□

2.3 Loss Elimination Ratio

For an insurance policy with a deductible d, the *loss elimination ratio* computes proportion of the expected loss that is not covered by the insurer:

Definition 2.6 The ***loss elimination ratio (LER)*** of a policy with loss variable X and (ordinary) deductible d is defined as

$$\boxed{\mathrm{LER}_X(d) := \frac{\mathbb{E}[X] - \mathbb{E}[(X-d)_+]}{\mathbb{E}[X]} = \frac{\mathbb{E}[X \wedge d]}{\mathbb{E}[X]}.}$$

Example 2.5 Let $X \sim \text{Pareto}(\alpha, \theta)$, where $\alpha > 1$. Compute $\mathrm{LER}_X(d)$.

Solution:
From Example 2.4, we have shown that

$$\mathbb{E}[X \wedge d] = \frac{\theta}{\alpha - 1}\left(1 - \left(\frac{\theta}{d+\theta}\right)^{\alpha-1}\right).$$

Since $\mathbb{E}[X] = \theta/(\alpha - 1)$, we have

$$\mathrm{LER}_X(d) = \frac{\mathbb{E}[X \wedge d]}{\mathbb{E}[X]} = \frac{\frac{\theta}{\alpha-1}\left(1 - \left(\frac{\theta}{d+\theta}\right)^{\alpha-1}\right)}{\frac{\theta}{\alpha-1}} = 1 - \left(\frac{\theta}{d+\theta}\right)^{\alpha-1}.$$

Example 2.6 (SOA EXAM STAM SAMPLE Q87 MODIFIED) Suppose that a loss variable X has a pdf given by

$$f_X(x) = \begin{cases} 0.01, & \text{if } 0 < x \leq 80; \\ 0.01\left(3 - \frac{x}{40}\right), & \text{if } 80 < x \leq 120; \\ 0, & \text{otherwise.} \end{cases}$$

Calculate the loss elimination ratio for a deductible of 20.

Solution:
We first compute $\mathbb{E}[X]$:

$$\begin{aligned} \mathbb{E}[X] &= \int_0^{120} x f_X(x) dx \\ &= \int_0^{80} 0.01x dx + \int_{80}^{120} 0.01x \left(3 - \frac{x}{40}\right) dx \\ &= 32 + \frac{56}{3} = \frac{152}{3}. \end{aligned}$$

Next, we compute $S_X(20)$ and $\mathbb{E}[X \wedge 20]$:

$$S_X(20) = 1 - F_X(20) = 1 - \int_0^{20} f_X(x)dx = 1 - \int_0^{20} 0.01dx = 0.8,$$

and

$$\begin{aligned}\mathbb{E}[X \wedge 20] &= \int_0^{20} x f_X(x)dx + 20S_X(20) \\ &= \int_0^{20} 0.01x dx + 20(0.8) \\ &= 2 + 16 = 18.\end{aligned}$$

Therefore, the LER is given by

$$\text{LER}_X(20) = \frac{\mathbb{E}[X \wedge 20]}{\mathbb{E}[X]} = \frac{18}{152/3} = 0.3553.$$

2.4 Coinsurance and Inflation

An insurance policy with a *coinsurance factor* $\alpha \in (0, 1]$ means the insurance company is paying α of the costs after accounting for deductibles and policy limit, if any.

Definition 2.7 The *payment per loss variable* Y^L for a loss X, with a deductible d, a ***maximum covered loss*** $u > d$, and a coinsurance factor $\alpha \in (0, 1]$, is defined as

$$\begin{aligned}Y^L &:= \begin{cases} 0, & \text{if } X < d; \\ \alpha(X-d), & \text{if } d \le X < u; \\ \alpha(u-d), & \text{if } X \ge u \end{cases} \\ &= \alpha\left(X \wedge u - X \wedge d\right).\end{aligned}$$

Remark 2.10. The value u is said to be the ***maximum covered loss***. This means that any loss exceeding u will not be covered. The ***maximum payment*** made is $\alpha(u - d)$.

Similar to the derivation of Equation (4), the k-th moment of Y^L can be computed as follows.

$$\mathbb{E}[(Y^L)^k] = \begin{cases} \displaystyle\int_d^u \alpha^k (x-d)^k f_X(x)dx + \alpha^k (u-d)^k S_X(u), & \text{if } X \text{ is continuous;} \\ \displaystyle\sum_{d<x\le u} \alpha^k (x-d)^k p_X(x) + \alpha^k (u-d)^k S_X(u), & \text{if } X \text{ is discrete.} \end{cases}$$

Sometimes we may want to multiply the loss variable by a growth factor $1 + r$, where $r > 0$. Indeed, X often represents the loss distribution at the current time point (e.g. at policy issuance). The actual loss in the future can be higher due to inflation, despite the shape of the loss distribution remains unchanged. Notice that, when adjusting for inflation for X, the deductible d and the maximum covered loss u remain unchanged.

Definition 2.8 The *payment per loss variable* Y^L for a loss X, with a deductible d, a maximum covered loss u, a coinsurance factor α, and an inflation rate r, is defined as

$$\begin{aligned} Y^L &:= \begin{cases} 0, & \text{if } (1+r)X < d; \\ \alpha((1+r)X - d), & \text{if } d \leq (1+r)X < u; \\ \alpha(u-d), & \text{if } (1+r)X \geq u \end{cases} \\ &= \alpha(1+r)\left(X \wedge \frac{u}{1+r} - X \wedge \frac{d}{1+r}\right). \end{aligned}$$

Example 2.7 The loss from an insurance policy for the year 2023 follows a Pareto distribution with shape parameter 3 and scale parameter 150. The insurance policy pays the loss above an ordinary deductible of 40, a maximum covered loss of 200, and a coinsurance factor of 90%. The loss size is expected to be 5% larger in 2024, but the insurance in 2024 has the same deductible, maximum covered loss, and coinsurance factor as in 2023. Find the percentage increase in the expected payment per loss from 2023 to 2024.

Solution:
In 2023, the loss distribution is $X \sim \text{Pareto}(3, 150)$, and the payment per loss variable is

$$Y^L_{2023} = 0.9\left(X \wedge 200 - X \wedge 40\right).$$

From Example 2.4, we know that

$$\mathbb{E}[X \wedge u] = \frac{150}{3-1}\left(1 - \left(\frac{150}{u+150}\right)^{3-1}\right) = 75\left(1 - \left(\frac{150}{u+150}\right)^2\right).$$

Hence,

$$\begin{aligned} \mathbb{E}[Y^L_{2023}] &= 0.9\left(\mathbb{E}[X \wedge 200] - \mathbb{E}[X \wedge 40]\right) \\ &= 0.9\left(75\left(1 - \left(\frac{150}{200+150}\right)^2\right) - 75\left(1 - \left(\frac{150}{40+150}\right)^2\right)\right) = 29.6727. \end{aligned}$$

In 2024, the payment per loss variable is

$$Y^L_{2024} = 0.9 \times 1.05 \times \left(X \wedge \frac{200}{1.05} - X \wedge \frac{40}{1.05}\right).$$

Hence,

$$\begin{aligned}
\mathbb{E}[Y^L_{2024}] &= 0.9 \times 1.05 \times \left(\mathbb{E}\left[X \wedge \frac{200}{1.05}\right] - \mathbb{E}\left[X \wedge \frac{40}{1.05}\right]\right) \\
&= 0.945 \left(75 \left(1 - \left(\frac{150}{200/1.05 + 150} \right)^2 \right) - 75 \left(1 - \left(\frac{150}{40/1.05 + 150} \right)^2 \right) \right) \\
&= 31.3171.
\end{aligned}$$

The percentage increase is thus

$$\frac{\mathbb{E}[Y^L_{2024}]}{\mathbb{E}[Y^L_{2023}]} - 1 = \frac{31.3171}{29.6727} - 1 = 5.5417\%.$$

3 Tails of Distributions

In Section 1, we often discussed about the *tail of a distribution*, which roughly means the weight the distribution puts on large values. A *heavy-tailed distribution*, e.g., the Pareto distribution, has a pdf which decays slowly with a considerable weight being put on large values. These distributions can be used to model catastrophic losses. In contrast, the pdf of a *light-tailed distribution* decays relatively fast. In this section, we provide different ways to compare tails of different distributions.

3.1 Comparison based on Moments

Recall that the k-th raw moment of a random variable X is defined as

$$\mu'_k = \mathbb{E}[X^k] = \int_{-\infty}^{\infty} x^k f_X(x) dx.$$

The integral exists if $x^k f_X(x)$ decays fast enough as $x \to \pm\infty$. We thus have the following classification:

μ'_k exists for all $k \in \mathbb{N} \Rightarrow$ a (relatively) light-tailed distribution
μ'_k exists only up to $k \leq N$ for some $N \Rightarrow$ a (relatively) heavy-tailed distribution

Example 3.1 The gamma, exponential, and the Weibull distributions have moments of all orders. They thus have lighter tails than the Pareto distribution (Pareto(α, θ)), where the k-th moment does not exist if $k > \alpha$.

3.2 Comparison based on Survival Functions

In Example 3.1, we see that all gamma, exponential, and the Weibull distributions have moments of all orders. If we want to compare the tails of distributions, solely based on moments would not be sufficient. To give more sophisticated comparisons of tails, we can make use of the survival functions.

Let X and Y be random variables with survival function S_X and S_Y, respectively. Suppose that the following limit exists:

$$\lim_{t\to\infty} \frac{S_X(t)}{S_Y(t)} = c \in [0, \infty].$$

We can make the following comparison:

- $c = 0 \Rightarrow S_Y$ decays much slower than S_X as $t \to \infty \Rightarrow Y$ has a *heavier tail* than X;
- $c = \infty \Rightarrow S_Y$ decays much faster than S_X as $t \to \infty \Rightarrow Y$ has a *lighter tail* than X;
- $c \in (0, \infty) \Rightarrow S_Y$ and S_X decays at similar rate as $t \to \infty \Rightarrow X$ and Y have *similar tails*.

By L'Hôpital's rule, the limit can also be computed as follows:

$$\boxed{c = \lim_{t\to\infty} \frac{S_X(t)}{S_Y(t)} = \lim_{t\to\infty} \frac{S'_X(t)}{S'_Y(t)} = \lim_{t\to\infty} \frac{f_X(t)}{f_Y(t)}.} \tag{7}$$

Example 3.2 Let $X \sim \text{Exp}(\theta)$, $Y \sim \text{Gamma}(\alpha, \theta)$, and $Z \sim \text{Weibull}(\alpha, \theta)$. Assume that $\alpha \neq 1$, compare the tails of X, Y, and Z based on their survival functions or pdfs.

<u>*Solution:*</u>

Using (7), it suffices to consider the rations of their pdfs. We first compare the tails of X and Y:

$$\lim_{t\to\infty} \frac{f_X(t)}{f_Y(t)} \propto \lim_{t\to\infty} \frac{e^{-\frac{t}{\theta}}}{t^{\alpha-1}e^{-\frac{t}{\theta}}} = \lim_{t\to\infty} t^{1-\alpha} = \begin{cases} \infty, & \text{if } \alpha < 1; \\ 0, & \text{if } \alpha > 1. \end{cases}$$

Hence, X has a heavier tail than Y if $\alpha < 1$, a lighter tail if $\alpha > 1$. Next, we compare X and Z:

$$\lim_{t\to\infty} \frac{f_Z(t)}{f_X(t)} \propto \lim_{t\to\infty} \frac{t^{\alpha-1}e^{-\left(\frac{t}{\theta}\right)^\alpha}}{e^{-\frac{t}{\theta}}} = \lim_{t\to\infty} t^{\alpha-1}e^{-\left(\frac{t}{\theta}\right)^\alpha + \frac{t}{\theta}} = \begin{cases} \infty, & \text{if } \alpha < 1; \\ 0, & \text{if } \alpha > 1. \end{cases}$$

Hence, Z has a heavier tail than X if $\alpha < 1$, a lighter tail if $\alpha > 1$.

To conclude:

- If $\alpha < 1$, Z has the heaviest tail, X has the second heaviest tail, and Y has the lightest tail;
- If $\alpha > 1$, Y has the heaviest tail, X has the second heaviest tail, and Z has the lightest tail;

3.3 Comparison based on Hazard Rate Functions

The *hazard rate function* defined below provides another way to describe the tail of a distribution.

Definition 3.1 The ***hazard rate function*** of a continuous random variable X is defined as

$$\boxed{h_X(x) := -\frac{d}{dx}\ln(S_X(x)) = \frac{f_X(x)}{S_X(x)}.}$$

The hazard rate function can be interpreted as follows. For any $x \in \mathbb{R}$, and any small increment $dx > 0$,

$$h_X(x)dx = \frac{f_X(x)dx}{S_X(x)} \approx \frac{\mathbb{P}(x < X \le x + dx)}{\mathbb{P}(X > x)} = \mathbb{P}(x < X \le x + dx | X > x).$$

In other words, it measures the likelihood that the severity X will be around x, given that it is at least x. We have the following definition depending on the monotonicity of h_X.

Definition 3.2 The distribution of X is said to have a ***decreasing failure rate*** **(DFR)** if h_X is a non-increasing function. It is said to have a ***increasing failure rate*** **(IFR)** if h_X is a non-decreasing function.

We have the following interpretations:

- If X has a DFR, the probability that $X \in (x, x + dx]$ given that $X > x$ is decreasing. Hence, as x increases, it is more likely that $X > x + dx$ given $X > x$. This indicates that X has a heavy tail.
- If X has an IFR, the probability that $X \in (x, x + dx]$ given that $X > x$ is increasing. Hence, as x increases, it is less likely that $X > x + dx$ given $X > x$. This indicates that X has a light tail.

To conclude, we have:

$$\boxed{\begin{array}{l}\text{Distribution of } X \text{ has a DFR} \Rightarrow X \text{ has a heavy tail,}\\ \text{Distribution of } X \text{ has an IFR} \Rightarrow X \text{ has a light tail.}\end{array}}$$

Example 3.3 Based on the hazard rate functions, classify the tailedness of the exponential, the Weibull, and the Pareto distributions, and whether each of these distributions has a DFR or IFR.

Solution:
For $X \sim \text{Exp}(\theta)$,

$$h_X(x) = \frac{f_X(x)}{S_X(x)} = \frac{\frac{1}{\theta}e^{-\frac{x}{\theta}}}{e^{-\frac{x}{\theta}}} = \frac{1}{\theta},$$

which is a constant function. In this case, we can say that the exponential distribution has a *medium tail*.

For $X \sim \text{Weibull}(\alpha, \theta)$,

$$h_X(x) = \frac{f_X(x)}{S_X(x)} \propto \frac{x^{\alpha-1}e^{-\left(\frac{x}{\theta}\right)^\alpha}}{e^{-\left(\frac{x}{\theta}\right)^\alpha}} = x^{\alpha-1},$$

which is an increasing function if $\alpha > 1$, and a decreasing function if $\alpha < 1$. Hence, the Weibull distribution has an IFR, and a light tail if $\alpha > 1$; and it has a DFR, and a heavy tail if $\alpha < 1$.

For $X \sim \text{Pareto}(\alpha, \theta)$,

$$h_X(x) = \frac{f_X(x)}{S_X(x)} \propto \frac{\frac{1}{(x+\theta)^{\alpha+1}}}{\left(\frac{\theta}{x+\theta}\right)^\alpha} \propto \frac{1}{x+\theta},$$

which is a decreasing function. Hence, the Pareto distribution has a DFR, and is a heavy-tailed distribution.

3.4 Comparison based on Mean Excess Loss

The mean excess loss function e_X can be used to discuss the tailedness of a distribution in a way similar to using the hazard rate function h_X: whether $e_X(\cdot)$ is an increasing/decreasing function.

Definition 3.3 The distribution of the loss variable X is said to have an ***increasing mean residual lifetime (IMRL)*** if $e_X(d)$ is a non-decreasing function of d. It is said to have a ***decreasing mean residual lifetime (DMRL)*** if $e_X(d)$ is a non-increasing function of d.

Following the same deduction as using the hazard rate function, we can decide the tailedness of X as follows:

Distribution of X has an IMRL $\Rightarrow$ Residual loss is non-decreasing $\Rightarrow$ X has a heavy tail,
Distribution of X has an DMRL $\Rightarrow$ Residual loss is non-increasing $\Rightarrow$ X has a light tail.

Example 3.4 Discuss the tailedness of the $X \sim \text{Pareto}(\alpha, \theta)$ with $\alpha > 1$ using the mean excess loss function.

Solution:
From Example 2.2, we have shown that $e_X(d) = \frac{d+\theta}{\alpha-1}$, which is an increasing function of d. Hence, the distribution has an IMRL, and thus a heavy tail.

The following result states the relationship between DFR/IFR and IMRL/DMRL, whose proof is omitted.

Theorem 3.1 If the distribution of X has a DFR, then it has an IMRL. If the distribution of X has an IFR, then it has a DMRL.

Remark 3.2.

1. As a consequence of Theorem 3.1, if the distribution of X has a heavy (resp. light) tail based on hazard rate function, then it also has a heavy (resp. light) tail based on MRL.
2. The converse of Theorem 3.1 is not true: IMRL $\not\Rightarrow$ DFR, and DMRL $\not\Rightarrow$ IFR.

4 Transformations of Distributions

In addition to the simple parametric models introduced in Section 1, we can apply transformations to those distributions to generate more distributions to model severity.

4.1 Scaling

New distributions can be generated by multiplying the old one with a scalar $c > 0$. Given the cdf of X, F_X, the cdf of $Y := cX$ can be derived as follows:

$$\boxed{F_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}(cX \le y) = \mathbb{P}\left(X \le \frac{y}{c}\right) = F_X\left(\frac{y}{c}\right),}$$

where the support of Y is given by $\text{Supp}(Y) = \{y : y/c \in \text{Supp}(X)\}$. If X is continuous, we can also obtain the pdf of Y by differentiation:

$$\boxed{f_Y(y) = \frac{1}{c} f_X\left(\frac{y}{c}\right), \ y \in \text{Supp}(Y).}$$

Definition 4.1 A random variable X belongs to a ***scale family*** of distributions, if the distribution of cX also belongs to the same family, perhaps with a different parameter.

Example 4.1 Show that each of the following distributions form a scale family: exponential, gamma, Weibull, and the Pareto distribution.

Solution:

- If $X \sim \text{Exp}(\theta)$, then for $c > 0$, the cdf of $Y = cX$ is given by

$$F_Y(y) = F_X\left(\frac{y}{c}\right) = 1 - e^{-\frac{y/c}{\theta}} = 1 - e^{-\frac{y}{c\theta}},\ y > 0,$$

  whence $Y \sim \text{Exp}(c\theta)$.
- If $X \sim \text{Gamma}(\alpha, \theta)$, then for $c > 0$, the pdf of $Y = cX$ is given by

$$f_Y(y) \propto f_X\left(\frac{y}{c}\right) = \left(\frac{y}{c}\right)^{\alpha-1} e^{-\frac{y/c}{\theta}} \propto y^{\alpha-1}e^{-\frac{y}{c\theta}},\ y > 0,$$

  whence $Y \sim \text{Gamma}(\alpha, c\theta)$.
- Similarly, if $X \sim \text{Weibull}(\alpha, \theta)$, then $Y = cX \sim \text{Weibull}(\alpha, c\theta)$. If $X \sim \text{Pareto}(\alpha, \theta)$, $Y = cX \sim \text{Pareto}(\alpha, c\theta)$.

Remark 4.1. Form Example 4.1, we see that the scale parameters for those new distributions being considered are c times the original ones, while the shape parameters (if any) remain unchanged. This also suggests why we call θ the *scale* parameter.

4.2 One-to-One Transformations

In general, we can construct new distributions by applying an one-to-one function on the old one. Let $g : \mathbb{R} \to \mathbb{R}$ be a strictly increasing or decreasing function, and X be a random variable with known distribution. Then, we can derive the cdf of $Y = g(X)$ as follows:

- If g is strictly increasing,

$$\boxed{F_Y(y) = \mathbb{P}(g(X) \le y) = \mathbb{P}(X \le g^{-1}(y)) = F_X(g^{-1}(y)),}$$

  where $y \in \text{Supp}(Y) = \{y = g(x) : x \in \text{Supp}(X)\}$.

- If g is strictly decreasing,

$$\boxed{F_Y(y) = \mathbb{P}(g(X) \le y) = \mathbb{P}(X \ge g^{-1}(y)).}$$

  If X is continuous, we further have

$$\boxed{F_Y(y) = S_X(g^{-1}(y)).}$$

If g is differentiable and X is continuous, we can obtain the pdf of Y by differentiation:

$$f_Y(y) = \begin{cases} f_X\left(g^{-1}(y)\right) \dfrac{dg^{-1}(y)}{dy}, & \text{if } g \text{ is strictly increasing,} \\ -f_X\left(g^{-1}(y)\right) \dfrac{dg^{-1}(y)}{dy}, & \text{if } g \text{ is strictly decreasing} \end{cases}$$
$$= f_X\left(g^{-1}(y)\right) \left| \frac{dg^{-1}(y)}{dy} \right|.$$

Here, the term $|dg^{-1}(y)/dy|$ is called the ***Jacobian*** of the transformation.

Example 4.2 Let $X \sim \text{Exp}(\theta)$. Determine the distribution of $Y := \theta(e^X - 1)$.

<u>*Solution:*</u>
Let $g(x) := \theta(e^x - 1)$. Notice that g is strictly increasing, with $g(0) = 0$ and $g(x) \to \infty$ as $x \to \infty$. Hence, the support of Y is $(0, \infty)$. Next, we deduce the cdf of Y:

$$\begin{aligned} F_Y(y) &= \mathbb{P}(\theta(e^X - 1) \le y) \\ &= \mathbb{P}\left(X \le \ln\left(\frac{y}{\theta} + 1\right)\right) \\ &= 1 - \exp\left(-\frac{\ln\left(\frac{y}{\theta} + 1\right)}{\theta}\right) \\ &= 1 - \left(\frac{1}{\frac{y}{\theta} + 1}\right)^{\frac{1}{\theta}} \\ &= 1 - \left(\frac{\theta}{y + \theta}\right)^{\frac{1}{\theta}}, \ y > 0. \end{aligned}$$

In other words, $Y \sim \text{Pareto}(1/\theta, \theta)$.

Example 4.3 Let $X \sim \text{Exp}(\theta)$. Find the pdf of $Y := X^\tau$, where $\tau > 0$.

<u>*Solution:*</u>
Let $g(x) := x^\tau$, which is strictly increasing with $g(0) = 0$ and $g(x) \to \infty$ as $x \to \infty$. Hence, the support of Y is $(0, \infty)$. Notice that $g^{-1}(y) = y^{1/\tau}$, we have

$$\frac{dg^{-1}(y)}{dy} = \frac{1}{\tau} y^{\frac{1}{\tau} - 1}.$$

By the change of variable formula, we arrive at

$$\begin{aligned} f_Y(y) &= f_X(y^{1/\tau}) \left(\frac{1}{\tau} y^{\frac{1}{\tau}-1} \right) \\ &= \left(\frac{1}{\theta} e^{-\frac{y^{1/\tau}}{\theta}} \right) \left(\frac{1}{\tau} y^{\frac{1}{\tau}-1} \right) \\ &= \frac{1/\tau}{(\theta^\tau)^{1/\tau}} y^{\frac{1}{\tau}-1} \exp\left(-\left(\frac{y}{\theta^\tau} \right)^{\frac{1}{\tau}} \right), \; y > 0. \end{aligned}$$

In other words, $Y \sim \text{Weibull}(1/\tau, \theta^\tau)$.

Example 4.4 (Log-normal Distribution) Let X be a continuous random variable that follows a *normal distribution* with mean μ and variance σ^2, i.e., $X \sim \mathcal{N}(\mu, \sigma^2)$ with the following pdf:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \; x \in \mathbb{R}.$$

The random variable $Y := e^X$ is said to follow a ***log-normal distribution*** with parameter μ and σ^2, denoted by $Y \sim \text{Lognormal}(\mu, \sigma^2)$. Find the pdf of Y.

Solution:

Since X is supported in $\mathbb{R}$, the support of $Y = g(X) := e^X$ is $(0, \infty)$. By writing $x = g^{-1}(y) := \ln y$, the pdf of Y is given by

$$\boxed{\begin{aligned} f_Y(y) &= f_X(\ln y) \frac{d \ln y}{dy} \\ &= \frac{1}{y\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}, \; y > 0. \end{aligned}}$$

One can also show that

$$\boxed{\mathbb{E}[Y] = e^{\mu+\frac{\sigma^2}{2}} \quad \text{and} \quad \text{Var}[Y] = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2}.}$$

5 Mixing

Mixing is another way to create a new distribution by aggregating a finite set of cdfs/pdfs/pmfs. Depending on the method of aggregation, we can divide the method into *discrete mixing* and *continuous mixing*. Writing the unconditional distribution of a loss variable X in terms of the conditional ones is an example of mixing.

5.1 Discrete Mixing

Definition 5.1 X is said to follow a (countable) ***mixture distribution*** if its cdf can be written as

$$\boxed{F_X(x) = \sum_{i=1}^{n} w_i F_{X_i}(x),} \tag{8}$$

where for each $i = 1, 2 \ldots, n$, F_{X_i} is the cdf of the random variable X_i and $w_i \geq 0$, with $\sum_{i=1}^{n} w_i = 1$.

Remark 5.1. Equation (8) does NOT imply $X = \sum_{i=1}^{n} w_i X_i$.

Depending on the support of each X_i, we can express the pmf/pdf of X based on Equation (8):

1. If the support of each X_i is continuous, then the pdf of X is given by

$$f_X(x) = \sum_{i=1}^{n} w_i f_{X_i}(x).$$

2. If the support of each X_i is discrete, then the pmf of X is given by

$$p_X(x) = \sum_{i=1}^{n} w_i p_{X_i}(x).$$

In Section 1, we have learnt a couple of parametric distributions, where the distribution of X depends on some parameters, say λ. The parameter λ therein is assumed to be known in advance. In practice, the parameter needs to be calibrated, and can be treated as a random variable, Λ. Given $\Lambda = \lambda$, we are able to model the distribution of X using a parametric distribution, $X|\Lambda = \lambda$. If Λ follows a discrete distribution with pmf $p_\Lambda(\lambda)$, then by the law of total probability, the *unconditional cdf* of X can be written as follows:

$$\boxed{\underbrace{F_X(x)}_{\text{unconditional cdf}} = \sum_i \underbrace{F_{X|\Lambda}(x|\lambda_i)}_{\text{parametric distribution}} \overbrace{p_\Lambda(\lambda_i)}^{\text{mixing distribution}}.} \tag{9}$$

Equation (9) can be considered as a mixture distribution, where $X_i = X|\Lambda = \lambda_i$, and $w_i = p_\Lambda(\lambda_i)$. Depending on the support of $X|\Lambda = \lambda_i$, we can express the pdf/pmf of X as follows:

$$f_X(x) = \sum_i^n f_{X|\Lambda}(x|\lambda_i)p_\Lambda(\lambda_i), \text{ if } X, X|\Lambda \text{ are continuous,}$$
$$p_X(x) = \sum_i^n p_{X|\Lambda}(x|\lambda_i)p_\Lambda(\lambda_i), \text{ if } X, X|\Lambda \text{ are discrete.}$$

Example 5.1 The loss X for an insurance coverage follows a distribution which is a mixture of an exponential distribution with mean 10 with 80% weight, and an exponential distribution with mean 100 with 20% weight.

(a) Calculate the probability that the loss is greater than 20.

(b) Find the mean of the loss.

(c) If the loss is covered by an insurance with a deductible 20. Find the expected payment per loss.

<u>*Solution:*</u>

(a) The cdf of the loss is given by

$$F_X(x) = 0.8e^{-\frac{x}{10}} + 0.2e^{-\frac{x}{100}}.$$

The required probability is thus

$$\mathbb{P}(X > 20) = 1 - F_X(20) = 1 - \left(0.8e^{-\frac{20}{10}} + 0.2e^{-\frac{20}{100}}\right) = 0.72799.$$

(b) The unconditional mean of X is the weighted sum of the conditional mean:

$$\mathbb{E}[X] = 0.8 \times 10 + 0.2 \times 100 = 28.$$

Indeed, let Θ be the random variable of the mean parameter. We have $p_\Theta(10) = 0.8$ and $p_\Theta(100) = 0.2$. Notice that $\mathbb{E}[X|\Theta] = \Theta$. Hence, by the law of iterated expectation,

$$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|\Theta]] = \mathbb{E}[\Theta] = 0.8 \times 10 + 0.2 \times 100 = 28.$$

(c) Notice that with probability 0.8, X has the same distribution as $X_1 \sim \text{Exp}(10)$, and with probability 0.2, X has the same distribution as $X_2 \sim \text{Exp}(100)$. Hence,

$$\mathbb{E}[(X - 20)_+] = 0.8\mathbb{E}[(X_1 - 20)_+] + 0.2\mathbb{E}[(X_2 - 20)_+].$$

To find $\mathbb{E}[(X_1 - 20)_+]$, we can condition on the event $X_1 \le 20$ and $X_1 > 20$:

$$
\begin{aligned}
&\mathbb{E}[(X_1-20)_+] \\
&= \mathbb{E}[(X_1-20)_+|X_1 \le 20]\mathbb{P}(X_1 \le 20) + \mathbb{E}[(X_1-20)_+|X_1 > 20]\mathbb{P}(X_1 > 20) \\
&= \mathbb{P}(X_1 > 20)\mathbb{E}[X_1 - 20|X > 20] \\
&= e^{-\frac{20}{10}}(10) \\
&= 10e^{-2},
\end{aligned}
$$

where the second last equality follows from the memoryless property of the exponential distribution. Similarly, we can derive

$$\mathbb{E}[(X_2-20)_+] = \mathbb{P}(X_2 > 20)\mathbb{E}[X_2 - 20|X > 20] = 100e^{-0.2}.$$

Therefore,

$$\mathbb{E}[(X-20)_+] = 0.8(10e^{-2}) + 0.2(100e^{-0.2}) = 17.4573.$$

5.2 Continuous Mixing

When the *mixing distribution* of the parameter Λ is continuous, the unconditional distribution of X can be expressed as an integral. In this case, we say that X follows a ***continuous mixture distribution***, whose unconditional cdf is given by

$$\boxed{F_X(x) = \int F_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda) d\lambda,}$$

where f_Λ is the pdf of the mixing distribution Λ. Depending on the support of X and $X|\Lambda$, the unconditional pdf/pmf of X is given by:

$$\boxed{\begin{aligned} f_X(x) &= \int f_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda) d\lambda, \ \text{ if } X, X|\Lambda \text{ are continuous,} \\ p_X(x) &= \int p_{X|\Lambda}(x|\lambda) f_\Lambda(\lambda) d\lambda, \ \text{ if } X, X|\Lambda \text{ are discrete.} \end{aligned}}$$

Example 5.2 Suppose that $X|\Theta \sim \text{Gamma}(2, \Theta)$, where Θ follows a single-parameter Pareto distribution with shape parameter 1 and minimum value 10. Calculate $\mathbb{P}(X \le 15)$.
<u>*Solution:*</u>
The pdf of $X|\Theta = \theta$ is given by

$$f_X(x) = \frac{1}{\theta^2} x e^{-\frac{x}{\theta}}, \ x > 0.$$

The conditional cdf is thus

$$\begin{aligned}F_{X|\Theta}(x|\theta) &= \int_0^x \frac{1}{\theta^2} t e^{-\frac{t}{\theta}} dt \\ &= -\frac{1}{\theta} x e^{-\frac{x}{\theta}} + \frac{1}{\theta}\int_0^x e^{-\frac{t}{\theta}} dt \\ &= 1 - e^{-\frac{x}{\theta}} - \frac{1}{\theta} x e^{-\frac{x}{\theta}}.\end{aligned}$$

On the other hand, the pdf of Θ is given by

$$f_\Theta(\theta) = \frac{10}{\theta^2},\ \theta > 10.$$

Hence, the unconditional cdf of X is given by

$$\begin{aligned}F_X(x) &= \int_{10}^\infty F_{X|\Theta}(x|\theta) f_\Theta(\theta) d\theta \\ &= \int_{10}^\infty \left(1 - e^{-\frac{x}{\theta}} - \frac{1}{\theta} x e^{-\frac{x}{\theta}}\right)\left(\frac{10}{\theta^2}\right) d\theta \\ &= \int_{10}^\infty \frac{10}{\theta^2} d\theta - 10\int_{10}^\infty \frac{e^{-\frac{x}{\theta}}}{\theta^2} d\theta - 10x \int_{10}^\infty \frac{e^{-\frac{x}{\theta}}}{\theta^3} d\theta.\end{aligned}$$

Notice that

$$\begin{aligned}&\int_{10}^\infty \frac{10}{\theta^2} d\theta = 1, \\ &\int_{10}^\infty \frac{e^{-\frac{x}{\theta}}}{\theta^2} d\theta = \frac{1}{x}\int_{10}^\infty d(e^{-\frac{x}{\theta}}) = \frac{1 - e^{-\frac{x}{10}}}{x}, \\ x &\int_{10}^\infty \frac{e^{-\frac{x}{\theta}}}{\theta^3} d\theta = \int_{10}^\infty \frac{d(e^{-\frac{x}{\theta}})}{\theta} = \left.\frac{e^{-\frac{x}{\theta}}}{\theta}\right|_{10}^\infty + \int_{10}^\infty \frac{e^{-\frac{x}{\theta}}}{\theta^2} d\theta = -\frac{e^{-\frac{x}{10}}}{10} + \frac{1 - e^{-\frac{x}{10}}}{x}.\end{aligned}$$

Hence, the unconditional cdf is given by

$$F_X(x) = 1 + e^{-\frac{x}{10}} - \frac{20(1 - e^{-\frac{x}{10}})}{x}.$$

Finally,

$$\mathbb{P}(X \le 15) = F_X(15) = 0.1873.$$